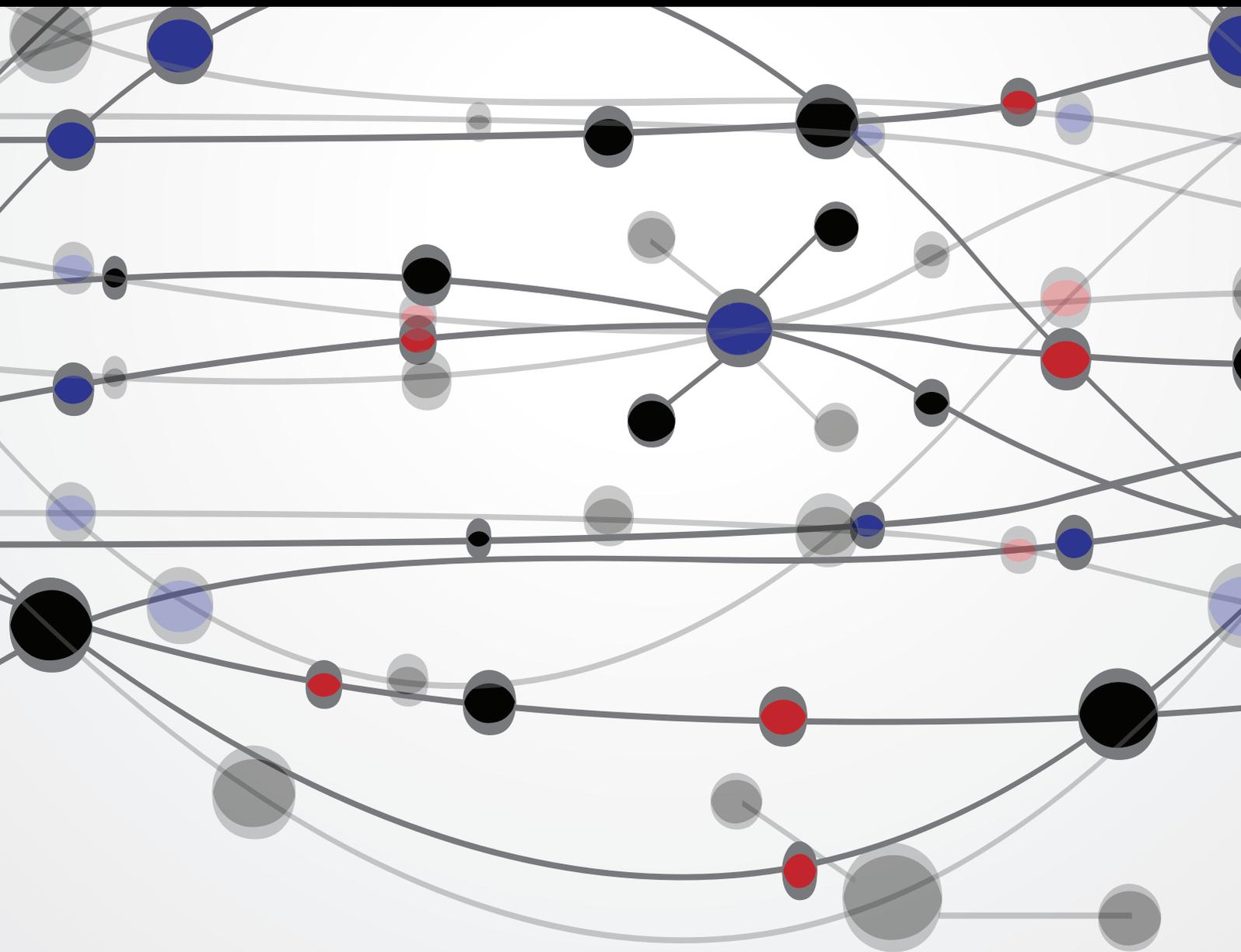


# Power, Control, and Optimization

Guest Editors: Pandian Vasant, Gerhard-Wilhelm Weber,  
Nader Barsoum, and Vo Ngoc Dieu





---

# **Power, Control, and Optimization**

The Scientific World Journal

---

## **Power, Control, and Optimization**

Guest Editors: Pandian Vasant, Gerhard-Wilhelm Weber,  
Nader Barsoum, and Vo Ngoc Dieu



---

Copyright © 2015 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in “The Scientific World Journal.” All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

# Contents

**Power, Control, and Optimization**, Pandian Vasant, Gerhard-Wilhelm Weber, Nader Barsoum, and Vo Ngoc Dieu  
Volume 2015, Article ID 964581, 2 pages

**Integer-Linear-Programming Optimization in Scalable Video Multicast with Adaptive Modulation and Coding in Wireless Networks**, Dongyul Lee and Chaewoo Lee  
Volume 2014, Article ID 769241, 7 pages

**Chaos Enhanced Differential Evolution in the Task of Evolutionary Control of Selected Set of Discrete Chaotic Systems**, Roman Senkerik, Ivan Zelinka, Michal Pluhacek, Donald Davendra, and Zuzana Oplatková Kominkova  
Volume 2014, Article ID 836484, 12 pages

**Energy Saving in Data Processing and Communication Systems**, Giuseppe Iazeolla and Alessandra Pieroni  
Volume 2014, Article ID 452863, 11 pages

**Power Quality Improvement by Unified Power Quality Conditioner Based on CSC Topology Using Synchronous Reference Frame Theory**, Rajasekaran Dharmalingam, Subhransu Sekhar Dash, Karthikrajan Senthilnathan, Arun Bhaskar Mayilvaganan, and Subramani Chinnamuthu  
Volume 2014, Article ID 391975, 7 pages

**Structural Optimization of a Knuckle with Consideration of Stiffness and Durability Requirements**, Geun-Yeon Kim, Seung-Ho Han, and Kwon-Hee Lee  
Volume 2014, Article ID 763692, 7 pages

**Mathematical Modelling of Thermal Process to Aquatic Environment with Different Hydrometeorological Conditions**, Alibek Issakhov  
Volume 2014, Article ID 678095, 10 pages

**Design of Heat Exchanger for Ericsson-Brayton Piston Engine**, Peter Durcansky, Stefan Papucik, Jozef Jandacka, Michal Holubcik, and Radovan Nosek  
Volume 2014, Article ID 138254, 6 pages

**Master-Slave Control Scheme in Electric Vehicle Smart Charging Infrastructure**, Ching-Yen Chung, Joshua Chynoweth, Chi-Cheng Chu, and Rajit Gadh  
Volume 2014, Article ID 462312, 14 pages

**Emission Controls Using Different Temperatures of Combustion Air**, Radovan Nosek, Michal Holubčík, and Štefan Papučík  
Volume 2014, Article ID 487549, 6 pages

**Heat Removal from Bipolar Transistor by Loop Heat Pipe with Nickel and Copper Porous Structures**, Patrik Nemeč, Martin Smitka, and Milan Malcho  
Volume 2014, Article ID 724740, 9 pages

**Structural Safety Analysis Based on Seismic Service Conditions for Butterfly Valves in a Nuclear Power Plant**, Sang-Uk Han, Dae-Gyun Ahn, Myeong-Gon Lee, Kwon-Hee Lee, and Seung-Ho Han  
Volume 2014, Article ID 743470, 9 pages

**Distributed Optimal Power and Rate Control in Wireless Sensor Networks**, Meiqin Tang, Jianyong Bai, Jing Li, and Yalin Xin  
Volume 2014, Article ID 580854, 8 pages

**Optimal Robust Motion Controller Design Using Multiobjective Genetic Algorithm**, Andrej Sarjaš, Rajko Svečko, and Amor Chowdhury  
Volume 2014, Article ID 978167, 15 pages

**Optimal Location through Distributed Algorithm to Avoid Energy Hole in Mobile Sink WSNs**, Li Qing-hua, Gui Wei-hua, and Chen Zhi-gang  
Volume 2014, Article ID 894018, 24 pages

**Global Optimization Ensemble Model for Classification Methods**, Hina Anwar, Usman Qamar, and Abdul Wahab Muzaffar Qureshi  
Volume 2014, Article ID 313164, 9 pages

**Absolute Stability Criteria for Large-Scale Lurie Direct Control Systems with Time-Varying Coefficients**, Fucheng Liao and Di Wang  
Volume 2014, Article ID 631604, 13 pages

**An Optimal Control Strategy for DC Bus Voltage Regulation in Photovoltaic System with Battery Energy Storage**, Muhamad Zalani Daud, Azah Mohamed, and M. A. Hannan  
Volume 2014, Article ID 271087, 16 pages

**Risk Intelligence: Making Profit from Uncertainty in Data Processing System**, Si Zheng, Xiangke Liao, and Xiaodong Liu  
Volume 2014, Article ID 398235, 16 pages

**Effect Analysis of Design Variables on the Disc in a Double-Eccentric Butterfly Valve**, Sangmo Kang, Da-Eun Kim, Kuk-Kyeom Kim, and Jun-Oh Kim  
Volume 2014, Article ID 305085, 6 pages

**Optimized Scheduling Technique of Null Subcarriers for Peak Power Control in 3GPP LTE Downlink**, Soobum Cho and Sang Kyu Park  
Volume 2014, Article ID 279217, 8 pages

**Balance Maintenance in High-Speed Motion of Humanoid Robot Arm-Based on the 6D Constraints of Momentum Change Rate**, Da-song Zhang, Rong Xiong, Jun Wu, and Jian Chu  
Volume 2014, Article ID 535294, 13 pages

**An Optimization Method for Condition Based Maintenance of Aircraft Fleet Considering Prognostics Uncertainty**, Qiang Feng, Yiran Chen, Bo Sun, and Songjie Li  
Volume 2014, Article ID 430190, 8 pages

**Joint Power and Multiple Access Control for Wireless Mesh Network with Rose Projection Method**, Meiqin Tang, Lili Shang, Yalin Xin, Xiaohua Liu, and Xinjiang Wei  
Volume 2014, Article ID 352809, 7 pages

**A High Performance Load Balance Strategy for Real-Time Multicore Systems**, Keng-Mao Cho, Chun-Wei Tsai, Yi-Shiuan Chiu, and Chu-Sing Yang  
Volume 2014, Article ID 101529, 14 pages

**A Game-Theoretical Approach to Multimedia Social Networks Security**, Enqiang Liu, Zengliang Liu, Fei Shao, and Zhiyong Zhang  
Volume 2014, Article ID 791690, 9 pages

**Energy Efficiency of Task Allocation for Embedded JPEG Systems**, Yang-Hsin Fan, Jan-Ou Wu, and San-Fu Wang  
Volume 2014, Article ID 718348, 8 pages

**Comparative Study of Popular Objective Functions for Damping Power System Oscillations in Multimachine System**, Naz Niamul Islam, M. A. Hannan, Hussain Shareef, Azah Mohamed, and M. A. Salam

Volume 2014, Article ID 549094, 8 pages

**Fault Location Based on Synchronized Measurements: A Comprehensive Survey**, A. H. Al-Mohammed and M. A. Abido

Volume 2014, Article ID 845307, 10 pages

**Design of an Optimal Preview Controller for Linear Discrete-Time Descriptor Noncausal Multirate Systems**, Mengjuan Cao and Fucheng Liao

Volume 2014, Article ID 965915, 11 pages

## Editorial

# Power, Control, and Optimization

**Pandian Vasant,<sup>1</sup> Gerhard-Wilhelm Weber,<sup>2</sup> Nader Barsoum,<sup>3</sup> and Vo Ngoc Dieu<sup>4</sup>**

<sup>1</sup>*Department of Fundamental and Applied Sciences, Universiti Teknologi Petronas, Malaysia*

<sup>2</sup>*Institute of Applied Mathematics, METU, Turkey*

<sup>3</sup>*School of Engineering and Science, Curtin University, Sarawak, Malaysia*

<sup>4</sup>*Department of Power Systems, University of Technology, Ho Chi Minh City, Vietnam*

Correspondence should be addressed to Pandian Vasant; [pvasant@gmail.com](mailto:pvasant@gmail.com)

Received 16 October 2014; Accepted 16 October 2014

Copyright © 2015 Pandian Vasant et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This special issue of this journal is devoted to the *7th Global Conference on Power, Control, and Optimization (PCO'2013)* (<http://ifors.org/web/seventh-global-conference-on-power-control-and-optimization-pco-2013/>), which was held on August 25–27, 2013, in Prague, Czech Republic. Best papers of this conference were covered. *PCO'2013* attracted participants from many countries and from across the *PCO* community in Europe and all over the world, for a meeting of vivid exchanges and lively debates. The *PCO'2013* conference provided an excellent floor for investigators and practitioners to promote their newest advances in power, control, and optimization to the wider community of scientists and practitioners, for identifying research challenges for their fields, as well as promising research developments in theory, methods, and applications, and for fostering given and creating new interactions with colleagues from related research areas of modern *PCO* areas and their emerging applications.

Applications of optimization techniques are well known in the research areas of power systems, control systems, and computer science. The current trends on application of metaheuristics techniques are very popular among the researchers in the areas of power systems, control systems, and computer science across the globe. In this regard, the *PCO* global conference organized a successful conference in the research areas of power systems, control systems, and optimization in Prague, the capital of Czech Republic. This special issue focuses on best and high quality selected papers which were presented at the *PCO* global conference. Well-known and new methodologies and techniques of optimization were used to solve some of the complicated and

hard problems in the areas of power systems, control systems, and optimization.

This special issue was open for interested authors in all power, control, and optimization areas, and they were welcome to submit their recent findings and best results. Both *PCO'2013* and our special issue focused on new development and contribution of the current research to the body of the knowledge in the area of *PCO*. The topics included are (i) optimal power; (ii) optimal control; (iii) optimization; (iv) cams, gears, wheels, and valve optimization; (v) android phone, mobile, WiMax, and wireless communications; (vi) smart grid, microgrid, distribution, and chaotic systems; (vii) scheduling and assignment problems; (viii) healthcare, bioinformatics, and signal processing; and (ix) future energy planning and the environment.

The subjects of power, control, and optimization as elaborated in our special issue are so closely related to the emerging questions and challenges on energy and electricity, from engineering, applied mathematics, operational research, computer science and data mining, education, and social sciences, both quantitatively and qualitatively, and they reach out very far and, eventually, to the living conditions of people on earth.

## Acknowledgments

The guest editors are very grateful to all the collaborators of the journal, for their welcoming and approval of this special issue and for their guidance, help, and close advice at every stage. We guest editors wish all the readers an enjoyable,

instructive, and inspiring study of the contributions of this journal special issue. Indeed, we cordially hope that our special issue will represent this journal as a rising and premium journal of science, which strongly supports research, education, and development everywhere on the globe.

*Pandian Vasant*  
*Gerhard-Wilhelm Weber*  
*Nader Barsoum*  
*Vo Ngoc Dieu*

## Research Article

# Integer-Linear-Programming Optimization in Scalable Video Multicast with Adaptive Modulation and Coding in Wireless Networks

**Dongyul Lee and Chaewoo Lee**

*School of Electrical and Computer Engineering, Woncheon Hall 432, Ajou University 206, World Cup-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do 443-749, Republic of Korea*

Correspondence should be addressed to Dongyul Lee; leedongyull@gmail.com

Received 6 March 2014; Accepted 7 July 2014; Published 3 September 2014

Academic Editor: Gerhard-Wilhelm Weber

Copyright © 2014 D. Lee and C. Lee. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The advancement in wideband wireless network supports real time services such as IPTV and live video streaming. However, because of the sharing nature of the wireless medium, efficient resource allocation has been studied to achieve a high level of acceptability and proliferation of wireless multimedia. Scalable video coding (SVC) with adaptive modulation and coding (AMC) provides an excellent solution for wireless video streaming. By assigning different modulation and coding schemes (MCSs) to video layers, SVC can provide good video quality to users in good channel conditions and also basic video quality to users in bad channel conditions. For optimal resource allocation, a key issue in applying SVC in the wireless multicast service is how to assign MCSs and the time resources to each SVC layer in the heterogeneous channel condition. We formulate this problem with integer linear programming (ILP) and provide numerical results to show the performance under 802.16 m environment. The result shows that our methodology enhances the overall system throughput compared to an existing algorithm.

## 1. Introduction

Rapid advancement in mobile wideband wireless network (i.e., 4G and 5G) supports real time services such as IPTV and live video streaming. Nevertheless, efficient resource allocation is required to satisfy delay requirements and to overcome the bandwidth shortage due to the increased number of users. Because of the sharing nature of the wireless medium, the wireless multicast has attracted much attention when multiple users want to receive the same video service at the same time [1].

Multicast service in wireless heterogeneous networks can adaptively choose a certain modulation and coding scheme (MCS) [2] based on the channel status and the hardware capability of the receivers. For the non-scalable coding (or called a single-layer coding) the multicast data rate and video quality are determined by the user who experiences the worst channel status [1, 3].

To alleviate the video performance degradation caused by the user who has the worst channel status, scalable video coding (SVC) [4] with adaptive modulation and coding

(AMC) provides an excellent solution. In SVC, a video stream is divided into multiple layers. SVC encodes video with the nested dependency: the base layer encodes the basic video quality and higher layers, called the enhancement layers, refine the visual quality from the base layer with smaller quantization granularities. To enhance the overall performance of wireless multicast video streaming, we assign a low MCS to base layer and high MCSs to enhancement layers, so that the users in bad channel conditions receive fewer enhancement layers and obtain basic video quality, while the users in good channel conditions receive more enhancement layers and obtain better video quality.

A key issue arising under this setting is how each layer selects its MCS under the constraint of available radio resources for multicast services. A lot of research has studied the resource allocation problem in wireless multicast streaming video services [5–10]. In particular, [8, 10] are most closely related to our research work, which provides solutions in both the single channel and slot TDMA-based networks. Reference [8] proposed an algorithm that guarantees full coverage which provides all destinations with at least a base

layer to accommodate them. It allocates other MCSs to the enhancement layers with a heuristic algorithm to maximize the sum of utilities of all the receivers. It assumed that the size of each video layer is the same and the utility function of a user is simply a log function of the received data. The size of each layer is usually different, such that SVC supports various video qualities. The video quality experienced by a user is peak signal-to-noise ratio (PSNR) not a simple log function, so PSNR should be used in the resource allocation.

Reference [10] formulated a model that provides different size for each layer to support various video qualities. It used PSNR as a utility function to reflect the video quality experienced by a user and utilized dynamic programming for optimal network design. While the PSNR depends on the number of received packets in a layer, it increases only when all the packets in a layer are received fully. This PSNR model cannot be applied to other encoding methods of which PSNR increases with respect to the received packets. Hence, the PSNR which is the function of the received packets has to be derived. To search for an optimal solution easily, it assumes that the packet size and the slot size are fixed. However, if data rate is varied, the fixed-size slot cannot be utilized fully. Thus, if the slot size is determined based on each MCS, more video data can be transmitted without idle time in a slot.

To overcome these problems, we analyze PSNR models of various SVC encoding methods and choose a utility function to accommodate all the PSNR models. We also formulate a model which varies the slot size based on each MCS and utilize integer linear programming (ILP) to obtain the optimal solution.

The rest of the paper is organized as follows. The background and system architecture are discussed in Section 2. In Section 3, we present the system model and utility formulation. The problem formulation is described in Section 4. Section 5 explains ILP modeling and Section 6 shows the numerical results. Finally, Section 7 concludes the paper.

## 2. Background

**2.1. Scalable Video Coding.** Scalable video coding compresses a raw video sequence into a base layer and one or multiple enhancement layers. The base layer provides video data whose resolution is low, while the higher layers refine the video generated by the lower layer. SVC can provide adaptive video quality in wireless networks, but its performance is poor in terms of video quality compared to single-layer coding. Recently, SVC has improved its coding efficiency. For example, the newly established MPEG-4 SVC provides equal visual quality comparable with H.264/AVC (single-layer coding) but with at most 10% higher bit rate.

The quality of video can be measured using PSNR, which is a nondecreasing function of the received video data [11–15]. The PSNR function has different shapes according to encoding schemes, and it can be classified into four groups as illustrated in Figure 1 [11–15]. From Figure 1, we can observe that the functions of Figures 1(a) and 1(d) are represented as piecewise linear functions directly. However, Figures 1(b) and 1(c) show nonlinear functions, and so we approximate them to piecewise linear functions in the next section.

**2.2. Mobile Wireless Systems.** In infrastructure-based wireless networks, we consider the video multicast. A video server is equipped in the layered video codec to encode the video into multiple layers as illustrated in Figure 2. The video server then streams the video from all layers to APs (or base stations). The number and the size of the layers corresponding to each video frame are delivered to APs for every video frame period. The connection speed between the video server and the APs (or base stations) is high and reliable. Video stream is usually transmitted at a constant bit rate and an AP or base station forwards the packets delivered by video server to users for fragmentation. So, the packet size is fixed and is transmitted in a wireless channel.

Scheduling (resource allocation) in a wireless network MAC is done once for every scheduling frame period whose length is identical to the length of the video frame in our system. A scheduling frame consists of multiple time slots. The size of a time slot is the minimum time required to transmit a packet with a MCS. It means that the maximum number of available slot sizes is equal to the number of MCSs.

We consider the heterogeneous wireless networks, and so the users who experience relatively bad channel status can select a relatively high rate that leads to the considerable packet loss. In our system, each user measures its SNR and finds which MCS indicates the maximum data rate that satisfies target BER (bit error ratio) to decode the received data. An AP or base station assigns the radio resource and MCSs to layers to maximize the sum of utilities of all users to find the optimal solution.

## 3. System Model and Utility Formulation

We consider the one-hop broadcasting in wireless networks for the problem formulation. There are  $J$  users, and  $F$  is the size of a scheduling frame period, which is equal to that of video frame period. We assume a set of possible MCSs  $\mathcal{M} = \{1, \dots, M\}$  (e.g., MCS  $m = 1$  indicates QPSK-1/2 and others). High MCS  $m$  means a high data rate. If a user can decode the data provided by MCS  $m$  by satisfying a target BER, it can also decode the data that is less than MCS  $m$ .  $M_j \in \mathcal{M}$  is the maximum MCS, which can be received by user  $j$  based on the channel status. User  $j$  transfers its MCS  $M_j \in \mathcal{M}$  to its AP (base station).  $t_m$  is the time slot size given by MCS  $m$ . We consider a single multicast session which has  $L$  layers. Each layer is divided into many slices, which is also split into the packets of fixed-size.  $N_l$  is the number of packets of layer  $l$ .

Now, we explain a utility function and the related variables. As aforementioned, the satisfaction of the users is proportional to the quality of multimedia defined as PSNR, which is a function of the received throughput. We assume that there is no packet loss in the wireless channel. That is, user  $j$  can decode a packet if it is transmitted by MCS  $m' \leq M_j$ . As illustrated in Figure 1, the PSNR function has four types. First, we are focusing on the second scheme of Figure 1(c) because the other schemes can be easily explained using this.

As we can see in Figure 1(c), it is nonlinear function and ILP cannot be formulated. In this paper, we approximate it

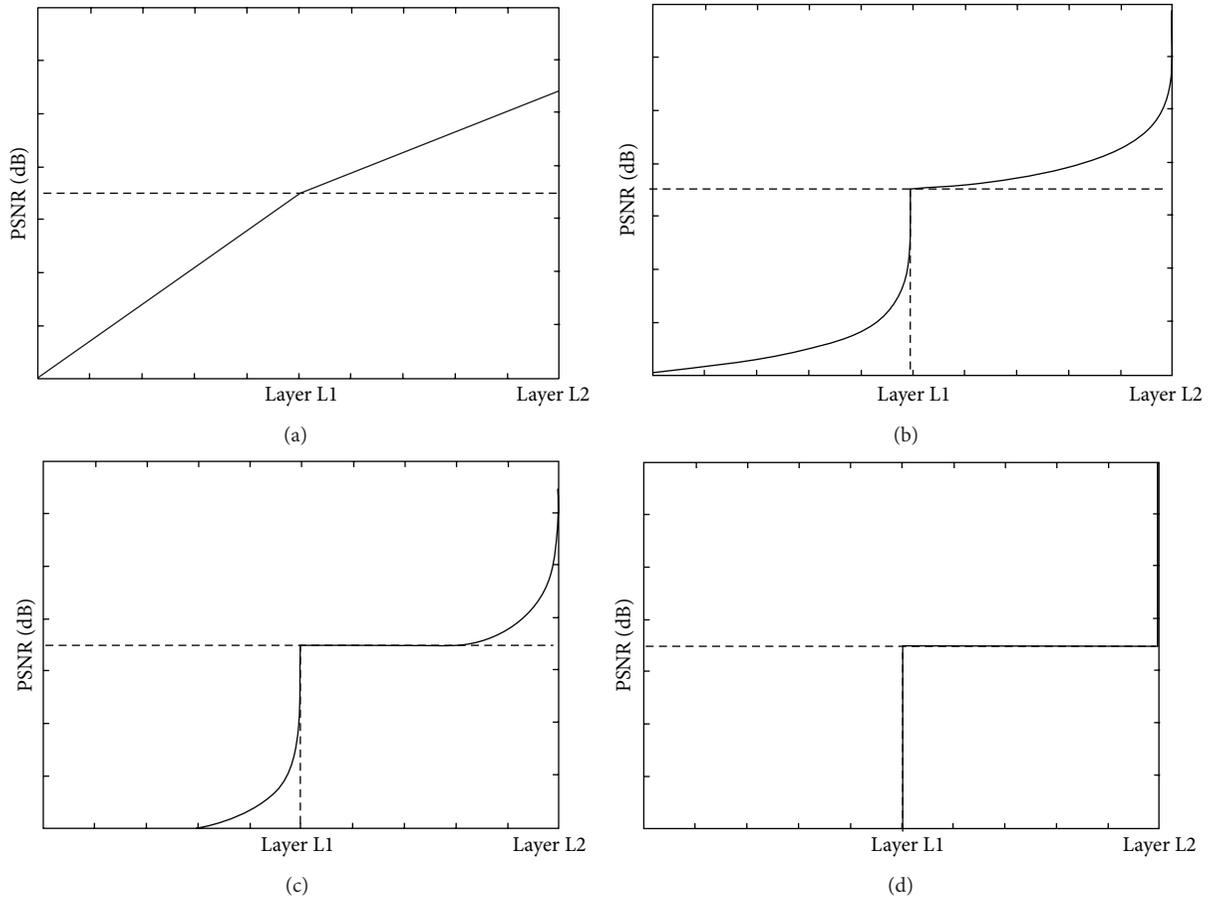


FIGURE 1: PSNR to the number of the packets received at a user in the four schemes.

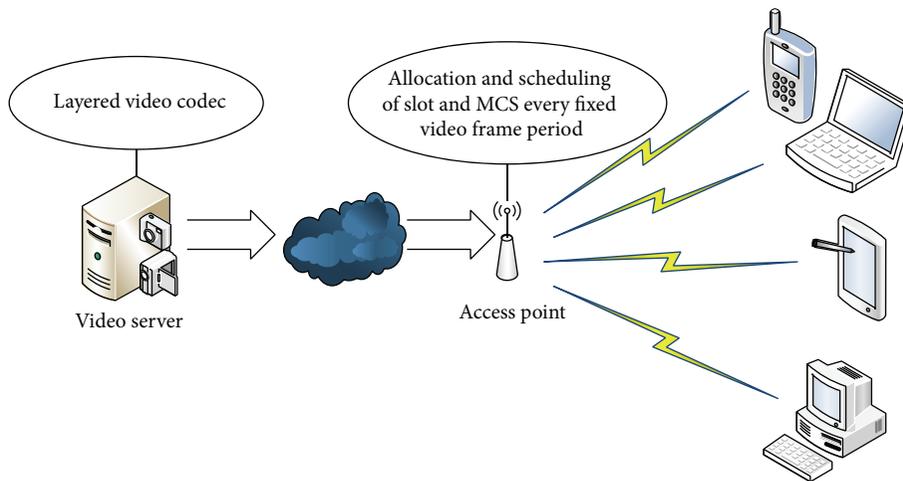


FIGURE 2: System architecture for scalable video multicast.

with piecewise linear function as illustrated in Figure 3. We can see that there is relatively large difference between the original PSNR and the approximated one with a few packet losses. However, in case of more packet loss, it shows little difference. In fact, the packet loss which shows relatively large difference is quite less among all the packets and so it is

not critical. Although this is incomplete, it can guarantee higher total utility than those of the existing methods [8–10] because the existing methods just consider a utility model that increases only when a user receives a layer completely. This approximation also can be formulated by ILP modeling. Figure 3 shows that the function has zero utility until it

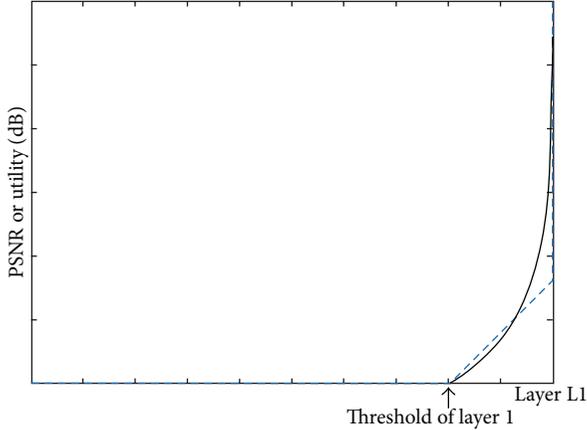


FIGURE 3: The difference between the original PSNR and approximated PSNR in a couple of packet losses.

reaches the threshold, indicating the number of packets  $\delta_l$  which does not increase the video quality in a layer. Let  $\alpha_{l,m}$  be the number of received packets which have zero utilities in layer  $l$ . The utility function increases linearly when the packets in a layer  $l$  are received from the threshold  $\delta_l$  to the number of packets  $N_l$  of layer  $l$ . Let  $\beta_{l,m}$  be the number of received packets which increases the utility linearly. Denote  $y_{l,m}$  as the indicator function that is 1 if layer  $l$  is transmitted by  $\forall m' \leq m$  and 0 otherwise:

$$y_{l,m} = \begin{cases} 1, & \text{if } \sum_{m'=1}^m \alpha_{l,m'} = \delta_l, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Denote  $z_{l,m}$  as the indicator function that is 1 if all the received packets  $\beta_{l,m}$  are modulated by  $\forall m' \leq m$  and 0 otherwise:

$$z_{l,m} = \begin{cases} 1, & \text{if } \sum_{m'=1}^m \beta_{l,m'} = N_l - \delta_l, \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

Here, we set that  $z_{l,0} = 0$  because  $\beta_{l,0} = 0$ . Hence, with these indicator variables, we utilize the approximated PSNR model as the utility function of the user  $j$  given as

$$U_j = \sum_{l=1}^L \sum_{m=1}^{M_j} (r_m t_m \phi_l \beta_{l,m} + \varphi_l (z_{l,m} - z_{l,m-1})) \quad (3)$$

with the constraints of layer transmission:

$$\begin{aligned} \sum_{m=1}^{M_j} \alpha_{l,m} &\leq \delta_l, \quad \forall l, \\ \sum_{m=1}^{M_j} \beta_{l,m} &\leq N_l - \delta_l, \quad \forall l, \\ 0 &\leq \sum_{m=1}^{M_j} \alpha_{l,m} + \sum_{m=1}^{M_j} \beta_{l,m} \leq N_l, \quad \forall l, \end{aligned} \quad (4)$$

where  $\phi_l$  is the PSNR weight of the packets which increases the utility of layer  $l$  linearly.  $\varphi_l$  is the PSNR weight of the complete reception of layer  $l$ .  $r_m t_m \phi_l \beta_{l,m}$  in (3) means the approximated PSNR to  $\beta_{l,m}$  in layer  $l$ .  $\varphi_l (z_{l,m} - z_{l,m-1})$  in (3) means the approximated PSNR by the complete reception of layer  $l$ . That is, if the last packet of layer  $l$  is transmitted by MCS  $m'$ ,  $(z_{l,m} - z_{l,m-1})$  has 1 only if  $m = m'$ , otherwise zero.

Three other methods of the utility function can be easily derived as follows. If we set  $\delta_l = 0$ , the first scheme can be formulated. If we set  $\delta_l = 0$  and  $\varphi_l = 0$ , the third scheme can be formulated. If we set  $\varphi_l = N_l$ , the final scheme can be formulated.

#### 4. Problem Formulation

In this paper, we formulate an ILP problem to assign modulation and coding scheme for each layer in terms of the sum of utilities under any given resource constraints. That is, we have to find out

$$\alpha = \begin{bmatrix} \alpha_{1,1} & \cdots & \alpha_{1,M} \\ \vdots & \ddots & \vdots \\ \alpha_{l,1} & \cdots & \alpha_{l,M} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_{1,1} & \cdots & \beta_{1,M} \\ \vdots & \ddots & \vdots \\ \beta_{l,1} & \cdots & \beta_{l,M} \end{bmatrix} \quad (5)$$

in order to

$$\text{Maximize } \sum_j U_j \quad (6)$$

following the total time constraints:

$$0 \leq \sum_{l=1}^L \sum_{m=1}^M t_m \alpha_{l,m} + \sum_{l=1}^L \sum_{m=1}^M t_m \beta_{l,m} \leq F \quad (7)$$

and layer transmission constraints (1), (2), and (4). Since this formulation does not consider the nested dependency of SVC, it cannot solve the optimal solution yet. We present three lemmas to get an optimal solution.

**Lemma 1.** *In the optimal solution of the multicast resource allocation problem in wireless networks with the encoding methods [11], if the last packet of layer  $l-1$  is transmitted using modulation  $m'$ , the packets that have zero utility in layer  $l$  use  $m$ , and then it must be  $m \geq m'$ . That is, the following condition should be satisfied:*

$$0 \leq \alpha_{l,m} \leq \delta_l z_{l-1,m} \quad (8)$$

with constraint of  $z_{l,m}$ , (2).

*Proof.* We assume that two layers  $l$  and  $l'$  use modulations  $m$  and  $m'$ , respectively. If  $l > l'$  and  $m < m'$ , one can improve the utility by swapping  $m$  and  $m'$  because the utility of packets of layer  $l$  using  $m$  is less than and equal to the utility of layer  $l$  using  $m'$ .  $\square$

**Lemma 2.** *In the optimal solution of the multicast resource allocation problem in wireless networks with the encoding methods [11], if the last packet which has zero utility in layer  $l$  is transmitted using modulation  $m'$ , the packets that have*

nonzero utilities in layer  $l$  use  $m$ ; then, it must be  $m \geq m'$ . That is, the following condition should be satisfied:

$$0 \leq \beta_{l,m} \leq (N_l - \delta_l) y_{l,m} \quad (9)$$

with constraint of  $y_{l,m}$  (1).

*Proof.* We assume that packets which have zero utilities and packets which increase the utility linearly of layer  $l$  use  $m$  and  $m'$ , respectively:  $\alpha_{l,m}$  and  $\beta_{l,m'}$ . If  $m < m'$ , one can improve the utility by swapping  $m$  and  $m'$  because the utility of  $\beta_{l,m}$  is less than and equal to the utility of  $\beta_{l,m'}$ .  $\square$

Lemmas 1 and 2 mean that a MCS can be assigned sequentially in the ascending order. We define full coverage property which guarantees that all the users receive the minimum video quality in Lemma 3.

**Lemma 3.** Let  $J_m$  be the set of users that can receive MCS  $m$ . To accommodate all users, the base layer should be transmitted by the MCS  $m_{base} = \max\{m : J_m = J\}$ :

$$\alpha_{1M_{base}} + \beta_{1M_{base}} = N_1. \quad (10)$$

*Proof.* The proof follows from the fact (1) that all users have to access the base layer and (2) the time resource required to fit the base layer will diminish with increasing modulation.  $\square$

Nevertheless, since (1) and (2) have nonlinear properties, they should be changed as the linear equations to formulate ILP problem.

## 5. ILP Modeling

We develop the remaining works to change conditions (1) and (2) as ILP model. The following lemma is proposed to change as the linear equations.

**Lemma 4.** For  $i \in \{0, 1, \dots, I\}$ ,  $c_i \in \{0, \mathbb{N}^+\}$ , if binary random variable  $x_i \in \{0, 1\}$  has the following condition:

$$x_i = \begin{cases} 1, & \text{if } \sum_{i=1}^I A_i B_i c_i = D, \\ 0, & \text{otherwise,} \end{cases} \quad (11)$$

where  $A_i$  and  $B_i$  are the elements of a constant vector and  $D$  is a constant and is always larger than  $\sum_{i=1}^I A_i B_i c_i$ , conditional equation (11) is given as

$$Dx_i \leq \sum_{i=1}^I A_i B_i c_i, \quad (12)$$

$$\sum_{i=1}^I A_i B_i c_i - D < x_i. \quad (13)$$

*Proof.* We can prove it easily. Equation (12) means that  $x_i$  should be 0 if  $\sum_{i=1}^I A_i B_i c_i \neq D$ , and it can be 1 or 0 otherwise. On the contrary to this, (13) means that  $x_i$  should be 1 if  $\sum_{i=1}^I A_i B_i c_i = D$ , and it can be 1 or 0 otherwise. If we use these

TABLE 1: MCS and number of users supporting them.

	$m$			
	1	2	3	4
Modulation	BPSK	QAM16	QAM64	QAM256
Code rate	3/4	1/2	2/3	3/4
Net PHY bit rate (Mbps)	2.12	5.64	11.29	13.39
The number of users supporting MCS	14	13	8	7

two equations, (12) and (13), we can guarantee condition (11). Using Lemma 4, the conditions in (1) and (2) are changed as follows.  $\square$

**Theorem 5.**  $y_{l,m}$  should be satisfied by

$$\delta_l y_{l,m} \leq \sum_{m'=1}^m \alpha_{l,m'}, \quad (14)$$

$$\sum_{m'=1}^m \alpha_{l,m'} - \delta_l < y_{l,m}.$$

*Proof.* This can be easily proved by Lemma 4.  $\square$

**Theorem 6.**  $z_{l,m}$  should be satisfied by

$$(N_l - \delta_l) z_{l,m} \leq \sum_{m'=1}^m \beta_{l,m'} \quad (15)$$

$$\sum_{m'=1}^m \beta_{l,m'} - (N_l - \delta_l) < z_{l,m}.$$

*Proof.* This is can be easily proved by Lemma 4.  $\square$

The ILP problem to maximize the sum of utilities of all users in the wireless multicast service streaming can be formulated as

$$\arg \max_{\alpha, \beta} \sum_j U_j \quad (16)$$

with the constraints (4), (7), (8), (9), (10), (14), and (15).

In the worst case, since our method utilizes the same domain as the existing method [10] to find out an optimal solution, the complexity obtained  $O(JL + LM^2 F/t_M)$ .  $O(JL)$  is the complexity which is precalculated to find out the highest MCS to each node.  $O(LM^2 F/t_M)$  is the complexity to solve  $\alpha$  and  $\beta$  for all  $1 \leq l \leq L$ ,  $1 \leq m \leq M$ , and  $1 \leq m \leq M$ .  $F/t_M$  is the maximum number of slots and  $t_M$  is the minimum size of a slot. The complexity of the proposed method is pseudopolynomial due to the factor related to the time resource and the number of nodes. Since the complexity is linearly proportional to the number of nodes, the network size does not deteriorate the performance in the proposed methods. Therefore, the complexity of the proposed method is acceptable.

TABLE 2: Slot size of used solution and the comparison of packet number between them.

	$m$			
	1	2	3	4
Slot size of the proposed solution ( $\mu s$ )	28.29	10.62	5.28	4.47
Slot size of fixed-slot solution ( $\mu s$ )	28.29	28.29	28.29	28.29
The number of packets which the proposed method can transmit during a slot of the fixed-slot method	1	2.67	5.33	6.33

TABLE 3: Performance difference between them.

	$m$			
	1	2	3	4
The sum of utilities of the proposed solution			7478	
Total used time in a frame			0.0299994	
The proposed solution (ms)	11.316	8.496	8.131	2.056
The number of packets of the proposed solution	400	800	1540	460
The sum of utilities of fixed-slot solution			6950	
Total used time in a frame			0.02979864	
Fixed-slot solution (ms)	11.31	5.658	9.051	3.771
The number of packets of fixed-slot solution	400	400	1600	800

## 6. Numerical Results

In this section, we present the numerical results for our proposed solution. We consider one base station with 14 mobile users distributed in the circular area from base station. The test is based on the 802.16 m evaluation methodology document [16] for a 3.5 MHz spectrum in 3.5 GHz range. We assume that there are four types of MCSs in the IEEE 802.16 m standard [17]. Table 1 shows the MCSs we use in our test and the number of users who can support each MCS. We assume that, for every scheduling period, the fixed-size packet of 60 bits is transmitted in a wireless channel. Other various settings including MAC/PHY header, ACK, interframe space time, and others are not considered. This is because we aim to provide abstraction of important features of our solutions. Other existing researches [8, 10] are also not considered for it. If the setting is required, our solution can use it easily.

In order to analyze the features of our solution, we consider a single video session. Each video session has a fixed layer rate of 800 kb/s, and the number of layers in each session is 8. This is based on the standard of the SVC extension [4] of H.264/MPEG4-AVC. So, each layer can handle 400 packets in a scheduling period because the size of the layer is 30 Kbyte in a scheduling period. We set that  $\phi_l = \{1, 0.95, 0.9, \dots, 0.65\}$  and  $\varphi_l = \{10, 10, \dots, 10\}$  for  $l = \{1, \dots, 8\}$  based on [13].

We compare our solution with the fixed-slot solution which was solved by dynamic algorithm [10], as the scheduling varies period from 20 ms to 30 ms with interval 1.25 ms. The software package CPLEX is used to find the solution. Table 2 shows the size of a slot to transmit a packet in both solutions and the number of packets that can be transmitted by the method during a slot of the fixed-slot method.

Figure 4 shows the sum of utilities of two solutions. The proposed method enhances the sum of utilities by about 1.9–7.6% compared to the other method. The difference between

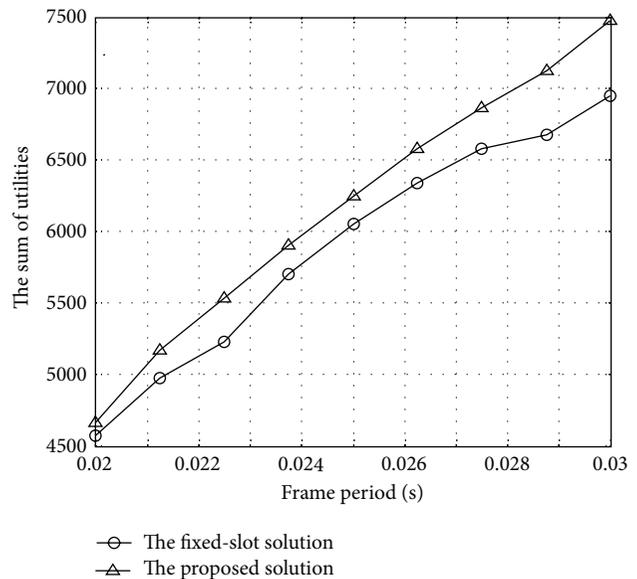


FIGURE 4: The sum of utilities with varying frame period.

the two solutions shows an increasing trend. This is because the proposed method assigns more time resource to the low MCSs compared to the other method by reducing the idle time. By detailed observation at 30 ms of frame period, we can understand why the proposed method can enhance the performance. Table 3 specifies the performance difference between the proposed solution and the fixed-size solution at 30 ms of frame period. The sum of utilities of our solution is 6950, while that of the fixed-slot solutions is 7478. Our solution is higher about 7.6%. This is because by eliminating idle time our solution provides second MCS ( $m = 2$ ) with

more slots than those provided by the fixed-slot solution even though both provide the transmission of all the layers.

## 7. Conclusion

In this paper, we considered and studied the resource allocation problem in SVC video multicast with AMC in wireless networks. First, we assume that all packets have equal length to consider the real video transmission environments. It is also assumed that each MCS has different slot length, which is the minimum time length required to transmit a packet with the corresponding MCS. We formulated a utility function as a piecewise function for ILP modeling using the existing results [11, 13]. Finally, we define ILP problem by proposing some lemmas and a theorem to formulate ILP problem. We provide numerical results to show performance difference with [10] under 802.16 m environments. The results show that our methodology enhances overall system throughput by eliminating radio resource waste.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

This research was supported by the Basic Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012R1A1A2007605).

## References

- [1] K. Obraczka, "Multicast transport protocols: a survey and taxonomy," *IEEE Communications Magazine*, vol. 36, no. 1, pp. 94–102, 1998.
- [2] J. Kim and D.-H. Cho, "Enhanced adaptive modulation and coding schemes based on multiple channel reportings for wireless multicast systems," in *Proceedings of the IEEE 62nd Vehicular Technology Conference (VTC-2005-Fall)*, vol. 2, pp. 725–729, September 2005.
- [3] L. Zhou, H. Wang, S. Lian, Y. Zhang, A. V. Vasilakos, and W. Jing, "Availability-aware multimedia scheduling in heterogeneous wireless networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 3, pp. 1161–1170, 2011.
- [4] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1103–1120, 2007.
- [5] Z. Liu, Z. Wu, P. Liu, H. Liu, and Y. Wang, "Layer bargaining: Multicast layered video over wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 3, pp. 445–455, 2010.
- [6] J. Kim, J. Cho, and H. Shin, "Layered resource allocation for video broadcasts over wireless networks," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 4, pp. 1609–1616, 2008.
- [7] H. M. Radha, M. van der Schaar, and Y. Chen, "The MPEG-4 fine-grained scalable video coding method for multimedia streaming over IP," *IEEE Transactions on Multimedia*, vol. 3, no. 1, pp. 53–68, 2001.
- [8] S. Deb, S. Jaiswal, and K. Nagaraj, "Real-time video multicast in WiMAX networks," in *Proceedings of the 27th IEEE Communications Society Conference on Computer Communications (INFOCOM '08)*, pp. 2252–2260, April 2008.
- [9] D. Wu, Y. T. Hou, and Y.-Q. Zhang, "Scalable video coding and transport over broadband wireless networks," *Proceedings of the IEEE*, vol. 89, no. 1, pp. 6–20, 2001.
- [10] P. Li, H. Zhang, B. Zhao, and S. Rangarajan, "Scalable video multicast with adaptive modulation and coding in broadband wireless data systems," *IEEE/ACM Transactions on Networking*, vol. 20, no. 1, pp. 57–68, 2012.
- [11] K. Stuhlmüller, N. Färber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 1012–1032, 2000.
- [12] S. Tang and P. R. Alfance, "Impact of packet loss on H. 264 scalable video coding," in *Proceedings of the 5th International Conferences on Advances in Multimedia (MMEDIA '13)*, pp. 67–73, 2013.
- [13] Z. Ma, Z. Liu, M. Xu, and Y. Wang, "Modeling the channel distortion of temporal scalable video," Tech. Rep., Polytechnic Institute of New York University, 2009.
- [14] J. Sun, W. Gao, D. Zhao, and W. Li, "On rate-distortion modeling and extraction of H.264/SVC fine-granular scalable video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 3, pp. 323–336, 2009.
- [15] M. R. Ardestani, A. A. B. Shirazi, and M. R. Hashemi, "Rate-distortion modeling for scalable video coding," in *Proceedings of the 17th International Conference on Telecommunications (ICT '10)*, pp. 923–928, April 2010.
- [16] R. Srinivasan, J. Zhuang, L. Jalloul, R. Novak, and J. Park, "IEEE 802.16 m evaluation methodology document (EMD)," IEEE 802.16 Broadband Wireless Access Working Group, 2008.
- [17] I. B. W. A. W. Group, "IEEE Standard for local and metropolitan area networks, part 16: air interface for broadband wireless access systems, amendment 3: advanced air interface," IEEE Std 802.16 mTM, 2011.

## Research Article

# Chaos Enhanced Differential Evolution in the Task of Evolutionary Control of Selected Set of Discrete Chaotic Systems

**Roman Senkerik,<sup>1</sup> Ivan Zelinka,<sup>2</sup> Michal Pluhacek,<sup>1</sup>  
Donald Davendra,<sup>2</sup> and Zuzana Oplatková Kominková<sup>1</sup>**

<sup>1</sup> Faculty of Applied Informatics, Tomas Bata University in Zlin, Náměstí T.G. Masaryka 5555, 760 01 Zlin, Czech Republic

<sup>2</sup> Faculty of Electrical Engineering and Computer Science, VŠB-Technical University of Ostrava, 17 Listopadu 15, Poruba, 708 33 Ostrava, Czech Republic

Correspondence should be addressed to Roman Senkerik; [senkerik@fai.utb.cz](mailto:senkerik@fai.utb.cz)

Received 6 March 2014; Accepted 7 July 2014; Published 26 August 2014

Academic Editor: Gerhard-Wilhelm Weber

Copyright © 2014 Roman Senkerik et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Evolutionary technique differential evolution (DE) is used for the evolutionary tuning of controller parameters for the stabilization of set of different chaotic systems. The novelty of the approach is that the selected controlled discrete dissipative chaotic system is used also as the chaotic pseudorandom number generator to drive the mutation and crossover process in the DE. The idea was to utilize the hidden chaotic dynamics in pseudorandom sequences given by chaotic map to help differential evolution algorithm search for the best controller settings for the very same chaotic system. The optimizations were performed for three different chaotic systems, two types of case studies and developed cost functions.

## 1. Introduction

In many engineering applications, one of the most innate tasks is the controlling of highly nonlinear dynamical systems in order to either eliminate or synchronize the chaos. The first pioneering approach to control chaotic dynamics by means of a simple analytical linearization method was introduced in 1990s by Ott et al. (i.e., OGY method) [1]. Subsequently, the rapid development of methods for stabilizing of nonlinear chaotic dynamics has arisen and many advanced techniques have been applied for chaos control and chaos synchronization including methods from the artificial intelligence field.

During recent years, usage of new intelligent systems in engineering, technology, modeling, computing, and simulations has attracted the attention of researchers worldwide. The most current methods are mostly based on soft computing, which is a discipline tightly bound to computers, representing a set of methods of special algorithms, belonging to the artificial intelligence paradigm. The most popular of these methods are neural networks, evolutionary algorithms (EAs), fuzzy logic, and tools for symbolic regression like genetic programming. Currently, EAs are known as a powerful set

of tools for almost any difficult and complex optimization problem.

The interest about the interconnection between evolutionary techniques and control of chaotic systems is rapidly spreading. The initial research was conducted in [2], whereas [3, 4] was more concerned with the tuning of parameters inside the chaos control technique based on the Pyragas method: extended delay feedback control (ETDAS) [5]. When compared to the aforementioned research, later works [6, 7] show a possibility as to how to generate the entire control law (not only how to optimize several parameters) for the purpose of stabilization of a chaotic system. Such approach also may overcome the possible sensitivity to initial conditions which may lead to stability issues. The synthesis of control is inspired by the Pyragas' delayed feedback control technique [8, 9]. This method is very advantageous for the evolutionary computation, due to its amount of easy accessible control parameters, which can be easily tuned by means of EAs.

Other approaches utilizing the EAs for stabilizing chaotic dynamics have mostly applied the particle swarm optimization (PSO) algorithm [10] and multi-interval gradient

method [11] or minimum entropy control technique [12]. EAs have been also frequently used in the task of synchronization of chaos [13–15]. In [16], an EA for optimizing local control of chaos based on a Lyapunov approach is presented.

Evolutionary techniques were also used for the synthesis of new complex discrete type structures with chaotic behavior [17] as well as the synthesis (identification) of a mathematical model of chaotic system based on the measured data [18].

Another example of interconnection between deterministic chaos and EAs represents the research focused on the embedding of chaotic dynamics into the EAs. Recent research in chaos driven heuristics has been fueled with the predisposition that, unlike stochastic approaches, a chaotic approach is able to bypass local optima stagnation. This one clause is of deep importance to EAs. A chaotic approach generally uses the chaotic map in the place of a pseudorandom number generator [19]. This causes the heuristic to map unique regions, since the chaotic map iterates to new regions. The task is then to select a very good chaotic map as the pseudorandom number generator.

The initial concept of embedding chaotic dynamics into EAs is given in [20]. Later, the initial study [21] was focused on the simple embedding of chaotic systems in the form of chaos pseudorandom number generator (CPRNG) for differential evolution (DE) [22] and self-organizing migrating algorithm (SOMA) [23] in the task of optimal PID tuning. Also, the PSO algorithm with elements of chaos was introduced as the CPSO [24]. This field of research was later extended with the successful experiments with chaos driven DE [25] in real domain as well as in combinatorial problems domain [26, 27].

At the same time, the chaos embedded PSO with inertia weigh strategy was closely investigated [28], followed by the introduction of a PSO strategy driven alternately by two chaotic systems [29] and novel chaotic multiple choice PSO (chaos MC-PSO) strategy [30]. Recently, the chaotic firefly algorithm was also introduced [31].

Finally, the last example represents the research focusing on the EAs and the edge of chaos. An unconventional approach of the edge of chaos and its application to discrete systems and evolutionary algorithms in terms of stagnation avoidance is presented in [32].

The organization of this paper is as follows: firstly, used evolutionary technique, which is DE, is described and is followed by the description of the ChaosDE concept. Thereafter, the problem design and appropriate corresponding cost functions are investigated and proposed. Results and conclusion follow afterwards.

## 2. Motivation

This paper extends the research of evolutionary chaos control optimization by means of ChaosDE algorithm [33–35]. Recent studies have shown that differential evolution [22] is one of the most potent heuristics and it has been used for a number of optimization tasks; [36–38] have explored DE for combinatorial problems; [39, 40] have hybridized DE whereas [41–43] have developed self-adaptive DE variants.

In this paper, the DE/rand/1/bin strategy driven by different chaotic maps (systems) was utilized to solve the issue of

evolutionary optimization of chaos control for the very same chaotic system used as a CPRNG in the particular case study. Thus, the idea was to utilize the hidden chaotic dynamics in pseudorandom sequences given by chaotic map to help differential evolution algorithm search for the best controller settings for the very same chaotic system. Since the very positive contribution of the chaotic dynamics to the performance of DE in the task of evolutionary chaos control optimization was proven in comparison with original canonical DE within the initial study [44], this paper is not primarily focused on the performance comparisons with different heuristic. But this research extends the initial work with the aforementioned idea and with the several case studies combining different chaotic systems and different utilized cost functions.

## 3. Used Heuristic: Differential Evolution (DE)

DE is a simple and powerful population-based optimization method that works either on real-number-coded individuals or with small modifications on discrete type individuals [22, 45, 46]. DE is quite robust, fast, and effective, with global optimization ability. This global optimization ability has been proven in many interdisciplinary researches. It works well even with noisy and time-dependent objective functions. The canonical basic principle is as follows.

For each individual  $\vec{x}_{i,G}$  in the current generation G, DE generates a new trial individual  $\vec{x}'_{i,G}$  by adding the weighted difference between two randomly selected individuals  $\vec{x}_{r1,G}$  and  $\vec{x}_{r2,G}$  to a randomly selected third individual  $\vec{x}_{r3,G}$ . The resulting individual  $\vec{x}'_{i,G}$  is crossed over with the original individual  $\vec{x}_{i,G}$ . The fitness of the resulting individual, referred to as a perturbed vector  $\vec{u}_{i,G+1}$ , is then compared with the fitness of  $\vec{x}_{i,G}$ . If the fitness of  $\vec{u}_{i,G+1}$  is greater than the fitness of  $\vec{x}_{i,G}$ , then  $\vec{x}_{i,G}$  is replaced with  $\vec{u}_{i,G+1}$ ; otherwise,  $\vec{x}_{i,G}$  remains in the population as  $\vec{x}_{i,G+1}$ .

Please refer to (1) for notation of crossover and to [22] for the detailed description of used DERand1Bin strategy and all other DE strategies:

$$u_{i,G+1} = x_{r1,G} + F \cdot (x_{r2,G} - x_{r3,G}). \quad (1)$$

## 4. Concept of ChaosDE

This section contains the description of discrete dissipative chaotic maps, which can be used as the chaotic pseudorandom generators for DE as well as the main principle of the ChaosDE concept. In this research, direct output iterations of the chaotic maps were used for the generation of real numbers in the process of crossover based on the user defined CR value and for the generation of the integer values used for selection of individuals.

*4.1. Chaotic Pseudorandom Number Generator.* The general idea of ChaosDE and CPRNG is to replace the default PRNG with the discrete chaotic map. As the discrete chaotic map is a set of equations with a static start position, we created a random start position of the map, in order to have different start position for different experiments (runs of EAs). This

random position is initialized with the default PRNG, as a one-off randomizer. Once the start position of the chaotic map has been obtained, the map generates the next sequence using its current position.

The first possible way is to generate and store a long data sequence (approximately 50–500 thousands numbers) during the evolutionary process initialization and keep the pointer to the actual used value in the memory. In case of the using up of the whole sequence, the new one will be generated with the last known value as the new initial one.

The second approach is that the chaotic map is not reinitialized during the experiment and any long data series is not stored; thus it is imperative to keep the current state of the map in memory to obtain the new output values.

As two different types of numbers are required in ChaosDE, real and integers, the use of modulo operators is used to obtain values between the specified ranges, as given in the following equations (2):

$$\text{rndreal} = \text{mod}(\text{abs}(\text{rndChaos}), 1.0) \tag{2}$$

$$\text{rndint} = \text{mod}(\text{abs}(\text{rndChaos}), 1.0) \times \text{Range} + 1,$$

where *abs* refers to the absolute portion of the chaotic map generated number *rndChaos*, *mod* is the modulo operator, and *Range* specifies the value (inclusive) till where the number is to be scaled.

**4.2. Selected Chaotic Systems.** This subsection contains the mathematical and graphical description of the three selected discrete dissipative systems, which served both as for CPRNGs and also as the examples of systems to be evolutionary controlled.

**4.2.1. Burgers Map.** The Burgers mapping is a discretization of a pair of coupled differential equations which were used by Burgers to illustrate the relevance of the concept of bifurcation to the study of hydrodynamics flows. The map equations are given in (3) with control parameters  $a = 0.75$  and  $b = 1.75$  as suggested in [47]:

$$\begin{aligned} X_{n+1} &= aX_n - Y_n^2, \\ Y_{n+1} &= bY_n + X_nY_n. \end{aligned} \tag{3}$$

**4.2.2. Delayed Logistic.** The delayed logistic is a simple two-dimensional discrete system similar to the one-dimensional logistic equation. The map equations are given in (4). The parameter used in this work is  $A = 2.27$  as also suggested in [47]:

$$\begin{aligned} X_{n+1} &= AX_n(1 - Y_n), \\ Y_{n+1} &= X_n. \end{aligned} \tag{4}$$

**4.2.3. Lozi Map.** The Lozi map is a simple discrete two-dimensional chaotic map. The map equations are given in (5). The parameters used in this work are as follows:  $a = 1.7$  and  $b = 0.5$  as suggested in [47]. For these values, the system exhibits typical chaotic behavior and, with this parameter

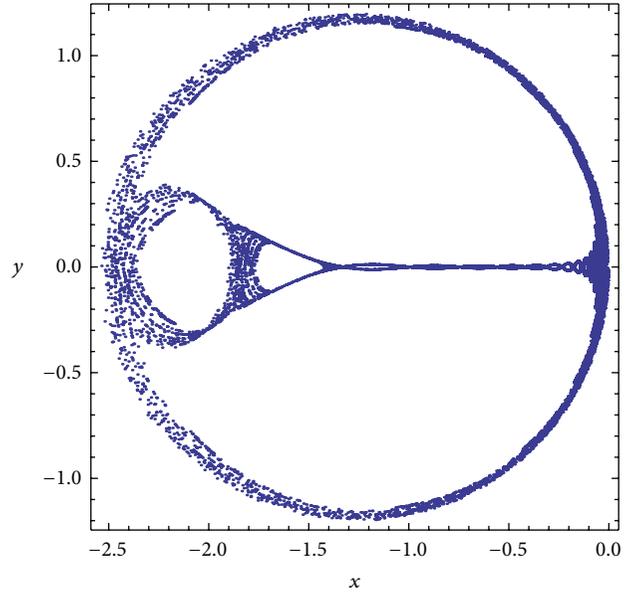


FIGURE 1: *x, y* plot of the Burgers map.

setting, it is used in most research papers and other literature sources [48]:

$$\begin{aligned} X_{n+1} &= 1 - a|X_n| + bY_n, \\ Y_{n+1} &= X_n. \end{aligned} \tag{5}$$

**4.3. Graphical Examples.** Three chaotic maps were selected for the CPRNG concept. The *x, y* plots of the selected maps are depicted in Figure 1 (Burgers map), Figure 4 (delayed logistic map), and Figure 7 (Lozi map). The chaotic behavior of the chaotic maps, represented by the examples of direct output iterations, is depicted in Figures 2, 5, and 8. Finally, the illustrative histograms of the distribution of real numbers transferred into the range  $\langle 0-1 \rangle$  generated by means of chaotic maps are in Figures 3, 6, and 9.

## 5. Design of Cost Functions for Chaotic System Stabilization

The proposal of the basic cost function ( $CF_{\text{Simple}}$ ) is in general based on the simplest CF, which could be used problem-free only for the stabilization of  $p-1$  orbit. The idea was to minimize the area created by the difference between the required state and the real system output on the whole simulation interval,  $\tau_i$  (6). This CF design is very convenient for the evolutionary searching process due to the relatively favorable CF surface. Nevertheless, this simple approach has one big disadvantage, which is the including of initial chaotic transient behavior of not stabilized system into the cost function value. As a result of this, the very tiny change of control method setting for extremely sensitive chaotic system (given by the very small change of CF value) can be

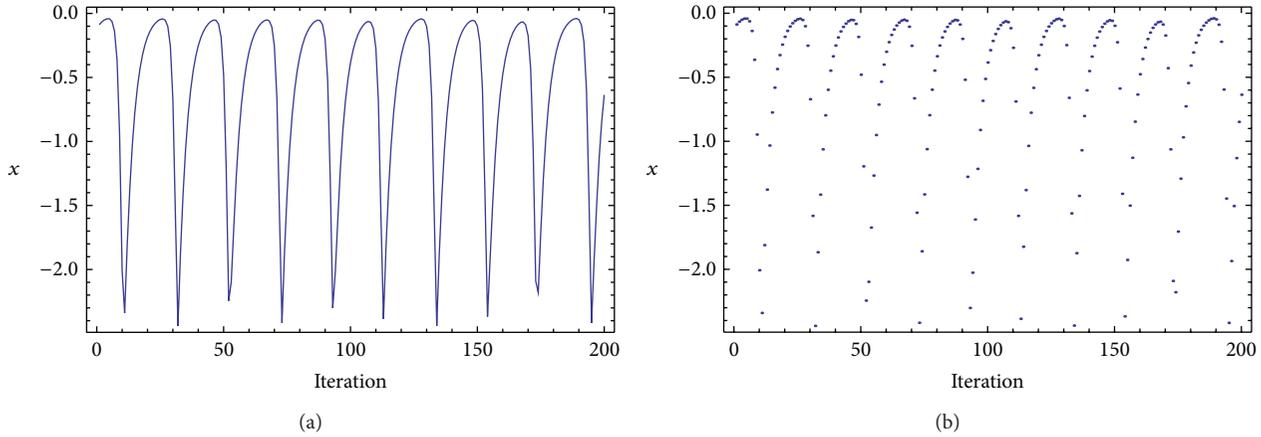


FIGURE 2: Iterations of the Burgers map (variable  $x$ ) (a), iterations of the Burgers map (variable  $x$ ), point-plot (b).

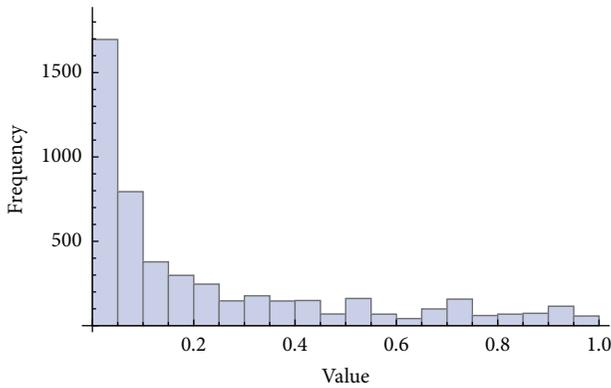


FIGURE 3: Histogram of the distribution of real numbers transferred into the range (0-1) generated by means of the chaotic Burgers map, 5000 samples.

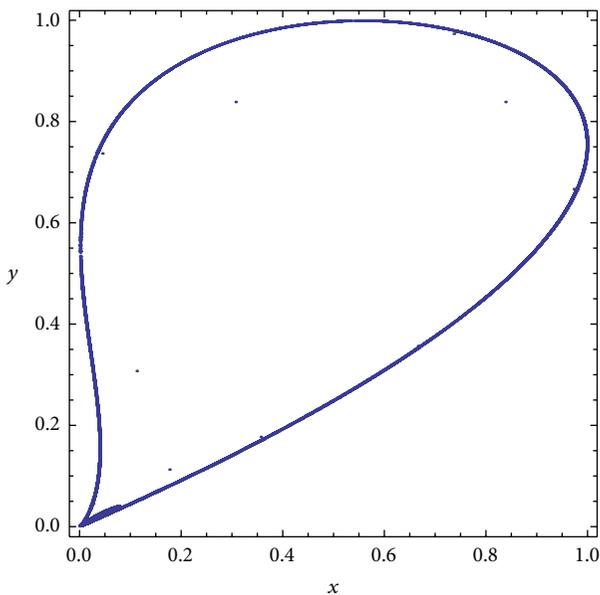


FIGURE 4:  $x, y$  plot of the delayed logistic map.

suppressed by the aforementioned inclusion of initial chaotic transient behavior. Consider

$$CF_{Simple} = \sum_{t=0}^{\tau_i} |TS_t - AS_t|, \quad (6)$$

where TS is target state and AS is actual state.

Different type of universal cost function without any selection rules is purely based on searching for the desired stabilized periodic orbit and thereafter calculation of the difference between desired and found actual periodic orbit on the short time interval,  $\tau_s$  (20 iterations), from the point where the first minimal value of difference between desired and actual system output is found (i.e., floating window for minimization; see Figure 10).

Such a design of universal CF should secure the successful stabilization of either p-1 orbit (stable state) or any higher periodic orbit anyway phase shifted. Furthermore, due to CF values converging towards zero, this CF also allows the use of decision rules, avoiding very time demanding simulations. This rule stops EA immediately, when the first individual with good parameter structure is reached; thus the value of CF is lower than the acceptable ( $CF_{acc}$ ) one. Based on the numerous experiments, typically  $CF_{acc} = 0.001$  at time interval  $\tau_s = 20$  iterations; thus the difference between desired and actual output has the value of 0.0005 per iteration, that is, successful stabilization for the used control technique. The  $CF_{UNI}$  has the following form:

$$CF_{UNI} = pen_1 + \sum_{t=\tau_1}^{\tau_2} |TS_t - AS_t|, \quad (7)$$

where  $\tau_1$  is the first min value of the difference between TS and AS and  $\tau_2$  is the end of optimization interval ( $\tau_1 + \tau_s$ ),  $pen_1 = 0$  if  $\tau_i - \tau_2 \geq \tau_s$ ;  $pen_1 = 10 * (\tau_i - \tau_2)$  if  $\tau_i - \tau_2 < \tau_s$  (i.e., late stabilization).

The issue of pure searching for periodic orbits causes very chaotic, erratic, and discrete type CF surfaces.

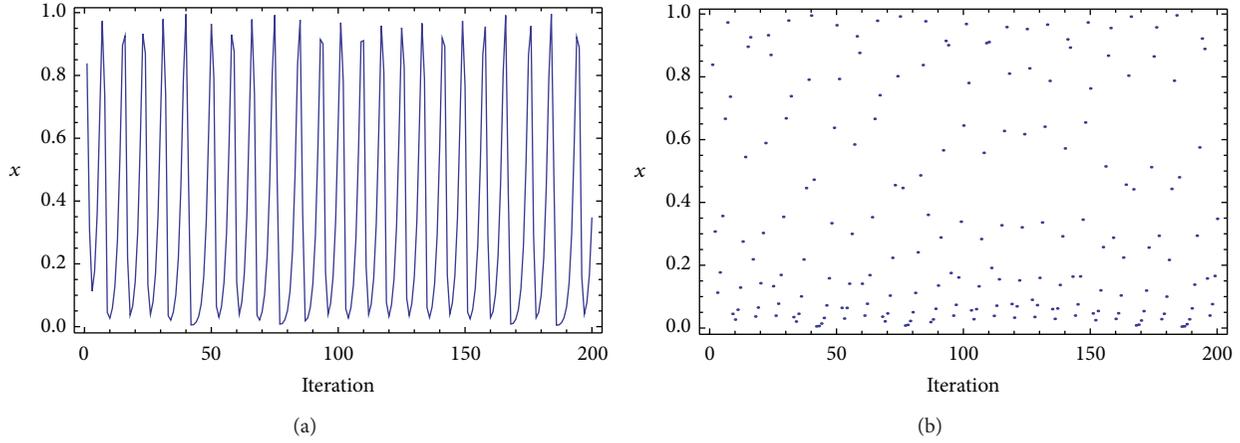


FIGURE 5: Iterations of the delayed logistic (variable  $x$ ) (a), iterations of the delayed logistic (variable  $x$ ), point-plot (b).

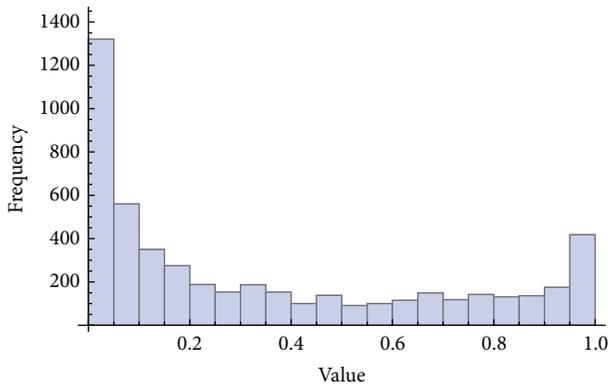


FIGURE 6: Histogram of the distribution of real numbers transferred into the range (0-1) generated by means of the chaotic delayed logistic map, 5000 samples.

## 6. Experimental Design

This research encompasses six case studies. Six different sets of discrete chaotic systems as CPRNGs/to be controlled and two different cost functions were combined in the following form:

- (i) case study 1: Burgers map as CPRNG/controlled system with  $CF_{Simple}$ ,
- (ii) case study 2: Burgers map as CPRNG/controlled system with  $CF_{UNI}$ ,
- (iii) case study 3: delayed logistic map as CPRNG/controlled system with  $CF_{Simple}$ ,
- (iv) case study 4: delayed logistic map as CPRNG/controlled system with  $CF_{UNI}$ ,
- (v) case study 5: Lozi map as CPRNG/controlled system with  $CF_{Simple}$ ,
- (vi) case study 6: Lozi map as CPRNG/controlled system with  $CF_{UNI}$ .

This work is focused on the utilization of the chaos driven DE for tuning of parameters for ET DAS control method

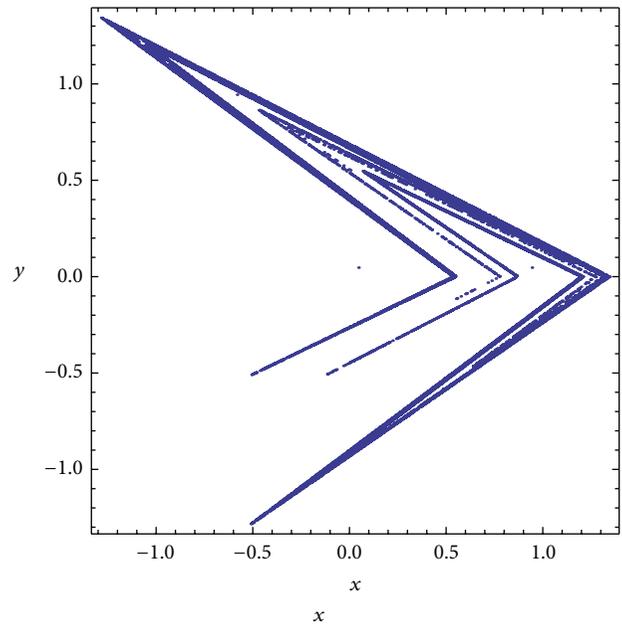


FIGURE 7:  $x, y$  plot of the Lozi map.

to stabilize desired unstable periodic orbits (UPO). In the described research, desired UPO was p-1 (stable state). The original control method, ET DAS, in the discrete form suitable for discrete chaotic maps has the following form:

$$F_n = K [(1 - R) S_{n-m} - x_n], \tag{8}$$

$$S_n = x_n + R S_{n-m},$$

where  $K$  and  $R$  are adjustable constants,  $F$  is the perturbation,  $S$  is given by a delay equation utilizing previous states of the system, and  $m$  is the period of  $m$ -periodic orbit to be stabilized. The perturbation  $F_n$  in (8) may have arbitrarily large value, which can cause the divergence of the system. Therefore,  $F_n$  should have a value between  $-F_{max}$  and  $F_{max}$ . The ranges of all evolutionary estimated parameters are given in Table 1.

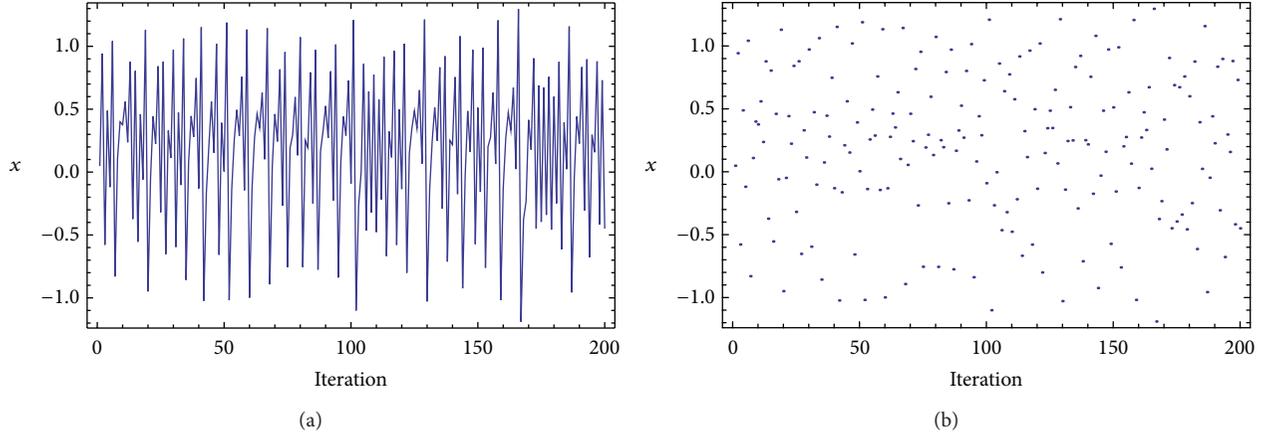


FIGURE 8: Iterations of the Lozi map (variable  $x$ ) (a), iterations of the Lozi map (variable  $x$ ), point-plot (b).

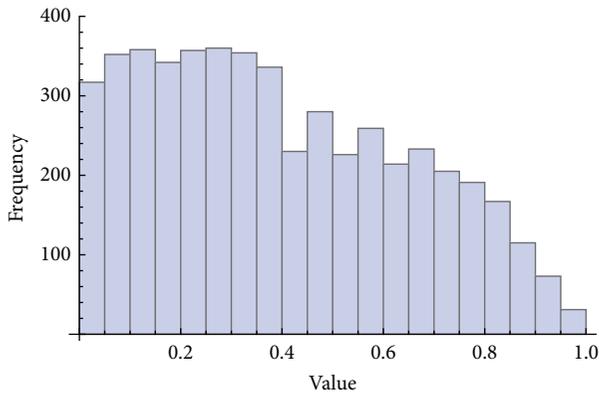


FIGURE 9: Histogram of the distribution of real numbers transferred into the range  $\langle 0-1 \rangle$  generated by means of the chaotic Lozi map, 5000 samples.

TABLE 1: Estimated parameters.

Parameter	Min	Max
$K$	-2	2
$R$	0	0.99
$F_{max}$	0	0.9

TABLE 2: DE settings.

Parameter	Value
Population size	25
$F$	0.8
Cr	0.8
Generations	250
Max cost function evaluations (CFE)	6250

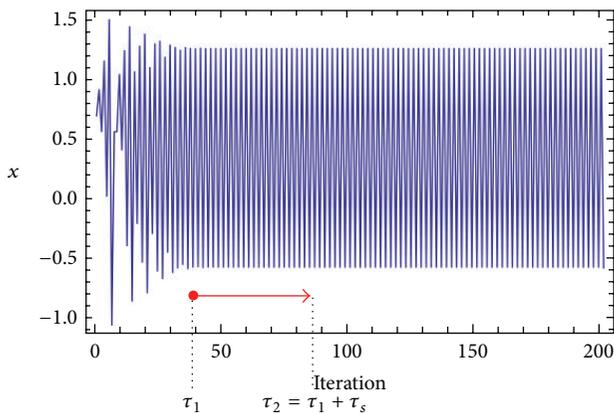


FIGURE 10: "Floating window" for minimization.

Within the research, a total number of 50 simulations with ChaosDE were carried out for each case study. The parameter settings for ChaosDE were obtained analytically based on numerous experiments and simulations (see

Table 2). Experiments were performed in an environment of *Wolfram Mathematica*; PRNG operations therefore used the built-in *Mathematica software* pseudorandom number generator. All experiments used different initialization; that is, different initial population was generated in each run of chaos driven DE.

### 7. Results

All simulations were successful and have given new optimal settings for ETDAS control method securing the fast stabilization of the chaotic system at required behaviour (p-1 orbit). Tables 4, 5, 6, 7, 8, and 9 contain the simple statistical overview of optimization/simulation results as well as the best founded individual solutions of parameters setup for ETDAS control method, corresponding final CF value, also, the Istab. Value representing the number of iterations required for stabilization on desired UPO, and further the average error between desired output value and real system output from the last 20 iterations.

Graphical simulation outputs of the best individual solutions for particular case studies are depicted in Figures 11, 13, and 15, whereas Figures 12, 14, and 16 show the simulation

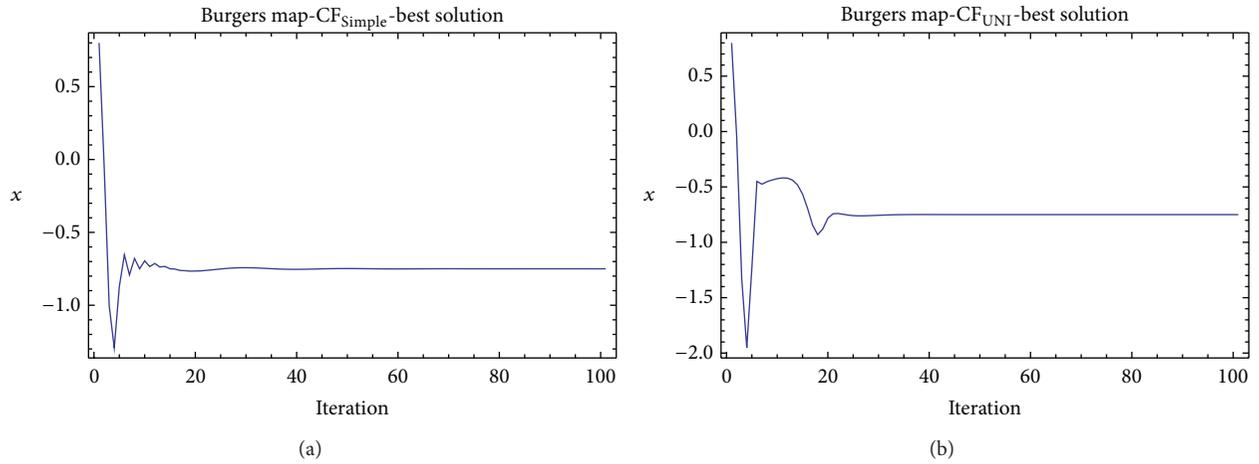


FIGURE 11: Simulation of the best individual solution, ChaosDE and Burgers map: case study 1,  $CF_{Simple}$  (a); case study 2,  $CF_{UNI}$  (b).

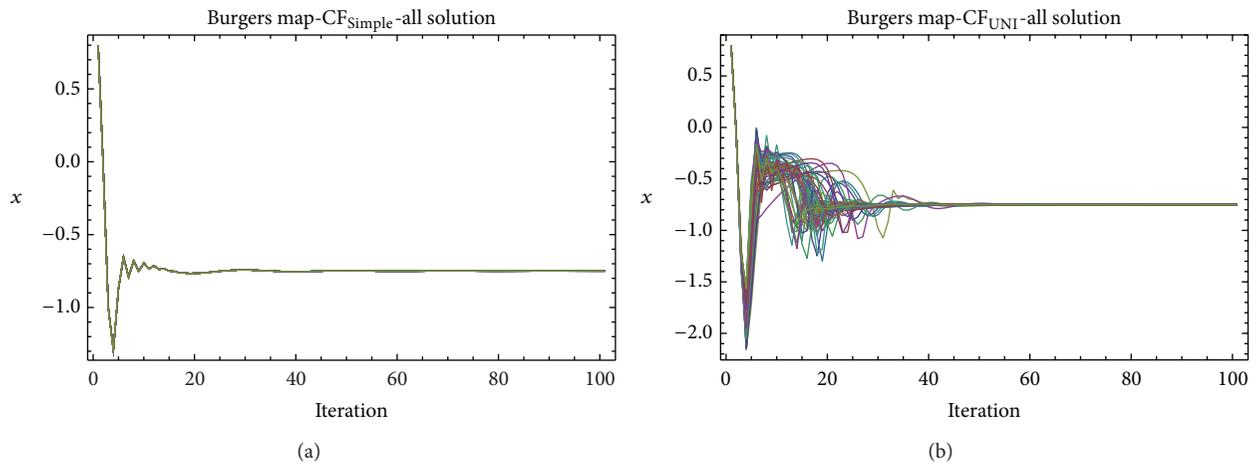


FIGURE 12: Simulation of all the 50 independent runs of EA, ChaosDE and Burgers map: case study 2,  $CF_{Simple}$  (a); case study 1,  $CF_{UNI}$  (b).

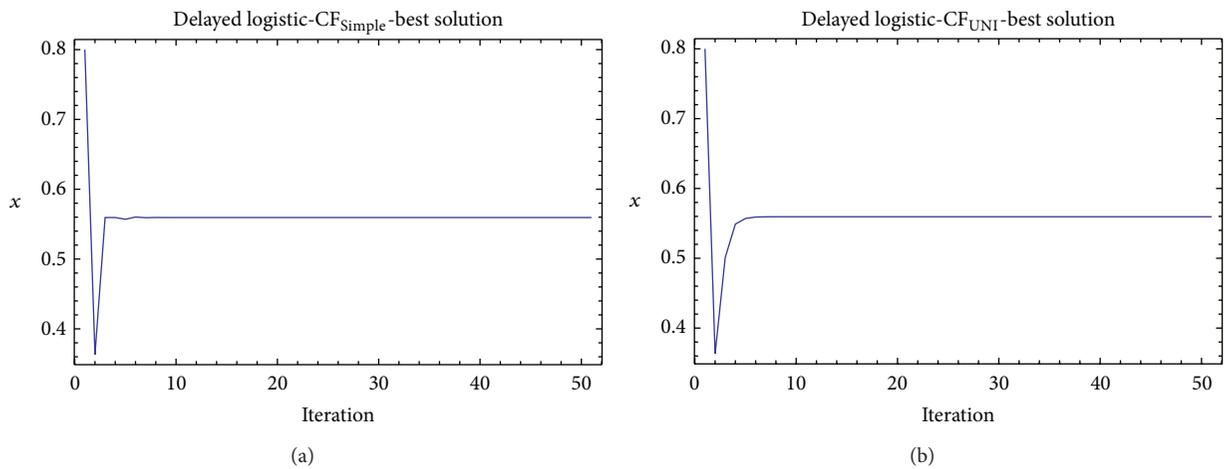


FIGURE 13: Simulation of the best individual solution, ChaosDE and delayed logistic: case study 3,  $CF_{Simple}$  (a); case study 4,  $CF_{UNI}$  (b).

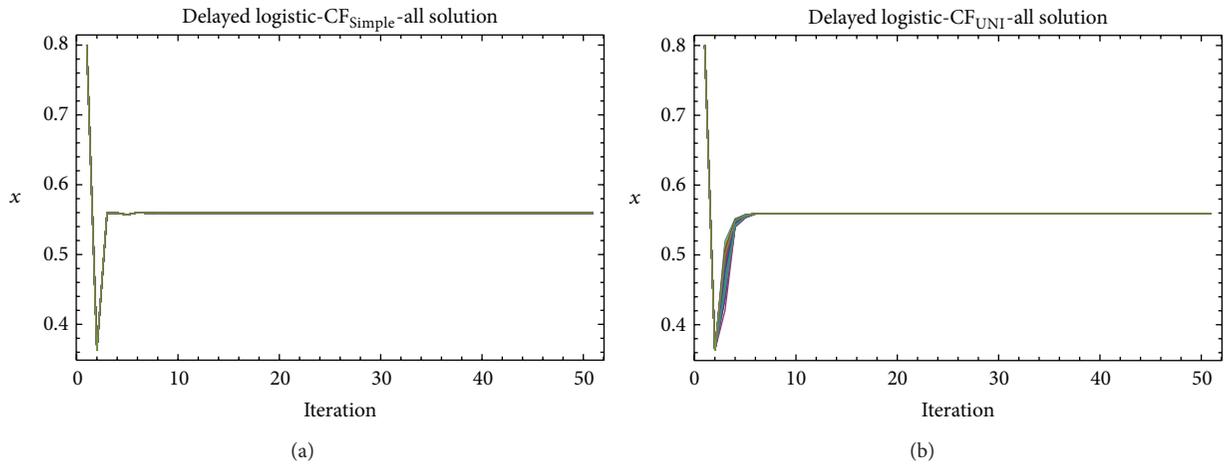


FIGURE 14: Simulation of all the 50 independent runs of EA, ChaosDE and Delayed logistic: case study 3,  $CF_{Simple}$  (a); case study 4,  $CF_{UNI}$  (b).

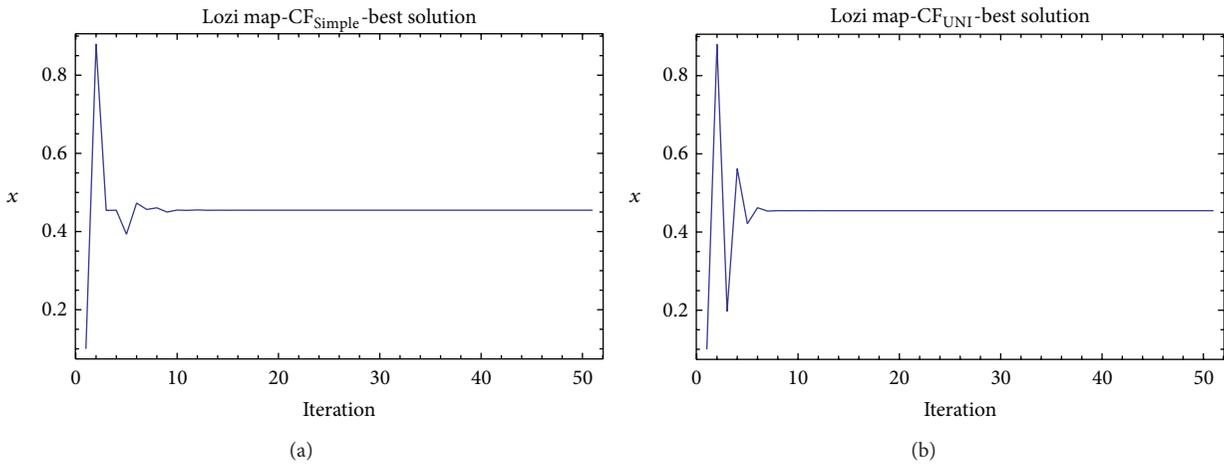


FIGURE 15: Simulation of the best individual solution, ChaosDE and Lozi map: case study 5,  $CF_{Simple}$  (a); case study 6,  $CF_{UNI}$  (b).

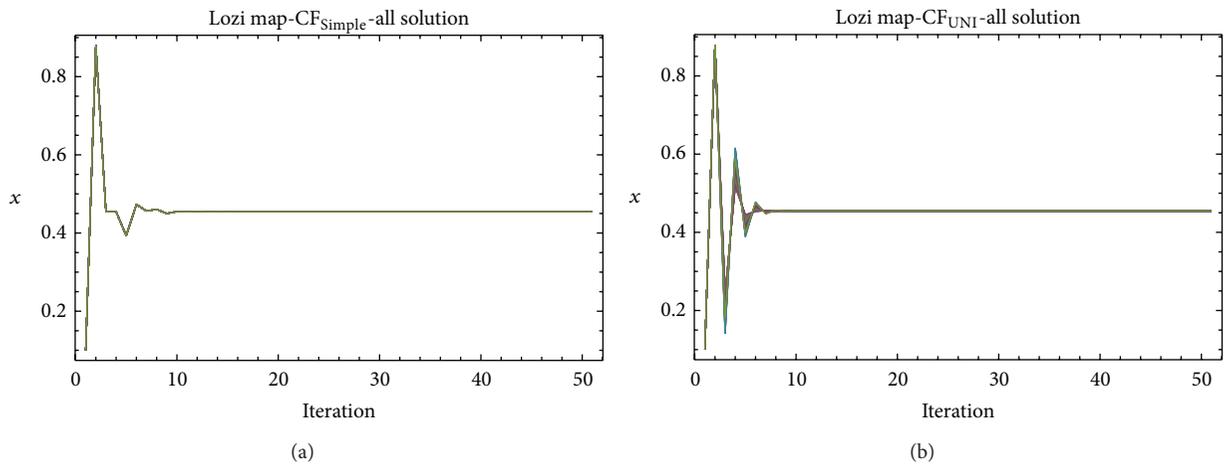


FIGURE 16: Simulation of the all 50 independent solution runs of EA, ChaosDE and Lozi map: case study 5,  $CF_{Simple}$  (a); case study 6,  $CF_{UNI}$  (b).

TABLE 3: The values for p-1 UPO.

Chaotic system	Values of p-1 UPO of unperturbed system
Burgers map	$x_F = -0.7499$
Delayed logistic map	$x_F = 0.559471$
Lozi map	$x_F = 0.454545$

TABLE 4: Simple CF statistics: joined case studies 1 and 2, Burgers map as CPRNG/controlled system.

Statistical data	Case study 1	Case study 2
	CF <sub>simple</sub> CF value	CF <sub>UNI</sub> CF value
Min	2.16199	$1.05 \cdot 10^{-6}$
Max	2.16199	0.0103
Average	2.16199	$6.67 \cdot 10^{-4}$
Median	2.16199	$5.32 \cdot 10^{-7}$
Std. dev.	$5.58 \cdot 10^{-11}$	$1.89 \cdot 10^{-3}$
Average full stab. (iteration)	45	35

output of all 50 runs of ChaosDE, thus confirming the robustness of this approach. For the illustrative purposes, all graphical simulations outputs are depicted only for the variable  $x$  of the chaotic systems.

The values for desired p-1 UPOs (fixed points) of unperturbed chaotic systems based on the mathematical analysis of the systems are given in Table 3.

From the results presented in the Tables 4–9, it follows that the CF-simple is very convenient for evolutionary process, which means that repeated runs of EA are giving identical optimal results (i.e., very close to the possible global extreme). This is graphically confirmed in Figures 12, 14, and 16, which show all 50 simulations. All the runs are basically merged into one line.

On the other hand, the disadvantage of including of initial chaotic transient behavior of not stabilized system into the cost function value and subsequent resulting very tiny change of control method setting for extremely sensitive chaotic system is causing suppression of stabilization speed and numerical precision.

Results obtained in the cases utilizing the CF<sub>UNI</sub> lend weight to the argument that the technique of pure searching for periodic orbits is advantageous for faster and more precise stabilization of chaotic system.

### 8. Conclusions

Based on obtained results, it may be claimed that the presented ChaosDE driven by selected discrete dissipative chaotic systems has given satisfactory results in the chaos control optimization issue.

The results show that the embedding of the chaotic dynamics in the form of chaotic pseudorandom number generator into the differential evolution algorithm may help to improve the performance and robustness of the DE. ChaosDE is able to obtain optimal solutions securing the very fast and precise stabilization for both convenient CF surface, in case of

TABLE 5: Characteristics of the best solution: joined case studies 1 and 2, Burgers map as CPRNG/controlled system.

Parameter	Case study 1	Case study 2
	CF <sub>simple</sub> Value	CF <sub>UNI</sub> Value
$K$	1.22847	0.732498
$F_{max}$	0.9	0.48495
$R$	0.574997	0.811742
CF value	2.16199	$1.05 \cdot 10^{-6}$
Istab. Value	45	25
Average error per iteration	$5.86 \cdot 10^{-5}$	$1.21 \cdot 10^{-8}$

TABLE 6: Simple CF statistics: joined case studies 3 and 4, delayed logistic as CPRNG/controlled system.

Statistical data	Case study 3	Case study 4
	CF <sub>simple</sub> CF value	CF <sub>UNI</sub> CF value
Min	0.199798	$2.3 \cdot 10^{-15}$
Max	0.199798	$2.7 \cdot 10^{-15}$
Average	0.199798	$2.44222 \cdot 10^{-15}$
Median	0.199798	$2.40551 \cdot 10^{-15}$
Std. dev.	$5.11 \cdot 10^{-16}$	$9.04629 \cdot 10^{-17}$
Average full stab. (iteration)	10	7.5

TABLE 7: Characteristics of the best solution: joined case studies 3 and 4, delayed logistic as CPRNG/controlled system.

Parameter	Case study 3	Case study 4
	CF <sub>simple</sub> Value	CF <sub>UNI</sub> Value
$K$	1.29837	1.31355
$F_{max}$	0.394579	0.336294
$R$	0.01	0.010219
CF value	0.199798	$2.3 \cdot 10^{-15}$
Istab. Value	10	8
Average error per iteration	$2.22 \cdot 10^{-17}$	0

TABLE 8: Simple CF statistics: joined case studies 5 and 6, Lozi map as CPRNG/controlled system.

Statistical data	Case study 5	Case study 6
	CF <sub>simple</sub> CF value	CF <sub>UNI</sub> CF value
Min	0.520639	$3.5331 \cdot 10^{-15}$
Max	0.520639	$4.0551 \cdot 10^{-15}$
Average	0.520639	$3.8063 \cdot 10^{-15}$
Median	0.520639	$3.6352 \cdot 10^{-16}$
Std. dev.	$2.41 \cdot 10^{-15}$	$1.19 \cdot 10^{-16}$
Average full stab. (iteration)	32	11

the CF-simple, as well as for the very chaotic and nonlinear CF surface in case of the CF-universal.

When comparing both the CF designs, the CF-simple is very convenient for evolutionary process (i.e., repeated

TABLE 9: Characteristics of the best solution: joined case studies 5 and 6, Lozi map as CPRNG/controlled system.

Parameter	Case study 5	Case study 6
	CF <sub>simple</sub> Value	CF <sub>UNI</sub> Value
$K$	-1.11259	-0.859989
$F_{\max}$	0.9	0.643099
$R$	0.289232	0.065669
CF value	0.520639	$3.5331 \cdot 10^{-15}$
Istab. Value	21	9
Average error per iteration	$7.21 \cdot 10^{-15}$	0

runs are giving identical optimal results), but it has many limitations.

The second universal CF design brings the possibility of using it problem free for any desired behavior of arbitrary chaotic systems but at the cost of the highly chaotic CF surface. Nevertheless, the embedding of the chaotic dynamics into the evolutionary algorithms helped to deal with such an issue.

The primary aim of this work was not to develop any new pseudorandom number generator, which should normally pass many statistical tests but to show that through embedding the hidden chaotic dynamics into the evolutionary process in the form of chaotic pseudorandom number generators may help to obtain better results and avoid problems connected with evolutionary computation such as premature convergence and stagnation in local extremes.

Future plans include testing of different chaotic systems, either manually or evolutionary tuning of parameters of chaotic maps, further complex comparisons with different heuristics and obtaining a large number of results to perform statistical tests.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work was supported by Grant Agency of the Czech Republic, GACR P103/13/08195S, partially supported by Grant of SGS nos. SP2014/159 and SP2014/170, VŠB-Technical University of Ostrava, Czech Republic and by the Development of Human Resources in Research and Development in Practice Project, Registry no. CZ.1.07/2.3.00/20.0072, funded by Operational Programme Education for Competitiveness, and cofinanced by ESF and state budget of the Czech Republic; it was further supported by the European Regional Development Fund under the Project CEBIA-Tech no. CZ.1.05/2.1.00/03.0089 and by the Internal Grant Agency of Tomas Bata University under the Project no. IGA/FAI/2014/010.

## References

- [1] E. Ott, C. Grebogi, and J. A. Yorke, "Controlling chaos," *Physical Review Letters*, vol. 64, no. 11, pp. 1196–1199, 1990.
- [2] I. Zelinka, "Real-time deterministic chaos control by means of selected evolutionary techniques," *Engineering Applications of Artificial Intelligence*, vol. 22, no. 2, pp. 283–297, 2009.
- [3] I. Zelinka, R. Senkerik, and E. Navratil, "Investigation on evolutionary optimization of chaos control," *Chaos, Solitons and Fractals*, vol. 40, no. 1, pp. 111–129, 2009.
- [4] R. Senkerik, I. Zelinka, D. Davendra, and Z. Oplatkova, "Utilization of SOMA and differential evolution for robust stabilization of chaotic Logistic equation," *Computers & Mathematics with Applications*, vol. 60, no. 4, pp. 1026–1037, 2010.
- [5] K. Pyragas, "Control of chaos via extended delay feedback," *Physics Letters A*, vol. 206, no. 5–6, pp. 323–330, 1995.
- [6] W. Just, "Principles of time delayed feedback control," in *Handbook of Chaos Control*, H. G. Schuster, Ed., Wiley-Vch, 1999.
- [7] R. Senkerik, Z. Oplatkova, I. Zelinka, and D. Davendra, "Synthesis of feedback controller for three selected chaotic systems by means of evolutionary techniques: analytic programming," *Mathematical and Computer Modelling*, vol. 57, no. 1–2, pp. 57–67, 2013.
- [8] Z. Kominkova Oplatkova, R. Senkerik, I. Zelinka, and M. Pluhacek, "Analytic programming in the task of evolutionary synthesis of a controller for high order oscillations stabilization of discrete chaotic systems," *Computers and Mathematics with Applications*, vol. 66, no. 2, pp. 177–189, 2013.
- [9] K. Pyragas, "Continuous control of chaos by self-controlling feedback," *Physics Letters A*, vol. 170, no. 6, pp. 421–428, 1992.
- [10] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, vol. 4, pp. 1942–1948, Perth, Wash, USA, December 1995.
- [11] M. Abedini, R. Vatankhah, and N. Assadian, "Stabilizing chaotic system on periodic orbits using multi-interval and modern optimal control strategies," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 10, pp. 3832–3842, 2012.
- [12] M. Sadeghpour, H. Salarieh, G. Vossoughi, and A. Alasty, "Multi-variable control of chaos using PSO-based minimum entropy control," *Communications in Nonlinear Science and Numerical Simulation*, vol. 16, no. 6, pp. 2397–2404, 2011.
- [13] L. D. S. Coelho and R. B. Grebogi, "Chaotic synchronization using PID control combined with population based incremental learning algorithm," *Expert Systems with Applications*, vol. 37, no. 7, pp. 5347–5352, 2010.
- [14] M. J. Shirazi, R. Vatankhah, M. Boroushaki, H. Salarieh, and A. Alasty, "Application of particle swarm optimization in chaos synchronization in noisy environment in presence of unknown parameter uncertainty," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 2, pp. 742–753, 2012.
- [15] I. Zelinka and A. Raidl, "Evolutionary synchronization of chaotic systems," in *Evolutionary Algorithms and Chaotic Systems*, I. Zelinka, S. Celikovskiy, H. Richter, and G. Chen, Eds., vol. 267, pp. 385–407, Springer, Berlin, Germany, 2010.
- [16] H. Richter and K. J. Reinschke, "Optimization of local control of chaos by an evolutionary algorithm," *Physica D: Nonlinear Phenomena*, vol. 144, no. 3–4, pp. 309–334, 2000.
- [17] I. Zelinka, G. Chen, and S. Celikovskiy, "Chaos synthesis by means of evolutionary algorithms," *International Journal of*

- Bifurcation and Chaos in Applied Sciences and Engineering*, vol. 18, no. 4, pp. 911–942, 2008.
- [18] I. Zelinka, M. Chadli, D. Davendra, R. Senkerik, and R. Jasek, “An investigation on evolutionary reconstruction of continuous chaotic systems,” *Mathematical and Computer Modelling*, vol. 57, no. 1–2, pp. 2–15, 2013.
- [19] I. Aydin, M. Karakose, and E. Akin, “Chaotic-based hybrid negative selection algorithm and its applications in fault and anomaly detection,” *Expert Systems with Applications*, vol. 37, no. 7, pp. 5285–5294, 2010.
- [20] R. Caponetto, L. Fortuna, S. Fazzino, and M. G. Xibilia, “Chaotic sequences to improve the performance of evolutionary algorithms,” *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 3, pp. 289–304, 2003.
- [21] D. Davendra, I. Zelinka, and R. Senkerik, “Chaos driven evolutionary algorithms for the task of PID control,” *Computers and Mathematics with Applications*, vol. 60, no. 4, pp. 1088–1104, 2010.
- [22] K. V. Price, R. M. Storn, and J. A. Lampinen, *Differential Evolution—a Practical Approach to Global Optimization*, Natural Computing Series, Springer, Berlin, Germany, 2005.
- [23] I. Zelinka, “SOMA—self-organizing migrating algorithm,” in *New Optimization Techniques in Engineering*, vol. 141 of *Studies in Fuzziness and Soft Computing*, pp. 167–217, Springer, Berlin, Germany, 2004.
- [24] L. D. S. Coelho and V. C. Mariani, “A novel chaotic particle swarm optimization approach using Hénon map and implicit filtering local search for economic load dispatch,” *Chaos, Solitons & Fractals*, vol. 39, no. 2, pp. 510–518, 2009.
- [25] R. Senkerik, D. Davendra, I. Zelinka, M. Pluhacek, and Z. Oplatkova, “An investigation on the chaos driven differential evolution: an initial study,” in *Proceedings of the 5th International Conference on Bioinspired Optimization Methods and Their Applications (BIOMA '12)*, pp. 185–194, 2012.
- [26] D. Davendra, I. Zelinka, R. Senkerik, and M. Bialic-Davendra, “Chaos driven evolutionary algorithm for the traveling salesman problem,” in *Traveling Salesman Problem, Theory and Applications*, D. Davendra, Ed., InTech, 2010.
- [27] D. Davendra, M. Bialic-Davendra, and R. Senkerik, “Scheduling the lot-streaming flowshop scheduling problem with setup time with the chaos-induced enhanced differential evolution,” in *Proceedings of the IEEE Symposium on Differential Evolution (SDE '13)*, pp. 119–126, Singapore, April 2013.
- [28] M. Pluhacek, R. Senkerik, D. Davendra, Z. Kominkova Oplatkova, and I. Zelinka, “On the behavior and performance of chaos driven PSO algorithm with inertia weight,” *Computers and Mathematics with Applications*, vol. 66, no. 2, pp. 122–134, 2013.
- [29] M. Pluhacek, R. Senkerik, I. Zelinka, and D. Davendra, “Chaos PSO algorithm driven alternately by two different chaotic maps—an initial study,” in *Proceedings of the IEEE Congress on Evolutionary Computation (CEC '13)*, pp. 2444–2449, June 2013.
- [30] M. Pluhacek, R. Senkerik, and I. Zelinka, “Multiple choice strategy based PSO algorithm with chaotic decision making—a preliminary study,” in *International Joint Conference SOCO'13-CISIS'13-ICEUTE'13*, Á. Herrero, B. Baroque, F. Klett et al., Eds., vol. 239 of *Advances in Intelligent Systems and Computing*, pp. 21–30, 2014.
- [31] L. D. S. Coelho and V. C. Mariani, “Firefly algorithm approach based on chaotic Tinkerbell map applied to multivariable PID controller tuning,” *Computers and Mathematics with Applications*, vol. 64, no. 8, pp. 2371–2382, 2012.
- [32] D. Davendra, “Evolutionary algorithms and the edge of chaos,” in *Evolutionary Algorithms and Chaotic Systems*, I. Zelinka, S. Celikovskiy, H. Richter, and G. Chen, Eds., vol. 267, pp. 145–161, Springer, Berlin, Germany, 2010.
- [33] R. Senkerik, M. Pluhacek, I. Zelinka, D. Davendra, Z. Oplatkova, and R. Jasek, “Evolutionary control of chaotic Lozi map by means of chaos driven differential evolution,” in *AETA 2013: Recent Advances in Electrical Engineering and Related Sciences*, I. Zelinka, V. H. Duy, and J. Cha, Eds., vol. 282 of *Lecture Notes in Electrical Engineering*, pp. 371–380, Springer, Berlin, Germany, 2014.
- [34] R. Senkerik, I. Zelinka, M. Pluhacek, and Z. Kominkova Oplatkova, “Evolutionary control of chaotic burgers map by means of chaos enhanced differential evolution,” *International Journal of Mathematics and Computers in Simulation*, vol. 8, pp. 39–45, 2014.
- [35] R. Senkerik, M. Pluhacek, I. Zelinka, D. Davendra, Z. Oplatkova, and R. Jasek, “Chaos driven differential evolution in the task of evolutionary control of delayed logistic chaotic system,” in *Proceedings of the 7th Global Conference on Power Control and Optimization (PCO '13)*, 2013.
- [36] M. F. Tasgetiren, P. N. Suganthan, and Q. K. Pan, “An ensemble of discrete differential evolution algorithms for solving the generalized traveling salesman problem,” *Applied Mathematics and Computation*, vol. 215, no. 9, pp. 3356–3368, 2010.
- [37] G. C. Onwubolu and D. Davendra, Eds., *Differential Evolution: A handbook for Permutation-based Combinatorial Optimization*, vol. 175 of *Studies in Computational Intelligence*, Springer, Berlin, Germany, 2009.
- [38] L. Wang, H. Qu, T. Chen, and F. P. Yan, “An effective hybrid self-adapting differential evolution algorithm for the joint replenishment and location-inventory problem in a three-level supply chain,” *The Scientific World Journal*, vol. 2013, Article ID 270249, 11 pages, 2013.
- [39] S. Das, A. Abraham, U. K. Chakraborty, and A. Konar, “Differential evolution using a neighborhood-based mutation operator,” *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 3, pp. 526–553, 2009.
- [40] T. J. Choi, C. W. Ahn, and J. An, “An adaptive cauchy differential evolution algorithm for global numerical optimization,” *The Scientific World Journal*, vol. 2013, Article ID 969734, 12 pages, 2013.
- [41] A. K. Qin, V. L. Huang, and P. N. Suganthan, “Differential evolution algorithm with strategy adaptation for global numerical optimization,” *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 2, pp. 398–417, 2009.
- [42] J. Zhang and A. C. Sanderson, “JADE: self-adaptive differential evolution with fast and reliable convergence performance,” in *Proceeding of the 2007 IEEE Congress on Evolutionary Computation (CEC '07)*, pp. 2251–2258, Singapore, September 2007.
- [43] J. Zhang and A. C. Sanderson, “Self-adaptive multi-objective differential evolution with direction information provided by archived inferior solutions,” in *Proceedings of the IEEE Congress on Evolutionary Computation (CEC '08)*, pp. 2801–2810, Hong Kong, China, June 2008.
- [44] R. Senkerik, D. Davendra, I. Zelinka, and Z. Oplatkova, “Chaos driven differential evolution in the task of chaos control optimization,” in *Proceedings of the IEEE Congress on Evolutionary Computation (CEC '10)*, July 2010.
- [45] R. Storn and K. Price, “Differential evolution—a simple and efficient heuristic for global optimization over continuous

- spaces,” *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.
- [46] K. V. Price, “An introduction to differential evolution,” in *New Ideas in Optimization*, D. Corne, M. Dorigo, and F. Glover, Eds., McGraw-Hill, 1999.
- [47] J. C. Sprott, *Chaos and Time-Series Analysis*, Oxford University Press, New York, NY, USA, 2003.
- [48] M. A. Aziz-Alaoui, C. Robert, and C. Grebogi, “Dynamics of a Hénon-Lozi-type map,” *Chaos, Solitons & Fractals*, vol. 12, no. 12, pp. 2323–2341, 2001.

## Review Article

# Energy Saving in Data Processing and Communication Systems

**Giuseppe Iazeolla and Alessandra Pieroni**

*Department of Enterprise Strategy and Applied Technology, "Guglielmo Marconi" University of Study, Roma, Italy*

Correspondence should be addressed to Alessandra Pieroni; [a.pieroni@unimarconi.it](mailto:a.pieroni@unimarconi.it)

Received 16 January 2014; Accepted 20 February 2014; Published 29 June 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 G. Iazeolla and A. Pieroni. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The power management of ICT systems, that is, data processing (Dp) and telecommunication (Tlc) systems, is becoming a relevant problem in economical terms. Dp systems totalize millions of servers and associated subsystems (processors, monitors, storage devices, etc.) all over the world that need to be electrically powered. Dp systems are also used in the government of Tlc systems, which, besides requiring Dp electrical power, also require *Tlc-specific* power, both for *mobile* networks (with their cell-phone towers and associated subsystems: base stations, subscriber stations, switching nodes, etc.) and for *wired* networks (with their routers, gateways, switches, etc.). ICT research is thus expected to investigate into methods to reduce Dp- and Tlc-specific power consumption. However, saving power may turn into waste of performance, in other words, into waste of ICT quality of service (QoS). This paper investigates the Dp and Tlc power management policies that look at compromises between power saving and QoS.

## 1. Introduction

The growth in ICT energy consumption is driven by the growth of demand for greater data processing (Dp) and larger access to telecommunications (Tlc), within almost every organization.

This growth has a number of important implications, including [1]

- (i) increased energy costs for business and government,
- (ii) increased emissions, including greenhouse gases, from electricity generation,
- (iii) increased strain on the existing power grid to meet the increased electricity demand,
- (iv) increased capital costs for expansion of data center capacity and construction of new data centers,
- (v) increased capital costs for expansion of wired and wireless access to communications.

Making 100 the total electrical power consumption for the ICT, around 14% is taken by the mobile Tlc, 74% by the wired Tlc, and the remaining 12% by the Dp technology.

Dp, however, is only apparently the less powered sector, since Tlc is itself a Dp consumer, and so any effort to

reduce the Dp power consumption may produce cascade effects that also reduce the Tlc one. Studying ways to save Dp power is thus central to any study for ICT power control and optimization.

In the US, power absorbed by data centers is estimated in more than 100 billion kW, for an expenditure of \$ 8 billion a year that corresponds to the expenditure in electricity of about 17 million homes [2].

This same US-local data-center problem becomes a global one when seeing power consumption by web companies, say Google, Yahoo, and so forth. The number of Google servers will reach an estimated 2,376,640 units by the end of 2013 [3].

Assuming a busy server absorbs around 240 W of power, Google will need about 600 MW of electrical power by the end of year 2013.

In Tlc systems, about 90% of Tlc-specific power consumption is concentrated in the routers. The links only absorb 10%. Current routers consume between 0.01 and 0.1 W/Mbps [4].

IT research is thus expected to investigate methods to reduce power absorbed by Dp and Tlc systems. To do that, one may decide to adopt policies to periodically switch off Dp servers or Tlc routers when they are in an idle state.

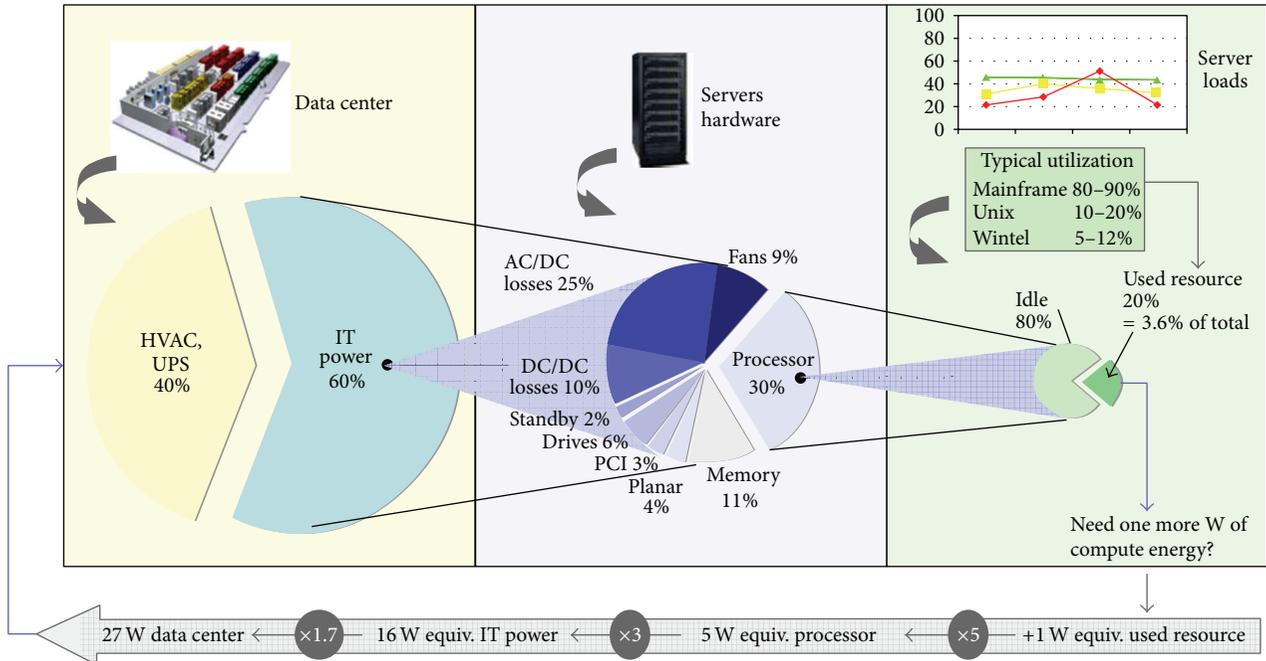


FIGURE 1: Data Center energy consumption sources [25].

Such policies, however, are to be sufficiently intelligent not to degrade the system quality of service (QoS). Indeed, returning an off server or an off router to its on state requires spending a nonnegligible amount of setup time that makes the server or router slower to respond to customer requests. This may turn into low-quality services such as low response to web queries and unsatisfactory VoIP communications and streaming of data. Any research in power management should thus look at compromises between power saving and QoS.

In this paper, Section 2 studies Dp power management policies and Section 3 studies Tlc power management policies.

## 2. Power Management in Dp Systems

Data centers have become common and essential to the functioning of business, communications, academic, and governmental systems.

During the past years, increasing demand for Dp services has led to significant growth in the number of data centers, along with an estimated doubling in the energy used by servers and the power and cooling infrastructure that supports them.

Figure 1 illustrates the way energy is spent in data centers. Heating and services for ventilation and air conditioning/backup (HVAC/UPS) absorb around 40% of electrical energy and Dp services the remaining 60%. The latter is in turn divided between AC/DC losses (25%), DC/DC losses (10%), fans, drives, PCI, and so forth, and memory consumptions (for a total 35%) and the remaining 30% is consumption in server processors.

In other words, the processors consumption totalizes  $0.30 \times 0.60$ , that is, 20% of total data center consumption. Such an amount, even though apparently negligible with respect to the total, is the main cause of the remaining 80%. Thus, any effort to reduce the processors 20% may produce cascade effects that also reduce the remaining 80%.

Figure 2 shows that 1 W savings at servers component level (processor, memory, hard disk, etc.) create a reduction in data center energy consumption of approximately 2.84 W.

For this reason, any research in ICT power saving should concentrate on policy to reduce Dp consumption at server components level.

Data centers can be seen as composed of a number of servers that can be organized into single farms or multifarms. In the following, Section 2.1 sees power saving policies in the single-farm case and Section 2.2 at the multifarm one.

**2.1. Energy Saving in Single-Farm Data Centers.** Most of power absorbed by the servers of a farm is wasted, since servers are busy (i.e., making processing work) only 20% to 30% of the time, on average. So, energy saving requires the adoption of management policies to avoid powering the servers when they are not processing. In other words, policies decide in which state (idle or off) to keep the servers when not busy. Two types of server management policies will be considered: static and dynamic policies.

**2.1.1. Energy Saving with Static Policies.** One may assume that a busy server in the on state absorbs around 240 W ( $P_{ON}$ ), an idle server about 160 W ( $P_{IDLE}$ ), and an off server 0 W ( $P_{OFF}$ ). So why not keep in the idle state or in the off state the servers when not busy? Just since switching a server from off to on consumes a time overhead. Thus, a power-saving policy

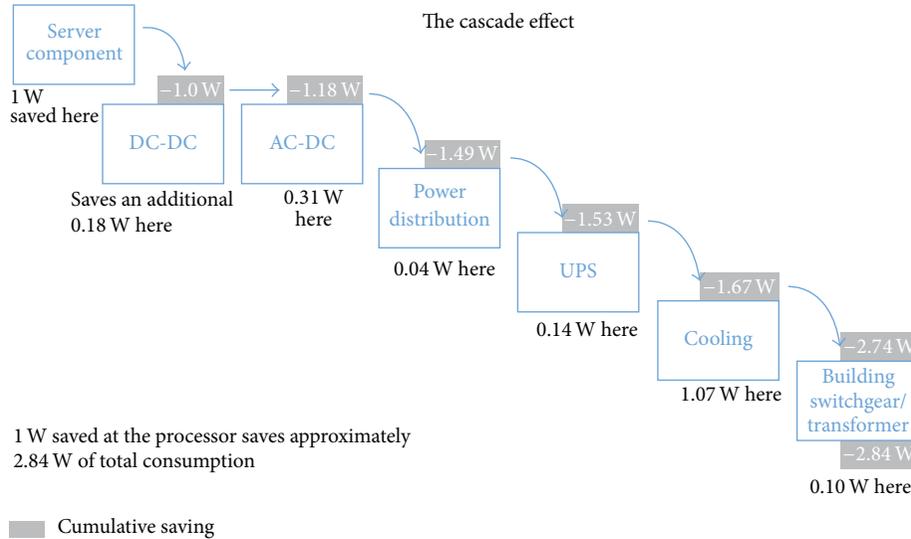


FIGURE 2: Cascade effect of energy saving in data centers [26].

may result in a time-wasting problem. As a consequence, the servers may lose performance (e.g., increased response time to the incoming jobs, lower throughput of communication packets, etc.) and its service may become unacceptable to customers.

To turn on an off server, we must first put the server in setup mode. During the setup period, the server cannot process jobs. The time spent in setup is called setup time. In [5] the authors consider server farms with a setup cost. Setup costs always take the form of a time delay, and sometimes there is also a power penalty, since during that entire period the server consumes the same power as being in the on state.

In [5] three different policies are studied to manage server farms: On/Idle policy, On/Off policy, and On/Off/Stag policy.

Under the On/Idle policy, servers are never turned off. All servers are either on or idle and remain in the idle mode when there are no jobs to serve. Assume that the farm consists of  $n$  servers; if an arrival finds a server idle, it starts serving on the idle server. Arrival that finds all  $n$  servers on busy, joins a central queue from which the servers pick jobs when they become idle.

The On/Off policy consists instead of immediately turning off the servers when not in use. As said above, however, there is a setup cost (in terms of time delay and of additional power penalty) for turning on an off server.

Finally, the On/Off/Stag policy is the same as the On/Off one, except that at most 1 server can be in setup at any point of time. This policy is known as the “staggered boot up” policy in data centers or “staggered spin up” in disk farms [5–7].

Figure 3(a) compares the On/Off and On/Idle policies for an example case.

The On/Idle policy proves to be better in terms of response time, because the incoming jobs do not suffer from setup time delays but involves a larger amount of power waste with respect to the On/Off policy, because of the amount of power an idle server absorbs.

Figure 4 compares the three server management policies in a farm consisting of  $k = 10$  servers, when the average setup

time changes from 1 to 100 sec and the average processing load  $\lambda$  (i.e., average job arrival rate) from 1 to 7 job/sec. The mean job size (service time) is assumed to be 1 sec.

Comparison is on the basis of the resulting mean response time  $E[T]$  to the incoming jobs and the average power consumption  $E[P]$ .

In the On/Idle case, when  $\lambda$  is low, there is no waiting and thus the mean response time  $E[T]$  is of about the mean job service time (1 sec) and increases for increasing  $\lambda$ .

A similar trend can be observed for the On/Off/Stag policy, since

$$E[T]_{\text{ON/OFF/STAG}} = E[T]_{\text{ON/IDLE}} + E[\text{setup time}], \quad (1)$$

as shown in [5].

For the On/Off policy, instead, the response time curve follows a bathtub behavior.

When the load  $\lambda$  is low, the mean response time is high, since almost every arrival finds servers in the off state, and thus every job incurs the setup time. For medium loads, servers are less frequently switched to the off condition and thus jobs are more likely served by available servers in the on state and do not incur in setup times. For high loads, finally, the mean response time increases due to large queuing in the system.

For the power consumption, one can show [5] that  $E[P]_{\text{ON/OFF/STAG}} < E[P]_{\text{ON/OFF}}$ , since at most one server can be in setup for the On/Off/Stag policy. There also results  $E[P]_{\text{ON/OFF}} < E[P]_{\text{ON/IDLE}}$ , since servers are turned off in the On/Off case. However, for loads  $\lambda$  above the medium, there results  $E[P]_{\text{ON/OFF}} > E[P]_{\text{ON/IDLE}}$  for medium setup time; that is,  $E[\text{setup time}] = 10$  sec, while for large setup time, that is,  $E[\text{setup time}] = 100$  sec, there always results  $E[P]_{\text{ON/OFF}} > E[P]_{\text{ON/IDLE}}$ , because of the large amount of power wasted in turning servers on in the On/Off policy.

Table 1 gives a synthetic comparison of the three considered policies in terms of response time and power consumption.

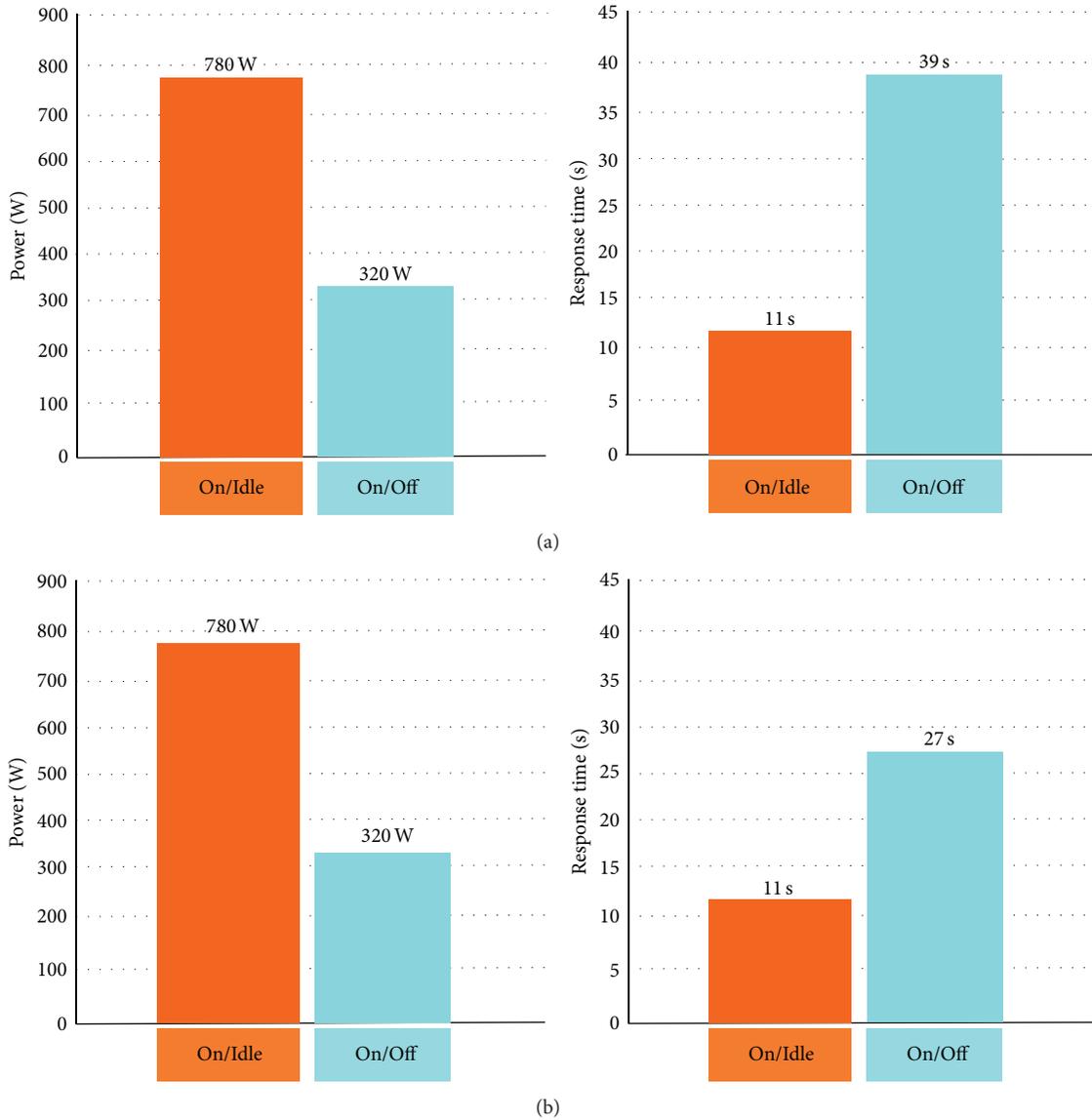


FIGURE 3: (a) Experimental results with 4 servers, utilization = 30% average setup time = 200 sec; average job size (service time) = 7 sec [2]. (b) Experimental results with same parameters of (a) except for SPTF or SJF queueing discipline.

TABLE 1: Synthetic view of the On/Off, On/Idle and On/Off/Stag power optimization policies.

	Response time	Power consumption
On/Idle	Small response times	High waste of power
On/Off	Medium response times	Medium waste for low setup times, high for increasing setup
On/Off/Stag	Large response times	Low waste of power

In conclusion, any reduction in power consumption is paid by an increase in response times. So, why not adopt queueing disciplines that minimize average response times? The SPTF (shortest processing time first) [8] or SJF (shortest

job first) [9] queueing discipline is known to perform better than the common FIFO (first in first out). Its use can then reduce the amount to pay in terms of response time to obtain a given power saving.

Figure 3(a) (that illustrates the FIFO queueing case) shows that to reduce the power consumption from 780 to 320 W we have to pay an increase from 11 to 39 sec in average response time.

Figure 3(b) illustrates that if the SPTF discipline is used instead; the debt to pay in response time is much smaller (from 39 to 27 sec) as proved in our simulations studies [10].

2.1.2. Seeking the Optimal  $(\pi, \delta)$  Strategy for the Single-Farm Data Centers. Under the assumption of Poisson arrivals, exponential service times, and deterministic setup times, authors in [11] prove that the optimal, or nearly optimal,

TABLE 2: Server farm results for low *setup*time ( $E[\text{setup}] = 1$  sec).

$\pi\delta$	$E[P](\pi, \delta)$ (W)		$E[T](\pi, \delta)$ (sec)		ERP( $\pi, \delta$ )	
	$\rho \leq 0.5$	$\rho \rightarrow 1$	$\rho \leq 0.5$	$\rho \rightarrow 1$	$\rho \leq 0.5$	$\rho \rightarrow 1$
On/Idle FIFO	6000	7100	1	1.8	6000	12780
On/Idle SPTF	6000	7100	1	1.3	6000	9230
On/Off FIFO	4200	7100	1.2	2	<b>5040</b>	14200
On/off SPTF	4200	7100	1.2	1.35	<b>5040</b>	9585

combination of  $(\pi, \delta)$ , with  $\pi$  being one policy from the set {On/Off, On/Idle} and  $\delta$  one queueing discipline from the set {FIFO, LIFO, RAND}, means minimizing a new metric called ERP (energy-response time product) and is defined as

$$\text{ERP}(\pi, \delta) = E[P](\pi, \delta) \times E[T](\pi, \delta). \quad (2)$$

Minimizing ERP( $\pi, \delta$ ) can be seen as maximizing the performance per Watt, with performance defined as the inverse of the mean response time [11].

In other words, according to their results, there is no need to consider other policies than the On/Off and the On/Idle policies.

They, however, only study the effects of moving from one policy  $\pi$  to another, without paying attention to the effects of also moving from a  $\delta = \text{FIFO}$  discipline to another time-independent discipline.

Under the FIFO assumption, however, they find that the On/Idle policy is typically superior to the remaining two in terms of ERP( $\pi, \delta$ ).

Our aim is to extend such results by studying the effects of the queueing discipline  $\delta$ , both on the ERP( $\pi, \delta$ ) index and on the  $E[P](\pi, \delta)$  and  $E[T](\pi, \delta)$  indices separately.

More precisely, the following four  $(\pi, \delta)$  strategies are investigated in the paper:

- (1) (On/Idle, FIFO),
- (2) (On/Idle, SPTF),
- (3) (On/Off, FIFO),
- (4) (On/Off, SPTF).

And for each of such strategy the ERP( $\pi, \delta$ ) product is studied besides the  $E[P](\pi, \delta)$  and  $E[T](\pi, \delta)$  indices. Two largely different setup times ( $E[\text{setup}] = 1$  s and  $E[\text{setup}] = 100$  s) will be used to stress the effect of the setup time on the On/Idle and On/Off polices.

Similarly, two largely different farm data center loads, low  $\rho$  ( $\rho \leq 0.5$ ) and high  $\rho$  ( $\rho \rightarrow 1$ ), will be used to stress the effect of the queueing disciplines.

The following farm data center characteristics are assumed: server mean setup time  $E[\text{setup}] = 1$  sec (or 100 sec), server  $P_{\text{ON}} = 240$  W, server  $P_{\text{SETUP}} = 240$  W, server  $P_{\text{IDLE}} = 150$  W, server  $P_{\text{OFF}} = 0$  W, mean job service time  $S = 1$  sec, and  $n = 30$  servers.

Table 2 shows simulation results [10] that compare the Dp power and QoS indices in the low setup case ( $E[T\text{setup}] = 1$  sec).

- (i) Seeing at the power consumption  $E[P]$ , we note that there is no effect by the queueing discipline  $\delta$  on the

power consumption  $E[P]$ , while there is an effect by the policy  $\pi$  for low  $\rho$ . Indeed, a drastic reduction can be seen (from 6000 W to 4200 W, for low  $\rho$ ) when moving from On/Idle to On/Off, since when  $\rho$  is low, the waiting queue is almost empty and thus a large number of servers is in the off state. For high  $\rho$  instead, the power consumption  $E[P]$  remains unchanged ( $E[P] = 7100$  W) with the discipline  $\delta$ , since the queue is always full and thus the servers remain always in the on state.

- (ii) Seeing at the response time  $E[T]$ , we note that there is an effect both by the queueing discipline  $\delta$  and by the policy  $\pi$ . The effects hold both for low  $\rho$  and for high  $\rho$ . In the On/Idle case, when  $\rho$  is low, there is no waiting in the Dp queue and thus the mean response time  $T$  is about the mean job service time ( $S = 1$  sec), while it increases ( $E[T] = 1.8$  sec for  $\delta = \text{FIFO}$  and  $E[T] = 1.3$  sec for  $\delta = \text{SPTF}$ ) for high  $\rho$ .

In the On/Off case, when  $\rho$  is low, the mean response time is higher ( $E[T] = 1.2$  sec with no effect by the discipline; the queue is empty), since almost every arrival finds servers in the off state, and thus every job incurs in the setup time. For high  $\rho$ , instead, the mean response time increases ( $E[T] = 2$  sec for  $\delta = \text{FIFO}$  and  $E[T] = 1.35$  sec for  $\delta = \text{SPTF}$ ) due to large queueing. As predicted above, we can see that the benefit in response time one may obtain from moving FIFO to SPTF is larger than the one obtainable from moving On/Off to On/Idle. Indeed (see high  $\rho$ ) moving from the (On/Off, FIFO) strategy to the (On/Idle, FIFO), the response time  $E[T]$  changes from 2 to 1.8 (a 10% reduction). Moving, instead, from the (On/Idle, FIFO) strategy to the (On/Idle, SPTF) the response time  $E[T]$  changes from 1.8 to 1.3 (an almost 30% reduction).

- (iii) Seeing at the ERP( $\pi, \delta$ ) index, its values are a consequence of the  $E[P]$  and the  $E[T]$  ones. Table 2 shows that the optimal ERP( $\pi, \delta$ ) is obtained for the (On/Off, FIFO) strategy and for the (On/Off, SPTF) strategy when  $\rho$  is low, while it is obtained for the (On/Idle, SPTF) strategy only when  $\rho$  is high.

For the high setup time case ( $E[\text{setup}] = 100$  sec), the reader is sent to [12].

In summary, making predictions of the Dp power management policies that optimizes

- (i) the Dp power consumption (minimum absorbed Watts) or

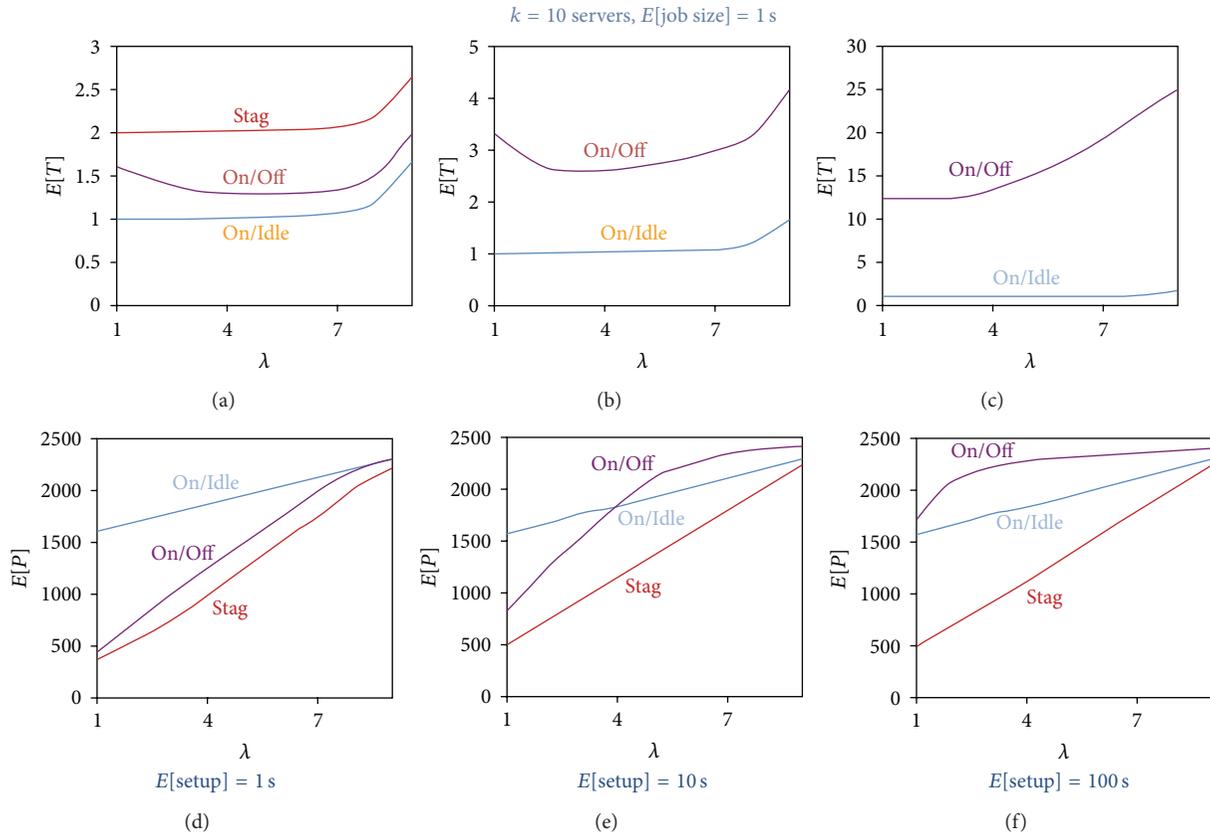


FIGURE 4: Effects of server management policies and setup time on response time and power consumption [2].

- (ii) the Dp performance (minimum response\_time) or
- (iii) the Dp performance-per-Watt (minimum response\_time-per-Watt)

is a nontrivial task. The most significant policies  $\pi$  are first to be drawn from the universe of all possible policies. Then, for each such policy, the effects of time-dependent and time-independent queueing disciplines are to be studied. On the other hand, once the modeling work has been done, the work the server-farm manager has to perform to direct his Dp is greatly simplified, since the universe of all possible  $(\pi, \delta)$  strategies he needs to choose from is drastically reduced to very small set of most significant strategies.

**2.1.3. Energy Saving with Dynamic Policies.** In any practical situation, the load  $\lambda$  changes over time, according to a given pattern  $\lambda(t)$ . One should then find policies that adapt themselves to changing load patterns. This is not the case of the policies introduced in Section 2.1.1, which are somehow static in nature and remain efficient only for given values of  $\lambda$ , while becoming inefficient for other values. Looking, for example, at Figure 4 case with  $E[\text{setup}] = 10$  sec, one can see that for changing values of  $\lambda$  there are situations in which the On/Off policy consumes less power than On/Idle and vice versa.

For this reason, two adaptive versions of the On/Off and On/Idle policies are known in the literature, respectively,

called DelayedOff and LookAhead, which dynamically adapt themselves to changing loads [11].

The DelayedOff policy is an improvement of the On/Off. According to DelayedOff, when a server goes idle, rather than turning off immediately, it sets a timer of duration  $t_{\text{wait}}$  and sits in the idle state for  $t_{\text{wait}}$  seconds. If a request arrives at the server during these  $t_{\text{wait}}$  seconds, the server goes back to the busy state (with zero setup cost); otherwise, the server is turned off.

The LookAhead policy is an improvement of the On/Idle. Under such a policy, the system fixes an optimally chosen number  $n^*$  of servers maintained in the on or idle states. According to the standard On/Idle, if an arrival finds a server idle it starts serving on the idle server. Arrivals that find all  $n^*$  server on busy, join a central queue from which servers pick jobs when they become idle.

The optimal  $t_{\text{wait}}$  and the optimal  $n^*$  of the two policies, respectively, are chosen to minimize the ERP index. As said above, minimizing ERP can be seen as maximizing the performance per Watt, with performance defined as the inverse of the mean response time [11].

In the LookAhead policy,  $n^*$  changes as a function of time. Indeed, the policy calculates  $n^*(t)$  for each time  $t$  basing on the forecast of the load  $l(t)$  at time  $t$ .

Figure 5 illustrates the autoscaling capabilities of the LookAhead and DelayedOff policies [11], with respect to the conventional On/Off, for Poisson arrivals with  $l(t)$  changing sinusoidally with time (period = 6 hrs).

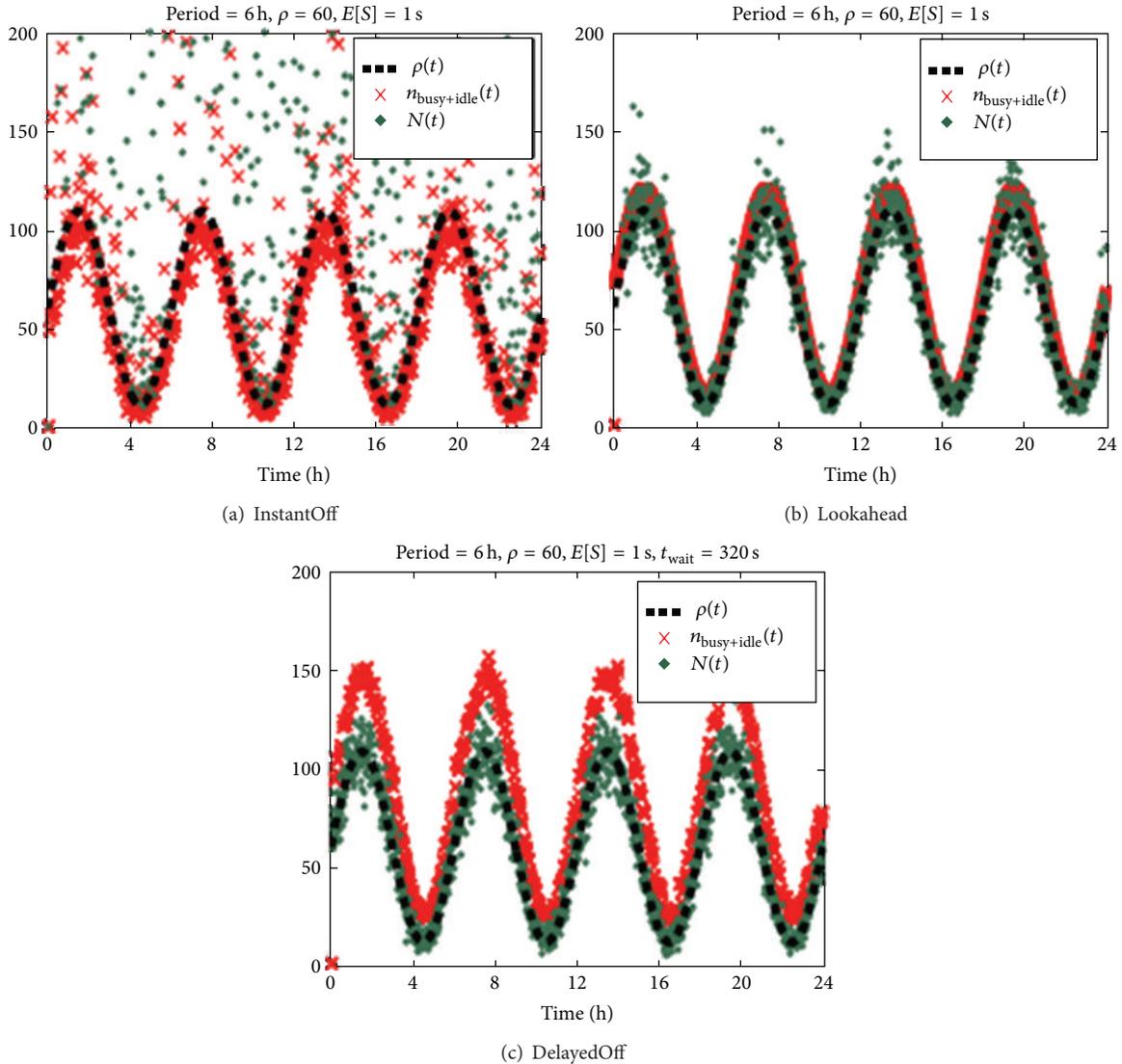


FIGURE 5: Effects of dynamic policies with respect to the static On/Off [11].

Figure 5(a) refers to the *On/Off* policy (called *InstantOff*), Figure 5(b) to the *LookAhead* and Figure 5(c) to the *DelayedOff*. The dashed line denotes the varying load at time  $t$ ,  $\lambda(t)$ . The crosses denote the number  $n_{\text{busy+idle}}(t)$  of servers that are busy or idle at time  $t$ , and the dots denote the number  $N(t)$  of jobs in the system at time  $t$ .

The illustration shows how, with the two dynamic policies, number  $n_{\text{busy+idle}}(t)$  and number  $N(t)$  almost completely follow the behavior of the demand pattern  $\lambda(t)$ , while in the *On/Off* case such numbers are somewhat dispersed; in other words, some servers remain in the idle state whereas they should be busy and vice versa, with the consequence of waste of power and worsened response time.

The two dynamic policies above simply try to optimize the  $E[P]$  by  $E[T]$  product.

In many practical situations, instead, the objective is to meet a given average response time, according to requirements dictated by specific service level agreements (SLAs).

In this case, specific dynamic policies have been introduced, which try to respect the  $E[T]$  requirement while minimizing the average power consumed by the servers ( $P_{\text{avg}}$ ) and the average number of used servers ( $N_{\text{avg}}$ ).

Such policies are known as the *AutoScale* policy [13], the *AlwaysOn* policy [14], the *Reactive* policy [15], and the *Predictive MWA* policy [16–19]. The latter will not be dealt with here, and we will only treat the *AutoScale* policy, which is an evolution of the remaining three.

The *AutoScale* policy generalizes the use of the  $t_{\text{wait}}$  time already seen for the *DelayedOff* policy. Differently from this latter, however, is that in the *AutoScale* case each server decides autonomously when to turn off, setting a timer of duration  $t_{\text{wait}}$  and sitting in the idle state for  $t_{\text{wait}}$  sec. As with the *DelayedOff*, however, if a request arrives at the server during these  $t_{\text{wait}}$  sec, then the server goes back to the busy state (with zero setup cost). Otherwise, the server is turned off.

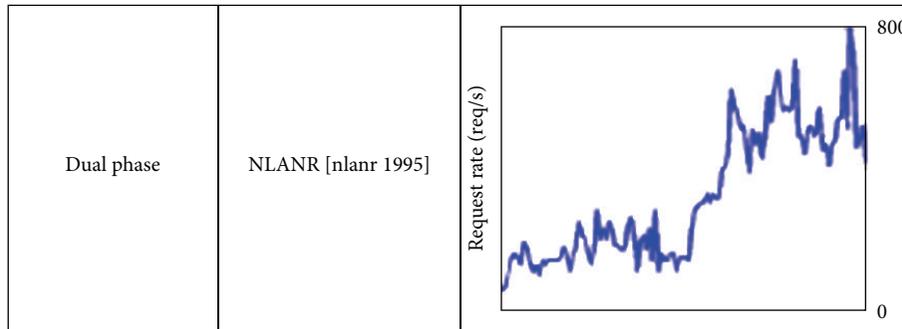


FIGURE 6: Dual phase pattern [13].

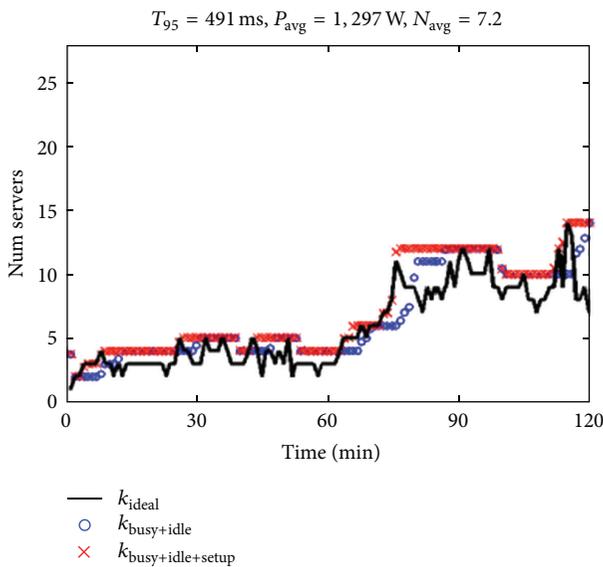


FIGURE 7: Effects of dynamic AutoScale policy [13].

The AutoScale and the three remaining policies have been evaluated in [13] according to a specific load pattern  $\lambda(t)$  varying over time between 0 and 800 req/s (see Figure 6). Such a pattern, known as dual phase pattern is used to represent the diurnal nature of typical data center traffic, where the request rate is low at the night time and high at day time.

Figure 7 illustrates the performance of the AutoScale policy, when the time requirement to meet is a 95-percentile response time  $T$  goal of 400 to 500 ms (denoted  $T_{95}$ ). In the illustration, the red lines denote the number  $k_{\text{busy+idle+setup}}(t)$  of busy+idle+setup servers and the blue lines the number  $k_{\text{busy+idle}}(t)$  of busy+idle servers at time  $t$ . The  $k_{\text{ideal}}$  line represents the number of servers that should be on at any given time to fully satisfy the demand  $\lambda(t)$ .

The illustration shows how, with the AutoScale policy, there is no dispersion in the available servers and the number  $k_{\text{busy+idle+setup}}(t)$  and number  $k_{\text{busy+idle}}(t)$  almost totally follow the demand pattern  $\lambda(t)$ , and a  $T_{95} = 491$  ms goal is achieved, with  $P_{\text{avg}} = 1,297$  W and  $N_{\text{avg}} = 7.2$  servers.

In the mentioned similar policies (AlwaysOn, Reactive, and Predictive MWA), instead, the  $T_{95}$  requirement can be

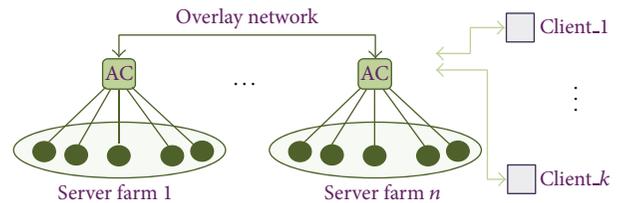


FIGURE 8: Example of a multifarm data center.

seen to be met only at the expense of server dispersion and/or at the expense of  $P_{\text{avg}}$  and  $N_{\text{avg}}$  [13].

Indeed, in the AlwaysOn case the  $T_{95}$  requirement is met ( $T_{95} = 291$  ms) but at the expense of a large dispersion in the available servers and large power consumption ( $P_{\text{avg}} = 2,322$  W and  $N_{\text{avg}} = 14$ ).

In the Reactive case, instead, a low dispersion of servers is achieved, with low power and low number of servers ( $P_{\text{avg}} = 1,281$  W,  $N_{\text{avg}} = 6.2$ ), but the time requirement is absolutely out of range ( $T_{95} = 11,003$  ms).

A better time performance ( $T_{95} = 7,740$  ms) is found in the Predictive MWA with similarly low dispersion of servers and similarly low power and number of servers ( $P_{\text{avg}} = 1,276$  W,  $N_{\text{avg}} = 6.3$ ).

**2.2. Energy Saving in Multifarm Data Centers.** Energy saving in multifarms is based on so-called self-organization and self-differentiation algorithms, whose goal is to transfer the load from a server to a less loaded one, to maximize the power efficiency of the whole data center.

These algorithms are widely adopted in the autonomic computing field. The term autonomic indicates systems able to self-manage, self-configure, self-protect, and self-repair; thus, systems have no need of external action to be managed [20].

Figure 8 illustrates the typical multifarm architecture that consists of a series of server farms (1 through  $n$ ) each farm controlled by a so-called autonomic component (AC), with the ACs interacting through an overlay network (ON). Each farm serves a number of clients (1 through  $k$ ).

The ON is a self-organized network, in other words a network which is created, maintained, and optimized through self-organization algorithms which cluster the ACs according to their properties or type [21].

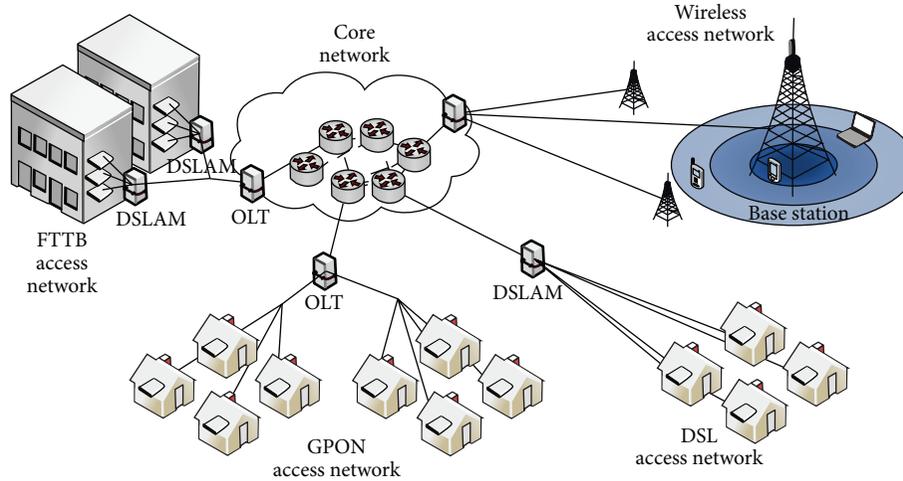


FIGURE 9: Typical Tlc network architecture [4].

The ACs, in turn, execute a particular kind of self-organization algorithm called self-differentiation algorithm, which takes decentralized decisions on the state and configuration of the ACs.

The AC aims to put state the servers in idle and transfer the load on the other servers to limit performance degradation. Three types of self-differentiation algorithms are known: Stand-by, Load Distribution, and Wake-up algorithms whose details can be found in [21].

The algorithms were evaluated by means of simulations of a use-case in which server farms are in charge to serve requests issued by a set of clients. Each client performs several requests, before terminating the connection. The percentage of energy that can be saved in a day goes from about 7% to about 12%, with a debt to pay in terms of response time from about 9 units of time (when the power saving is 7%) to about 11 units of time (when the power saving is 12%).

### 3. Energy Management in Tlc Systems

Tlc systems may consist of wired or wireless access networks or of a combination thereof.

In addition to the basic Dp infrastructure, Tlc systems also include Tlc-specific subsystems: cell-phone, towers with associated base stations, subscriber stations, switching nodes, and so forth, for the wireless part, and communication processors, routers, gateways, switches, and so forth, for the wired part.

Power management in Tlc systems, thus, includes not only power optimization of their Dp infrastructure, but also power optimization of Tlc-specific subsystems.

In this section we will only deal with Tlc-specific subsystems, since the power optimization of Dp infrastructure is dealt with is already seen in Section 2.

Figure 9 describes a typical Tlc architecture, which combines wired and wireless communication networks.

In the wired part, three main types of connections are found: (1) the twisted pair copper cable connection based on the DSL (digital subscriber line) technology; (2) the coax cable connection based on the DOCS (data over cable service)

technology; and (3) the optical fiber connection based on the GPON (gigabit passive optical network) technology, used when higher bit rates are required. The illustration also shows the DSLAM (DSL access multiplexer) nodes, the OLT (optical line termination) nodes, and the FTTB (fiber to the building) nodes.

In order to interconnect different user areas, a core network is used, that consists of a number of core nodes that are interconnected through wavelength-division multiplexed (WDM) optical fiber links, usually in a mesh or ring topology.

In the wireless part of the network we find base stations (BS) to which the user's devices are connected by means of radio signals. Each BS is further connected to the core network through a so-called backhaul network. Different technologies can be found, from WiMAX (worldwide interoperability for microwave access) [22], to HSPA (high speed packet access), and to the most recent LTE (long term evolution).

In such a system, about 90% of Tlc-specific power consumption is concentrated in the routers (with 75% the line cards, 10% the power supply and fans, and 10% the switch fabric) [23]. Current routers consume between 0.01 and 0.1 W/Mbps. One can calculate that at ADSL access rates (8 Mbps) the power absorbed per subscriber is of about 0.24 W/subs, while at 100 Mbps becomes of about 3 W/subs [4].

Currently, Tlc networks are designed to handle the peak loads. Designing adaptable networks, where one can switch off line cards when the demand is lower and can lead to lower power consuming networks.

In core networks this can be achieved by use of dynamic topology optimization algorithms: from all possible topologies that satisfy the required traffic demand, the topologies with lower overall power consumption are chosen. By such algorithms, reductions of power consumption for more than 50% during off-peak hours can be achieved [24].

Base stations (BS) with differentiated cell sizes are the key in wireless networks optimization if the so-called hybrid hierarchical BS deployment is used.

A low layer access network is first created, providing a low bit rate (but large cell sizes) to the users. In the higher layers, BS with higher bit rates (but smaller cell sizes) is utilized to provide high bandwidth connections when required. The advantage is that the higher layers can be switched to the idle and only switched on with high traffic demand.

Tlc power optimization also tries to minimize the power consumption of the home gateways. These are individual devices that only need to be on when the user is active. At other times, they could be switched off. In reality this is rarely operated, but legislations concerning standby power consumption standards of 0.5 W are emerging [4].

#### 4. Conclusions

The power management of ICT systems, that is, data processing (Dp) and telecommunication (Tlc) systems, is a complex issue with implications in economical terms.

The paper has illustrated methods to optimize Dp power consumption by use of power management policies (static and dynamic policies) that yield electrical power saving while maintaining the system QoS at acceptable levels.

The paper has also illustrated methods to optimize Tlc power consumption by use of power management policies to be adopted in wired and wireless Tlc systems. This achieves electrical power saving without compromising the service quality.

#### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

#### Acknowledgments

This work was partially supported by the Ph.D. research program of the University of Roma TorVergata funded by Telecom Italia and by the Guglielmo Marconi University of Studies in Roma.

#### References

- [1] U.S. Environmental Protection Agency ENERGY STAR Program, "Report to Congress on Server and Data Center Energy Efficiency Public Law 109-431," 2007.
- [2] M. Harchol-Balter, "Open problems in power management of data centers," in *Proceedings of the IFIP WG7.3 Workshop*, University of Namur, Namur, Belgium, November 2010.
- [3] J. Dean, *Designs, Lessons and Advice from Building Large Distributed Systems*, Google Fellow, 2009.
- [4] W. Vereecken, W. Van Heddeghem, M. Deruyck et al., "Power consumption in telecommunication networks: overview and reduction strategies," *IEEE Communications Magazine*, vol. 49, no. 6, pp. 62-69, 2011.
- [5] A. Gandhi, M. Harchol-Balter, and I. Adan, "Server farms with setup costs," *Performance Evaluation*, vol. 67, no. 11, pp. 1123-1138, 2010.
- [6] Intel Corporation, "Serial ATA staggered spin-up," White Paper, 2004.
- [7] M. W. Storer, K. M. Greenan, E. L. Miller, and K. Voruganti, "Pergamum: replacing tape with energy efficient, reliable, disk-based archival storage," in *Proceedings of the 6th USENIX Conference on File and Storage Technologies (FAST '08)*, pp. 1-16, San Jose, Calif, USA, February 2008.
- [8] R. W. Conway, W. L. Maxwell, and L. W. Miller, *Theory of Scheduling*, Addison-Wesley, 1967.
- [9] M. Harchol-Balter, *Performance Modeling and Design of Computer Systems*, Cambridge University Press, 2013.
- [10] G. Iazeolla, A. Pieroni, and G. Scorzini, "Simulation study of server farms power optimization," RI.01.13 T.R., Software Engineering Laboratory, University of Roma Tor Vergata, Rome, Italy, 2013.
- [11] A. Gandhi, V. Gupta, M. Harchol-Balter, and M. A. Kozuch, "Optimality analysis of energy-performance trade-off for server farm management," *Performance Evaluation*, vol. 67, no. 11, pp. 1155-1171, 2010.
- [12] G. Iazeolla and A. Pieroni, "Power management of server farms," in *Proceedings of the International Conference on Power Science and Engineering (ICPSE '13)*, Paris, France, December 2013.
- [13] A. Gandhi, M. Harchol-Balter, R. Raghunathan, and M. A. Kozuch, "Autoscale: dynamic, robust capacity management for multi-tier data centers," *ACM Transactions on Computer Systems*, vol. 30, no. 4, article 14, 2012.
- [14] G. Chen, W. He, J. Liu et al., "Energy-aware server provisioning and load dispatching for connection-intensive internet services," in *Proceedings of the 5th USENIX Symposium on Networked Systems Design and Implementation (NSDI '08)*, pp. 337-350, 2008.
- [15] B. Urgaonkar, P. Shenoy, A. Chandra, and P. Goyal, "Dynamic provisioning of multi-tier Internet applications," in *Proceedings of the 2nd International Conference on Autonomic Computing (ICAC '05)*, pp. 217-228, June 2005.
- [16] P. Bodik, R. Griffith, C. Sutton, A. Fox, M. Jordan, and D. Patterson, "Statistical machine learning makes automatic control practical for internet datacenters," in *Proceedings of the Conference on Hot Topics in Cloud Computing (HotCloud '09)*, 2009.
- [17] D. Grunwald, C. B. Morrey III., P. Levis, M. Neufeld, and K. I. Farkas, "Policies for dynamic clock scheduling," in *Proceedings of the 4th Conference on Symposium of Operating System Design and Implementation (OSDI '00)*, vol. 4, Berkeley, Calif, USA, 2000.
- [18] T. Pering, T. Burd, and R. Brodersen, "The Simulation and evaluation of dynamic voltage scaling algorithms," in *Proceedings of the International Symposium on Low Power Electronics and Design (ISLPED '98)*, pp. 76-81, August 1998.
- [19] A. Verma, G. Dasgupta, T. K. Nayak, P. De, and R. Kothari, "Server workload analysis for power minimization using consolidation," in *Proceedings of the Conference on USENIX Annual Technical Conference (USENIX '09)*, 2009.
- [20] D. Barbagallo, E. Di Nitto, D. J. Dubois, and R. Mirandola, "A bio-inspired algorithm for energy optimization in a self-organizing data center," in *Proceedings of the 1st International Conference on Self-Organizing Architectures (SOAR '09)*, pp. 127-151.
- [21] S. S. Martínez, J. S. Pareta, B. Otero, L. Ferrari, A. Manzalini, and C. Moiso, "Self-organized server farms for energy savings," in *Proceedings of the 6th International Conference Industry Session on Autonomic Computing and Communications Industry Session (ICAC-INDST '09)*, pp. 39-40, Barcelona, Spain, June 2009.

- [22] G. Iazeolla, A. Pieroni, A. D'Ambrogio, and D. Gianni, "A distributed approach to wireless system simulation," in *Proceedings of the 6th Advanced International Conference on Telecommunications (AICT '10)*, pp. 252–262, Barcelona, Spain, May 2010.
- [23] G. Iazeolla and A. Pieroni, "Power control and optimization in ICT," in *Proceedings of the Power Control and Optimization Conference (PCO '13)*, Prague, Czech Republic, August 2013.
- [24] B. Puype, W. Vereecken, D. Colle, M. Pickavet, and P. Demeester, "Power reduction techniques in multilayer traffic engineering," in *Proceedings of the 11th International Conference on Transparent Optical Networks (ICTON '09)*, pp. 1–4, Azores, Portugal, July 2009.
- [25] F. Renzi, *Business E<sup>3</sup>: Energy, Efficiency, Economy*, IBM Corporation, 2007.
- [26] A White Paper from the Experts in Business-Critical Continuity, *Energy Logic: Reducing Data Center Energy Consumption by Creating Savings that Cascade Across Systems*, Emerson Network Power, 2012.

## Research Article

# Power Quality Improvement by Unified Power Quality Conditioner Based on CSC Topology Using Synchronous Reference Frame Theory

Rajasekaran Dharmalingam,<sup>1</sup> Subhransu Sekhar Dash,<sup>2</sup> Karthikrajan Senthilnathan,<sup>3</sup>  
Arun Bhaskar Mayilvaganan,<sup>3</sup> and Subramani Chinnamuthu<sup>2</sup>

<sup>1</sup> Department of Electrical and Electronics Engineering, RMD Engineering College, Chennai, India

<sup>2</sup> Department of Electrical and Electronics Engineering, SRM University, Chennai, India

<sup>3</sup> Department of Electrical and Electronics Engineering, Velammal Engineering College, Chennai, India

Correspondence should be addressed to Karthikrajan Senthilnathan; [karthik.svkk@gmail.com](mailto:karthik.svkk@gmail.com)

Received 27 February 2014; Accepted 19 March 2014; Published 11 June 2014

Academic Editors: N. Barsoum, P. Vasant, and G.-W. Weber

Copyright © 2014 Rajasekaran Dharmalingam et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper deals with the performance of unified power quality conditioner (UPQC) based on current source converter (CSC) topology. UPQC is used to mitigate the power quality problems like harmonics and sag. The shunt and series active filter performs the simultaneous elimination of current and voltage problems. The power fed is linked through common DC link and maintains constant real power exchange. The DC link is connected through the reactor. The real power supply is given by the photovoltaic system for the compensation of power quality problems. The reference current and voltage generation for shunt and series converter is based on phase locked loop and synchronous reference frame theory. The proposed UPQC-CSC design has superior performance for mitigating the power quality problems.

## 1. Introduction

The main impact in the power distribution system is the quality of power, which causes more distortion in the source due to using nonlinear loads (power electronics loads). The main cause for distortion is harmonics, notching, and interharmonics. Distortion is that the fundamental frequency sine wave is represented as super position of all harmonic frequency sine waves on fundamental sine wave. The usage of power electronics loads is increased day by day, while considering that industries power electronics drives are used for the automation of the industries. To compensate the distortion in the system, passive filters were used and while using the passive filters particular harmonic range is only eliminated. In order to overcome the drawbacks of passive filter, for the elimination of power quality problems, active filters were used.

Power quality problems are harmonics, sag, and swell which are mitigated by the active filters by the configuration of dynamic voltage restorer (DVR), distribution-static synchronous compensator (D-STATCOM), and unified power quality conditioner (UPQC) [1]. In this paper UPQC [2] is used for the mitigation of the power quality problems which is the combination of series and shunt active filters. The series and shunt active power filters are voltage and current source converters which are controlled by the PWM signals which are generated by the controllers.

## 2. Unified Power Quality Conditioner (UPQC)

The unified power quality conditioner is commonly called UPQC. The design configuration is based on the connection of series and shunt inverters. In this, the design configuration

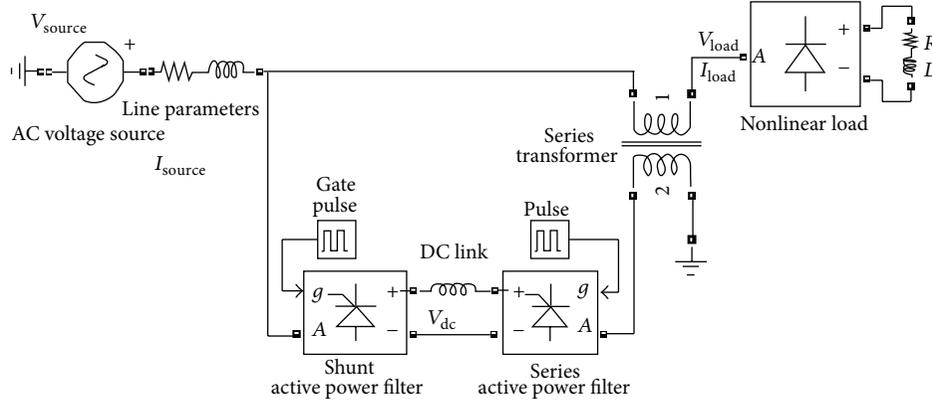


FIGURE 1: The design configuration of UPQC-CSC.

is right series and left shunt with the current source converter (CSC) [3, 4]. In this paper, UPQC-CSC [5, 6] is designed and analysis of the results has been done. Unified power quality conditioner (UPQC) for nonlinear and voltage sensitive load has following facilities.

- (i) It reduces the harmonics in the supply current, so that it can improve utility current quality for nonlinear loads.
- (ii) UPQC provides the VAR requirement of the load, so that the supply voltage and current are always in phase; therefore, no additional power factor correction equipment is required.
- (iii) UPQC maintains load end voltage at the rated value even in the presence of supply voltage sag.

The design configuration of UPQC-CSC [7] is shown in Figure 1.

### 3. Synchronous Reference Frame (SRF) Theory

The control strategy for the unified power quality conditioner is based on the synchronous reference frame (SRF) [8, 9] theory. In this theory controlling of the three-phase converters using the rotating frame theory by converting the source voltage and current to direct and quadrature axis is done. The voltage is converted to  $dq$  in the series controller and current is converted to  $dq$  in the series controller. Consider

$$\begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \sin(\omega t) & \sin\left(\omega t - \frac{2\pi}{3}\right) & \sin\left(\omega t + \frac{2\pi}{3}\right) \\ \cos(\omega t) & \cos\left(\omega t - \frac{2\pi}{3}\right) & \cos\left(\omega t + \frac{2\pi}{3}\right) \end{bmatrix} \begin{bmatrix} V_d \\ V_q \\ V_0 \end{bmatrix} \quad (1)$$

The  $dq$  transform is again converted to the  $V'_{abc}$  in order to get the reference signal which is used for the generation of the pulse for the three-phase converter in the system. Consider

$$\begin{bmatrix} V_d \\ V_q \\ V_0 \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} \frac{1}{\sqrt{2}} & \sin(\omega t) & \cos(\omega t) \\ \frac{1}{\sqrt{2}} & \sin\left(\omega t - \frac{2\pi}{3}\right) & \cos\left(\omega t - \frac{2\pi}{3}\right) \\ \frac{1}{\sqrt{2}} & \sin\left(\omega t + \frac{2\pi}{3}\right) & \cos\left(\omega t + \frac{2\pi}{3}\right) \end{bmatrix} \begin{bmatrix} V'_a \\ V'_b \\ V'_c \end{bmatrix} \quad (2)$$

The shunt converter performs the process of elimination of harmonics and series converter performs process of elimination of the voltage related problems. The control block diagram for the synchronous reference frame theory is shown in Figure 2.

**3.1. Series Controller.** The control strategy of the series controller is achieved through the synchronous reference frame theory. In this, the series controller gets the reference signal for the generation of pulse for the three-phase converter, by comparing the source voltage with distortion and constant voltage. The source voltage  $V_{sabc}$  and constant voltage  $V_{refabc}$  are converted to the  $V_{sdq0}$  and  $V_{refdq0}$  transform. The  $V_{sdq0}$  and  $V_{refdq0}$  are compared to get the error signal which is again converted to  $V'_{labc}$ . The  $V'_{labc}$  is the reference signal for the pulse generator. The simulation diagram for synchronous reference frame theory based series controller is shown in Figure 3.

**3.2. Shunt Controller.** The shunt converter has the function of compensating the current related problems. Along with the shunt controller, DC link voltage is maintained. The  $abc$  to  $dq0$  transform is inverted and converted to  $abc$ ; that signal is given as the reference signal and the measured signal is given to the hysteresis band PWM to produce the pulse signals for the operation of shunt converter. The simulation diagram for shunt controller is shown in Figure 4.



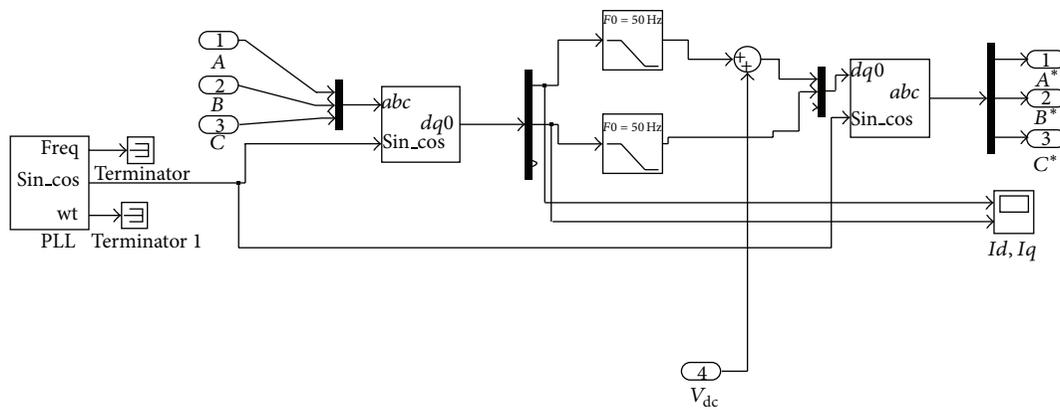


FIGURE 4: Simulation of shunt controller.

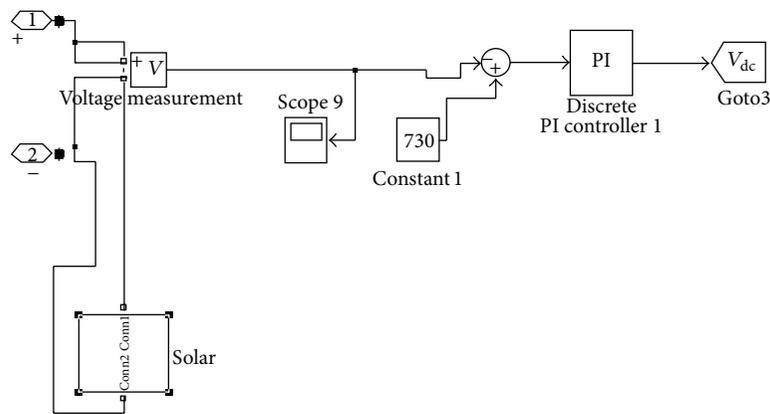


FIGURE 5: DC link controller.

3.3. *DC Link Controller.* The direct current link controller has the PI controller in which the constant voltage is given as the set point and the measured voltage is given for the comparison to maintain the constant voltage. The PV array is attached with the DC link for injection. The DC link controller is shown in Figure 5.

**4. Simulation and Results**

The UPQC-CSC has the reactor as the DC link for the series and shunt converter and is controlled by the synchronous reference frame (SRF) theory and the pulse is generated by the hysteresis band controller. The shunt and series converters have the function of compensating current and voltage problems, respectively. The simulation of UPQC-CSC is shown in Figure 6. The output of UPQC-CSC is shown in Figure 7 which shows the voltage with sag, current with harmonics, and compensated voltage and current. The compensation of sag is shown in Figure 8. The shunt compensation is shown in Figure 9. The series compensation is shown in Figure 10.

4.1. *System Parameters.* Consider

- source voltage: 415 V, 50 Hz;
- load parameters:

- resistive load: 10 K $\Omega$ ;
- inductive load: 2 mH;
- RLC load: 10 KW;

shunt inverter side:

- LC filter: 3.5 mH, 5  $\Omega$ , and 10  $\mu$ F;

series inverter side:

- LC filter: 12  $\mu$ H, 5  $\Omega$ , and 10  $\mu$ F;

DC link reactor:

- for UPQC-CSC: 200 mH;
- solar voltage: 727.1 V.

Figure 7 shows the simulation output of the UPQC-CSC simulation for voltage sag mitigation. The sudden addition of load in the system causes voltage sag for the time duration of 0.04 to 0.08 s. The compensation for the sag is by the series active filter using the SRF theory for the reference signal generated and pulse generated by the hysteresis band and given to the IGBTs in the filter.

The compensation of the voltage related problems is done by the series active filter to maintain the system voltage 1

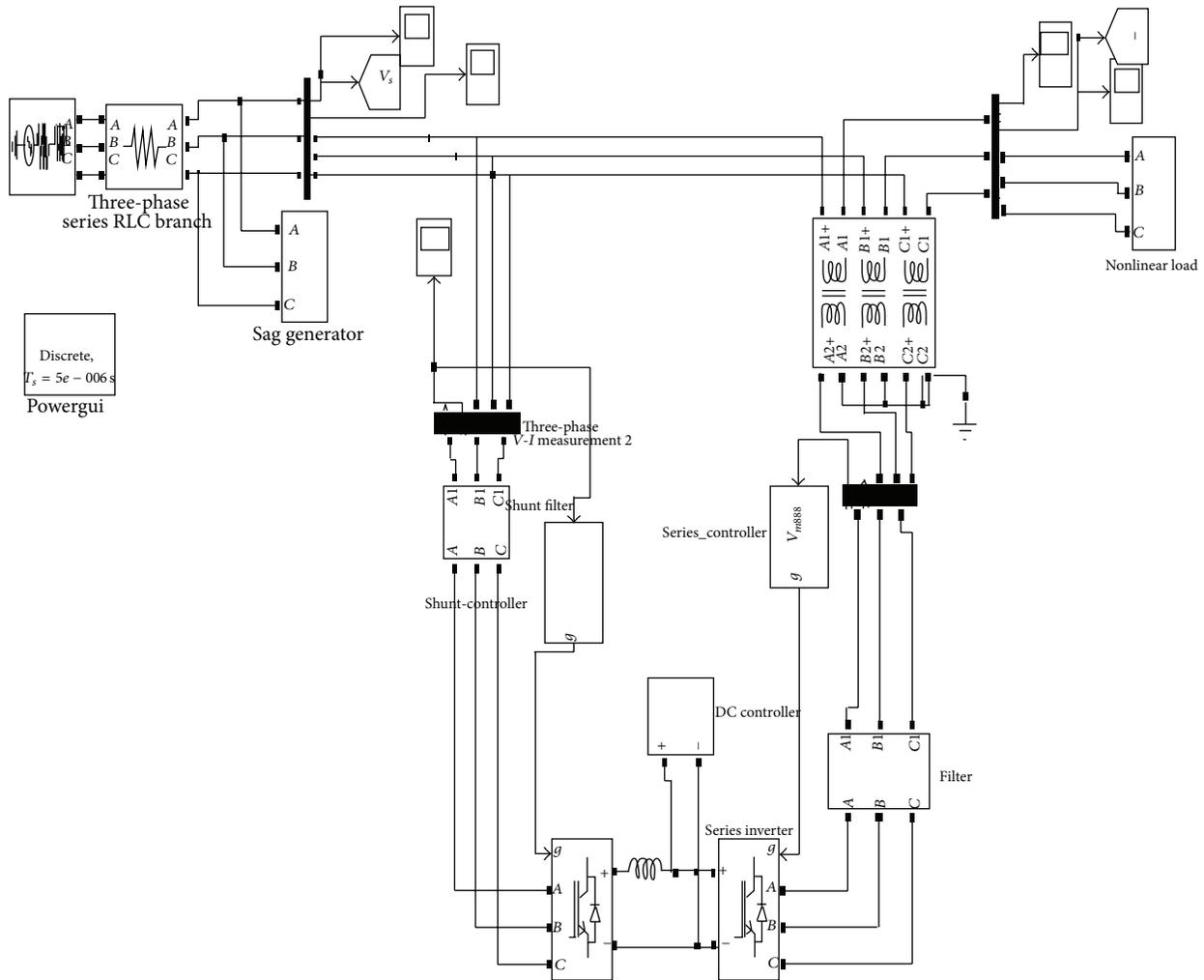


FIGURE 6: UPQC-CSC simulation diagram.

P.U. By using the SRF theory even a minor disturbance in the system is sensed and compensation is done; Figure 10 shows the series compensation for the system. The harmonics compensation is done by the shunt active filter along with the DC link voltage controller. Total harmonics distortion (THD) for the current source converter is shown in Table 1. Figure 9 shows the compensation given for reducing the harmonics.

The Fourier fast transform analysis graph for the source voltage THD of about 0.89% is shown in Figure 11.

The Fourier fast transform analysis graph for the load voltage with the nonlinear loading conditions of about 0.45% is shown in Figure 12.

The Fourier fast transform analysis graph for the load current with the nonlinear loading conditions of about 0.17% is shown in Figure 13.

### 5. Conclusion

In this paper, synchronous reference frame theory based control method is implemented to control the working of unified power quality conditioner based on current source

TABLE 1: Total harmonics distortion (THD in %).

Hn order	Current source converter		
	Source voltage ( $V_s$ ) in %	Load voltage ( $V_L$ ) in %	Load current ( $I_L$ ) in %
H	0.97	0.99	0.04
H3	0.28	0.08	0.08
H5	0.09	0.05	0.06
H7	0.03	0.06	0.02
H9	0.03	0.04	0.04
H11	0.02	0.03	0.06
THD	0.89	0.45	0.17

converter topology. The simulation results show that the device is capable of compensating the current harmonics under unbalanced and nonlinear load conditions, simultaneously mitigating voltage sag and swell. The proposed UPQC-CSC design has superior performance for mitigating

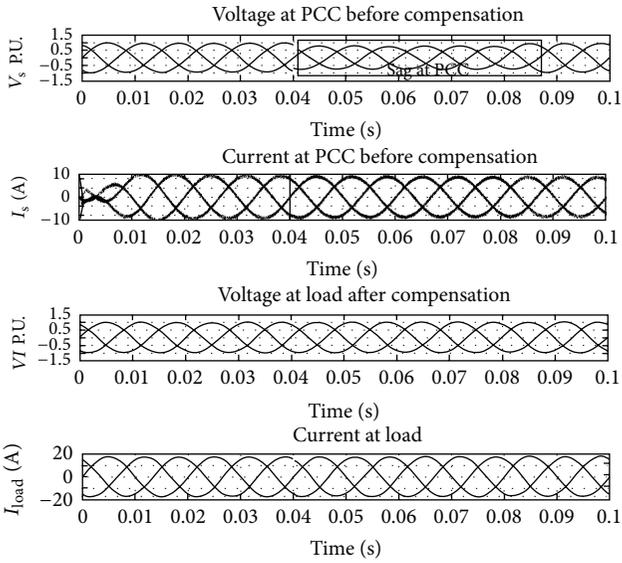


FIGURE 7: Output of source voltage and current and load voltage and current waveform.

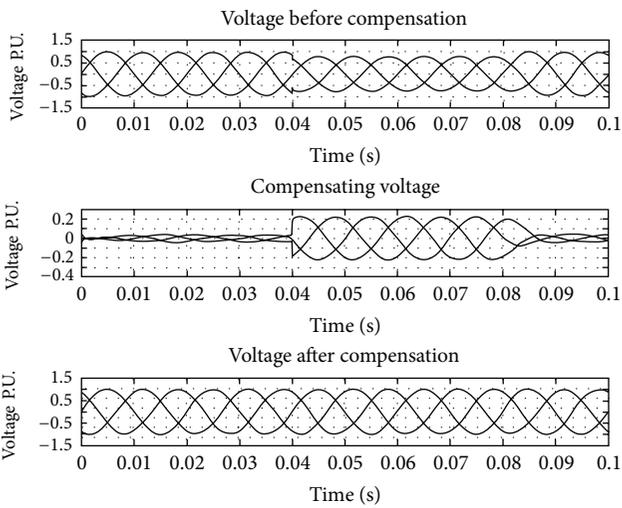


FIGURE 8: PCC voltage with sag, compensating voltage and voltage after compensation.

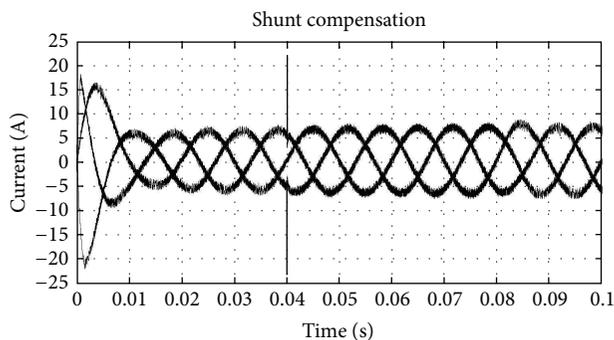


FIGURE 9: Shunt injection for THD compensation.

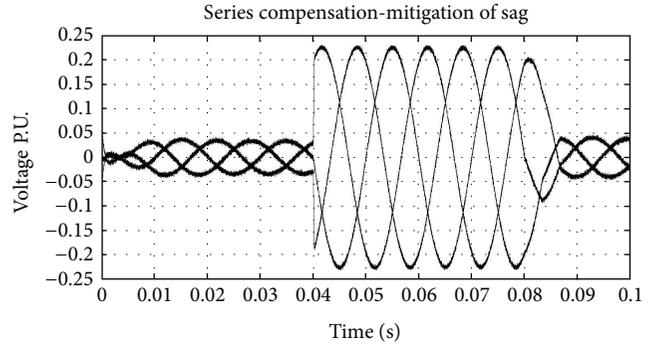


FIGURE 10: Series injection for sag compensation.

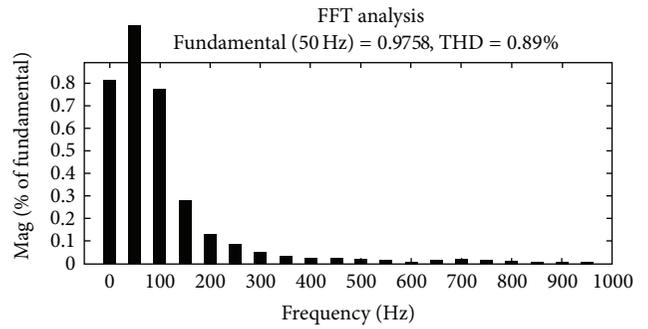


FIGURE 11: Source voltage THD graph.

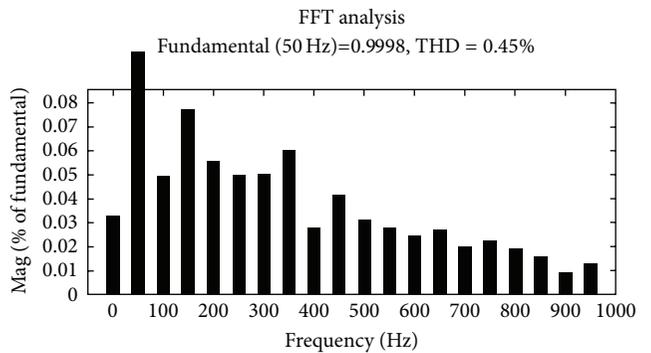


FIGURE 12: Load voltage THD graph.

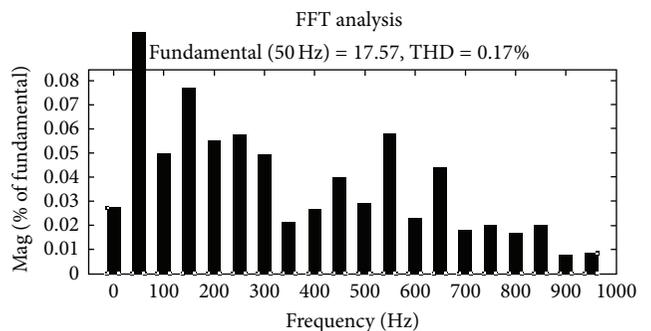


FIGURE 13: Load current THD graph.

the power quality problems. The series converter is capable of mitigating the voltage related problems and shunt converter is capable of mitigating the harmonics.

### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

### References

- [1] V. Khadkikar, "Enhancing electric power quality using UPQC: a comprehensive overview," *IEEE Transactions on Power Electronics*, vol. 27, no. 5, pp. 2284–2297, 2012.
- [2] M. Kesler and E. Ozdemir, "Synchronous-reference-frame-based control method for UPQC under unbalanced and distorted load conditions," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 9, pp. 3967–3975, 2011.
- [3] N. Zhu, D. Xu, B. Wu, F. Liu, N. R. Zargari, and M. Kazerani, "Common-mode voltage reduction methods for current-source converters in medium-voltage drives," *IEEE Transactions on Power Electronics*, vol. 28, no. 2, pp. 995–1006, 2013.
- [4] P. E. Melin, J. R. Espinoza, L. A. Moran et al., "Analysis, design and control of a unified power-quality conditioner based on a current-source topology," *IEEE Transactions on Power Delivery*, vol. 27, no. 4, pp. 1727–1736, 2012.
- [5] A. Terciyani, M. Ermis, and I. Cadirci, "A selective harmonic amplification method for reduction of kVA rating of current source converters in shunt active power filters," *IEEE Transactions on Power Delivery*, vol. 26, no. 1, pp. 65–78, 2011.
- [6] V. Kinhal, P. Agarwal, and H. O. Gupta, "Performance investigation of neural-network-based unified power-quality conditioner," *IEEE Transactions on Power Delivery*, vol. 26, no. 1, pp. 431–437, 2011.
- [7] R. El Shatshat, M. M. A. Salama, and M. Kazerani, "Artificial intelligent controller for current source converter-based modular active power filters," *IEEE Transactions on Power Delivery*, vol. 19, no. 3, pp. 1314–1320, 2004.
- [8] C. H. da Silva, R. R. Pereira, L. E. B. da Silva, G. Lambert-Torres, B. K. Bose, and S. U. Ahn, "A digital PLL scheme for three-phase system using modified synchronous reference frame," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 11, pp. 3814–3821, 2010.
- [9] J. M. Espí Huerta, J. Castelló-Moreno, J. R. Fischer, and R. García-Gil, "A synchronous reference frame robust predictive current control for three-phase grid-connected inverters," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 3, pp. 954–962, 2010.

## Research Article

# Structural Optimization of a Knuckle with Consideration of Stiffness and Durability Requirements

**Geun-Yeon Kim, Seung-Ho Han, and Kwon-Hee Lee**

*Department of Mechanical Engineering, Dong-A University, Busan 604-714, Republic of Korea*

Correspondence should be addressed to Kwon-Hee Lee; [leekh@dau.ac.kr](mailto:leekh@dau.ac.kr)

Received 27 January 2014; Accepted 9 February 2014; Published 8 June 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Geun-Yeon Kim et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The automobile's knuckle is connected to the parts of the steering system and the suspension system and it is used for adjusting the direction of a rotation through its attachment to the wheel. This study changes the existing material made of GCD45 to Al6082M and recommends the lightweight design of the knuckle as the optimal design technique to be installed in small cars. Six shape design variables were selected for the optimization of the knuckle and the criteria relevant to stiffness and durability were considered as the design requirements during the optimization process. The metamodel-based optimization method that uses the kriging interpolation method as the optimization technique was applied. The result shows that all constraints for stiffness and durability are satisfied using Al6082M, while reducing the weight of the knuckle by 60% compared to that of the existing GCD450.

## 1. Introduction

The linking parts of the steering system and the suspension system of automobiles have a direct impact on the performance of the vehicle's ride, durability, and steerability. Therefore, the performance of these parts is directly related to the quality of the vehicle. This paper examines the structural design of the knuckle, which can adjust the directional rotation due to its connection to the parts of the suspension system and steering system, as well as the wheel. When designing the structure of the knuckle, it is common to consider durability and stiffness [1–5].

For this purpose, the strength of the knuckle under the vehicle's service loads is calculated and the durability is examined. Under the same load condition, the strain or deformation is also calculated and it is verified that this value is within the allowable value. If the design requirement for stiffness is not satisfied, it can be considered that the quality targets such as ride and steerability are not satisfied. In this study, the finite element analysis was used to examine the vehicle's performance of stiffness and durability.

Because the knuckle arm has a greater weight compared to the outer tie rod, inner tie rod, control arm, and ball joint, it can have a greater impact in reducing the overall weight

of the steering parts compared to the other parts. To lessen the weight of the knuckle, a material less dense than steel is used and the design methodology is applied. In this study, aluminum was used as an alternate material to steel and the structural optimization was applied as a design technique.

The shape of the knuckle arm is complex compared to the other parts. Additionally, casting and forging are mainly used in the production process of the knuckle arm. Thus, the knuckle is modeled as a solid element in the finite element analysis, corresponding to the shape optimization in the structural optimization category. The iteration can be stopped due to the mesh distortion during the optimization process in the structural optimization with complex shape. In this study, the metamodel-based optimization technique was applied to solve this problem. The optimization technique using the metamodel is suitable for the following design problems: (1) the long and extensive calculations required by the analysis; (2) the difficulty to mathematically define the shape parameter due to the presence of the complex surfaces; or (3) the severe or broken distortion of the finite element during the process of optimization [3, 4].

The following is a recent research trend related to the knuckle: Triantafyllidis et al. revealed that the fracture mechanism of the knuckle is mainly due to bending fatigue through

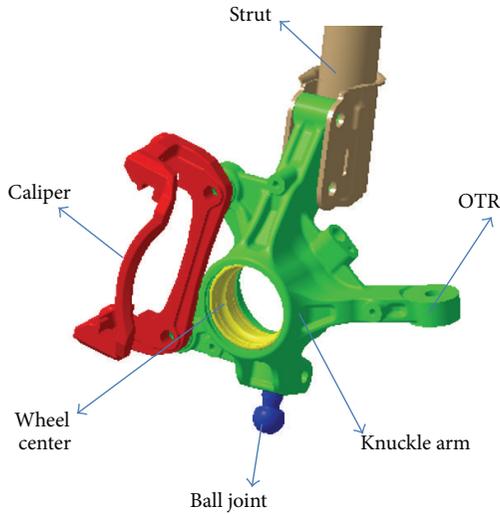


FIGURE 1: Structure of steering parts.

scanning electron microscopy (SEM) images [1]. D'Ippolito et al. suggested that the response surface model for the knuckle's fatigue life was built and then an optimization scheme was introduced considering structural reliability [2]. In [3], the fatigue life of the knuckle was predicted and the reliability was calculated by considering uncertainties. Vijayarangan et al. suggested that a new material made of aluminum alloy is reinforced with titanium carbide particulate to replace spheroidal graphite iron [4]. The aforementioned papers are the studies that mostly dealt with the development of a new material or the reliability of fatigue analysis. In addition, Park et al. [5] have proposed the shape optimization that applies the design of experiments using the orthogonal array to the structural design of the knuckle. However, this method has severe limitations in selecting the optimum design variables among very limited values.

In this study, the lightweight design method that can be applied was suggested at the early stage in the development of the knuckle. First, the base design was completed and six shape design variables were defined. During this process, the shapes of the parts such as the strut, OTR, and spring were fixed. Al6082M, which was developed in the existing Reference [6], was used as the material of the knuckle. In addition, the kriging interpolation method [7–10] was applied as the metamodel technique. MSC/Nastran and MSC/Fatigue were used for durability analysis and Abaqus was used for stiffness analysis.

## 2. Initial Finite Element Analysis of a Knuckle

The knuckle's shape, designed at the early stage of development, is shown in Figure 1. The material used was GCD450, a type of spherical graphite cast iron. The finite element model consisted of the tetrahedron element was created by using Hypermesh. The number of the nodes of the finite element model is 67,128, and the number of the elements is 38,837. The outer tie rod, strut, caliper, ball joint, and hard point of the wheel center, which are the peripheral parts of the knuckle,

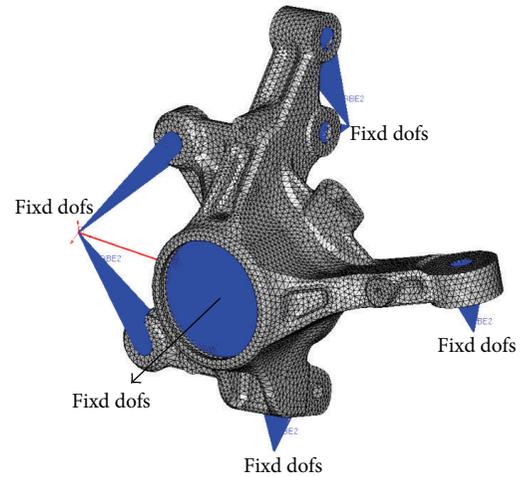


FIGURE 2: FE model of knuckle.

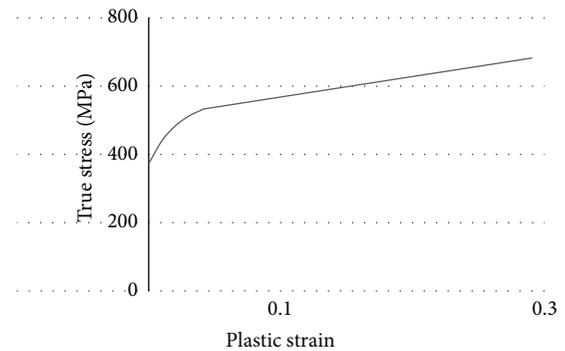


FIGURE 3: Plastic strain and true stress curve.

TABLE 1: Material properties of GCD450.

Property	Value
Elasticity (MPa)	170,000
Density ( $\text{g/cm}^3$ )	7.9
Poisson's ratio	0.29
Tensile strength (MPa)	450
Yield strength (MPa)	370

were modeled such that they were connected by a rigid bar with the joint part of the knuckle. This modeling is shown in Figure 2. The material properties of GCD450 are shown in Table 1 and Figure 3.

**2.1. Stiffness Analysis.** The load that is delivered from the road surface to a car during the operation is transferred to the parts of the suspension and steering systems through the tire and wheel and this load affects the stiffness and strength of the vehicle. If the stiffness of each part is reduced then it induces the excessive deformation that has negative effect on the ride, handling and NVH performance. Thus, each automobile maker sets its own allowable value for the amount of deformation on each car model. In this study, the loading condition and the design criterion that Company A uses were

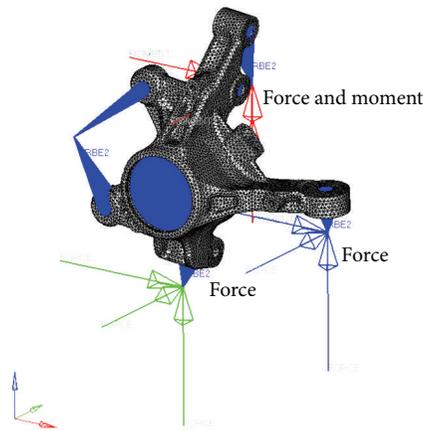


FIGURE 4: FE model for stiffness analysis of forward-braking case.

TABLE 2: Loading condition and case for stiffness.

Number	Loading condition	Number of load case
1	Braking	2
2	Pothole	6
3	Wheel bump	2
4	Cornering	2
Total		12

applied. The number of entire loading cases is 12 and the equivalent plastic strain for each number is calculated and checked that it is within the allowable value.

The equivalent plastic strain,  $\bar{\epsilon}_{pl}$ , is defined by the following equation [11]:

$$\bar{\epsilon}_{pl} = \bar{\epsilon}_{pl}|_0 + \int_0^t \sqrt{\frac{2}{3} \dot{\epsilon}^{pl} : \dot{\epsilon}^{pl}} dt, \quad (1)$$

where the initial equivalent plastic strain, which is the first term of the right-hand side, was set to 0.

As the boundary condition for the stiffness analysis all degrees of freedom of the wheel center, which is the center part of the knuckle hole, were constrained. Twelve load cases for the stiffness analysis can be determined by considering the ultimate load that can be received when the car operates. These load cases are shown in Table 2. A force in each loading case is acted on the hard point of outer tie rod, strut, and ball joint. Additionally, the moment is acted on the strut. The finite element model for the forward-braking load among the loading cases for the stiffness analysis is shown in Figure 4.

The severe result of the stiffness analysis of the knuckle arm came from within the Pothole loading case. The maximum  $\bar{\epsilon}_{pl}$  occurred at the connection area with the strut and its value is 0.0034. This is shown in Figure 5. The calculated maximum  $\bar{\epsilon}_{pl}$  value fully satisfies the acceptance criterion.

**2.2. Durability Analysis.** Automobile parts can fail by fatigue if repeated loading is applied. Therefore, it is essential to review the durability of an entire car unit or parts unit when a new car is developed. In this study, the fatigue life caused by the

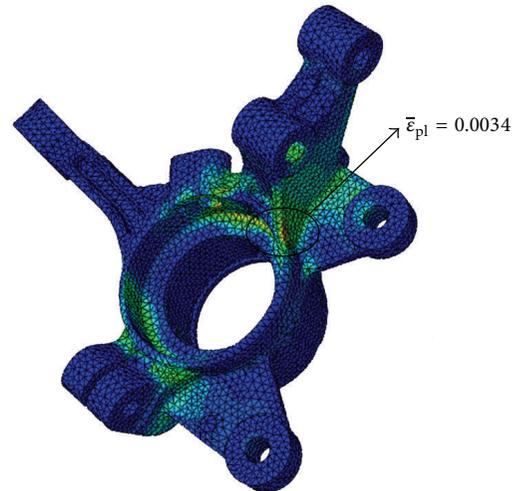


FIGURE 5: Stiffness analysis result at pothole loading condition.

repeated loads acting on the knuckle parts was calculated and it is concluded that this value was less than the allowable value.

Stress-life method and strain-life method are the methods that calculate the fatigue life. The stress-life method is suitable only when stress and strain exist in the elastic area. On the other hand, the strain-life method is suitable for the problems that occur due to stress concentration causing plastic strain. In this study, the strain-life method was used to predict the fatigue life of the knuckle [6, 12].

The loading case for determining the fatigue life of the knuckle arm is 23 loading cases as suggested by the Company A. These loading cases are divided into the nonbraking loading condition and the braking loading condition. The boundary condition of the nonbraking loading condition is the same as the case of the stiffness analysis. On the other hand, in the case of the braking loading condition, all degrees of freedom except for the rotational degree of freedom of Y direction in the wheel center of the knuckle arm were fixed. Also, the element between the caliper and wheel center was modeled as a spring. The FE model of the nonbraking loading condition and the applied load as well as the FE analysis results of this particular case are shown in Figure 6. As the result analyzing the durability by considering all loading cases, the minimum lifetime was calculated to be 635,000 cycles. This lifetime exists within the allowable value.

### 3. Shape Optimization of a Knuckle Using Kriging Metamodel

The results of the stiffness and durability analyses of the knuckle made of material GCD450 show that both results satisfy the criteria and have the marginal safety of about 10 times. In this study, the material of the knuckle was replaced with GCD450 from A16082M and the lightweight design was implemented by applying the metamodel-based optimization using the kriging model. The initial design of the optimization

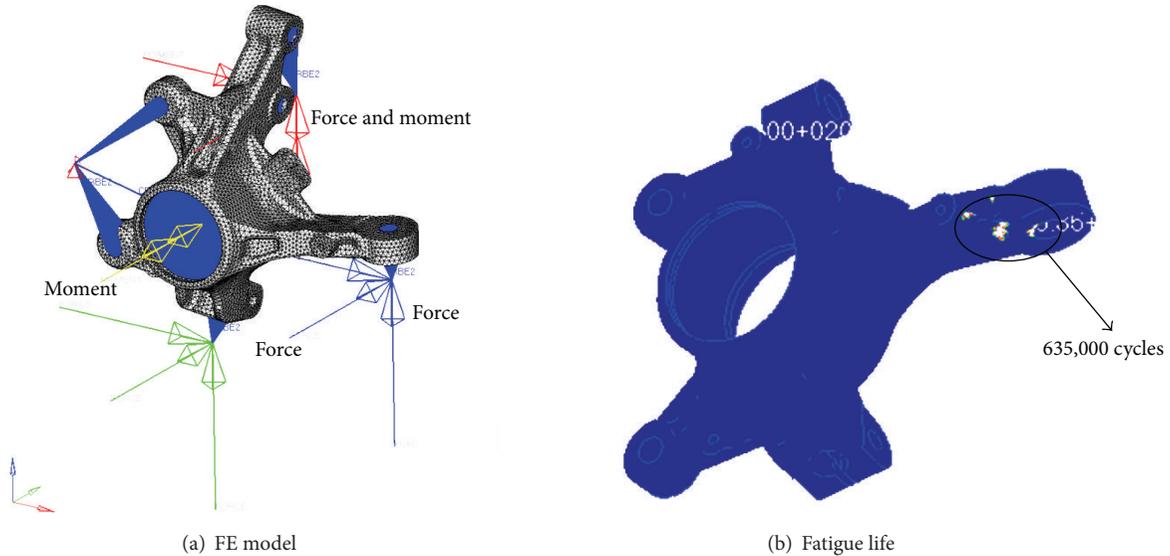


FIGURE 6: FE analysis for durability analysis of nonbraking condition.

is the shape of a knuckle made of GCD450 material. Also, during the optimization process, the most vulnerable loading case for the stiffness analysis and the durability analysis was included on each one. The material property of A16082M is the same as shown in Table 3.

**3.1. Shape Design Variables and Formulation.** The areas that are the most vulnerable in the stiffness and durability analyses are the joint of the outer tie rod and knuckle and the joint of the strut and knuckle. Design variables  $t_1$  and  $t_2$  were defined, respectively, to include the joint between the outer tie rod and knuckle to the shape design variables and design variables  $t_3, t_4, t_5,$  and  $t_6$  were defined, respectively, to include the joint between the strut and knuckle to the variables. These variables are shown in Figure 7.

The formulation for the structural optimization of a knuckle is expressed as follows:

$$\begin{aligned}
 & \text{Minimize } W(\mathbf{t}) \\
 & \text{Subject to } \bar{\epsilon}_{pl}(\mathbf{t}) \leq \bar{\epsilon}_{pla} \\
 & \quad \text{SF}(\mathbf{t}) \geq \text{SF}_a \\
 & \quad \mathbf{t}_L \leq \mathbf{t} \leq \mathbf{t}_U,
 \end{aligned} \tag{2}$$

where  $W(\mathbf{t})$  means the weight of a knuckle, and  $\bar{\epsilon}_{pl}(\mathbf{t})$  means the allowable equivalent plastic strain, and  $\text{SF}(\mathbf{t})$  means the safety factor for fatigue life, and  $\text{SF}_a$  means the allowable safety factor. In addition,  $\mathbf{t} = [t_1, t_2, t_3, t_4, t_5, t_6]$  is the design variable vector, and  $\mathbf{t}_L$  and  $\mathbf{t}_U$  are the lower and upper bounds of the design variable vectors, respectively, that are set as  $\mathbf{t}_L = [18, 16, 20, 20, 20, 20]$  mm and  $\mathbf{t}_U = [28, 24, 32, 32, 32, 32]$  mm.

The lower and upper bounds of each design variable are established by considering the quality of the geometrical shape and the mesh for the finite element analysis of a knuckle. The first inequality equation in (2) is the constraint

TABLE 3: Material properties of A16082M.

Property	Value
Elasticity (MPa)	72,000
Density (g/cm <sup>3</sup> )	2.71
Poisson's ratio	0.29
Tensile strength (MPa)	380
Yield strength (MPa)	340

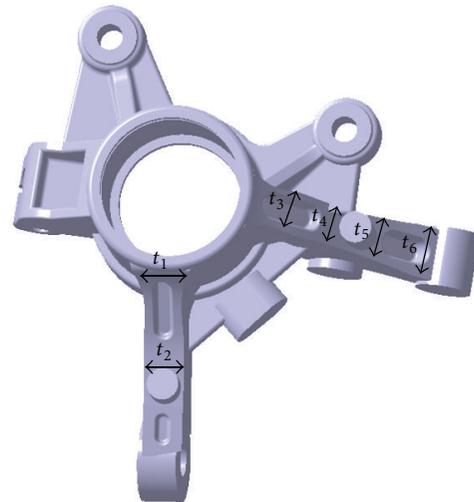


FIGURE 7: Definition of shape design variables.

function related to the stiffness and the second inequality equation is the constraint function for fatigue life. In order to solve (2) by using the metamodel-based optimization method, the surrogate model for  $W, \bar{\epsilon}_{pl},$  and  $\text{SF}$  in (2) should be generated. Then, because (2) was expressed as simple expressions of design variables, obtaining the optimum

design becomes a very easy process. In this study,  $W$ ,  $\bar{\epsilon}_{pl}$ , and SF were approximated as the kriging metamodel.

3.2. *Optimization Using Kriging Interpolation Method.* For global optimization, the kriging interpolation method is introduced. Kriging is an interpolation method named after a South African mining engineer named D. G. Krige, who developed the method while trying to increase the accuracy in predicting the ore reserves. Kriging interpolation for approximation model is well explained in [6–12]. In general, the response function  $f(\mathbf{t})$  is represented as

$$f(\mathbf{t}) = \beta + z(\mathbf{t}), \tag{3}$$

where  $\beta$  is a constant and  $z(\mathbf{t})$  is the realization of a stochastic process with mean zero and variance  $\sigma^2$  following the Gaussian distribution.

If  $\hat{f}(\mathbf{t})$  is designated as the approximation model and the mean squared errors of  $f(\mathbf{t})$  and  $\hat{f}(\mathbf{t})$  are minimized to satisfy the unbiased condition,  $f(\mathbf{t})$  can be estimated as

$$\hat{f}(\mathbf{t}) = \hat{\beta} + \mathbf{r}^T(\mathbf{t}) \mathbf{R}^{-1}(\mathbf{y} - \hat{\beta} \mathbf{i}), \tag{4}$$

where  $\mathbf{R}$  is the correlation matrix,  $\mathbf{r}$  is the correlation vector,  $\mathbf{y}$  is the observed data, and  $\mathbf{i}$  is the unit vector.

In this research,  $W$ ,  $\bar{\epsilon}_{pl}$ , and SF are considered as  $f(\mathbf{t})$ , respectively. Correlation matrix and correlation vector are defined as

$$\mathbf{R}(\mathbf{t}^j, \mathbf{t}^k) = \text{Exp} \left[ -\sum_{i=1}^n \theta_i |t_i^j - t_i^k|^2 \right], \tag{5}$$

$$(j = 1, \dots, n_s, k = 1, \dots, n_s),$$

$$\mathbf{r}(\mathbf{t}) = [R(\mathbf{t}, \mathbf{t}^{(1)}), R(\mathbf{t}, \mathbf{t}^{(2)}), \dots, R(\mathbf{t}, \mathbf{t}^{(n_s)})]^T. \tag{6}$$

The unknown parameters  $\theta_1, \theta_2, \dots, \theta_n$  are obtained by maximizing the following equation:

$$\frac{\left[ n_s \ln(\hat{\sigma}^2) + \ln |\mathbf{R}| \right]}{2}, \tag{7}$$

where  $\theta_i$  ( $i = 1, 2, \dots, n$ )  $> 0$ .

To assess the kriging model, a few metrics can be utilized. In this study, the CV called the cross validation is used. The CV is defined as

$$\text{CV} = \sqrt{\frac{1}{n_s} \sum_{i=1}^{n_s} (f_i - \hat{f}_{-i})^2}, \tag{8}$$

where  $\hat{f}_{-i}$  is the  $i$ th estimator of kriging model constructed without the  $i$ th observation.

The optimization process applied in this study is shown in Figure 8. First, the base design of a knuckle was completed through CATIA and then the finite element modeling was

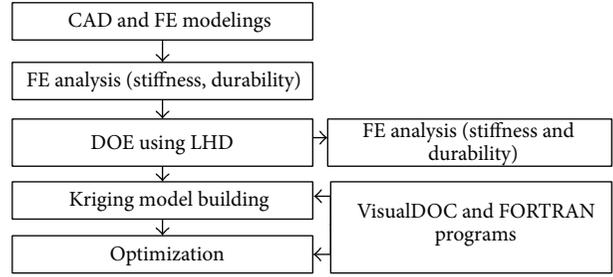


FIGURE 8: Flowchart of the suggested design process.

performed. It was verified that each criterion was met after carrying out the stiffness and durability analyses for the initial model of steel material. The optimization was performed after changing the material into aluminum in order to reduce the weight of the knuckle. First, the shape design variables were defined and then each design variable from the CAD model was parameterized. The Latin hypercube design using the command “*lhsdesign*,” built-in in MATLAB, was used as the sampling method. At this time, the sampling points  $n_s$  were 30. This number is to be empirically determined from previous studies [6, 7, 12]. Then, for each experiment point, the finite element analysis for the stiffness and durability analyses is performed. Each approximated function for  $W$ ,  $\bar{\epsilon}_{pl}$ , and SF is generated by using the kriging interpolation method based on the results of finite element analysis. The optimization problem of (2) was solved by the algorithm of the method of modified feasible direction built in VisualDOC.

3.3. *Optimization Results.* Table 4 shows the results of the finite element analysis of the stiffness analysis and durability analysis for the sampling point generated by the “*lhsdesign*” command. The responses of  $W$ ,  $\bar{\epsilon}_{pl}$ , and SF with respect to each sampling point are listed. The kriging model for  $W$ ,  $\bar{\epsilon}_{pl}$ , and SF were generated based on those values and the parameter values for each kriging model are summarized in Table 5 including the CV value of (8) for each response. In addition, the optimum design obtained from (2) is shown in Table 6 where  $W$ ,  $\bar{\epsilon}_{pl}$ , and SF values are the values predicted from the kriging model. The results of the confirmation analysis through the finite element analysis for the optimal solution were calculated as 1.140 kg for weight, 0.02 for  $\bar{\epsilon}_{pl}$ , and 3.04 for SF. If it is assumed that the result of the finite element analysis in the optimum design is the true value, then the kriging predicted values of  $W$ ,  $\bar{\epsilon}_{pl}$ , and SF have the errors of 0.2%, 0.5%, and 7.3%. The active constraint of two constraint functions for the stiffness and fatigue of (2) is the constraint for stiffness and the constraint for fatigue, which was confirmed as an inactive constraint. In other words, the constraint function for stiffness played a big role in determining this optimum design and was more influential than the constraint for fatigue.

As shown in Table 6, the constraint function for stiffness is not satisfied in the initial design. Therefore, the value of the weight in the initial design is meaningless. In this study, a design that satisfies all of the stiffness and durability

TABLE 4: Design of experiments using LHD.

Number	Design variable (mm)						Response		
	$t_1$	$t_2$	$t_3$	$t_4$	$t_5$	$t_6$	W (kg)	PEEQ	SF
1	21.5	22.53	26.24	27.65	27.33	26.24	1.189	0.0097	2.67
2	23.17	21.73	28.7	29.69	27.06	25.96	1.197	0.0122	3.38
3	21.17	22	25.96	26.63	25.69	24.05	1.176	0.0100	1.94
4	27.83	23.87	28.97	25.27	25.42	24.87	1.196	0.0095	2.94
5	22.5	16.13	24.87	25.95	22.14	23.5	1.147	0.0215	3.09
6	20.17	17.73	28.15	23.23	29.24	22.68	1.167	0.0184	2.82
7	24.17	19.33	30.06	24.59	24.32	28.7	1.175	0.0100	3.18
8	22.83	18.27	25.69	26.97	26.78	24.32	1.173	0.0357	2.67
9	25.17	20.93	24.32	29.01	28.7	25.69	1.197	0.0089	2.82
10	20.5	17.47	29.24	23.57	30.06	26.51	1.175	0.0335	3.16
11	26.17	23.33	27.6	29.35	28.15	27.88	1.211	0.0149	3.09
12	24.5	19.87	24.05	21.53	24.87	28.42	1.165	0.0141	2.99
13	23.83	16.67	29.79	23.91	28.42	30.06	1.18	0.0272	3.23
14	20.83	23.6	26.78	25.61	22.41	26.78	1.17	0.0250	3.18
15	18.5	18.53	22.41	30.03	23.78	27.06	1.161	0.0120	2.5
16	19.83	16.4	29.52	28.33	27.6	22.96	1.171	0.0187	2.91
17	27.17	16.93	22.96	22.55	26.24	29.24	1.169	0.0242	3.04
18	19.17	20.13	24.6	27.99	28.97	24.6	1.179	0.0198	2.82
19	18.17	21.47	23.78	28.67	25.14	29.79	1.174	0.0351	3.04
20	26.83	17.2	25.14	20.51	29.79	25.42	1.174	0.0242	3.13
21	18.83	20.67	22.14	21.19	22.96	22.41	1.139	0.0228	2.89
22	25.5	21.2	22.68	24.93	23.5	28.97	1.173	0.0155	3.18
23	24.83	22.8	27.88	26.29	29.52	27.33	1.203	0.0103	2.74
24	22.17	18	25.42	20.85	24.6	27.6	1.153	0.0367	2.94
25	21.83	19.6	27.33	27.31	23.23	23.23	1.165	0.0091	2.75
26	19.5	20.4	27.06	22.89	24.05	22.14	1.153	0.0206	2.99
27	26.5	23.07	28.42	20.17	25.96	28.15	1.183	0.0357	3.09
28	23.5	18.8	23.5	24.25	22.68	29.52	1.159	0.0257	3.33
29	27.5	19.07	23.23	21.87	27.88	23.78	1.174	0.0136	2.89
30	25.83	22.27	26.51	22.21	26.51	25.14	1.181	0.0249	2.99

TABLE 5: Optimum parameter and validation of kriging model.

Responses	Optimum parameter*							Validation
	$\beta^*$	$\theta_1^*$	$\theta_2^*$	$\theta_3^*$	$\theta_4^*$	$\theta_5^*$	$\theta_6^*$	CV
W	1.1749	4.4769	7.9588	0.0341	0.2377	2.2620	51.316	0.014
PEEQ	0.0170	2.0487	1.9495	0.3696	0.0483	0.2500	0.2140	0.084
SF	2.8378	20.705	28.048	0.0006	5.7071	0.0021	0.0003	3.129

requirements of a knuckle with the minimum weight is proposed using an optimization process. The final shape determined through this study is shown in Figure 9. Though this study focused on the specific car, the proposed design process can be applicable for all kinds of knuckle.

**4. Conclusion and Future Work**

In this study of a lightweight design of the knuckle mounted to a small car, the material was changed from steel to aluminum and the metamodel-based optimization was applied. The results are as follows.

(1) When aluminum was first adapted to the existing design, the constraint function for stiffness was not met. In the proposed design, the shape of the knuckle was redesigned and the optimal minimum weight was calculated, achieving a weight reduction of about 60% (as compared to the initial design) without sacrificing the stiffness and durability requirements. The weight reduction of the knuckle can make a direct contribution to improved fuel efficiency and reduced emissions.

(2) It could be seen that the approximated optimization that uses kriging for the lightweight design of a knuckle is very effective for the shape optimization that is difficult to

TABLE 6: Responses at initial and optimum designs.

Design	Deign variable (mm)						Response		
	$t_1$	$t_2$	$t_3$	$t_4$	$t_5$	$t_6$	$W$ (kg)	$\bar{\epsilon}_{pl}$	SF
Initial (steel)	18.0	22.0	23.0	24.0	21.5	27.5	3.4	0.0034	12.7
Initial (Al)	18.0	22.0	23.0	24.0	21.5	27.5	1.160	0.025	3.1
Optimum	19.3	20.6	20.0	21.8	20.6	26.2	1.142	0.0199	2.8

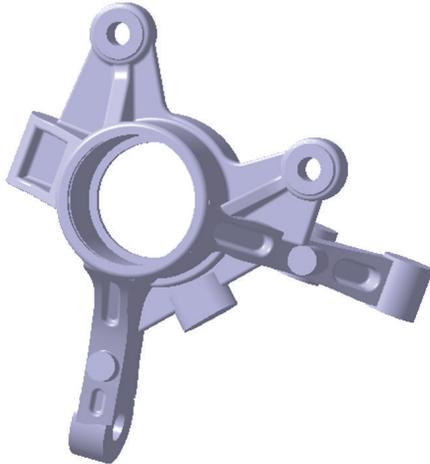


FIGURE 9: Suggested optimum design of a knuckle.

implement in existing commercial software. In addition, the validation of the kriging model was carried out through the cross validation and CV index and the predicted values of the kriging model for the weight, equivalent plastic strain, and safety factor of fatigue life had no significant differences compared to the results from the finite element analysis. For future work, the forging of the optimum shape of the knuckle proposed is scheduled to be reviewed.

### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

### Acknowledgments

This research was financially supported by the Ministry of Education Science and Technology (MEST) and the National Research Foundation of Korea (NRF) through the Human Resource Training Project for Regional Innovation (2012H1B8A2026078).

### References

- [1] G. K. Triantafyllidis, A. Antonopoulos, A. Spiliotis, S. Fedonos, and D. Repanis, "Fracture characteristics of fatigue failure of a vehicle's ductile iron steering knuckle," *Journal of Failure Analysis and Prevention*, vol. 9, no. 4, pp. 323–328, 2009.
- [2] R. d'Ippolito, M. Hack, S. Donders, L. Hermans, N. Tzannetakis, and D. Vandepitte, "Improving the fatigue life of a vehicle

knuckle with a reliability-based design optimization approach," *Journal of Statistical Planning and Inference*, vol. 139, no. 5, pp. 1619–1632, 2009.

- [3] E. A. Azrulhisham, Y. M. Asri, A. W. Dzuraidah, N. M. Nik Abdullah, A. Shahrum, and C. H. Che Hassan, "Evaluation of fatigue life reliability of steering knuckle using pearson parametric distribution model," *International Journal of Quality, Statistics, and Reliability*, vol. 2010, Article ID 816407, 8 pages, 2010.
- [4] S. Vijayarangan, N. Rajamanickam, and V. Sivananth, "Evaluation of metal matrix composite to replace spheroidal graphite iron for a critical component, steering knuckle," *Materials & Design*, vol. 43, pp. 532–541, 2013.
- [5] Y. C. Park, K. H. Lee, D. H. Lee, and K. Y. Lee, "Shape optimization design of the knuckle using orthogonal array and the finite element analysis," *Transactions of the Korean Society of Automotive Engineers*, vol. 11, pp. 138–144, 2003 (Korean).
- [6] B.-C. Song, Y.-C. Park, S.-W. Kang, and K.-H. Lee, "Structural optimization of an upper control arm, considering the strength," *Proceedings of the Institution of Mechanical Engineers D: Journal of Automobile Engineering*, vol. 223, no. 6, pp. 727–735, 2009.
- [7] X. G. Song, J. H. Jung, H. J. Son, J. H. Park, K. H. Lee, and Y. C. Park, "Metamodel-based optimization of a control arm considering strength and durability performance," *Computers and Mathematics with Applications*, vol. 60, no. 4, pp. 976–980, 2010.
- [8] J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn, "Design and analysis of computer experiments," *Statistical Science*, vol. 4, pp. 409–435, 1989.
- [9] A. Guinta and L. Watson, "A comparison of approximation modeling techniques: polynomial versus interpolating models," in *Proceedings of the 7th AIAA/USAF/NASA/ISSMO Symp on Multid Anal and Optim (AIAA '98)*, vol. 2, pp. 392–440, 1998.
- [10] K. T. Fang, R. Li, and A. Sudjianto, *Design and Modeling for Computer Experiments*, Chapman & Hall/CRC, 2006.
- [11] Simulia, "Abaqus 6.10 analysis user's manual: materials," vol. 3.
- [12] J. K. Kim, Y. J. Kim, W. H. Yang, Y. C. Park, and K.-H. Lee, "Structural design of an outer tie rod for a passenger car," *International Journal of Automotive Technology*, vol. 12, no. 3, pp. 375–381, 2011.

## Research Article

# Mathematical Modelling of Thermal Process to Aquatic Environment with Different Hydrometeorological Conditions

**Alibek Issakhov**

*Al-Farabi Kazakh National University, Almaty 050040, Kazakhstan*

Correspondence should be addressed to Alibek Issakhov; [alibek.issakhov@mail.ru](mailto:alibek.issakhov@mail.ru)

Received 30 January 2014; Accepted 10 February 2014; Published 2 June 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Alibek Issakhov. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents the mathematical model of the thermal process from thermal power plant to aquatic environment of the reservoir-cooler, which is located in the Pavlodar region, 17 Km to the north-east of Ekibastuz town. The thermal process in reservoir-cooler with different hydrometeorological conditions is considered, which is solved by three-dimensional Navier-Stokes equations and temperature equation for an incompressible flow in a stratified medium. A numerical method based on the projection method, divides the problem into three stages. At the first stage, it is assumed that the transfer of momentum occurs only by convection and diffusion. Intermediate velocity field is solved by fractional steps method. At the second stage, three-dimensional Poisson equation is solved by the Fourier method in combination with tridiagonal matrix method (Thomas algorithm). Finally, at the third stage, it is expected that the transfer is only due to the pressure gradient. Numerical method determines the basic laws of the hydrothermal processes that qualitatively and quantitatively are approximated depending on different hydrometeorological conditions.

## 1. Introduction

Many years in the study of hydrodynamics of lakes and reservoirs were two independent directions; one of them is the analysis of data and the other is mathematical modelling. Natural experiments—observations, although made in a variety of complex conditions, were passive, as they do not enable you to manage the experiment; they do not seem possible based on their prediction of hydrophysical processes. Performed experiments and calculations were not convincing enough, because the models were not calibrated and are usually not verified. Most often, operated models are available to the developer or the user but are not optimal from the point of view of the problem.

The lakes are observed not only by the wave movement but also by the vortices which have spatial scales that could be compared to the size of the reservoir, or to the portions thereof, mushroom formations, meanders, jet currents, trails, and torches waters of different origins. As shown by observations of flows and a number of indirect factors, there is significant variability of dynamic processes typical for areas

with complex geometry of bottom and coasts. These topographic features are manifested in the formation of eddies and meandering streams. Spectrum observation of lakes' vortices and wave motions widens and extends from microvortices leading to energy dissipation to large-scale eddies, covering the entire lake. Description of large-scale topographic eddies could be done by using the conservation law of potential vortices.

One of the most effective methods of studying the hydrodynamics of the lake is a method of mathematical modeling. In some cases, this may be the only tool to predict changes in the hydrological regime and lake ecosystems, for example, when studying the changes that may occur in spatial redistribution of water, while constructing waterworks and other events associated with the use of water objects.

Mathematical models can be classified according to several criteria. It can be classified according to their model "dimension": one-dimensional (vertical or horizontal), two-dimensional (horizontal or vertical plane), and three-dimensional model.

The most simple are one-dimensional models that are commonly used for modeling of currents in rivers. Two-dimensional models are used to study wind and seiche flows, storm surges, and so forth.

Lick [1] proposes to consider the following types of mathematical models of wind currents like (1) integrated model (full flow), in which vertical integration over the flow is accounted but vertical profile of flows is not modeled, (2) stationary models of wind currents for constant and variable density of water, and (3) nonstationary model for barotropic and baroclinic lakes.

To study the wind currents, Shang et al. [2] divided the models to Ekman, integrated by vertical direction, multilevel, and multilayered. In addition to these models, the dynamic method and a variety of three-dimensional thermohydrodynamic models are used for calculation of lakes' flows.

It is accepted to divide modelling into two classes—prognostic and diagnostic. In the first class, the formation of interconnected flow fields, temperature, and boundary layers of the atmosphere and the lake are simulated. Solution of this problems class involves great difficulties of a numerical simulation of unsteady nonlinear partial differential equations. The use of data obtained from the observation of temperature and wind fields greatly simplifies the problems of circulation in lakes. This is the meaning of diagnostic solutions, which are widely used in the class of oceanographic problems and subsequently for the study of flows in deep stratified lakes.

It can be attributed to the special class models that simulate mesoscale movements such as inertial oscillations, seiche flows, Kelvin, topographic and Poincare waves, the flow in the zone of coastal upwelling, and coherent structure.

The model originally developed for the ocean is widely used in the modeling of the dynamics of lake, especially large and deep. At the same time, taking into account the specific features of the thermodynamics of lakes and their small size compared to the ocean, the presence of the maximum density at about 4°C is remarkable property of the largest stratified lakes that leads to the formation of the thermal bar, which mostly determines the dynamics of the lake as well as its chemical and biological mode. Moreover, it is important that the low tide, the beta effect, and absence of Rossby waves that are typical for the ocean could be related to the small size lakes too. However, the heterogeneity of the bottom topography in the lakes contributes to the formation of topographic Rossby waves, analogues oceanic Rossby waves. In closed basins, standing waves, seiche, which, if sufficiently large, lakes split into fashion with the formation of waves Kelvin and Poincare, play an important role.

## 2. Background

Getting electricity from thermal power plants (TPP) had a higher priority than its impact on the environment. Technology of production of electrical energy from power plant is connected with a lot of waste heat released into the environment. Today, the problem of influence on the nature by power is particularly acute because the pollution of the atmosphere and hydrosphere increases each year.

The energy consumption scale is increasing year by year; as a result, negative impact of energy on the environment also increases. Before, to get energy, primarily guided feasibility was economic costs, but now, in the construction and operation of energy, the most important issue is their impact on the environment.

Another problem of TPP is thermal pollution to the reservoir or lake. Dropping hot water-push chain reaction that begins reservoir overgrown with algae, it violates the oxygen balance, which in turn is a threat to the life of all its inhabitants. Thermal power plants with cooling water shed 4–7 kJ of heat for 1 kW/h electricity generation. Meanwhile, according to the Health Standards, discharged warm water from TPP should not rise higher than 3°C in the summer and 5°C in the winter from the initial temperature of the reservoir.

Spread of harmful emissions from TPP depends on several factors: the terrain, environmental temperature, wind speed, cloud cover, precipitation intensity, and so forth. Existing meteorology conditions like wind velocity and so forth increase the thermal pollution area.

Large proportion of electricity (81.3%) in the world is produced by thermal power plants. Therefore, emissions of this type of power plants to the atmosphere and hydrosphere provide the greatest amount of anthropogenic contaminants in it.

Thermal pollution of reservoirs or lakes water that cause multiple violations of their state is one representation of environment danger. Thermal power plants generate energy through turbines, driven by hot steam, while the exhaust steam is cooled by water. Therefore, from the power plants in the reservoirs or lakes, this process is continuously transferred from the water flow temperature at 8–12°C above the temperature of the water in the reservoir. Large TPP sheds till 90 m<sup>3</sup>/s of heated water. For example, according to estimates of German and Swiss scientists, the possibility of rivers of Switzerland and the upper flows of the Rhine on the heating have been exhausted. Hot water at any place of the river should not exceed more than 3°C maximum temperature of the river water, which is assumed to be 28°C. Following these conditions, the power station of Germany, constructed on the Rhine, Inna, Weser, and Elbe, is limited by 35 000 MW. Thermal pollution can lead to tragic consequences. Scientists predict change in the characteristics of the environment in the next 100–200 years.

Let us consider hydrosphere pollution. Heat from TPP mainly is given to the environment from the water-cooled condenser steam turbines. The value of heat released to the environment depends on the capacity of thermal power plants. If we consider power plants, from 40 to 70% of the amount of diverted energy to the environment is taken from thermal energy released by combustion fuel. Cooling water in and direct-flow-back scheme of intake and discharge of water are limited by the local allowable increase in the temperature of the source water (river, lake, or reservoir) in the thermal effects. Water supply system has a number of features of TPP. Almost all of the water up to 95% of total cost is applied to cool the condenser coils and auxiliary steam turbines. With up to 5% of the total value of the water supply to the thermal

power plant equipment is generally irreversible consumption. As a rule, the main building of the condensing power plant is located directly at the shore line of the river-, lake-, or reservoir-cooler. Water is supplied to the main unit of heat removal to the environment pumping stations. After heating it in condensers and heat exchangers, water is discharged to the surface of the water. However, this amount of water is heated.

### 3. Study Area

In this work, Ekibastuz SDPP-1 is considered, which is located in Pavlodar region, 17 km to the north-east of the Ekibastuz city, Kazakhstan. It is taken as an example of the impact of TPP on the reservoir-cooling. Technical water supply of SDPP-I was carried on the recircuit with cooling water circulation. The surface of the reservoir is at 158.5 m above the sea, the area is 19.6 km<sup>2</sup>, the maximum size is 4 × 6 km, the average depth is 4.6 m, maximum depth is 8.5 m at the intake, and the volume of the reservoir is 80 million m<sup>3</sup>. Moreover, combined type of selective intake and discharge is used in the body of the water. Discharged water enters the prechannel mixer and then through a filtration dam uniformly enters the reservoir-cooling. Water intake is at a distance of 40 m from the dam and the depth is 5 m. Design flow of water is 120 m<sup>3</sup>/s, and the actual flow rate varies depending on the mode of TPP within 80–120 m<sup>3</sup>/s.

### 4. Materials and Methods

Now hydrophysical problems associated with discharge of heated water into reservoirs in the operation of thermal or nuclear power plant become important. The discharged heated water is an important hydrologic and environmental problem. That is the reason why we have to predict and control the temperature of the water and the content of impurities in the reservoirs and rivers. Successful study of the processes occurring in the reservoir involves a complex study of the problem: instead of measurements taken from hydrothermal parameters, it is better to use mathematical modeling of the processes and then compare results of the modelling of physical process in laboratory and natural conditions. To construct mathematical models, it is necessary to consider the main characteristic of the flow in a reservoir-turbulent fluid motion. This in turn merges with one of the major problems of hydrodynamics-theory of turbulence.

Thermal and nuclear power plants both require reservoirs. Electricity production is increasing worldwide and especially from postwar period, doubling every 7–10 years. Large amounts of water are required to operate these plants for cooling units, in average 35–40 m<sup>3</sup>/s for 1 million kW of installed capacity. Hence, it is evident that for the thermal power plants of 2.4 million kW 70–160 m<sup>3</sup>/s of water is required. Therefore, water supply becomes important when we choose to build coal and nuclear power plants. Naturally, the large thermal power plants should be located on the banks of large rivers, ponds, lakes, or artificial reservoirs. The creation of artificial reservoirs requires large investment,

so power stations tend to have existing reservoirs and lakes. Often, industrial facilities located on the shores of lakes and reservoirs disposed of with warm water waste products in the form of impurities. If we consider that in the most developed countries for 2010–2020 years cooling thermal power plants and industrial facilities will use more than 10% of water resources, the problems of optimal and efficient use of water reservoir for cooling are in great importance. In solving these problems, there is a need to be able to predict and control the temperature of the water and the spread of passive pollutants of reservoir.

In recent years, association put tough restrictions to protection of the environment. The designer of artificial reservoir usually has to follow the rules, which limit the size of “zone transfer” by fault of hot water so that it does not exceed half the width of the river and occupy no more than half of the total cross-sectional area and flow. If these rules are not followed, it may lead to short-term or long-term stop of power plant. That is why the accuracy requirements to constructive analysis are very strict. In fact, emerging with the hydrodynamic problem, the process can be described as fully three-dimensional, with irregular boundaries, with the presence of buoyancy and the velocity of the main flow, which can vary by an order, sometimes so fast that the important role plays the effects of nonstationary. In addition, there are large areas of recycling if certain combinations of conditions were applied when the fault-heated water is almost drawn into the upstream region of cooling water. The result could be a significant loss of total operating efficiency of the system.

### 5. Mathematical Model

From the above, it follows that the construction of a mathematical model relevant to real processes in the reservoir-cooler is quite a challenge.

There are many mathematical and numerical models that have been developed to simulate distribution temperature after launching TPP in reservoir-cooler [3–5].

The impact of thermal or nuclear power plant on the hydrological and biological conditions of the reservoir is various. Most of all, thermal pollution often reaches 30–35 degrees of heated water. This increases the water temperature and adversely affects hydrobiological condition, self-purification of water quality of the reservoir. In the reservoir-cooler, spatial change of temperature is small. Therefore, stratified flow in the reservoir-cooler can be described by equations in the Boussinesq approximation. For the mathematical modeling of the system motion, continuity and temperature equations are considered. The development of spatial turbulent stratified flows in reservoir-cooler is considered in [6–9]. Three-dimensional model is used for mathematical modelling of temperature distribution in the reservoir-cooler [10–12]. Consider

$$\frac{\partial u_i}{\partial t} + \frac{\partial u_j u_i}{\partial x_j} = -\frac{\partial p}{\partial x_i} + \nu \frac{\partial}{\partial x_j} \left( \frac{\partial u_i}{\partial x_j} \right) - \frac{\partial \tau_{ij}}{\partial x_j} + \beta g_i (T - T_0), \quad (1)$$

$$\frac{\partial u_j}{\partial x_j} = 0, \tag{2}$$

$$\frac{\partial T}{\partial t} + \frac{\partial u_j T}{\partial x_j} = \frac{\partial}{\partial x_j} \left( \chi \frac{\partial T}{\partial x_j} \right), \tag{3}$$

$$\text{where } \tau_{ij} = \overline{u_i u_j} - \overline{u_i} \overline{u_j}, \tag{4}$$

where  $g_i$  is the gravity acceleration,  $\beta$  the coefficient of volume expansion,  $u_i$  velocity components,  $\chi$  thermal diffusivity coefficient,  $T_0$  the equilibrium temperature, and  $T$  deviation of temperature from the balance.

This system of equations was filtered by using large eddy simulation (LES) method. The basic idea of LES method is a mathematical division of the large and small universal vortices. This procedure can be performed through spatial averaging, that is, to define the field of large-scale quantities by filter. Consider

$$\overline{u}(x, t) = \int_V G(\mathbf{r}, \mathbf{x}) u(x - r, t) dr, \tag{5}$$

where  $\mathbf{u} = (u_1, u_2, u_3)$ —vector of velocity components, sign “dash” denotes averaging,  $\mathbf{x} = (x_1, x_2, x_3)$ —vector of coordinate system,  $\mathbf{r} = (r_1, r_2, r_3)$ —vector of coordinate system by which the integration is done,  $V$  is volume of integration, and  $G(\mathbf{r}, \mathbf{x})$  is filter function with characteristic length scale such that

$$\int_V G(\mathbf{r}, \mathbf{x}) d\mathbf{r} = 1. \tag{6}$$

Small-scale fluctuations are as follows:

$$u'(x, t) = u(x, t) - \overline{u}(x, t), \tag{7}$$

and, in many cases, depending on the filter function, it looks like

$$\overline{\mathbf{u}'}(\mathbf{x}, t) \neq 0. \tag{8}$$

There are different approaches to characterize the filter:

(i) “Box” filter:

$$G(\mathbf{r}) = \frac{1}{\Delta^3} \begin{cases} 1, & |r| \leq \frac{\Delta}{2} \\ 0, & |r_i| > \frac{\Delta}{2}, \end{cases} \quad \forall i = 1, 2, 3. \tag{9}$$

(ii) Gaussian filter:

$$G(\mathbf{r}) = \prod_{i=1}^3 \left( \frac{6}{\pi \Delta^2} \right)^{1/2} \exp\left(-\frac{6r_i^2}{\Delta^2}\right). \tag{10}$$

(iii) Cut filter:

$$G(\mathbf{r}) = \prod_{i=1}^3 \frac{\sin(\pi r_i / \Delta)}{\pi r_i}. \tag{11}$$

$\Delta$  is characteristic length of the filter, which is the order of the mesh size. It is usually taken like [13]

$$\Delta = (h_1 h_2 h_3)^{1/3},$$

$$\Delta = (h_1^2 + h_2^2 + h_3^2)^{1/2}, \tag{12}$$

$$\Delta = \min(h_1, h_2, h_3),$$

where  $h_i$  is step size corresponding to axes of a Cartesian coordinate system.

We start with regular LES corresponding to a “bar-filter” of  $\Delta x$  width and an operator associating with the function  $\overline{f}(\overline{x}, t)$ . Then we define a second “test filter” tilde of large  $2\Delta x$  width associating with  $\widetilde{f}(\overline{x}, t)$ . Let us first apply this filter product to the Navier-Stokes equation. The subgrid-scale tensor of the field  $\widetilde{u}_i$  is obtained from (4) by replacing filter bar to double filter and tilde filter:

$$\tau_{ij} = \widetilde{\overline{u_i u_j}} - \widetilde{\overline{u_i}} \widetilde{\overline{u_j}}, \tag{13}$$

$$l_{ij} = \widetilde{\overline{u_i u_j}} - \widetilde{\overline{u_i}} \widetilde{\overline{u_j}}. \tag{14}$$

Now we apply the tilde filter to (4), which leads to

$$\widetilde{\tau}_{ij} = \widetilde{\overline{u_i u_j}} - \widetilde{\overline{u_i}} \widetilde{\overline{u_j}}. \tag{15}$$

Adding (14) and (15) and using (13), we obtain

$$l_{ij} = \tau_{ij} - \widetilde{\tau}_{ij}. \tag{16}$$

We use Smagorinsky model expression for the subgrid stresses related to the bar filter and tilde-filter to get

$$\widetilde{\tau}_{ij} - \frac{1}{3} \delta_{ij} \widetilde{\tau}_{kk} = 2C \widetilde{A}_{ij}, \quad \text{where } A_{ij} = (\Delta x)^2 \left| \overline{S} \right| \overline{S}_{ij}. \tag{17}$$

Further on, we have to determine  $\tau_{ij}$ , the stress resulting from the filter product. This is again obtained using Smagorinsky model, which yields to

$$\tau_{ij} - \frac{1}{3} \delta_{ij} \tau_{kk} = 2CB_{ij}, \quad \text{where } B_{ij} = (2\Delta x)^2 \left| \widetilde{S} \right| \widetilde{S}_{ij}. \tag{18}$$

Subtracting (17) from (18) and using Germano’s identity, we get

$$l_{ij} - \frac{1}{3} \delta_{ij} l_{kk} = 2CB_{ij} - 2C \widetilde{A}_{ij}, \tag{19}$$

$$l_{ij} - \frac{1}{3} \delta_{ij} l_{kk} = 2CM_{ij}, \tag{20}$$

$$\text{where } M_{ij} = B_{ij} - \widetilde{A}_{ij}. \tag{21}$$

All the terms of (21) may now be determined by using  $\overline{u}$ . Unfortunately, there are five independent equations for only one variable  $C$  and thus the overdetermined problem. The first solution was proposed by Germano to multiply (21) tensorially by  $\overline{S}_{ij}$  to get

$$C = \frac{1}{2} \frac{l_{ij} \overline{S}_{ij}}{M_{ij} \overline{S}_{ij}}. \tag{22}$$

This provides finally dynamical evaluation of  $C$ , which can be used in the LES for the bar field  $\bar{u}$  [13, 14].

Initial and boundary conditions are defined so that they satisfy the nonstationary three-dimensional equations of motion, continuity, and temperature.

### 6. Numerical Algorithm

Numerical solution of (1)-(2) is carried out on the staggered grid using the scheme against a stream of the second type. Moreover, compact approximation is used for convective terms [15, 16]. In view of the above with the proposed model of turbulence scheme of splitting on physical parameters is used to solve the problem. At the first stage, the transfer of momentum occurs only through convection and diffusion. Intermediate velocity field is solved by using fractional step method through the tridiagonal matrix method (Thomas algorithm) [16, 17]. The second stage is for pressure which is found by intermediate velocity field. Three-dimensional Poisson equation for pressure is solved by Fourier method in combination with the tridiagonal matrix method (Thomas algorithm) that is applied to determine the Fourier coefficients [16, 18]. At the third stage, it is supposed that the transfer is carried out only by the pressure gradient. The algorithm was parallelized on the high-performance system [16]. And we can mathematically propose this algorithm like

$$\begin{aligned} \text{(i)} \quad & \frac{\bar{u}^* - \bar{u}^n}{\tau} = -(\nabla \bar{u}^n \bar{u}^* - \nu \Delta \bar{u}^*), \\ \text{(ii)} \quad & \Delta p = \frac{\nabla \bar{u}^*}{\tau}, \\ \text{(iii)} \quad & \frac{\bar{u}^{n+1} - \bar{u}^*}{\tau} = -\nabla p. \end{aligned} \tag{23}$$

For the first stage, intermediate velocity field is solved by using fractional step method through the tridiagonal matrix method (Thomas algorithm):

$$\begin{aligned} \frac{f^{n+1/3} - f^n}{\tau} &= \frac{1}{2} \Lambda_1 f^{n+1/3} + \frac{1}{2} \Lambda_1 f^n + \Lambda_2 f^n + \Lambda_3 f^n, \\ \frac{f^{n+2/3} - f^{n+1/3}}{\tau} &= \frac{1}{2} \Lambda_2 f^{n+2/3} - \frac{1}{2} \Lambda_2 f^n, \\ \frac{f^* - f^{n+2/3}}{\tau} &= \frac{1}{2} \Lambda_3 f^* - \frac{1}{2} \Lambda_3 f^n, \end{aligned} \tag{24}$$

where the operators are like

$$\begin{aligned} \Lambda_1 f &= -\frac{\partial(\bar{u}_1^n f)}{\partial x_1} + \nu \frac{\partial^2 f}{\partial x_1^2} - \frac{\partial \tau_{i1}}{\partial x_1}, \\ \Lambda_2 f &= -\frac{\partial(\bar{u}_2^n f)}{\partial x_2} + \nu \frac{\partial^2 f}{\partial x_2^2} - \frac{\partial \tau_{i2}}{\partial x_2}, \\ \Lambda_3 f &= -\frac{\partial(\bar{u}_3^n f)}{\partial x_3} + \nu \frac{\partial^2 f}{\partial x_3^2} - \frac{\partial \tau_{i3}}{\partial x_3}, \\ f &= \bar{u}_i, \quad i = 1, 2, 3. \end{aligned} \tag{25}$$

The second stage is for pressure which is found by intermediate velocity field. Three-dimensional Poisson equation for pressure is solved by Fourier method for one coordinate in combination with the tridiagonal matrix method (Thomas algorithm). It is applied to determine the Fourier coefficients [16, 18]. The numerical algorithm for Poisson equation was parallelized on the high-performance system [16]. Mathematically, we can write Fourier method as follows:

$$p_{i,j,k} = \frac{2}{N_3} \sum_{l=0}^{N_3} \rho_l a_{i,j,l} \cos \frac{\pi kl}{N_3}, \tag{26}$$

$$F_{i,j,k} = \frac{2}{N_3} \sum_{l=0}^{N_3} \rho_l b_{i,j,l} \cos \frac{\pi kl}{N_3},$$

where

$$a_{i,j,l} = \sum_{k=0}^{N_3} \rho_k p_{i,j,k} \cos \frac{\pi kl}{N_3}, \tag{27}$$

$$b_{i,j,l} = \sum_{k=0}^{N_3} \rho_k F_{i,j,k} \cos \frac{\pi kl}{N_3}.$$

Substituting the equation above into three-dimensional Poisson equation for pressure, we obtain the following expression:

$$\begin{aligned} \frac{H^2}{L_1^2} \frac{2}{N_3} \sum_{l=0}^{N_3} \rho_l \frac{(a_{i+1,j,l} - 2a_{i,j,l} + a_{i-1,j,l})}{\Delta x_1^2} \cos \frac{\pi kl}{N_3} + \frac{H^2}{L_2^2} \frac{2}{N_3} \\ \times \sum_{l=0}^{N_3} \rho_l \frac{(a_{i,j+1,l} - 2a_{i,j,l} + a_{i,j-1,l})}{\Delta x_2^2} \\ \times \cos \frac{\pi kl}{N_3} + \frac{2}{N_3} \\ \times \sum_{l=0}^{N_3} \frac{\rho_l a_{i,j,l}}{\Delta x_3^2} \left( \cos \frac{\pi(k+1)l}{N_3} - 2 \cos \frac{\pi kl}{N_3} + \cos \frac{\pi(k-1)l}{N_3} \right) \\ = \frac{2}{N_3} \sum_{l=0}^{N_3} \rho_l b_{i,j,l} \cos \frac{\pi kl}{N_3}. \end{aligned} \tag{28}$$

Using the expression below,

$$\cos \frac{\pi(k+1)l}{N_3} + \cos \frac{\pi(k-1)l}{N_3} = 2 \cos \frac{\pi kl}{N_3} \cos \frac{\pi l}{N_3}, \tag{29}$$

we can write (28) in the following form:

$$\begin{aligned} & \frac{H^2}{L_1^2} \frac{2}{N_3} \sum_{l=0}^{N_3} \frac{\rho_l (a_{i+1,j,l} - 2a_{i,j,l} + a_{i-1,j,l})}{\Delta x_1^2} \cos \frac{\pi kl}{N_3} \\ & + \frac{H^2}{L_2^2} \frac{2}{N_3} \sum_{l=0}^{N_3} \frac{\rho_l (a_{i,j+1,l} - 2a_{i,j,l} + a_{i,j-1,l})}{\Delta x_2^2} \\ & \times \cos \frac{\pi kl}{N_3} + \frac{2}{N_3} \sum_{l=0}^{N_3} \frac{\rho_l a_{i,j,l}}{\Delta x_3^2} \left( 2 \cos \frac{\pi l}{N_3} - 2 \right) \\ & \times 2 \cos \frac{\pi kl}{N_3} = \frac{2}{N_3} \sum_{l=0}^{N_3} \rho_l b_{i,j,l} \cos \frac{\pi kl}{N_3}. \end{aligned} \tag{30}$$

The last expression can be written at a fixed value  $l$  and divided by  $(2/N_3)\rho_l \cos(\pi kl/N_3)$  and then we get

$$\begin{aligned} & \frac{H^2}{L_1^2} \frac{a_{i+1,j} - 2a_{i,j} + a_{i-1,j}}{\Delta x_1^2} + \frac{H^2}{L_2^2} \frac{a_{i,j+1} - 2a_{i,j} + a_{i,j-1}}{\Delta x_2^2} \\ & + \frac{a_{i,j}}{\Delta x_3^2} \left( 2 \cos \frac{\pi l}{N_3} - 2 \right) = b_{i,j}. \end{aligned} \tag{31}$$

Furthermore, this equation is transformed to the following form:

$$\begin{aligned} & -\frac{H^2}{L_2^2} \frac{a_{i,j-1}}{\Delta x_2^2} + \left[ \left( \frac{H^2}{L_1^2} \frac{2}{\Delta x_1^2} + \frac{H^2}{L_2^2} \frac{2}{\Delta x_2^2} - \frac{1}{\Delta x_2^2} \right. \right. \\ & \left. \left. \times \left( 2 \cos \frac{\pi l}{N_3} - 2 \right) \right) a_{i,j} \right. \\ & \left. - \frac{H^2}{L_1^2} \frac{a_{i+1,j} + a_{i-1,j}}{\Delta x_1^2} \right] - \frac{H^2}{L_2^2} \frac{a_{i,j+1}}{\Delta x_2^2} = -b_{i,j}. \end{aligned} \tag{32}$$

In vector form, this equation can be written as follows:

$$-A_j \vec{a}_{j-1} + B_j \vec{a}_j - C_j \vec{a}_{j+1} = \vec{F}_j, \tag{33}$$

where  $A_j, B_j, C_j$  matrices and  $\vec{F}_j, \vec{a}_j$  vectors are taken in that form:

$$\begin{aligned} \vec{a}_j &= \begin{bmatrix} a_{0,j} \\ \vdots \\ a_{N_1,j} \end{bmatrix}, \quad A_j = \begin{bmatrix} \frac{H^2}{L_2^2} \frac{1}{\Delta x_2^2} & & 0 \\ & \ddots & \\ 0 & & \frac{H^2}{L_2^2} \frac{1}{\Delta x_2^2} \end{bmatrix}, \\ B_j &= \begin{bmatrix} \frac{H^2}{L_1^2} \frac{2}{\Delta x_1^2} + \frac{H^2}{L_2^2} \frac{2}{\Delta x_2^2} - \frac{1}{\Delta x_2^2} \left( 2 \cos \frac{\pi l}{N_3} - 2 \right) - \frac{H^2}{L_1^2} \frac{2}{\Delta x_1^2} & & 0 \\ & -\frac{H^2}{L_1^2} \frac{1}{\Delta x_1^2} & \frac{H^2}{L_1^2} \frac{1}{\Delta x_1^2} \\ 0 & & -\frac{H^2}{L_1^2} \frac{2}{\Delta x_1^2} \frac{H^2}{L_1^2} \frac{2}{\Delta x_1^2} + \frac{H^2}{L_2^2} \frac{2}{\Delta x_2^2} - \frac{1}{\Delta x_2^2} \left( 2 \cos \frac{\pi l}{N_3} - 2 \right) \end{bmatrix}, \\ C_j &= \begin{bmatrix} \frac{H^2}{L_2^2} \frac{1}{\Delta x_2^2} & & 0 \\ & \ddots & \\ 0 & & \frac{H^2}{L_2^2} \frac{1}{\Delta x_2^2} \end{bmatrix}, \quad \vec{F}_j = - \begin{bmatrix} b_{0,j} \\ \vdots \\ b_{N_1,j} \end{bmatrix}. \end{aligned} \tag{34}$$

Tridiagonal matrix algorithm (Thomas algorithm) for (33) looks like

$$\begin{aligned} & \alpha_{j+1} = (C_j - A_j \alpha_j)^{-1} B_j, \\ & j = 1, 2, \dots, N_2 - 1, \quad \alpha_1 = C_0^{-1} B_0, \\ & \vec{\beta}_{j+1} = (C_j - A_j \alpha_j)^{-1} (\vec{F}_j + A_j \vec{\beta}_j), \\ & j = 1, 2, \dots, N_2, \quad \vec{\beta}_1 = C_0^{-1} \vec{F}_0, \end{aligned}$$

$$\vec{a}_j = \alpha_{j+1} \vec{a}_{j+1} + \vec{\beta}_{j+1},$$

$$j = N_2 - 1, \dots, 0, \quad \vec{a}_{N_2} = \vec{\beta}_{N_2+1}.$$

(35)

The third is a correction stage; it is supposed that the transfer is carried out only by the pressure gradient. After calculating  $a_{i,j,k}$ , pressure field values are found from (33). To calculate the sum (26), it is necessary to apply the fast Fourier transformation. That allows calculating the method by  $O(N \ln N)$ . And finally the temperature equation (3) is

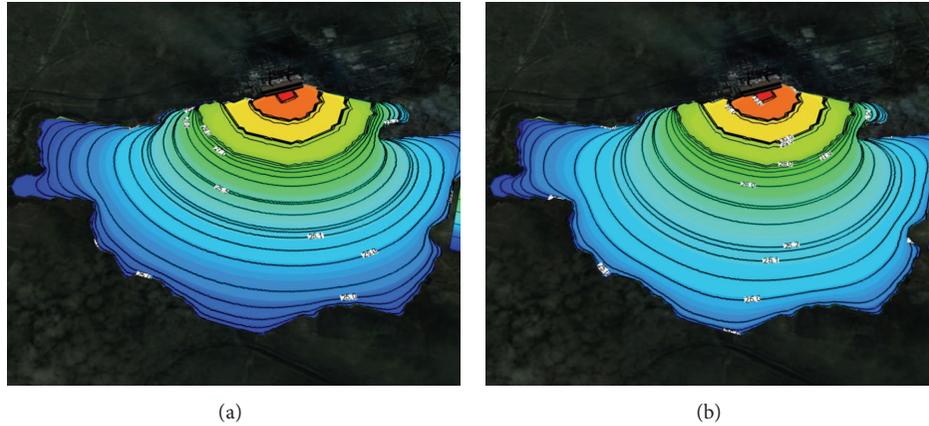


FIGURE 1: Outline and contours of temperature at 15 and 24 hours after launch of Ekibastuz SDPP-1 on the surface and the side view.

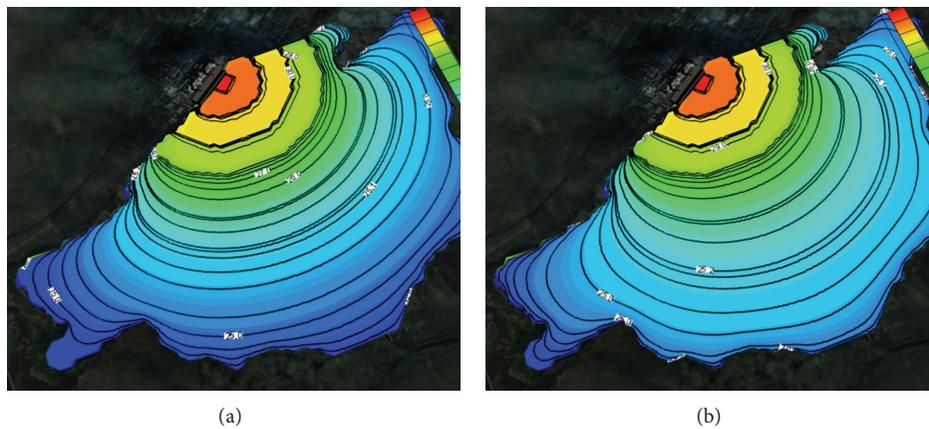


FIGURE 2: Outline and contours of temperature at 15 and 24 hours after launch of Ekibastuz SDPP-1 on the surface and the top view.

also solved by using fractional step method through the tridiagonal matrix method (Thomas algorithm).

### 7. Results of Numerical Modelling

In the simulation, the mesh size of  $200 \times 200 \times 200$  was used. Figures 1 and 2 show the solved three-dimensional spatial outline and contour of the temperature distribution at different times after the launch of Ekibastuz SDPP-1, on the surface, from different angles of view. Figures 3 and 4 show the solved spatial contour, contour of temperature at different times at the west wind after the launch of Ekibastuz SDPP-1, on the surface, from different angles of view. Figures 5 and 6 show the solved spatial contour, contour of temperature and velocity vectors at different times at the north-west wind after the launch of Ekibastuz SDPP-1, on the surface, from different angles of view. In all figures, we can see that temperature varies from  $25^{\circ}\text{C}$  to  $33^{\circ}\text{C}$ . Moreover, it can be observed that temperature on the surface of reservoir-cooler near Ekibastuz SDPP-1 is higher than that at a far distance from Ekibastuz SDPP-1. It means that mathematical model qualitatively describes the physical process. All the figures show that temperature distribution in some distance from

Ekibastuz SDPP-1 approaches isothermal distribution. The results show that the temperature distribution is spread over the larger area of the reservoir-cooler. In all the figures, the simulation was done with different hydrometeorological conditions merged with real relief, which was taken from satellite pictures.

### 8. Discussions

LES is a more universal approach to close the system of equations which was filtered by Favre approach. A necessary condition for the performance of turbulent closures is “subgrid” model that correctly describes the dissipation of the kinetic energy of smoothed velocity fluctuations and the ability to simulate the circuit direct energy cascade from large to small eddies. This stage is the primary mechanism for the redistribution of energy in the inertial range of three-dimensional homogeneous isotropic turbulence. The principal advantage of the LES from RANS is that, due to the relative homogeneity and isotropy of the small-scale turbulence, plotting a subgrid model is much simpler than the used turbulence models for RANS, when it is necessary to model the full range of turbulence. For the same reason, the hope for

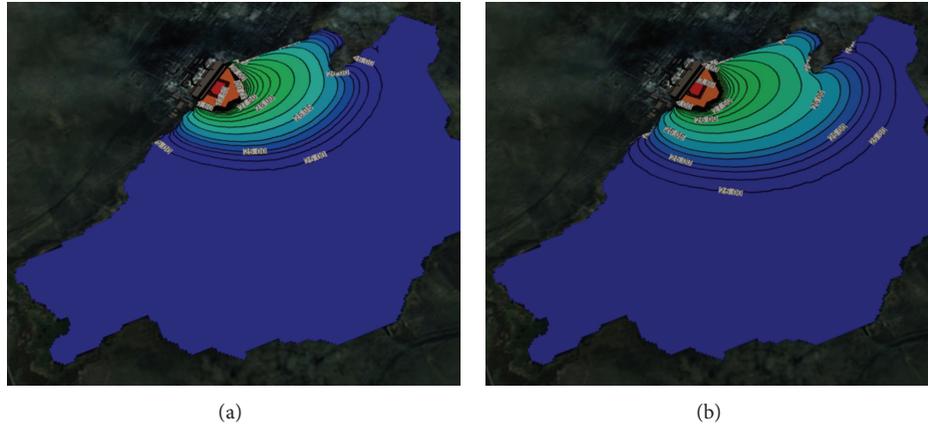


FIGURE 3: Outline and contours of temperature at 15 and 24 hours at the west wind after the launch of Ekibastuz SDPP-I on the surface and the side view.

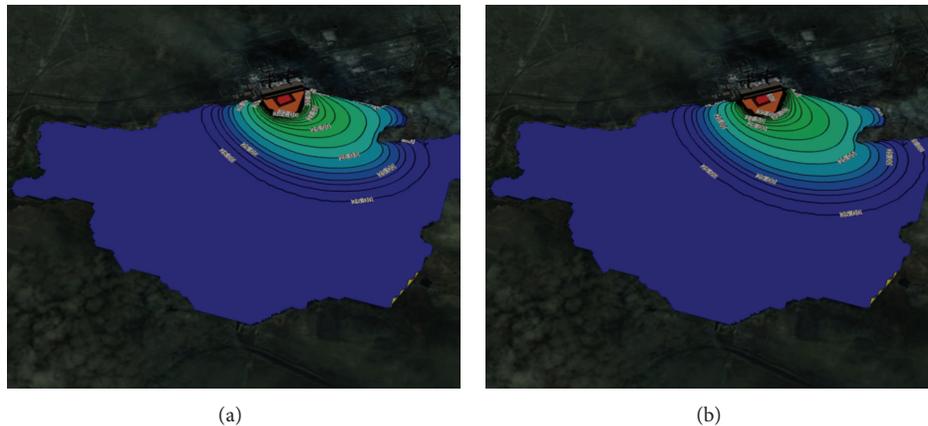


FIGURE 4: Outline and contours of temperature at 15 and 24 hours at the west wind after the launch of Ekibastuz SDPP-I on the surface of the water and the top view.

a “universal” subgrid model for LES is much more reasonable than a similar model for RANS. These important benefits of LES increase significantly computational cost associated with the need (also for direct numerical simulation (DNS) case) of three-dimensional time-dependent calculations on sufficiently fine grids, even also in cases where direct interest in the practice of the average flow is two-dimensional and stationary. On the other hand, for obvious reasons, the computational resources which are required to implement the LES are much smaller than those for the DNS. The degree of influence of different processes governing the formation of stratified flows and hydrothermal conditions in the entire body of water can be divided into two zones. The first (near) zone is directly adjacent to the water of outlet structures. The second is for the major part of the reservoir. In the near zone, formation of the stratified flow is influenced by the processes of mixing discharged water with water from the reservoir. It should be regulated by creating a specific hydraulic regime in the outfalls. In the second, zone of hydrothermal regime is formed primarily by the processes of heat transfer. The propagation of heat in this part of the reservoir is more dependent on the wind (direction and speed). When you

spread the heated water in a cold environment, density difference between the upper layer of warm water and bottom layer of cold water appears. This allows the use of a combined intake and discharge instead of building costly diversion canals to the discharge. Accordingly, this raises the problem of optimal choice of the geometrical and operational parameters of the reservoir-cooler for efficient work of power plant.

## 9. Conclusions

Thus, the usage of a mathematical model of three-dimensional stratified turbulent flow gives us a possibility which approximately qualitatively and quantitatively determines the basic laws of the hydrothermal processes occurring in the reservoir-cooler. Performed earlier from computational studies of hydrothermal regime of Ekibastuz SDPP-I reservoir-cooler, the velocity and temperature fields measurements were done. These fields have revealed the basic laws of hydrothermal and thermal fields in the water with stationary and under various hydrometeorological conditions. The distribution of temperature and passive scalar affect not

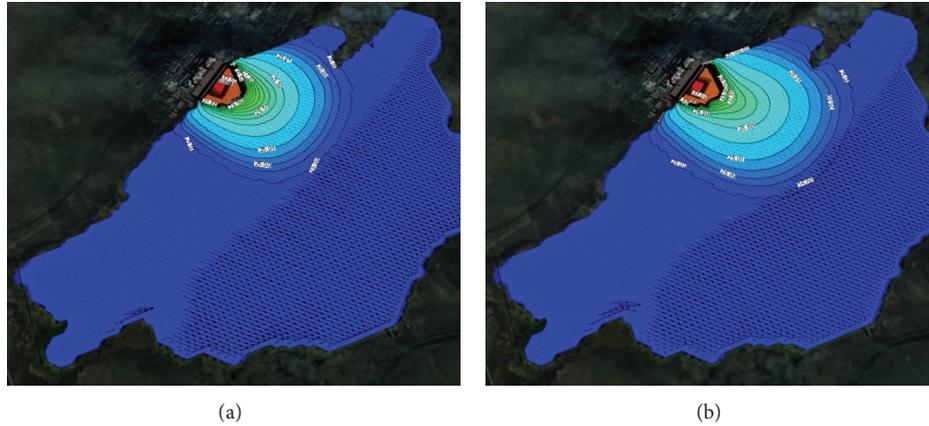


FIGURE 5: Outline and contours of temperature at 15 and 24 hours at the north-west wind after the launch of Ekibastuz SDPP-1 on the surface and the top view.

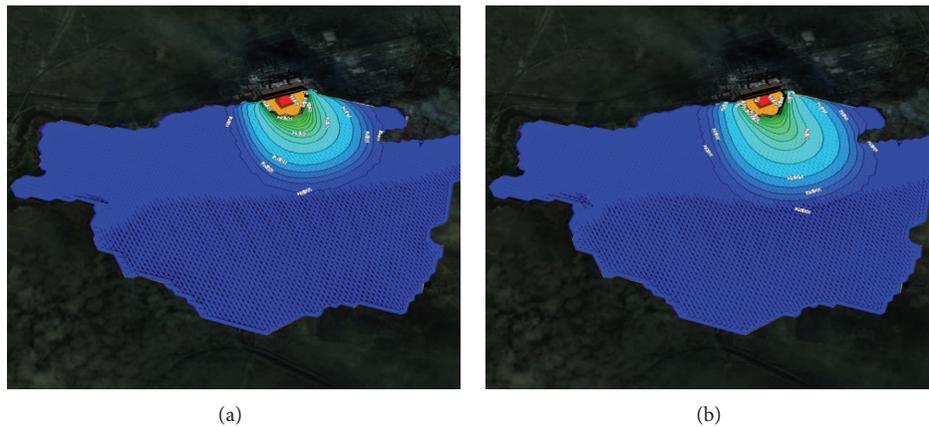


FIGURE 6: Outline and contours of temperature at 15 and 24 hours at the north-west wind after the launch of Ekibastuz SDPP-1 on the surface of the water and the side view.

only the processes of heat and mass but also the density stratification. Stratification appears in connection with the difference between the density of discharged water and the density of surrounding water in the pond, or the presence of impurities in the discharged water. For example, the heated water is light, so it is in the form of a jet or standing stretches near the free surface. Sustainable density stratification of water reduces turbulent exchange between the vertical layers of fluids, especially in the area with big difference. In general, hydrothermal regime of the reservoir is formed under the influence of uncontrollable natural factors (solar and atmospheric radiation, wind, convective heat transfer, evaporation, etc.) and the factors which may be adjusted (the amount and temperature of the discharged water, the presence of impurities, selective sampling, etc.).

More detailed mathematical model and data analysis are necessary to simulate more accurate thermal process in the reservoir-cooler. That is also relevant for future research direction too.

## Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

## References

- [1] W. Lick, "Numerical models of lakes currents," *Annual Review of Earth and Planetary Sciences*, vol. 4, pp. 49–74, 1976.
- [2] Y. P. Shang, W. Lick, R. T. Gedsey, and F. B. Molls, "Numerical computation of three-dimensional circulation in Lake Erie: a comparison of a free-surface model and a rigid-lid model," *Journal of Physical Oceanography*, vol. 8, no. 4, pp. 713–727, 1978.
- [3] Y. L. Zhang, A. M. Baptista, and E. P. Myers, "A cross-scale model for 3D baroclinic circulation in estuary-plume-shelf systems: I. Formulation and skill assessment," *Continental Shelf Research*, vol. 24, no. 18, pp. 2187–2214, 2004.
- [4] L. Y. Oey, "An OGCM with movable land-sea boundaries," *Ocean Modelling*, vol. 13, no. 2, pp. 176–195, 2006.

- [5] R. T. Cheng and V. Casulli, "Modeling a three-dimensional river plume over continental shelf using a 3D unstructured grid model," in *Proceedings of the 8th International Conference on Estuarine and Coastal Modeling*, pp. 1027–1043, Monterey, Calif, USA, 2004.
- [6] C. A. Fletcher, *Computational Techniques for Fluid Dynamics: Specific Techniques for Different Flow Categories*, vol. 2, Springer, Berlin, Germany, 1988.
- [7] P. J. Roache, *Computational Fluid Dynamics*, Hermosa, Albuquerque, NM, USA, 1972.
- [8] R. Peyret and T. D. Taylor, *Computational Methods for Fluid Flow*, Springer, Berlin, Germany, 1983.
- [9] J. C. Tannehill, D. A. Anderson, and R. H. Pletcher, *Computational Fluid Mechanics and Heat Transfer*, McGraw-Hill, New York, NY, USA, 2nd edition, 1997.
- [10] S. A. Lowe, F. Schuepfer, and D. J. Dunning, "Case study: three-dimensional hydrodynamic model of a power plant thermal discharge," *Journal of Hydraulic Engineering*, vol. 135, no. 4, pp. 247–256, 2009.
- [11] A. Issakhov, "Mathematical modelling of the influence of thermal power plant to the aquatic environment by using parallel technologies," in *Proceedings of the 6th Global Conference on Power Control and Optimization*, vol. 1499, pp. 15–18, AIP, Las Vegas, Nev, USA, 2012.
- [12] A. Issakhov, "Mathematical modelling of the influence of thermal power plant on the aquatic environment with different meteorological condition by using parallel technologies," in *Power, Control and Optimization*, vol. 239 of *Lecture Notes in Electrical Engineering*, pp. 165–179, 2013.
- [13] M. Lesieur, O. Metais, and P. Comte, *Large Eddy Simulation of Turbulence*, Cambridge University Press, New York, NY, USA, 2005.
- [14] H. Tennekes and J. L. Lumley, *A First Course in Turbulence*, The MIT Press, 1972.
- [15] A. I. Tolstykh, *Compact Difference Scheme and Their Applications to Fluid Dynamics Problems*, Nauka, Moscow, Russia, 1990.
- [16] A. Issakhov, "Large eddy simulation of turbulent mixing by using 3D decomposition method," *Journal of Physics*, vol. 318, no. 4, Article ID 42051, 2011.
- [17] N. N. Yanenko, "The method of fractional steps," in *Space Matrix Computations*, J. B. Bunch and D. J. Rose, Eds., Academic Press, Springer, New York, NY, USA, 1979.
- [18] A. Issakhov, "Development of parallel algorithm for numerical solution of three-dimensional Poisson equation," *Journal of Communication and Computer*, vol. 9, no. 9, pp. 977–980, 2012.

## Research Article

# Design of Heat Exchanger for Ericsson-Brayton Piston Engine

**Peter Durcansky, Stefan Papucik, Jozef Jandacka, Michal Holubcik, and Radovan Nosek**

*University of Zilina, Univerzitná 8215/1, 010 26 Zilina, Slovakia*

Correspondence should be addressed to Peter Durcansky; [peter.durcansky@fstroj.uniza.sk](mailto:peter.durcansky@fstroj.uniza.sk)

Received 28 February 2014; Accepted 2 April 2014; Published 28 May 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Peter Durcansky et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Combined power generation or cogeneration is a highly effective technology that produces heat and electricity in one device more efficiently than separate production. Overall effectiveness is growing by use of combined technologies of energy extraction, taking heat from flue gases and coolants of machines. Another problem is the dependence of such devices on fossil fuels as fuel. For the combustion turbine is mostly used as fuel natural gas, kerosene and as fuel for heating power plants is mostly used coal. It is therefore necessary to seek for compensation today, which confirms the assumption in the future. At first glance, the obvious efforts are to restrict the use of largely oil and change the type of energy used in transport. Another significant change is the increase in renewable energy—energy that is produced from renewable sources. Among machines gaining energy by unconventional way belong mainly the steam engine, Stirling engine, and Ericsson engine. In these machines, the energy is obtained by external combustion and engine performs work in a medium that receives and transmits energy from combustion or flue gases indirectly. The paper deals with the principle of hot-air engines, and their use in combined heat and electricity production from biomass and with heat exchangers as primary energy transforming element.

## 1. Introduction

*Microcogeneration Unit with Nonconventional Engine.* Combustion engines with gas as fuel, with the mostly used being natural gas, are most used as power unit of microcogeneration devices. Losses in electricity generation are mainly associated with imperfect energy transformation in burning fuel in an imperfect transformation of energy working medium in the turbine. There are also represented mechanical losses and loss of energy in transmission lines. The minimal losses have cogeneration plants. Cogeneration unit is a technical device, which is manufactured by electric and thermal energy simultaneously. As an example cogeneration unit may be mentioned with an internal combustion gas engine. The engine burns the gas, thereby gaining the mechanical power on the shaft to drive an electric generator. The engine has no classic cooler but has the heat exchanger from which we obtain thermal energy. Used heat exchangers are connected in series circuits, where the working medium, usually water, is preheated and heated in several stages. Overall efficiency of CHP unit can be increased with multistage heat recovery and this reduces also the total cost of fuel [1]. Unconventional

engines are a possible alternative to the internal combustion engines. They work with external combustion, which allows, unlike conventional internal combustion engines, controlling the course of combustion and therefore its quality, which is reflected in the composition of air pollutants emitted to the atmosphere. The most known hot-air engines are Stirling and Ericsson engines. Ericsson engine is also an external combustion engine. In contrast to Stirling engine, it has two possible alternatives—open and closed [2]. In the case of Stirling engine dual function of regenerator is immediately apparent. Regenerator works as heater and cooler while in Ericsson engine cooler and heater are separated. Figure 1 presents Ericsson-Brayton engine with open cycle.

The air is compressed in the compressor and flows through the heat exchanger, where at constant pressure it is receiving heat. Consequently, it is led to the expansion cylinder, which expands adiabatically. Part of this work will be used to drive the compressor and part is used as mechanical work to drive an electric generator. As the heat source can be used almost any fuel for burning, as it is an external combustion engine. Fuel is burned in a separate combustion chamber and heat energy is transformed through

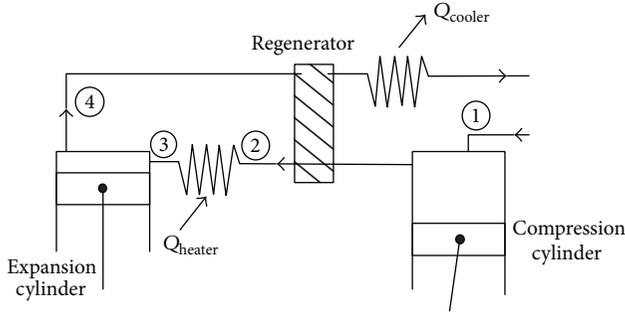


FIGURE 1: Scheme of Ericsson-Brayton hot-air engine with open cycle.

a heat exchanger to the working media. The working medium in open cycle, mostly dry air, after passing the cycle, is discharged into the atmosphere. In a closed cycle the medium after each cycle cools in refrigerant heat exchanger, where it gives heat energy and is fed back into the cycle. With the use of closed cycle we can improve the efficiency of heating equipment [3, 4].

The proposed microgeneration unit uses two heat exchangers: cooler and heater (see Figure 2). A different purpose sets other requirements for the heat exchangers. The first requirement is to ensure optimal heat transfer between flowing media. The heat transfer is characterized by a heat transfer coefficient. This summary represents the characteristics of the heat exchanger, its layout, and the flowing media. Coefficient depends on the characteristics of the flowing media, from the heat capacity, and the selected construction option and in some cases is significantly influenced by the material used and the heat exchanger. The requirement is that the coefficient is the highest while respecting the chosen solutions. Further requirements are then asked to compact size exchanger and the total pressure loss and also maintenance options are required [3, 4].

## 2. Heat Exchanger Design

As the first step, the working conditions of the CHP were set. The experimental application with Ericsson-Brayton hot-air engine sets a wide range of specifications, not only on the heat exchanger, but also on the whole system. The whole unit should supply energy for household. In the determination of the operating conditions we have preliminary set the highest temperatures from 500°C up to 620°C, according to [3, 5]. In this paper, the authors presented the highest temperature of 600°C. Mr. Creyx [1] has presented systems with different working fluids and also different hot-air engine configurations. The system presented in this paper should work with closed cycle, with dry air as working fluid [4]. The closed cycle enables heat recovery from working fluid, so the regenerated heating power is bigger than that in opened cycle, where the most part of heat energy is used to preheat the air after compression. It is assumed that temperature of the working fluid after expansion is in the range of 240°C–320°C [4, 5]. For each working fluid, the dry air in the tubes and the

TABLE 1: Coefficient  $k_T$  for dry air by pressure  $10^5$  Pa.

$T$ [K]	$t$ [°C]	$k_T$
373,15	100	1,054403
393,15	120	1,066696
413,15	140	1,075804
433,15	160	1,087817
453,15	180	1,103398

TABLE 2: Coefficient  $k_p$  for dry air by temperature 273 K.

$p$ [Pa]	$p$ [bar]	$k_p$
10	$10^{-4}$	0,464348
$10^2$	$10^{-3}$	0,880435
$10^3$	$10^{-2}$	0,984783
$10^4$	$10^{-1}$	0,993333
$10^5$	1	1,000000
$5 \cdot 10^5$	5	1,003509

exhaust gases outside the tubes were set as the characteristic temperatures and physical properties.

There are many ways for how to calculate the properties of flowing media. In order to determine the heat transfer, it was necessary to know the thermodynamic properties of flowing gas. It is important to determine the dynamic and kinematic viscosity. For heat transfer the thermal conductivity of the gas is also needed to be known. The following equations were used for calculation [6, 7].

Dynamic viscosity is

$$\eta_{TP} = 1.0607 \cdot 10^{-6} \cdot T^{0,5} \cdot k_T \cdot k_p. \quad (1)$$

Kinematic viscosity is

$$\nu_{TP} = 304.52344 \cdot 10^{-6} \frac{T^{1,5}}{p} k_T k_p, \quad (2)$$

and the thermal conductivity is

$$\lambda = 1513.8151 \cdot 10^{-6} T^{0,5} (k_T k_p)^{1,5}. \quad (3)$$

The coefficients  $k_T$  and  $k_p$  are based on temperatures from 0°C up to 1000°C. The main difference to real values of parameters is up to 3%, so it is possible to say that the computation is accurate. Table 1 shows the values of coefficient  $k_T$ .

The values of the coefficient  $k_p$  are set for dry air by constant temperature. In Table 2, some values of coefficient  $k_p$  can be seen.

There are many methods for calculation of flue gas density. Two of them were used in this work. At first it is possible to read the right values in the tables, which are calculated or measured. In Table 3 there are some values of air properties. The second column is density. The values that are not in the table can be calculated.

Density can be calculated based on known parameters. In the following equation density is expressed as a function of dynamic and kinematic viscosity:

$$\rho = \frac{\mu}{\nu}. \quad (4)$$

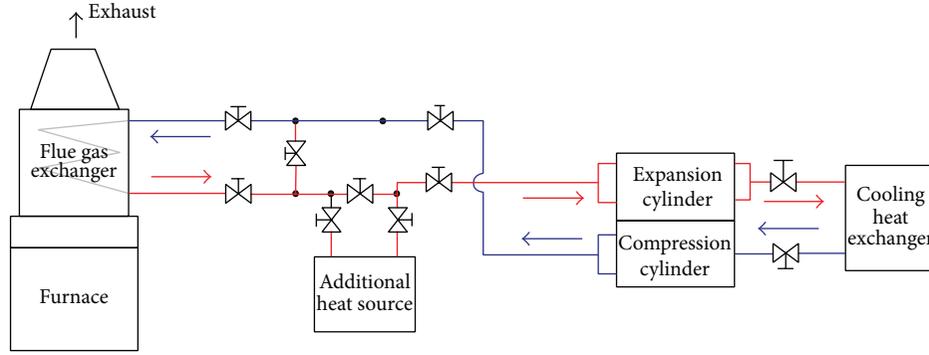


FIGURE 2: Experimental scheme of Ericsson-Brayton hot-air engine with open cycle.

TABLE 3: Physical properties for dry air by pressure 100 kPa.

$t$ [°C]	$\rho$ [kg/m <sup>3</sup> ]	$c$ [J/kg·K]	$\lambda \cdot 10^2$ [W/(m·K)]	$a \cdot 10^6$ [m <sup>2</sup> /s]
0	1,275	1005	2,37	18,5
10	1,23	1005	2,45	19,82
20	1,188	1010	2,52	21
40	1,112	1013	2,65	23,53
60	1,046	1017	2,8	26,32
80	0,986	1020	2,93	29,13
100	0,934	1022	3,07	32,16
120	0,886	1024	3,2	35,27
140	0,843	1027	3,33	38,46
160	0,804	1030	3,44	41,54
180	0,769	1034	3,57	44,9
200	0,736	1037	3,7	48,48

So in this way we can define the properties of flowing medium. Specifying the geometrical properties or features of the chosen type of heat exchanger is also very important. There are many basic concepts of heat exchangers. Based on the geometrical features or heat transfer, methods can be classified into many classes. For the purpose of this work, pipe exchanger was selected. The heat exchanger in this category differs in the arrangement of tubes. The tubes can be organized straight or staggered or partly staggered. It is characterized with the dimensionless constants “a” and “b.”

If the tube bundle has horizontal spacing “s<sub>1</sub>” and vertical spacing “s<sub>2</sub>,” as in Figure 3, the bundle can be characterized with these constants:

$$\begin{aligned}
 a &= \frac{s_1}{d_0}, \\
 b &= \frac{s_2}{d_0}, \\
 \psi &= 1 - \frac{\pi}{4 \cdot a}.
 \end{aligned}
 \tag{5}$$

The streamed length “l” can be expressed as length of flow path across over a single tube [7]:

$$l = \frac{\pi}{2} \cdot d_0.
 \tag{6}$$

Another difference is in the nondimensional criteria. Reynolds number is characterizing the flowing medium and the type of flow. It depends on flow velocity and also on the geometry. For heat transfer through tubes in bundle the following Reynolds number criteria were used:

$$Re = \frac{w \cdot l}{\psi \cdot \nu}.
 \tag{7}$$

Nusselt number is characterizing the heat transfer. If the turbulence in the inflowing medium is low, deviations in the Nusselt number may occur. The average Nusselt number in a cross-flow over a bundle of smooth tubes can be calculated from that in a cross-flow over a single tube. For the purpose of this work, the criteria equation was used according to [7, 8]. The heat transfer is described by the 2 parts of flow, the turbulent part and the laminar part of the flow near the walls, as follows:

$$\begin{aligned}
 Nu_{l,lam} &= 0.664 \cdot \sqrt{Re_{\psi,l}} \cdot \sqrt[3]{Pr}, \\
 Nu_{l,turb.} &= \frac{0.037 \cdot Re_{\psi,l}^{0,8} \cdot Pr}{1 + 2.443 \cdot Re_{\psi,l}^{-0,1} \cdot (Pr^{2/3} - 1)}.
 \end{aligned}
 \tag{8}$$

Turbulent flow in pipe sets in at  $Re > 10^4$ . In the transition region of Reynolds number from 2300 to  $10^4$  the type of flow

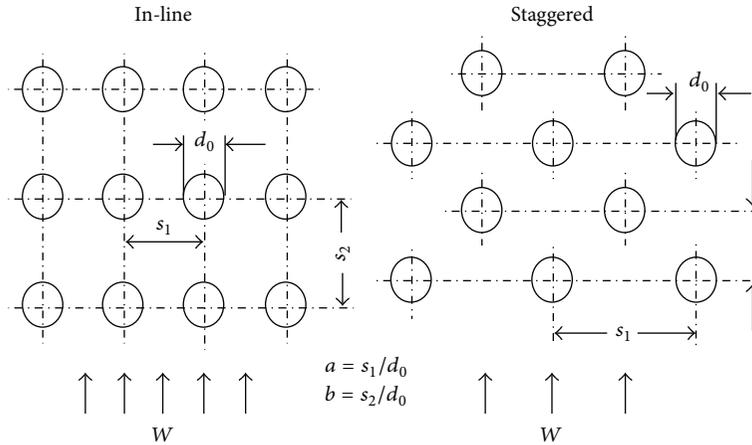


FIGURE 3: Lateral and longitudinal spacing in tube bundles.

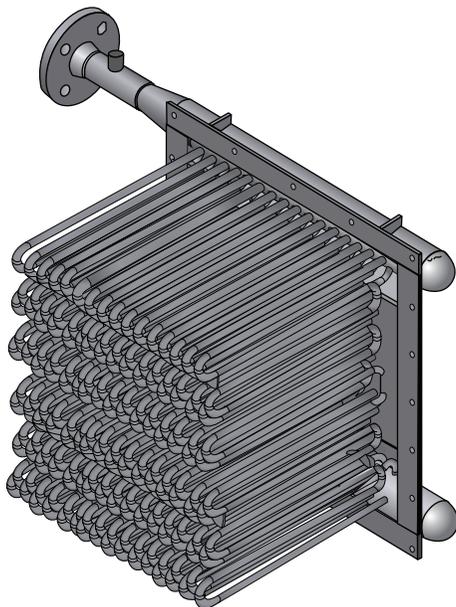


FIGURE 4: 3D model of proposed heat exchanger.

is also influenced by the nature of inlet stream and the form of pipe inlet. Tube bundles with in-line tubes behave more like parallel channels, which are formed by the tube rows. An expected increase in heat transfer coefficient due to the turbulence enhancement caused by the tube rows does not occur [7].

The application for hot-air Ericsson-Brayton engine will use as primary heat exchanger tube with staggered tubes.

The average Nusselt number for this type of heat transfer through tube bundle is defined according to the following equation [7]:

$$Nu_{0,bundle} = \frac{1 + (n - 1) \cdot fa}{n} \cdot Nu_{l,0}, \tag{9}$$

where

$$f_{a,stag} = 1 + \frac{2}{3b}, \tag{10}$$

$$Nu_{l,0} = 0.3 + \sqrt{Nu_{l,lam}^2 + Nu_{l,turb}^2}.$$

Then followed the estimation of overall coefficient of heat transfer, which is depending on the Nusselt number, is

$$\alpha = \frac{Nu_{bundle} \cdot \lambda_{TM}}{l}. \tag{11}$$

When both sides of equation are known, the overall heat transfer coefficient and the required heat transfer surface can be estimated. Subsequently, 3D model of heat exchanger was created. The model was in the first step created with wall thickness of tubes and inlet tube. But this solution sets major requirements for computing hardware, so a simplified model with tubes as full material was created, the proposal is in Figure 4.

### 3. Heat Exchanger Verification Using Ansys Fluent

The model for Ansys Fluent was created using 3D modeling software. By the creation of the model the substitution of all the construction elements by simple geometrical features was very important [8].

This means that the whole exchanger was created as one volume with tubes as full material. The tubes have multiple collectors at inlet and outflow. No construction tolerances are reflected.

The exterior of the heat exchanger was created by cutting out the material from volume. In the first step, the tetrahedron mesh was used to fill the whole volume. The generated mesh is displayed in Figure 5. Quality of the generated mesh is determined by skewness of elements and by minimal orthogonal quality. Skewness of the model was  $7.1833 \cdot 10^{-7}$ , where lower value is representing worse quality [8]. Based on this, the tetrahedron mesh was converted to polyhedral mesh.

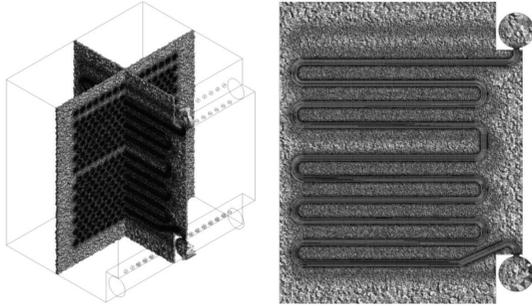


FIGURE 5: Generated tetrahedral mesh.

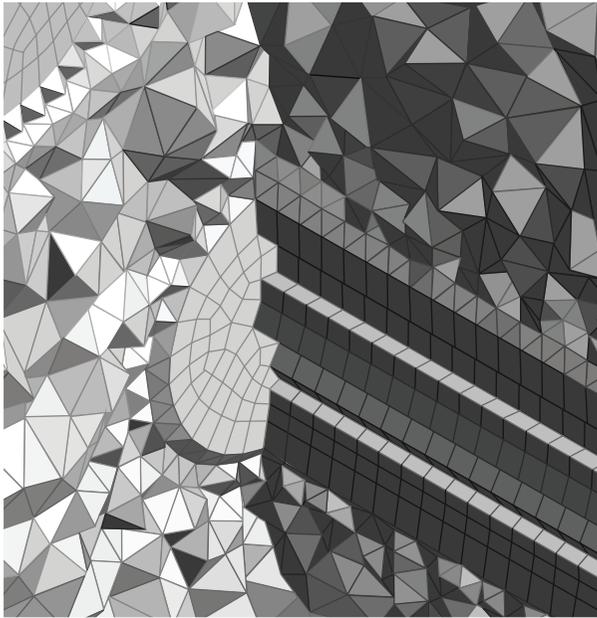


FIGURE 6: Details of polyhedral mesh.

The details of the generated mesh are in Figure 6. The model was solved with polyhedral mesh and  $K-\epsilon$  model. The flow was predicted as turbulent. In Figure 7 the velocity contour is shown and in Figure 8 the temperature field can be observed. The current model has confirmed the mathematical model and also accuracy of chosen geometry.

**4. Conclusion**

Hot-air Ericsson-Brayton engine used in cogeneration unit is a nonconventional engine, which produces electric energy by using different types of fuel, for example, biomass, wood pellets, and so forth. Heat exchanger design for hot-air Ericsson-Brayton engine sets a wide range of specifications. At the first step, the working conditions of whole unit and the required power and temperatures for each part of this device were defined. The basic dimensions of heat exchanger were set using criterion formula. With this calculation, the inlet and outlet temperatures of the heat exchanger were verified. Then, the calculation using Ansys Fluent has followed. As the

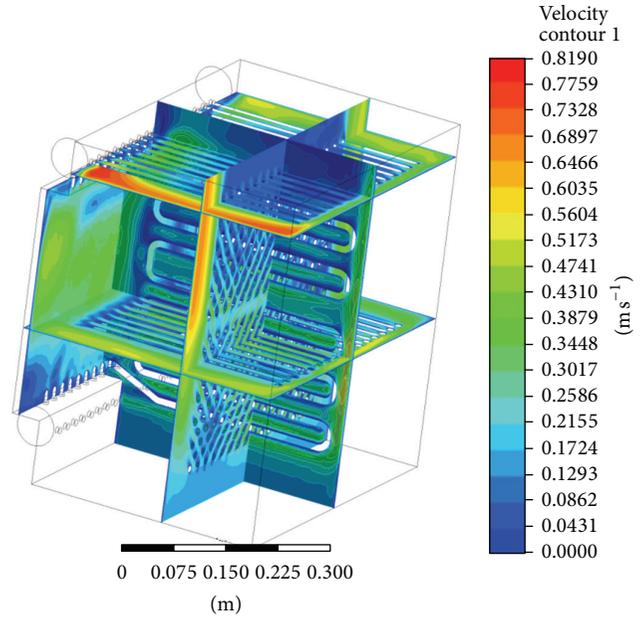


FIGURE 7: Velocity contour in the proposed heat exchanger.

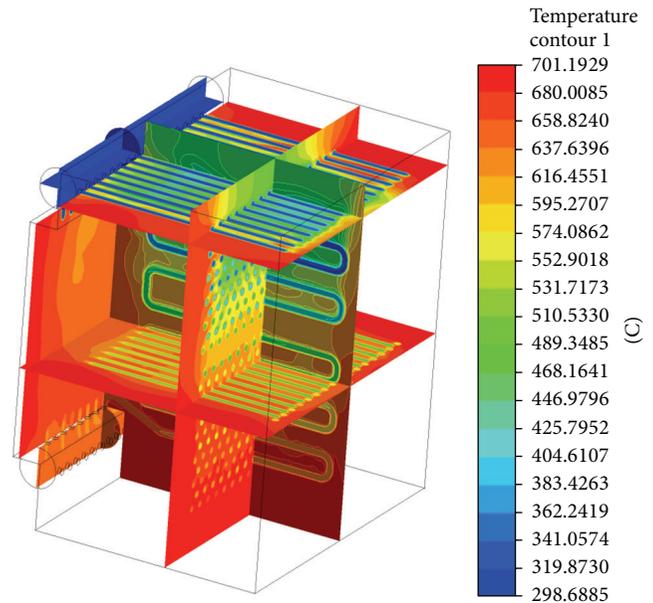


FIGURE 8: Temperature contour in the proposed heat exchanger.

next step was to complete the construction documentation and finish all design fundamentals, the construction and real measurements can follow.

**Conflict of Interests**

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

This work is supported by the financial assistance of the Project “Výskum Nových Spôsobov Premeny Tepla z OZE na Elektrickú Energiu Využitím Nových Progresívnych Cyklov” ITMS 26220220117.

## References

- [1] M. Creyx, *Energetic Optimization of the Performances of a Hot Air Engine for Micro-CHP Systems Working with a Joule or an Ericsson Cycle*, Elsevier, Paris, France, 2012.
- [2] J. Kalčík and K. Sýkora, *Technická Termodynamika*, Academia Praha, Prague, Czech Republic, 1973.
- [3] S. Bonnet, M. Alaphilippe, and P. Stouffs, *Energy, Exergy and Cost Analysis of a Micro-Cogeneration System Based on an Ericsson Engine*, Elsevier, Paris, France, 2011.
- [4] P. Ďurčanský, J. Jandačka, A. Kapjor, and Š. Papučík, “Návrh výmenníka tepla pre Ericsson-Braytonov motor,” in *SKMTaT*, K. Kaduchova, Ed., pp. 21–25, 2013.
- [5] P. Nemeč and J. Huzvar, *Proposal of Heat Exchanger in Micro Cogeneration Unit, Configuration with Biomass Combustion*, Materials Science and Technology, Žilina, Slovakia, 2011.
- [6] P. Stehlík, *Tepelné Pochody, Výpočet Výmenníku Tepla*, VUT Brno, Brno, Czech Republic, 1991.
- [7] Verein Deutscher Ingenieure, *VDI Heat Atlas*, Springer, Berlin, Germany, 2010.
- [8] R. Lenhard and M. Malcho, “Numerical simulation device for the transport of geothermal heat with forced circulation of media,” *Mathematical and Computer Modelling*, vol. 57, no. 1-2, pp. 111–125, 2013.

## Research Article

# Master-Slave Control Scheme in Electric Vehicle Smart Charging Infrastructure

**Ching-Yen Chung, Joshua Chynoweth, Chi-Cheng Chu, and Rajit Gadh**

*Department of Mechanical and Aerospace Engineering, University of California at Los Angeles, 420 Westwood Plaza, Los Angeles, CA 90095-1594, USA*

Correspondence should be addressed to Ching-Yen Chung; [chingyenchung@ucla.edu](mailto:chingyenchung@ucla.edu)

Received 28 February 2014; Accepted 18 March 2014; Published 26 May 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Ching-Yen Chung et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

WINSmartEV is a software based plug-in electric vehicle (PEV) monitoring, control, and management system. It not only incorporates intelligence at every level so that charge scheduling can avoid grid bottlenecks, but it also multiplies the number of PEVs that can be plugged into a single circuit. This paper proposes, designs, and executes many upgrades to WINSmartEV. These upgrades include new hardware that makes the level 1 and level 2 chargers faster, more robust, and more scalable. It includes algorithms that provide a more optimal charge scheduling for the level 2 (EVSE) and an enhanced vehicle monitoring/identification module (VMM) system that can automatically identify PEVs and authorize charging.

## 1. Introduction

Every plug-in electric vehicle (PEV) that is purchased instead of a dedicated fossil fuel burner is a good step in the direction of energy independence and lower greenhouse gas emissions. As ever more of these PEVs hit the road, sufficient charging infrastructure becomes ever more important in furthering the proliferation of PEVs in the car market. In order to maximize the charging infrastructure that can be installed on a given electrical grid, optimization needs to not only account for energy production, but also account for constraints that may appear in the system at any level. PEVs not only burden the energy production system, but also pockets of PEVs in certain areas may strain the local grid and transformers. Furthermore, each new current electric vehicle supply equipment (EVSE) requires a dedicated electrical circuit that incurs expenses that limit the number of EVSEs that will be installed. WINSmartEV [1–3] is a software based PEV monitoring, control and management system that not only incorporates intelligence at every level so that charge scheduling can avoid grid bottlenecks, but it also multiplies the number of PEVs that can be plugged into a single circuit. This combination of optimizing the use of the electrical grid while multiplying the number of PEVs per circuit is

a one-two punch against the limits of the electrical infrastructure in charging PEVs.

Simple commercial charging stations such as Leviton and ClipperCreek, which simply provide basic charging function without network features, do not provide network services for smart charging purposes. One exception is Coulomb Inc. Coulomb devised its own proprietary network-controlled charging system through a remote server [4, 5], but these stations are not suitable for current sharing purposes because they only have one or two outlets. A method of electrical circuit sharing for charging stations is proposed by Coulomb [6]; however, no details of charging algorithms are provided.

Since WINSmartEV is software based, intelligent charging algorithms can be implemented and updated when needed. The algorithms can be developed based on user's time, energy price, or energy amount. A charging algorithm that relies on a smart phone interface for entering PEV data, such as arrival and departure times and initial and final state of charge (SOC), is proposed in [7]. The scheduling algorithm proposed in [8] requires the initial energy states of a PEV as the input. These approaches are not valid unless the user provides the actual SOC data. To solve this problem, the authors of [9] propose a custom-built module, named vehicle monitoring/identification module (VMM),

which reads the in-vehicle controller area network (CAN) data bus and transmits SOC data via a ZigBee wireless link to a charging station and then onto the charging controller. However, without insider knowledge of the PEV manufacturers, identifying data location on the CAN bus could present a challenge for obtaining the SOC data. Several charging algorithms are presented in [10–12]; however, none of them mentioned a method to achieve variable current and multiplexing control, let alone the collaboration between the control center and the charging stations.

In order to fully utilize the power resource on the local grid, collaboration between the master controller (server) and the slave controllers (charging stations) in the PEV charging infrastructure is required to manage the charging sessions and/or control the current to the PEVs. This ability is not incorporated into the current WINSmartEV design. In this paper, a master-slave control scheme for the electric vehicle smart charging infrastructure is proposed to enhance the performance and features of this smart charging infrastructure. These improvements include hardware upgrades that will enable better collaboration between EVSE and server, enhanced smart charging algorithms, improved safety requirements, and incorporating RFID authentication and authorization into the VMM system. This paper is structured in the following way. First the current version of WINSmartEV is outlined in Section 2. Next, the proposed upgrades to the smart charging control scheme is discussed in 2 sections, Section 3 for level 1 EVSEs and Section 4 for level 2 EVSEs. Then an RFID authentication and authorization scheme is discussed in Section 5.

## 2. Existing WINSmartEV Infrastructure

There are three subsystems in UCLA WINSmartEV smart charging infrastructure including the control system, the communication system, and the metering system. Some special features such as smart charging algorithms [13], safety requirement integration [14], and RFID mesh network system for user authentication and authorization [15] are developed based on the existing hardware and software. Figure 1 shows the network architecture of WINSmartEV.

In order to implement the electrical power sharing concept, a four-outlet smart charging station connected to a single circuit is designed and implemented in [16]. The one-circuit-to-four-outlet design is based on the limitation of normal circuit installation (30 A continuous) and the minimum PEV charging current (6 A) defined in J1772 standard. Theoretically, the number of outlets could be 5 in order to fully utilize the maximum capacity of the circuit. However, in real practice, it will easily trip the circuit breaker if any one of five PEVs draws a bit more than specified current. Figure 2 shows the installation of a level 2 smart charging station and a level 1 smart charging station in a UCLA parking lot.

The details of the subsystems including control system, communication system, and metering system are described in the subsections.

**2.1. Control System.** In this section, the functionality of master controller (central server), the slave controller (charging stations) is described in the following. There are two types of charging stations, level 1 charging stations that connect to standard 120 V household circuits and level 2 charging stations that connect to 208 V or 240 V circuits for faster charging. The level 1 charger controls four EVSmartPlug outlets to provide power to the PEVs. Because the PEV user plugs the PEV's trickle charging cable into the outlet to charge PEV, the control system switches the outlet on and off in order to control the 120 V power to the trickle charging cable.

The level 1 charging stations are currently controlled by a server-based central controller equipped with smart scheduling algorithms. Different algorithms, including real time algorithms and scheduling algorithms, can be implemented to control charging. Round-robin and FCFS (first-come, first-served) are examples of real-time algorithms. Scheduling algorithms can be developed to include many factors such as time, energy price, energy amount, or SOC. For example, a round-robin algorithm, which only turns on 1 channel and charges 1 PEV at a time, is currently used to schedule charging in the level 1 EVSmartPlug station to share a single 120 V power source with four PEVs. This algorithm only takes into account how long the PEVs have been present. As more sophisticated algorithms are developed, the central controller has the flexibility and extensibility to be updated to include these new algorithms.

Because it incorporates J1772 standards with cables that can be plugged directly into the PEVs, the level 2 charger is required to turn on and off the power to the PEVs by controlling relays and rate the power that each PEV pulls by controlling the duty cycle of the pilot signal. The hardware and firmware of the level 2 smart charging station local controller is designed and implemented in [16].

In order to accelerate the response of the smart charging stations by reducing the traffic between the smart PEV charger and the control center, a power information collector (PIC) in [17] is designed to collect the power information locally and relay it to the control center periodically. The response time can be further reduced by pushing the information to the control center; thus, a fast response smart PEV charging infrastructure is achieved. However, because the control scheme is server-based, the server will need to wait for  $T_{\text{wait}}$  due to the communication delay and the response time of the charging station and the PEV. In order to accelerate the performance of the system, a master-slave control scheme is required as proposed in Section 3.

**2.2. Communication System.** In [16], there is a multiple protocol gateway inside the smart charging station to provide communication services for multiple functions. To connect to the internet, there are three types of methods including 3G, Ethernet, and WiFi. 3G communication is required due to its flexibility and accessibility to be everywhere as long as the cellular signal exists, especially where wired or WiFi communication is unavailable. When using Ethernet for communication, the gateway can directly connect to the internet with a static IP or a dynamic IP assigned from

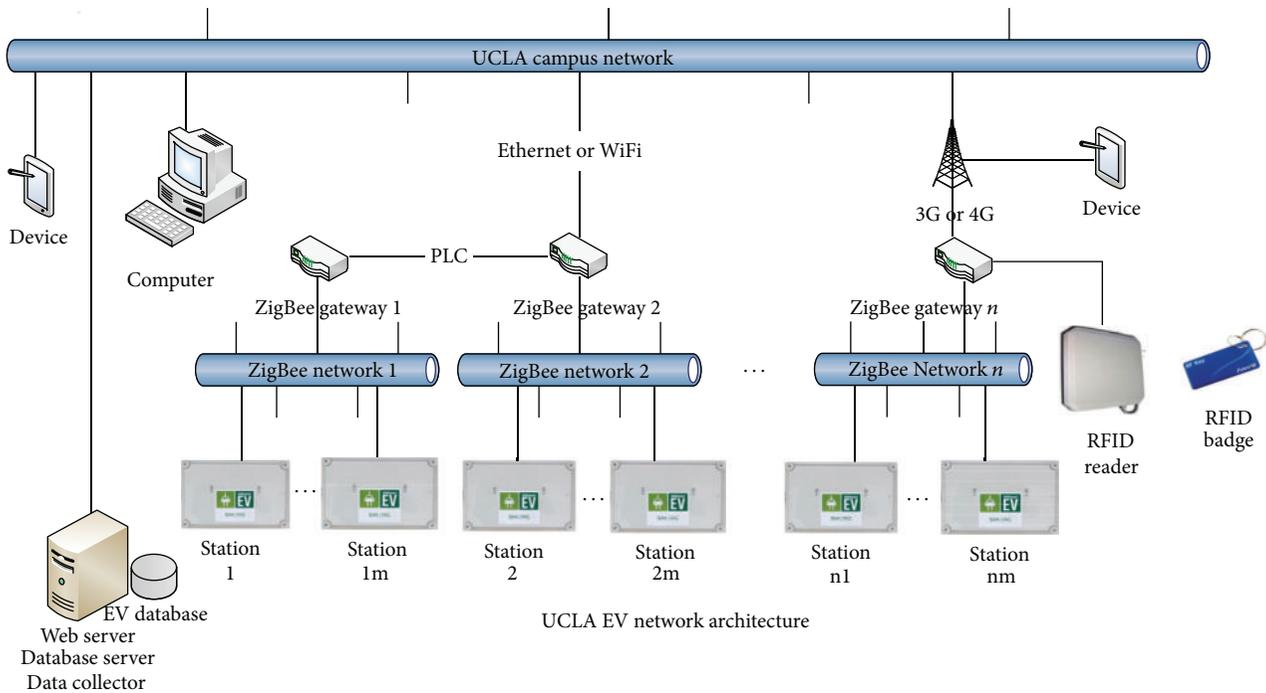


FIGURE 1: Network architecture of WINSmartEV.



FIGURE 2: Installation of smart charging stations.

a DHCP router. When Ethernet connection and 3G service are unavailable in a parking area, WIFI or PLC can be used to connect to another gateway or router that does have an internet connection. The EVSE's gateway can use a PLC module on its Ethernet port to connect to other gateways or routers connected to any electrical circuit on the same transformer. When using WiFi for local communication, the gateway needs to be setup as a client to connect to other gateways or routers. In this case, a port forwarding method is used on the other gateway or router so that the server can access the client gateway.

The information interchanged between the gateway, meters, and the control unit are through ZigBee communication. The function of the ZigBee coordinator on the gateway is to handle the messages between the gateway and the end devices or routers including the meters, control units, and vehicle monitoring/identification modules (VMMs) [9] on the PEVs. In order to dispatch the commands and parameters to the desired devices, the ZigBee coordinator needs to recognize and register the unique MAC addresses of the end devices or routers. Since a number of devices communicate

using ZigBee mesh network capabilities, only one gateway is required in a geographic locale.

The current system has two types of controllers, one with and the other without ZigBee communication. The controller without a ZigBee module talks to the gateway directly through USB port with RS232 communication. On the other hand, the controller set with ZigBee communication consists of ZigBee coordinator and ZigBee end-device. The gateway talks to the ZigBee coordinator to dispatch or receive response from the ZigBee end-device. Both types of controllers require a RS232-USB adapter cable in between the gateway and controllers. In order to ensure proper functionality, the RS232-USB adapter cable must be compatible with the gateway. When the 3G dongle is used for communication, both the 3G dongle and RS232-USB adapter cable must be assigned a USB port and only the assigned USB ports should be used.

**2.3. Metering System.** The metering system in both level 1 and level 2 charging stations consists of a gateway and four meters. The meter inside the charging station returns its power information, including voltage, current, frequency, power factor, and energy consumption, to the gateway upon receiving the command of power information retrieval through ZigBee communication. The meters need to join the ZigBee mesh network created by the ZigBee coordinator embedded in the gateway. The function of the metering system also requires the association of the meters' ID and the physical outlet numbers. The detailed schematics of a four-outlet metering system are shown in Figure 3.

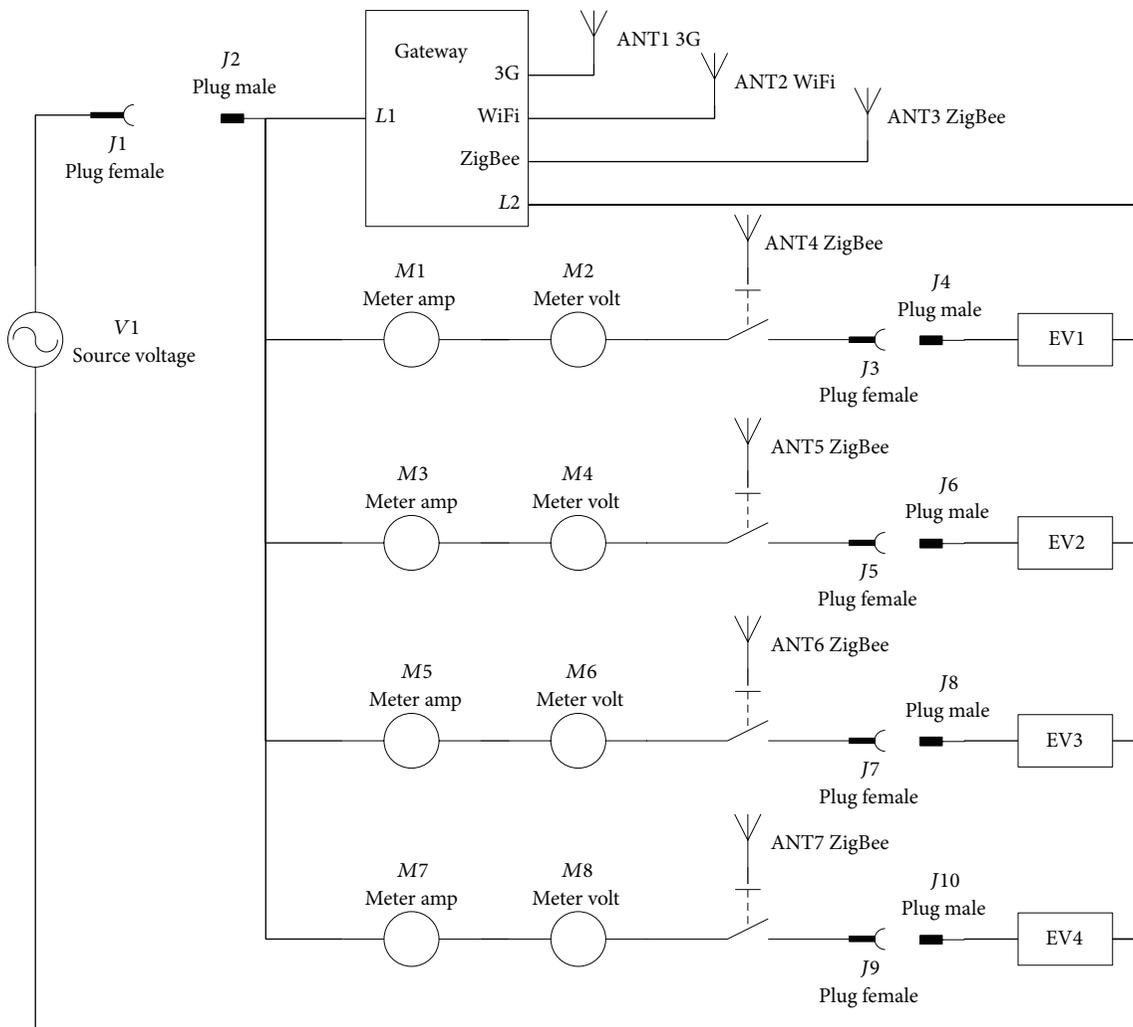


FIGURE 3: Schematics of metering system.

### 3. Proposed Control Scheme and Results for Level 1 EVSE

The processes involving the collaboration between the server and the charging stations such as the smart charging algorithm [13], safety requirement [14], and the RFID authentication and authorization [15] are the prior arts in publications. In order to fully utilize the power resource on local grid and improve the performance of the PEV charging infrastructure in the management of charging sessions or current control, the collaboration between the master (server) and the slave (local controller) is required. Therefore, a master-slave control scheme for the electric vehicle smart charging infrastructure is proposed to enhance the performance of the features including smart charging algorithms, safety integration, and RFID authentication and authorization. The proposed control scheme involves a server-based central controller and local controllers inside the charging stations. The details of the collaboration scheme for the level 1 EVSE is presented in this section and the details for the level 2 EVSE will be presented in the next section. RFID authentication and authorization will be discussed in Section 5.

**3.1. Local Controller Design.** In current level 1 charging station design [1–3], there is no local controller inside the charging station. When the data pull method sends a power information request command from the server to a charging station, the signal must pass through the internet and through a 3G network before it reaches the gateway of the charging station. Then the gateway relays the power information request command to the specific meter it is meant for. When the gateway receives a reply from the meter, it relays the response back to the server where the information travels back in reverse order. With multiple meters requiring multiple requests each, the aggregated round trip times cause slow performance. In order to enhance the system's performance and shorten the response time of the system, a device named the power information collector (PIC) [17] collects the power information locally in order to send it in to the server together as one packet. By decreasing the number of communications required for status reports and control operations, the PIC significantly decreases the delay time for switching PEV charging sessions or changing current to the PEVs. In order to accelerate the response time of the charging station, based on the design of the PIC, a local controller



FIGURE 4: Local controller with controllability over power to outlets.

with the controllability over the power on the outlets is implemented as shown in Figure 4.

Because the local controller is responsible for turning on/off the outlet, it reduces the round trips of command and response between the server and the charging stations. Thus, it accelerates the response time of the charging station. Since the power information is available on the local charging station, local charging algorithms can be realized on this design. Certain charging algorithms can be implemented at the local level. With less communication traffic between the server and the charging stations, the charging algorithms at the local level could be more efficient than that at the server level. By reducing the traffic between the server and the charging stations, the improvements allow the control center to serve a larger system, which enhances the capability of the existing WINSmartEV framework.

Nevertheless, considering the calculation power of the microprocessor, only certain simple charging algorithms, such as round-robin and Schedule Time, can be implemented in the microprocessor. Thus, with local charging algorithms implemented on the charging station level, the server will only need to select the mode of the charging algorithms of each charging station. This will save significant server computing resources. With local charging algorithms implemented, the control center could handle a larger smart charging system due to the reduction of traffic between the control center and the smart charging stations.

However, because of the lack of computing power at the local level, more complex charging and scheduling algorithms still need to be implemented on the server. No matter where the charging algorithm sits, two major operation flows including ENABLE CHARGING and DISABLE CHARGING involve in the smart charging algorithms at server level. In each operation flow, two subprocesses including READ OUTLET ON/OFF STATUS and READ POWER INFORMATION need to be done.

This smart charging infrastructure can use its metering system to monitor the simple commercial charging stations such as Leviton, ClipperCreek, and Schneider so that the server can have control over these charging stations. These charging stations can run switching type charging algorithms at the server level including a round-robin algorithm or a fair charging algorithm [13]. In order to have fair usage of power resource for every PEV user, currently, a round-robin algorithm is used to schedule charging in the multiplex charging system WINSmartEV. In current practice, the round-robin

algorithm is developed at the server level for the lack of local controller.

**3.2. Fair Charging Algorithm.** In order to appeal to more users, a fair charging algorithm [12] is proposed to maximize fairness in the allocation of charge time for the smart plug charger. The fair charging algorithm is designed for a switching type of charging stations such as the EVSmartPlug, where only one PEV can charge at a time. In this case, each user's charge ratio  $\tau$  is defined as the ratio of the charging time  $T_{\text{Charge}}$  and the stay time  $T_{\text{Stay}}$  in

$$\tau \equiv \frac{T_{\text{Charge}}}{T_{\text{Stay}}}. \quad (1)$$

The fairness of the charging system depends on how close all the users' charge ratios  $\tau$  are to each other. The fairness system can also be stated as follows: for a given charging event, every user's mean charge ratio  $\mu(\tau)$  should be close to the mean charge ratio of every other user charging at the time  $\mu[\mu(\tau)]$ . Therefore, both  $\sigma[\mu(\tau)]$  and  $\mu[\sigma(\tau)]$  must approach 0 as the system approaches complete fairness. The fairness index  $\alpha$  is defined in (2) to indicate the fairness of the system:

$$\alpha \equiv 1 - \frac{\{\sigma[\mu(\tau)] + \mu[\sigma(\tau)]\}}{2}. \quad (2)$$

Even though the round-robin algorithm seems fair, the fairness index [13] shows that it favors the first user that starts charging over the users that arrive later. In the fair charging algorithm, when the second user's charging session overlaps that of the first user's charging session, the server predicts the second user's charge time and the first user's stay time in order to create a fair charging schedule. The schedule will allot each of the two PEVs the charge time required so that they leave with same charge time ratio. The schedule will switch charging enough to avoid the risk of a large charge imbalance, but not so much as to take up too much time switching. When a third or a fourth PEV arrives, then new fair charge schedules are created that takes into account the new user. If the time to switch charging between PEVs is close to zero, then the optimization algorithm can be executed. The fairness could be maximized by continuously switching charging power between PEVs. However, there is a noticeable time delay in switching charging sessions between PEVs due to data pull method. Thus, the period of time to switch charging from one PEV to the next can be as high as minutes, given hardware and network constraints. If the system switched continuously between users, much charge time would be wasted in the switching process, causing all users to be worse off. Fairness maximization can be obtained while only switching charging once if exact stay time of the PEV is known. If PEV's stay time is unknown, fewer switching may often leave the charge time for each PEV lopsided and unfair. Therefore, the optimization of the fairness algorithm needs to take into account both the confidence of stay time and the time wasted in switching. Therefore, the Fair Charging Algorithm counts heavily on the accuracy of the prediction of user's stay time. A forecast of users to the PEV charging station in [18] can be used

for more accuracy on user's stay time. If the prediction of the user's stay time is accurate, fairness maximization can be obtained while only switching charging once. There is no way for the charging infrastructure to retrieve a PEV's SOC status for the purpose of user's stay hour prediction unless extra devices, VMM [9], for SOC data retrieval are equipped on the PEV. The PEV's SOC status should not be considered available by using the current PEV charging station standard J1772. Therefore, based on a user's historical charging records, for predictable people, either  $u(T_{\text{Stay}})$  or a linear regression function is used to predict  $T_{\text{Stay}}$  of the user. For unpredictable people, the average stay time of all users  $u[u(T_{\text{Stay}})]$  is used for prediction. A number of switches may be required if the prediction of the stay time is not accurate enough.

Currently, the fair charging algorithm is implemented at server level. In order to have better performance, the calculations which rely on the historical data in the data base should still be finished at the server level. The local controller inside the charging station is responsible to execute the charging schedule calculated by the server. After the server calculates the charging schedule according to the selected charging algorithm, it sends the charging schedule to the stations. The charging stations control the charging sessions based on the schedule. When some charging events happen during the charging sessions, the local controller requests the server to update the charging schedule. When the server receives the request of charging schedule change, it calculates a new charging schedule for the charging stations.

**3.3. Safety Requirement.** Because the control of pilot signal for the level 1 charging station, EVSmartPlug, takes place within the PEV's trickle charge cable, the automatic reset of GFCI is not required by UL certification. Therefore, a single commercial breaker with GFCI on the power source fulfills all the safety requirements.

#### 4. Proposed Control Scheme and Results for Level 2 EVSE

The possible efficient control scheme of smart charging algorithms can be developed based on user preference and the local power capacity. The central controller, or server, equipped with smart charging algorithms sends commands or schedule to the charging station through a multiple protocol gateway. The central server functions as the master controller while the local controllers, embedded in the charging station, serve as the slave controller in the infrastructure.

**4.1. Local Controller Design.** The current level 2 chargers have a ZigBee-based local (slave) controller with multiple functions including the pilot signal generator, pilot signal monitor, safety relay controller, and autoreset function as implemented in [16]. In this design, three microprocessors are utilized to fulfill the functionality. In the current practice, because each charging station is equipped with a multiple protocol gateway, the information exchange between the charging stations can be fulfilled by WiFi or Ethernet. Therefore, the ZigBee function on the local controller is indeed a redundant

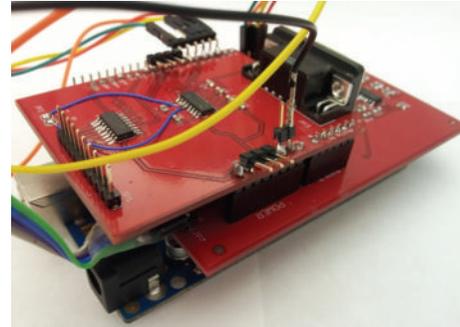


FIGURE 5: Simplified J1772 EVSE controller without ZigBee function. (1) RS232 module. (2) Pilot signal generator/monitor. (3) Aduino Mega2560.

communication channel. In order to simplify the design and enhance the features and functionalities, the ZigBee module has been removed and a more powerful microprocessor, ATmega2560, with more input and output pins (I/Os) has been added as shown in Figure 5. The pilot signals are created by ATmega2560's internal timer and monitored by its analog input pins. In this design, only one microprocessor is used to fulfill the aforementioned functionality.

With enhanced processing power, simple charge scheduling algorithms can be implemented on local device. These simple current sharing algorithms can be designed and implemented by revising the firmware-based state machine as shown in Figure 6. In the simple Current Sharing Algorithm, the local controller assigns the available power to the designated outlet by setting up the duty cycle of the pilot signal before the Monitor EV stage.

The process of setting up the duty cycle is inserted in between the processes of Run Pilot Flow and Monitor EV. The current sharing algorithm is based on the configuration of the box. If there is no PEV charging in an adjacent channel, the firmware will set the maximum available current to the given outlet. Otherwise, the firmware will divide the current for the PEVs to share. Note that experimental results have shown a five-second delay in the PEV response time [16]. Once a PEV is unplugged, the local controller restores the power to other PEVs. In [16], the Monitor EV stage is handled periodically based on the timer interrupt flag. In order to fulfill the J1772 standard to handle a faster PEV unplug detection, the Monitor EV stage is moved to the main loop for continuously checking unplug status. Figure 7 shows an example configuration of a simple Current Sharing Algorithm. In this configuration, Ch1 and Ch2 share one power circuit while Ch3 and Ch4 share another power circuit.

The firmware sets up the maximum duty cycle 30% (18 A) for Ch1 initially. When the adjacent channel Ch2 is plugged in, the maximum duty cycle is divided by 2 and becomes 15% (9 A). When the PEV at Ch1 is fully charged or unplugged, the duty cycle of Ch2 is set back to the maximum duty cycle. Ch3 and Ch4 present the same result.

Simple commercial charging stations can also be modified as smart charging stations by connecting a metering system and local controllers. In the ClipperCreek case,

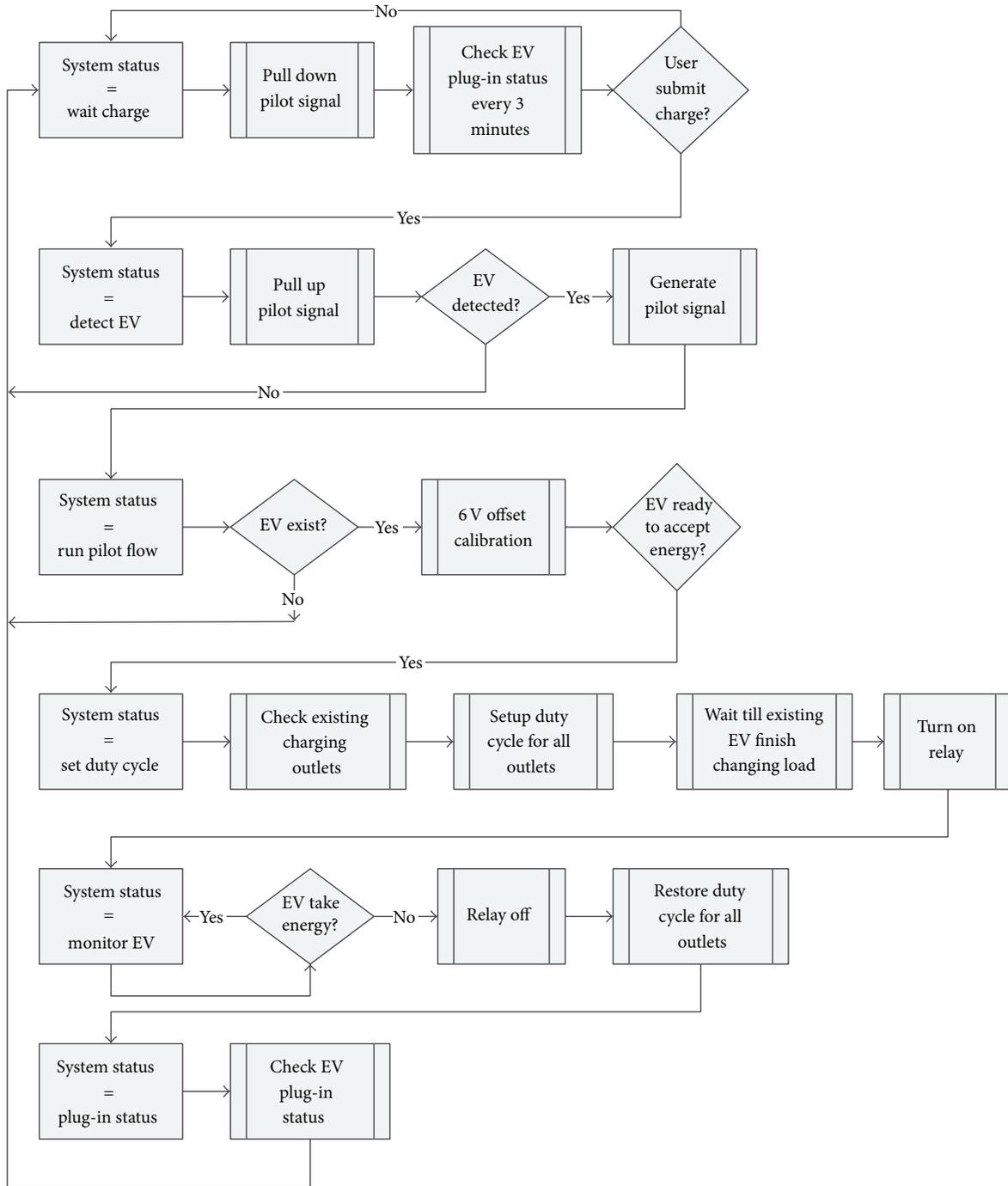


FIGURE 6: State machine of simple Current Sharing Algorithm.

the charging station model CS-40 provides the terminals for three stage-current control (30 A, 6 A, and 0 A) [19]. To control the connectivity of the terminals, a ZigBee-based local controller with relay module is designed and implemented as shown in Figure 8.

With ZigBee communication, local charging stations can exchange information with each other in a local area. With this design, only one gateway is required per locale; thus, it saves in communication costs.

In order to implement a variable continuous current control on the simple commercial charging stations, an extra circuit would need to be installed in between the charging station and the PEV to emulate the behavior of the PEV. The design requirements for this circuit would include the ability to generate a pilot signal with variable duty cycle in response to the command from the server. The design of the extra circuit, which is beyond the scope of this paper, is not addressed here.

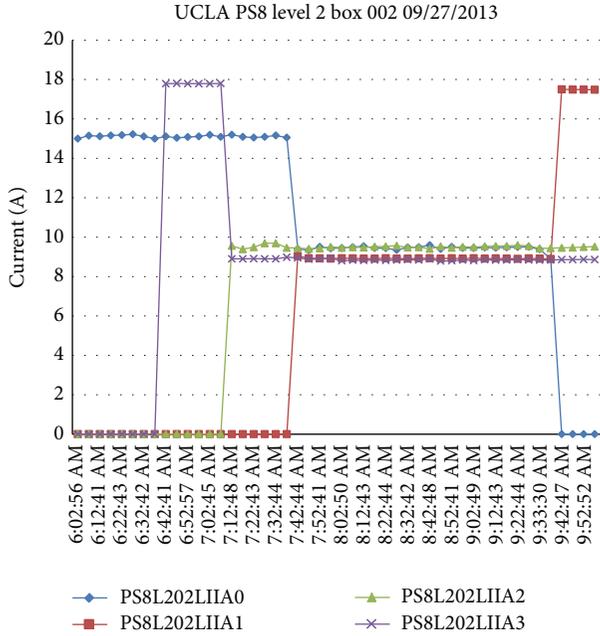


FIGURE 7: Example of simple current sharing algorithm.

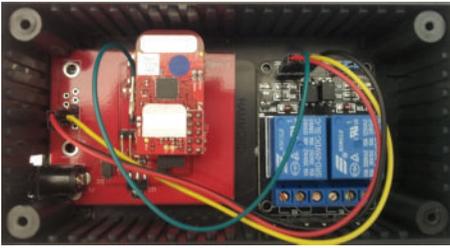


FIGURE 8: ZigBee-based local controller for ClipperCreek charging station.

**4.2. Fair Current Sharing Algorithms.** To involve the smart charging algorithms at server level, the server can select or disable local charging algorithms embedded in the firmware of the local controller. No matter where the charging algorithm sits, three major operation flows at server level need to be engaged including ENABLE CHARGING, DISABLE CHARGING, and PILOT SIGNAL DUTY CYCLE CHANGE. In each operational flow, there are three subprocesses including READ METER ON/OFF STATUS, READ METER'S POWER INFORMATION, and READ OUTLET'S STATUS. The power information includes the voltage, current, and active power. The outlet's status includes pilot signal's duty cycle, safety relay on/off status, PEV plug-in status, and firmware-based state machine's stage.

To deal with a variable current charging station, a fair current sharing algorithm is proposed. Considering a variable current control charging station, the fairness index is now related to the energy consumption  $E$  during the stay time  $T_{\text{Stay}}$ , which equals the average power  $P$  for the user in

$$P \equiv \frac{E}{T_{\text{Stay}}} = V(t) I(t) \times \frac{t}{T_{\text{Stay}}}. \quad (3)$$

For a completely fair system, the average power  $P$  of each user should be the same, which means each user's charge rate is the same during the stay time. For a fair enough system, each user's charge rate should be close enough. Assuming  $V(t)$  is a constant, a current share ratio  $\rho$  can be defined in

$$\rho \equiv \frac{I(t) \times t}{I_{\text{MAX}} \times T_{\text{Stay}}}, \quad (4)$$

where  $I_{\text{MAX}}$  is the current capacity.

The time sharing type of fair charging algorithm required for switching chargers can now be viewed as a special case of a current sharing algorithm with a discrete current instead of a variable current. If the system is fair, every user's  $\mu(\rho)$  should be close to  $\mu[\mu(\rho)]$ . Therefore, both  $\sigma[\mu(\rho)]$  and  $\mu[\mu(\rho)]$  approach 0 if and only if the system approaches complete fairness for each user. Here, a new fairness index  $\beta$  is defined in

$$\beta \equiv 1 - \frac{\{\sigma[\mu(\rho)] + \mu[\sigma(\rho)]\}}{2}. \quad (5)$$

The fairness index  $\beta$  approaches 1 if and only if both  $\sigma[\mu(\rho)]$  and  $\mu[\mu(\rho)]$  approach 0, which is used to indicate the fairness of the system. The parameter  $\sigma[\sigma(\rho)]$  is viewed as the convergence of the system;  $\sigma[\sigma(\rho)]$  converges to 0 when the system is fair.

Figure 9 shows the flow of Fair Current Sharing Algorithm.

When the second user's charging session overlaps that of the first user, the server still predicts the second user's charging time and the first user's stay time. However, instead of calculating charge time allocation for each PEV, the server calculates the maximum current each PEV is allowed to draw current  $I$  based on the remaining energy consumption and the current share ratio  $\rho$  in (6). Instead of switching charging between the users' charging sessions, the server communicates the current  $I$  that each PEV is allowed by changing the duty cycles  $D$  of the pilot signal:

$$I_i = D_i \times 0.6$$

$$I_{1,i} = \frac{(\rho_{\text{Target}} \times I_{\text{MAX}} \times T_{1,\text{Stay}} - I_{1,i-1} \times t)}{(T_{1,\text{Stay}} - t)} \quad (6)$$

$$I_{2,i} = \rho_{\text{Target}} \times I_{\text{MAX}}.$$

$I_{1,i}$  and  $I_{2,i}$  represent the maximum allowed current for the first and second user, respectively.  $\rho_{\text{Target}}$  is the target value of the current share ratio of the system. The third user's session is treated as an overlap of the second user's session, and the fourth user's turn is treated as an overlap of the third user's session. Note that, in practical implementation, the maximum allowed current drawn is a discontinuous function of  $D_i$  based on J1772 standards:

$$I_i = 0, \quad 0 < D_i < 10$$

$$I_i = 0.6 \times D_i, \quad 10 \leq D_i < 85$$

$$I_i = 2.5 \times (D_i - 64), \quad 85 \leq D_i < 96$$

$$I_i = 0, \quad D_i \geq 96. \quad (7)$$

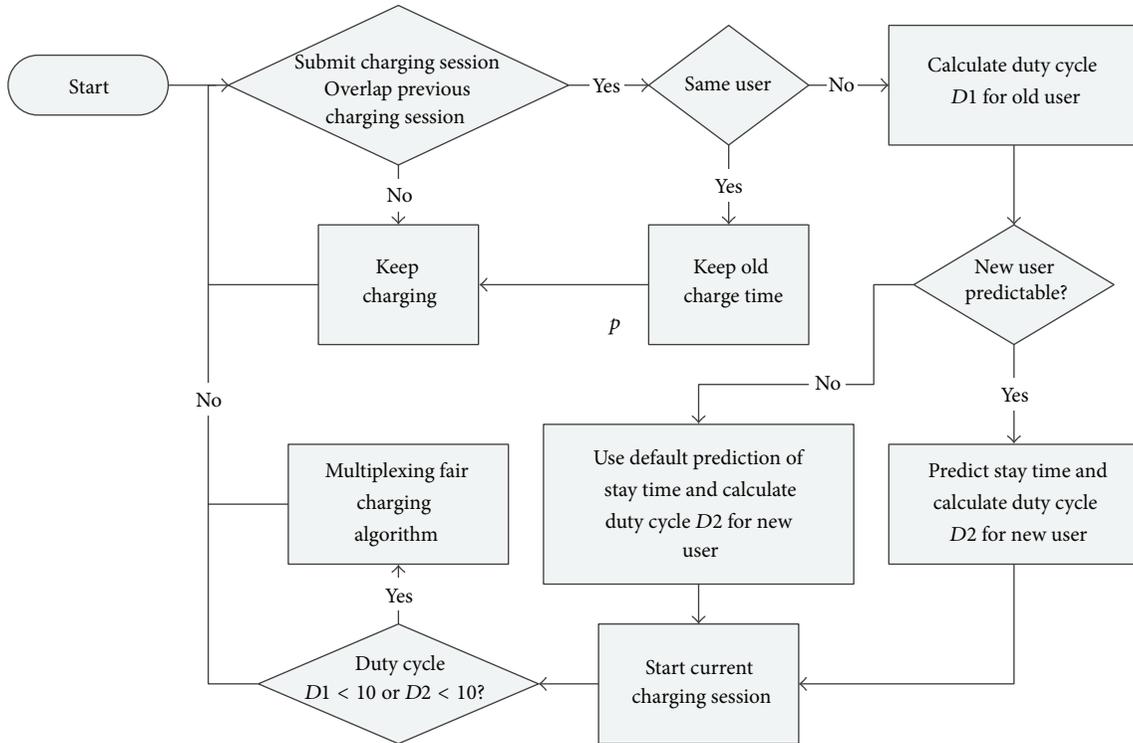


FIGURE 9: Fair Current Sharing Algorithm.

Because 240 V with 30 A is the most common installation, only the conditions in (8) are taken into account:

$$\begin{aligned}
 I_i &= 0, & 0 < D_i < 10 \\
 I_i &= 0.6 \times D_i, & 10 \leq D_i < 30.
 \end{aligned}
 \tag{8}$$

If the result of maximum current drawn  $I_i$  is less than 6 A, the Fair Charging Algorithm with a 6 A maximum will be used instead of the Current Sharing Algorithm.

Similar to the Fair Charging Algorithm proposed in [13], the Fair Current Sharing Algorithm also counts on the accuracy of the prediction of user’s stay time because the duty cycle calculation in (6) is based on the predicted stay time.

The command sets for the server and return values from the charging station are summarized in [16]. Based on the experiments in [9, 16], the server waiting time for the command sets in [15, 16] is formulized and summarized in Table 1.

In order to accelerate the server’s performance, the server’s waiting time should be set to different values according to the command set sent from Table 1. In the case of duty cycle change, to shorten the server’s waiting time  $T_{\text{waiting}}$ , it can be set to be variable values based on  $I_{\text{init}}$  and  $I_{\text{final}}$  rather than a fixed value to satisfy all conditions. Moreover, the simulation of pilot signal duty cycle change in [16] shows that it takes 30 ms to reach steady state. This value needs to be compensated in the firmware of the local controller.

The authors in [18] concluded that charging algorithms with power information retrieval can be implemented locally in the charging station. The system with embedded charging

algorithms, in which the traffic between the charging station and control center is reduced, is faster than that with remote charging algorithms which are implemented on the server. However, from the experimental result, the Fair Current Sharing Algorithm performs better than the Simple Current Sharing Algorithm with more users’ information. Because the Fair Current Sharing Algorithm is based on the user’s historical charging records, it is more proper to implement it on the server. With the charging stations equipped with PICs set on data pushing mode, the Fair Current Sharing Algorithm becomes more practical due to the improved response time of the system caused by the PIC.

Nevertheless, it is also possible to obtain the user’s stay time, if the user’s PEV is equipped with the VMM proposed in [9], which facilitates using the battery’s state of charging (SOC) to predict the user’s stay time. In this way, it is possible to implement the Fair Current Sharing Algorithm in the charging station.

**4.3. Safety Requirement.** As for the safety requirement for the level 2 charging station, since the charging station controls PEV charging by the pilot signal, the charging station is required to handle the GFCI function in both J1772 and UL standard. Although the authors in [20] claim the GFCI of a networked charging station can be reset remotely, no details of control methods or schematics are presented. To provide a safe smart technology for charging PEVs, a design for the safety system is presented in [14], which is implemented on all levels of control. The administrator can turn the relays of the charging stations on or off and check their status

TABLE I: Waiting time of command sets.

Comd.	Description and waiting time
atrs	Auto-reset the charging station $T_{\text{Waiting}} > 0.5T_{3G} + T_{\text{GatewayStartUp}}$
duty	Change the duty cycle of pilot signal $T_{\text{Waiting}} > \begin{cases} 1000 + 400 \times I_{\text{final}} - 0.5 \times T_{3G}, & I_{\text{init}} = 0 \\ 110 - 0.5 \times T_{3G}, & I_{\text{final}} = 0 \\ 310 - 0.5 \times T_{3G}, & I_{\text{init}} < I_{\text{final}} \\ 5060 - 0.5 \times T_{3G}, & I_{\text{init}} > I_{\text{final}} \end{cases}$
enab	Enable EV charging $T_{\text{Waiting}} > T_{3G} + T_{\text{EnableChargingProcess}} + T_{\text{EVResp}}$
rely	Turn on/off relay manually $T_{\text{Waiting}} > T_{3G}$
rest	Disable EV charging $T_{\text{Waiting}} > T_{3G}$
resp	ZigBee handshake response $T_{\text{Waiting}} = 0$
rgst	Return all registered ZigBee MAC address $T_{\text{Waiting}} > T_{3G} + T_{\text{GatewayUSBTimeout}}$
stat	Charging station status request $T_{\text{Waiting}} > T_{3G} + T_{\text{GatewayUSBTimeout}}$
test	ZigBee handshake request $T_{\text{Waiting}} = 0$

by sending out commands. The charging stations can be reset manually or automatically on schedule as long as the connection between the server and the charging stations exists. The pilot signal monitor can reset the whole system by turning off the switch on the power source of the charging station upon receiving the system reset command from the server. After the charging station loses power, the switch on the power source of the charging station is back to its normal position such that the charging station turns on again. Any emergency action taken at the top level will have a delay time that depends on the condition of the wireless communication including 3G, WiFi, ZigBee, and Cloud. Thus, a fast acting local unit is implemented to stop charging in case of an emergency.

In order to prevent electrical hazards, there should be no voltage on the handle of the charging cable until it is plugged into a PEV. The detection of the PEV plug-in status is implemented in the state machine of the firmware of the control unit based on the J1772 standard. The voltage of the pilot signal pin on the handle should be +12 V when there is no PEV connected to the charging station. After plugging in the PEV, the voltage will be +9 V or +6 V depending on whether or not the PEV is ready to accept energy. The PEV plug-in status detection is implemented in the state machine in the firmware of the control unit. Furthermore, the charging station is required to shut off the power immediately to prevent the hazard of electric shock when there is an abnormal diversion of current from one of the hot wires.

The ground fault circuit interrupter (GFCI) detects the difference of current between two hot wires and shuts off the safety relay when the difference has crossed the threshold amperage. Unlike a traditional GFCI which requires manually pressing the reset button, a pure hardware GFCI with a remote reset function is used to increase the reliability. The power to the PEV can be controlled by the server, the control unit of the charging station, and the GFCI circuit.

In [14], one leg of the contactor is controlled by the pilot signal monitor, while the other leg is controlled by the GFCI. The pilot signal monitor can reset the GFCI by toggling the switch on the power source of the GFCI. However, in real world applications, in order to have independent control over each GFCI channel, instead of controlling the power source of the GFCI board, the microprocessor generates the reset signal for the SR latch as shown in Figure 10.

Every time when one outlet is tripped, the microprocessor is able to reset the GFCI independently after the user unplugs the PEV without affecting other PEV's charging session.

Because the GFCI board is sensitive, false alarms are easily triggered due to the glitch at the rising edge in the output signal of GFCI board. In order to avoid triggering false alarms, the local controller deglitches the output from GFCI board and shuts off the contactor. Figure 11 shows the new relay control method to avoid GFCI false alarms.

In the new design, the GFCI board feeds its output to the local controller. Instead of controlling one leg of the contactor directly by GFCI board, the local controller controls both legs of the contactor for different conditions. The implementation of Non-ZigBee Level 2 J1772 local controller with GFCI function is shown in Figure 12.

In order to have the fastest response time, the interaction between the GFCI board and the microprocessor is handled by the interrupt routine in the firmware as shown in Figure 13.

In the implementation of the GFCI function, four interrupt pins are used to monitor the outputs from the GFCI board with a rising edge trigger. In order to avoid the false alarms caused by the glitch at the rising edge, after one interrupt trigger, the outputs of the GFCI board are monitored by the digital input pins with 500 us delay in the interrupt loop. If the controller detects that the output pin is HIGH, it terminates the power to the PEV by turning off the specific relays, thereby shutting off the contactor.

The GFCI of the outlet will be reset after the user unplugs the PEV. When the GFCI of the outlet is triggered, the system status of the state machine for the outlet will go to UNPLUG CHECK, in which the local controller keeps monitoring the unplug status until the user unplugs the PEV.

The design of GFCI in [14] is triggered by 14 mA difference between two hot wires on the positive cycle. Since the design of GFCI only works for positive cycle of the AC, if an abnormal diversion of current from a hot wire happens on the negative cycle, the GFCI trigger will be delayed by 8.3 ms, which is half cycle of 60 Hz. In addition, the GFCI circuit itself has approximately 1 ms delay. In our proposed GFCI design, compared to the maximum delay time of 8.3ms after the GFCI triggers, which is half cycle of 60 Hz, a 500 us delay time is acceptable. To satisfy the safety requirement by UL (Underwriters Laboratories) certification, the total

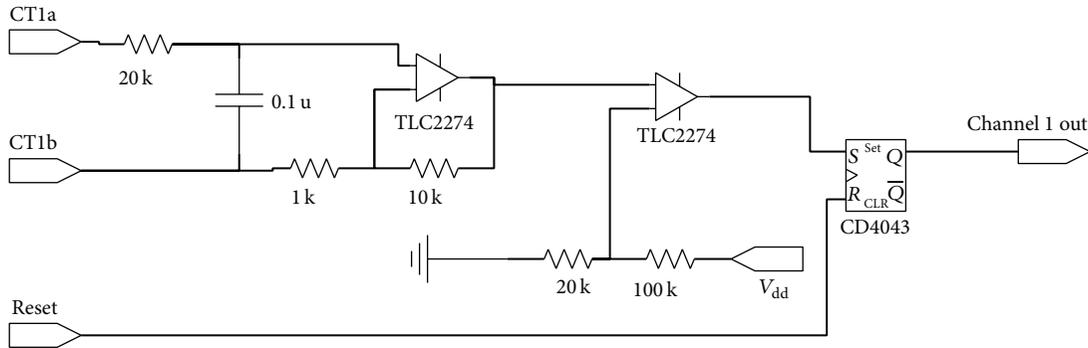


FIGURE 10: Schematic of GFCI.

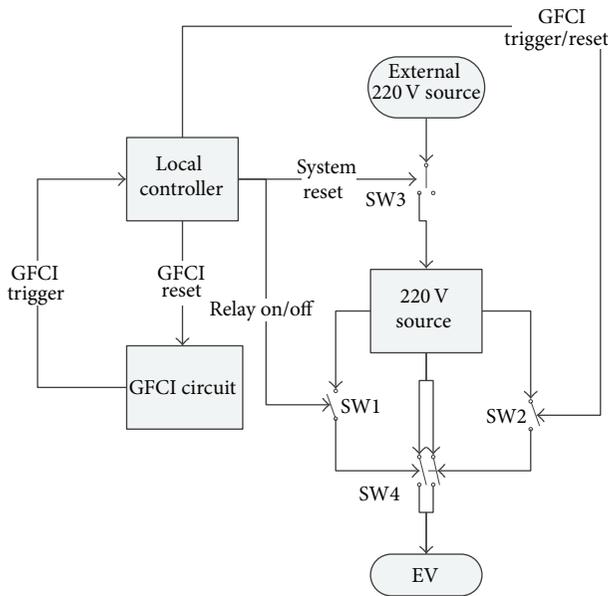


FIGURE 11: Schematic of safety control for the relay.

delay of GFCI function should meet the requirement of the maximum value, which is 24.9 ms as per UL standard. Overall, the maximum delay in our proposed GFCI is 9.8 ms, which satisfies the safety requirement by UL standard.

Besides the time delay requirement for GFCI, the UL requires extra circuits such as GFCI tester and voltage monitor on the contactors to fulfill the safety requirement. The GFCI tester is the circuit to test the GFCI function before energizing the contactor. A solution for the GFCI tester is to add an extra wire in the current sensor from 12 V dc source. Then, the local controller turns on a specified small current on this wire by using a digital output with a FET transistor and a power resistor. The voltage detector will check the circuit to see if the contactor is welded before enabling the contactor. If the contactor is welded, which means the charging station cannot stop power to the outlet, the system should stop providing service. A possible solution is to use a voltage divider with power resistors to obtain small AC voltage. Then, a transformer is inserted in between to isolate the AC and DC voltage. Later, the local controller detects this



FIGURE 12: Local controller for level 2 charging station with GFCI.

DC voltage with a Schmitt trigger through its digital input pin to see if there is voltage on the outlet of the charging station.

As mentioned before, because the GFCI board is sensitive, in order to reduce the magnetic field disturbance from the electromagnetic relay, the GFCI board needs to be enclosed in a grounded metal surface box.

Once the safety feature is certified, the UL does not allow firmware change by checking the CRC code of the firmware. Since the safety function is not charging all the time, in order to have the flexibility to add new features in the future, the separation of safety feature from the other functions is needed. Therefore, one possible solution is to use extra microprocessor to handle the safety features while the original one deals with other features. Thus, the charging station can keep updating with new features while satisfying the UL certification.

### 5. RFID Authentication and Authorization Scheme

In current WINSmartEV system, users are able to authenticate themselves through a mobile app [1-3]. A concept of mesh network radio frequency identification (RFID) charging authorization system in [15], which facilitates

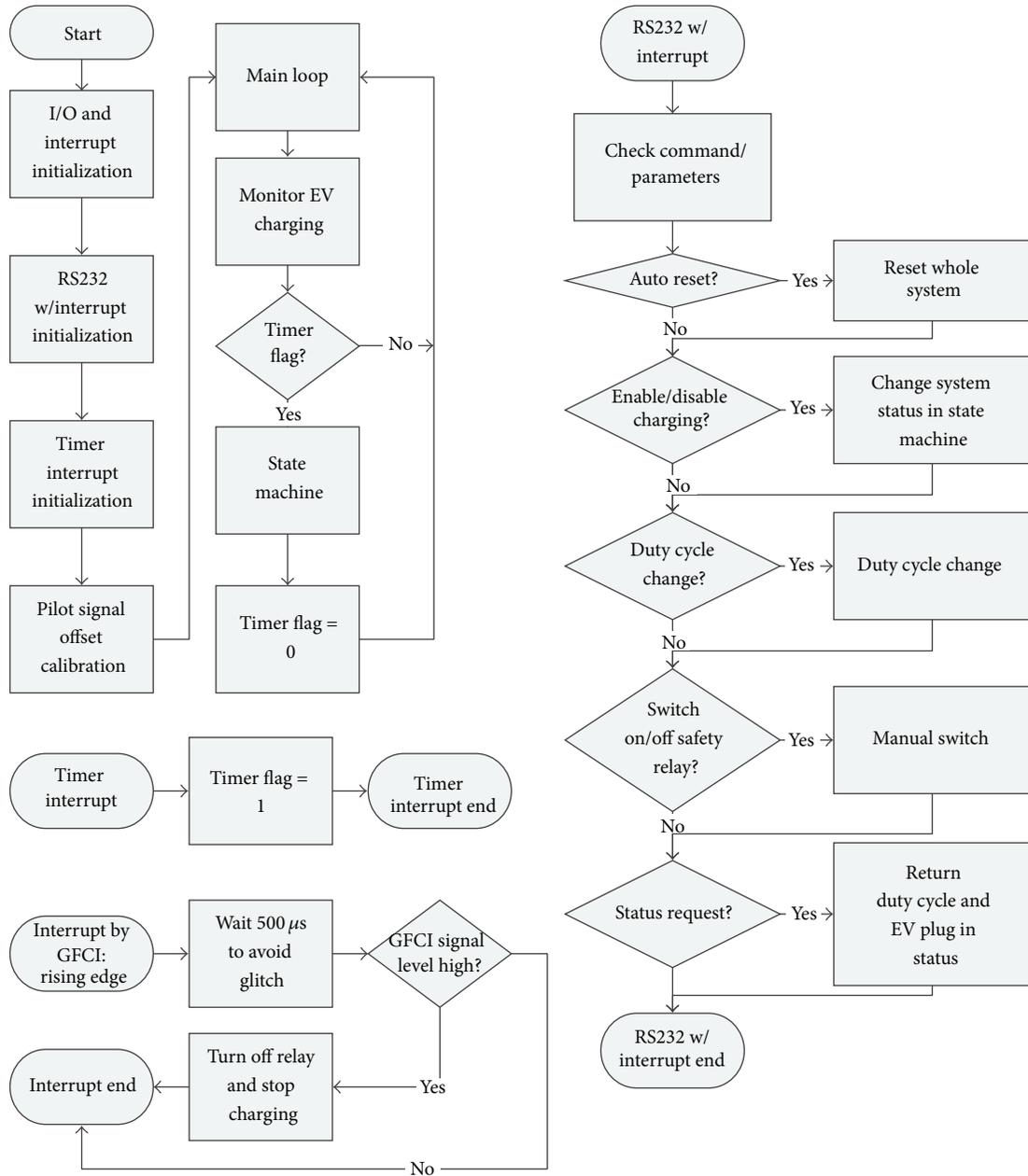


FIGURE 13: Local controller firmware flow chart.

the authentication process at a smart charging station, allows charging authorization to take place at the moment of PEV arrival without user involvement. The mesh network provides robust connections between PEVs and charging stations in a real world environment subject to signal blocking conditions. The ZigBee routers in the VMMS serve as RFID tags while the ZigBee coordinator, attached to the Gateway in the charging station, serves as the RFID reader. The unique 64-bit MAC address of each ZigBee device is utilized as an RFID tag. The charging authentication process includes ZigBee MAC address retrieval, user authorization, and PEV plug-in status detection. When the charging station detects the PEV at a distance, the received signal strength indication (RSSI) of the handshake serves as the metric for identifying

a PEV approaching a charging station. The PEV plug-in status detection is used to identify the presence of a PEV at a charging station and to associate the vehicle's ID with a particular channel. The mesh network RFID is developed based on existing hardware without additional cost and provides traditional RFID benefits while adding mesh network capability.

In order to add authorization/identification capability, the firmware of the ZigBee coordinator inside the charging station and the software on the server need to be redesigned. Rough processes in PEV charging authentication via RFID, including ZigBee MAC address retrieval, user authorization, and PEV plug-in status detection, are presented in [15]. However, more details of the collaboration between the

master controller (server) and the local controller (ZigBee coordinator) need to be addressed for implementation. The details of the proposed master-slave control scheme for RFID authentication and authorization are presented in the following.

In [15], the authentication and authorization processes are periodically handled by the server. In the authentication process, new PEV arrivals are checked by the ZigBee coordinator inside the charging station. The server sends out the “rgst” command to check if new tag IDs have been registered after RFID reader initialization. The “stat” command is later sent out to identify which charging station a newly arrived PEV is plugged into.

In real practice, in order to accelerate the performance of the system, the system needs to be modified to data push system. Instead of periodically sending “rgst” command to retrieve the new tag IDs, the ZigBee coordinator pushes tag IDs to the database once new tag ID is detected. Notice that the local controller inside the charging station serves the trigger signal of PEV plug-in status. Once the PEV plug-in status is detected, the local controller pushes the status to the database. In the authorization process, if the tag ID corresponds to an authorized user account in the database, the command to enable charging is sent out to begin PEV charging.

In the collaboration between master and slave controllers in RFID mesh network feature, several issues including handshake request interval, PEV approaching and leaving determination, and exception condition handling are discussed in the following.

About the handshake request interval, the maximum time for a two-hop response is 2 seconds in [9], which means  $T_{\text{ZigBee}}$  has maximum value of 2 seconds. Considering  $T_{\text{CAN\_read}}$  is with maximum value of 0.1 second,  $T_{\text{wait}}$  will have to be greater than 2.1 seconds per equation (1). As a result, 2.1 second minimum waiting interval must be incorporated on the local controller. Therefore, an interval much larger than 2.1 seconds needs to be incorporated for detecting an approaching PEV. Taking 3G communication delay presented in [10] into account, the maximum round trip time of 3G is around 5 seconds, which means the server will need to wait 7.1 seconds to receive a response to a data request.

As for PEV approaching and leaving determination, in most cases, the accepted speed limit in parking lot is 5 mph, which means a PEV approaches a charging station by 4.5 meters every 2 seconds. Assuming that the PEV parks 5 meters away from the charging station, after a PEV is detected at a distance of 50 meters, the station will have a maximum of 10 handshakes to determine whether the PEV is approaching or leaving.

Considering the exception condition handling, when more than two PEVs come to the same charging station around the same time, the charging station might not have a way to associate the IDs with the corresponding outlets. In this case, the server needs an exception handling process to handle the charging sessions. If the arriving PEVs have different size of on-board chargers, the charging station is able to associate the IDs with outlets due to different current when charging. However, if the on-board chargers are the same

size, the server cannot associate IDs with the outlets. In this case, the server can later associate the charging sessions with IDs and outlets when the PEVs leave by detecting the PEVs' RSSI; the server can also associate the IDs and outlets by SOC when PEVs are fully charged before they leave. If the PEVs with same size on board chargers arrive and leave around the same time without being fully charged, there is no need to distinguish the charging session because their drivers will be billed for the same energy consumption.

## 6. Conclusion

In this paper, we have proposed, designed, and implemented a master-slave control scheme for the PEV smart charging infrastructure. This scheme includes adding a power information collector (PIC) to the level one EVSE that not only makes it faster and more scalable, but also it enables the level 1 charger to execute operations within the EVSE itself. With these features, the level 1 EVSE can execute simple charging algorithms such as round-robin locally. Furthermore, these enhanced capabilities allow the server to control the level 1 charger as a slave, making the network structure more robust.

The hardware for level 2 EVSE has been updated in order to simplify the design and enhance its features and functionalities. This update includes removing the redundant ZigBee communication system and updating the microprocessor to a more powerful one. These updates allow the implementation of a power sharing algorithm locally. This is the simplest algorithm for level 2 charging. Furthermore, a fair charging algorithm appropriate for the level 2 EVSE is proposed. This charging algorithm is calculated on the server side and executed with minimal instructions on the EVSE side. Enhancements to the GFCI system have also been proposed that runs a system check to ensure that the GFCI system is operating properly and will shut off power when required in order to prevent hazards.

Algorithms have also been proposed to enhance the capability of the VMM system. These enhancements will include the capability to automatically authenticate and authorize each PEV as it approaches the EVSE. This allows the user to drive up to the EVSE and connect the cable. These algorithms control the system to automatically do the rest, so the user does not have to log into the server to begin charging.

Every PEV that is purchased instead of a dedicated fossil fuel burner is a good step in the direction of energy independence and lower greenhouse gas emissions. As ever more of these PEVs hit the road, sufficient charging infrastructure becomes even more important in proliferating PEVs throughout the car market. With its new enhanced ability to multiply the number of PEVs serviced for a given electrical infrastructure, WINSmartEV is poised to not only serve as a part of the nationwide smart grid system but as part of the larger push for PEV proliferation.

## Disclosure

Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer,

or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work has been sponsored in part by a Grant from the LADWP/DOE fund 20699 & 20686, (Smart Grid Regional Demonstration Project). This material is based upon work supported by the United States Department of Energy under Award no. DE-OE000012 and the Los Angeles Department of Water and Power. Neither the United States Government nor any agency thereof, the Los Angeles Department of Water and Power, nor any of their employees make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights.

## References

- [1] R. Gadh, S. Mal, S. Prabhu et al., "Smart electric vehicle (ev) charging and grid integration apparatus and methods," US patent, PCT international patent, no. US20130179061A1, WO2011156776A2, WO2011156776A3, PCT/US2011/040077, US13/693, 2010.
- [2] R. Gadh, A. Chattopadhyay, C.-Y. Chung et al., "Intelligent electric vehicle charging system," US patent, PCT international patent, no. WO2013019989A2, WO2013019989A3, PCT/US2012/049393, 2011.
- [3] R. Gadh, C. Y. Chung, L. Qi, and C.-C. Chu, "Network based management for multiplexed electric vehicle charging," US patent, PCT international patent, no. US20130154561A1, US13/691, 709, 2011.
- [4] R. Lowenthal, D. Baxter, H. Bhade, and P. Mandal, "Network-controlled charging system for electric vehicles," US patent US7956570B2, 2011.
- [5] D. Baxter, H. Bhade, R. Lowenthal, and P. Mandal, "Network-controlled charging system for electric vehicles through use of a remote server," US patent US8138715B2, 2012.
- [6] D. Baxter, C. F. Hagenmaier Jr., M. T. Tormey, and R. Lowenthal, "Electrical circuit sharing for electric vehicle charging stations," US patent US8013570B2, 2011.
- [7] S. Mal, A. Chattopadhyay, A. Yang, and R. Gadh, "Electric vehicle smart charging and vehicle-to-grid operation," *International Journal of Parallel, Emergent and Distributed Systems*, vol. 27, no. 3, 2012.
- [8] Y. He, B. Venkatesh, and L. Guan, "Optimal scheduling for charging and discharging of electric vehicles," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1095–1105, 2012.
- [9] A. Shepelev, C. Chung, C. Chu, and R. Gadh, "Mesh network design for smart charging infrastructure and electric vehicle remote monitoring," in *Proceedings of the International Conference on ICT Convergence*, pp. 14–16, Jeju, Republic of Korea, October 2013.
- [10] P. Kulshrestha, L. Wang, M.-Y. Chow, and S. Lukic, "Intelligent energy management system simulator for PHEVs at municipal parking deck in a smart grid environment," in *Proceedings of the IEEE Power & Energy Society General Meeting (PES '09)*, Calgary, Canada, July 2009.
- [11] S. Yang, W. Cheng, Y. Hsu, C. Gan, and Y. Lin, "Chargescheduling of electricvehicles in highways," *Mathematical and Computer Modelling*, vol. 57, no. 11-12, pp. 2873–2882, 2013.
- [12] W. Su and M. Chow, "Computational intelligence-based energy management for a large-scale PHEV/PEV enabled municipal parking deck," *Applied Energy*, vol. 96, pp. 171–182, 2012.
- [13] C. Chung, J. Chynoweth, C. Qiu, C. Chu, and R. Gadh, "Design of fair charging algorithm for smart electrical vehicle charging infrastructure," in *Proceedings of the International Conference on ICT Convergence*, pp. 14–16, Jeju, Republic of Korea, October 2013.
- [14] C. Chung, E. Youn, J. Chynoweth, C. Qiu, C. Chu, and R. Gadh, "Safety design for smart electric vehicle charging with current and multiplexing control," in *Proceedings of the IEEE International Conference on Smart Grid Communications*, pp. 21–24, Vancouver, Canada, October 2013.
- [15] C. Chung, A. Shepelev, C. Qiu, C. Chu, and R. Gadh, "Design of RFID mesh network for electric vehicle smart charging infrastructure," in *Proceedings of the IEEE International Conference on RFID Technologies and Applications*, Johor Bahru, Malaysia, September 2013.
- [16] C. Chung, P. Chu, and R. Gadh, "Design of smart charging infrastructure hardware and firmware design of the variable current multiplexing charging system," in *Proceedings of the 7th Global Conference on Power Control and Optimization (PCO '13)*, pp. 25–27, August 2013.
- [17] C. Chung, J. Chynoweth, C. Qiu, C. Chu, and R. Gadh, "Design of fast response smart electric vehicle charging infrastructure," in *Proceedings of the IEEE Green Energy and Systems Conference (IGESC '13)*, Long Beach, Calif, USA, November 2013.
- [18] M. Majidpour, C. Qiu, C. Chung, P. Chu, R. Gadh, and H. R. Pota, "Fast demand forecast of electric vehicle charging stations for cell phone application," in *Proceedings of the IEEE Power & Energy Society General Meeting*, National Harbor, Md, USA, July 2014.
- [19] *User's Manual Model CS-40*, Clipper Creek, Auburn, Calif, USA, 2014, <http://www.clippercreek.com/pdf/CS-40%20User%20Manual%20DLP%20IP%20081205%20v03.pdf>.
- [20] H. Bhade, M. Tormey, D. Baxter, R. Lowenthal, and P. Mandal, "Overcurrent and ground fault protection in a networked charging station for electric vehicles," US patent US8072184B2, 2011.

## Research Article

# Emission Controls Using Different Temperatures of Combustion Air

**Radovan Nosek, Michal Holubčík, and Štefan Papučík**

*University of Zilina, Univerzitna 8215/1, 010 26 Zilina, Slovakia*

Correspondence should be addressed to Radovan Nosek; [radovan.nosek@fstroj.uniza.sk](mailto:radovan.nosek@fstroj.uniza.sk)

Received 28 February 2014; Accepted 19 March 2014; Published 21 May 2014

Academic Editors: N. Barsoum, P. Vasant, and G.-W. Weber

Copyright © 2014 Radovan Nosek et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The effort of many manufacturers of heat sources is to achieve the maximum efficiency of energy transformation chemically bound in the fuel to heat. Therefore, it is necessary to streamline the combustion process and minimize the formation of emission during combustion. The paper presents an analysis of the combustion air temperature to the heat performance and emission parameters of burning biomass. In the second part of the paper the impact of different dendromass on formation of emissions in small heat source is evaluated. The measured results show that the regulation of the temperature of the combustion air has an effect on concentration of emissions from the combustion of biomass.

## 1. Introduction

The main intention of European Union is to exploit the potential of energy savings and renewable sources. In Slovakia the most promising renewable energy source seems to be biomass. Its use has growing importance. The most common form of biomass is wood, either in pieces or as wood waste. During the combustion process of renewable fuels pollutants are generated into the atmosphere and have a negative impact on human health. The most monitored pollutants are particulate matter, carbon monoxide, nitrogen oxides, and sulphur dioxide [1, 2].

Emissions emitted during combustion are mainly constituted of gaseous and particulate pollutants. The aim is to reduce the concentration of these substances to acceptable levels, since the emissions have a significant proportion of air pollution [3].

The solid particles are entrained with flue gas stream from the combustion chamber of boiler. Particulate matter (PM) consists of soot, inorganic matter (ash), and organic matter (nonvolatile flammable). Particles are imported into the flue gas by ash, nonvolatile, and combustible soot.

Particulate matter formation during fuel combustion depends on many factors, including flame temperature, composition and concentration of combustion reactants, and residence time within the reaction zone [4]. Although PM

formation from combustion is not fully understood, it is suspected that the process involves both nucleation and condensation mechanisms [5].

The size of particles formed during combustion is dependent on the time spent in the formation and oxidation zones. The size of a biomass exhaust particle can span a range from less than  $0.01\ \mu\text{m}$  to greater than  $100\ \mu\text{m}$ . However, the majority of biomass combustion aerosol is typically smaller than  $1\ \mu\text{m}$  in diameter [6].

Today is the greatest attention paid to the size of particles (aerodynamic diameter) less than  $10\ \mu\text{m}$  (PM<sub>10</sub>), which may penetrate into the respiratory tract. Particles of this fraction are divided into two groups based on different sizes, the mechanism, the composition, and behaviour of the atmosphere.

The first group is made up of particles of size below  $2.5\ \mu\text{m}$  (fine respirable fraction—PM<sub>2,5</sub>), arising from chemical reactions nucleation, condensation of gaseous emissions generated at the surface of particles, or coagulation of the finest particles.

The second group created particles in the range of the size from  $2.5$  to  $10\ \mu\text{m}$  (coarse fraction—PM<sub>2,5</sub> to 10).

Finest particles with a diameter below  $2.5\ \mu\text{m}$  (PM<sub>2,5</sub>) are considered to cause the greatest harm to human health. They deposit deep in the lungs and block the reproduction of cells [7–9].

Various types of wood have different composition and properties such as calorific value and ash melting behavior of temperature, which greatly affect the production of PM.

In this work, experimental measurements were carried out and focused on the formation of PM during combustion of different types of dendromass in a small heat source. The effect of various temperatures of the primary combustion air to the emission parameters is also evaluated.

## 2. Measurement of Emission Parameters

Methods for measuring emissions of pollutants can be divided in principle into measuring of particulate matter and gaseous substances. Methods and measurement principles are based on the emission properties of the fluid medium. One of the method for measuring particulate matters is presented below.

*Gravimetric Method.* Gravimetric method is the manual single method with sampling of the flow gas by probe. It is based on determination of the median concentrations by sampling from multiple points of measurements cross-section and their subsequent gravimetric assessment. Solid contaminants are usually separated by an external filter.

Representative sampling is performed by sampling probe suitable shape and the correct speed under isokinetic condition [10].

Concentration of particulate matter in the flue gas is covered to standard conditions and can be determined for wet or dry flue gas. Measured volume of sample taken on the volume gas meter should be converted to standard conditions, that is, 101325 Pa pressure and temperature of 273.15 K (0°C). Therefore, the temperature and pressure of measured sample are measured before gas meter.

The cumulative collection can provide in the cross-section average concentration but not concentration profile. Flow velocity or flow of the sample gas is measured by ensuring of isokinetic conditions, for example, by aperture track and a total collected amount of gas by gas meter [11, 12].

In gravimetric method, the taking of representative samples is realized by probe with appropriate shape right from the flowing gas [13].

To meet the increasing requirements toward the fine particulate determination, the multistage impactor probe was used in these experiments. Impactor separation system is intended to filtrate and separate solid emissions in three-stage impactor. The construction of device allows parallel separation of solid elements PM 10 and PM 2,5 (Figure 1).

The advantage of the gravimetric method is its simplicity and the relatively low sampler costs.

## 3. Experimental Measurement

As the heat source was used fireplace rated at 6 kW, which is designed for burning of piece wood. Bottom of the combustion chamber is topped with grate and the container where the ash falls. Access to the combustion chamber is

through the doors that are glazed with high heat resistant glass.

*3.1. Cooling/Heating of Combustion Air.* Changing the temperature of the combustion air inlet was performed on the primary combustion air. The heat exchangers are plugged to pipe of primary air supply for heating/cooling of combustion air. This way is the temperature of the incoming primary combustion air heated/cooled to the desired temperature level. The minimal supply air temperature was  $-5^{\circ}\text{C}$  and gradually increased up to  $40^{\circ}\text{C}$ . The increase in temperature between the measurements was  $5^{\circ}\text{C}$  and was regulated by the heat exchanger, which is located behind the fan in a duct. Temperature control for the heat exchanger was ensured by circulatory thermostat Julabo F40.

The scheme of experimental stand for the heating/cooling air supply is shown in Figure 2.

In order to evaluate the quality of combustion process, the gas composition was measured by analyzer.

*3.2. Dendromass.* During the experiment, the different types of wood were tested as well. Every measurement lasted 1 hour and was burned to about 1,5 kg of fuel. For the experimental measurements the following types of wood that are listed in Table 1 were used.

*3.3. Position of Secondary Air.* Modern modifications allow an increase of heating efficiency and reduction of emission concentration. The amount of emissions can be affected by several factors. One of the important factors is the position of secondary combustion air.

The experimental heat source has the following air inlets:

- (i) primary (frontal)—airflow through the grate and ashtray towards fuel,
- (ii) secondary (back)—process using residual combustible gases that would normally escape through the chimney. There is an increase in efficiency and thus lower fuel consumption,
- (iii) tertiary (top)—used for blowing off the windshield, preventing clogging, also contributing to improvement of combustion process, and reducing emissions. Fireplace is designed for burning of piece wood (see Figure 3).

In this task, the different positions of secondary air inlet were investigated. The aim was to evaluate whenever the location of air inlet has influence on the formation of particulate matter.

## 4. Results and Discussion

During the measurements concentrations of following emissions were recorded: CO, CO<sub>2</sub>, and NO and particulate matters in the flue gas.

*4.1. Effect of Air Temperature on Formation of Emission.* The temperature of the primary combustion air supplied to

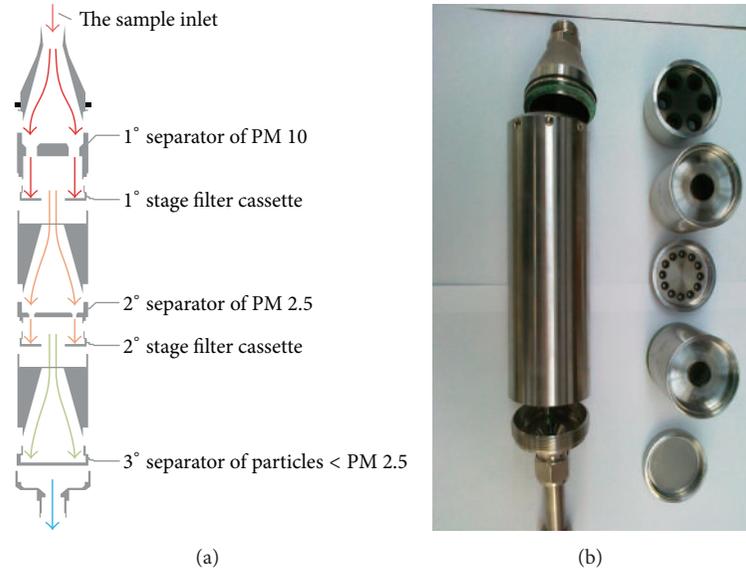


FIGURE 1: Multistage separation impactor.

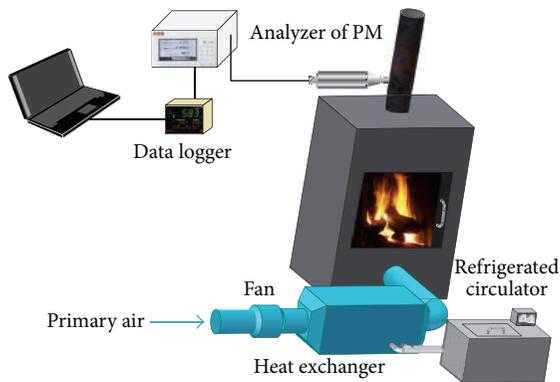


FIGURE 2: Scheme of experimental setup.

TABLE 1: Types of wood and their parameters.

Type of fuel	Calorific value (MJ/kg)	Humidity (%)
Beech	17,5	7,64
Spruce	19,3	7,87
White birch with bark	28	18,15
White birch without bark	19	18,15

fireplace varied by changing the setting temperature on the refrigerated circulator.

Different temperatures of the primary combustion air have impact on formation of gaseous emissions and particulate matter.

Figure 4 shows the results of the measurement of carbon dioxide according to the set temperature of the primary combustion air.

The highest average CO<sub>2</sub> was recorded at 35°C of inlet air, while at 15°C of supplied air the lowest average value of 3.20% was registered. Carbon dioxide formation has a trend

to increase with increasing temperature of the primary combustion air.

Figure 5 shows the results of the measurement of carbon monoxide.

The highest average values reached 7193 mg·m<sup>-3</sup> of CO and were recorded at 10°C inlet air, while at 30°C supply air reached the lowest average value of 5051 mg·m<sup>-3</sup>. The results indicate that formation of carbon monoxide has a trend to decrease with increasing temperature of the primary combustion air.

Dependence of NO<sub>x</sub> formation on the different temperatures of the primary combustion air to the experimental heat source shows Figure 6.

The highest average values of the measured NO<sub>x</sub> (111.65 mg·m<sup>-3</sup>) were achieved at 10°C, and the lowest average values were measured at 20°C with a value of 80.16 mg·m<sup>-3</sup>. NO<sub>x</sub> production has a trend to decrease with increasing temperature of the primary combustion air.

The results of PM concentration depending on the temperature of primary combustion air are shown in Figures 7 and 8.

Measurement of particulate matter with a change of temperature of combustion air has reached the maximum concentration of 202 mg·m<sup>-3</sup>. Minimum concentration of PM emission was generated at 35°C of combustion air.

4.2. *Different Type of Dendromass.* The second part of the work deals with the effect of different dendromass to formation of solid particles. Generation of emissions is largely influenced by type of fuel that is burned in heat source. Every fuel has different properties and chemical composition, which ultimately affects the combustion process, the amount of actual emissions, and ash content. During the experimental measurements the same combustion conditions were secured, that is, uniform supply of primary, secondary

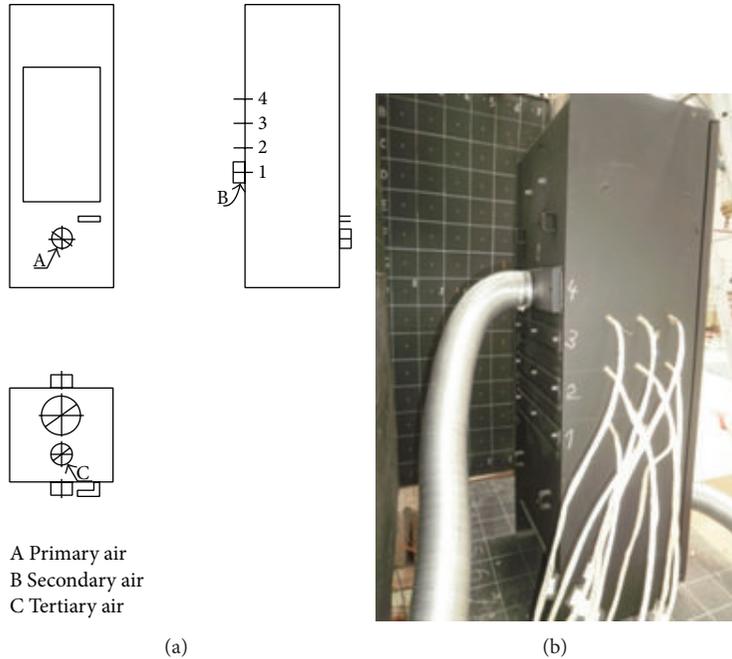


FIGURE 3: Position of combustion air inlets.

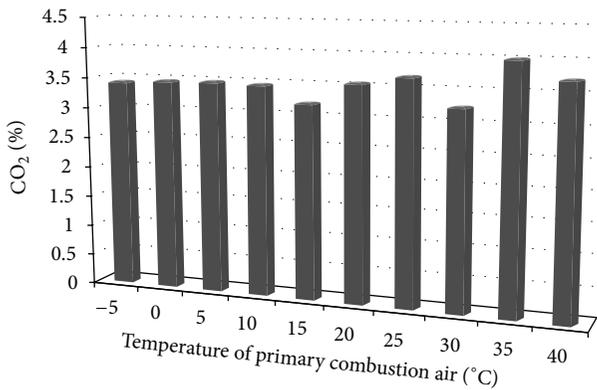


FIGURE 4: Average CO<sub>2</sub> emissions depending on the temperature change of the primary combustion air.

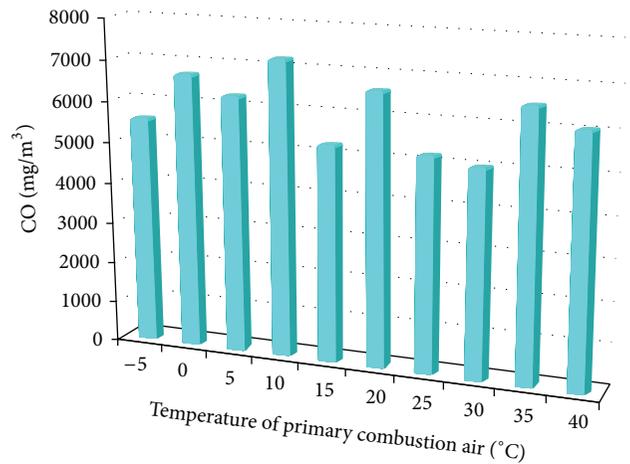


FIGURE 5: Average CO emissions depending on the temperature change of the primary combustion air.

and tertiary air, the same pressure in chimney (12 Pa), and a maximum dose of 1.5 kg of fuel.

Particulate measurements were conducted on all types of wood for 30 minutes. During this time were captured PM to the filters from each sample. These were subsequently stripped of moisture and weighed. Concentrations of particulate matter were determined by difference weight of the filter before and after the measurement. The highest amount of particulate matter was observed in measurements of white birch with bark and beech (Figure 9).

4.3. *Different Positions of Secondary Air.* The final part presents the most effective location of secondary air inlet in relation to formation of particulate matters. The influence of three air inlets position was analyzed.

On Figure 10 the minimal and maximal values of measured particulate matter concentrations (PM) are shown. Measurement of PM for all fully open combustion air reached concentration 21 mg·m<sup>-3</sup>. Minimum concentration of PM was registered with the involvement of secondary supply in the second row, where only 13,09 mg·m<sup>-3</sup> was measured.

It can be concluded that in terms of PM it is advantageous to supply the combustion air through second row.

### 5. Conclusion

The aim of this work was to demonstrate the impact of the primary combustion air temperature on emissions parameters.

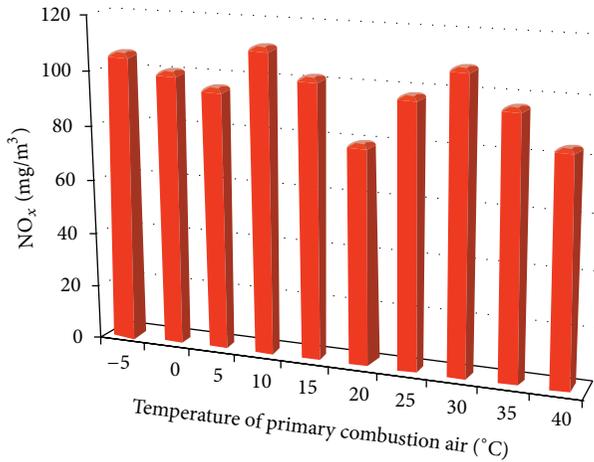


FIGURE 6: Average NO<sub>x</sub> emissions depending on the temperature change of the primary combustion air.

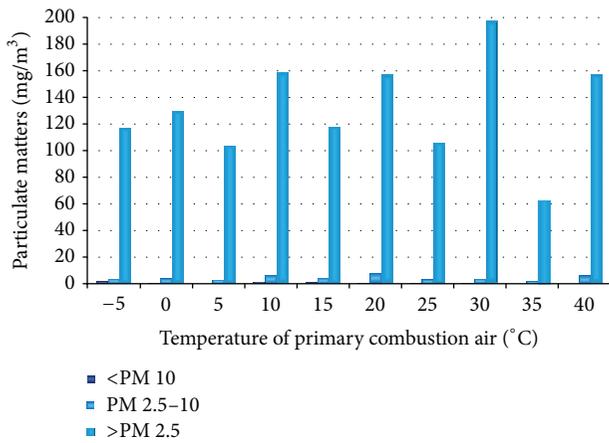


FIGURE 7: Concentrations of particulate matter for different temperatures of air.

Presented results of emissions depending on the temperature of the primary combustion air do not indicate the most suitable setting of temperature. For each type of emission the lowest value at different temperatures of the primary combustion air has been reached.

From the experimental measurements of solid emissions it is clear that in terms of the lowest value of PM it is preferred to supply the primary combustion air into the combustion process at a temperature of 35°C.

It can be argued that the production of carbon monoxide (CO) decreases with increasing temperature at the expense of higher production of carbon dioxide (CO<sub>2</sub>). The formation of CO is influenced by several factors and therefore its different concentration during the measurements cannot be attributed to changing temperatures of the combustion air.

In this research work analysis of the impact of different types of dendromass on the formation of particulate matters during the combustion process was carried out. The results of measurements indicate that the type of fuel has a considerable influence on the combustion process and the formation of

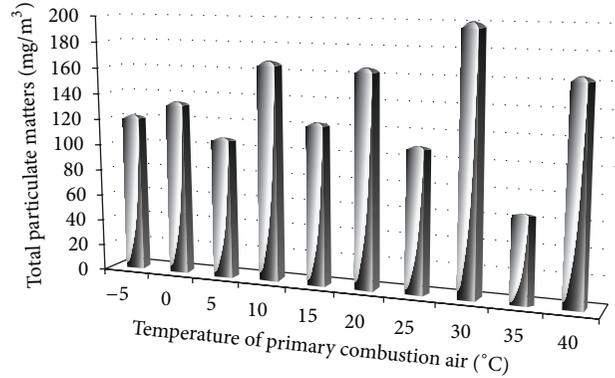


FIGURE 8: Dependence of total particulate matter on the temperature.

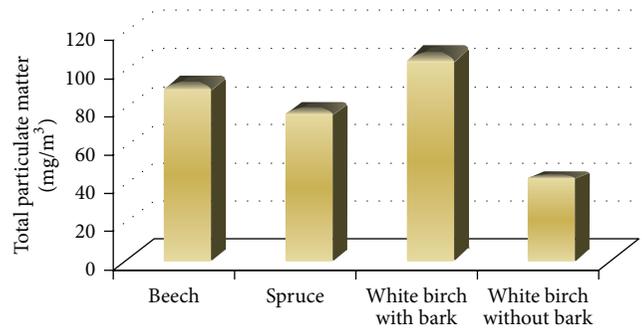


FIGURE 9: Concentrations of PM for different types of dendromass.

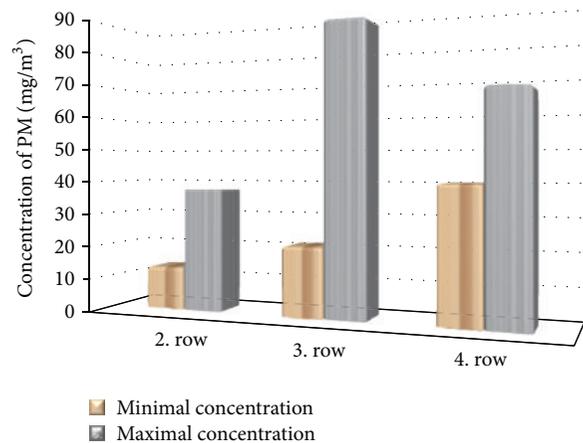


FIGURE 10: Dependence of PM on position of air inlet.

particulate matters. This phenomenon is largely influenced by the different properties and chemical composition of different types of dendromass.

In the case of birch without bark, the lowest values of PM were measured, suggesting that the bark of firewood has a significant proportion on the formation of solid particles.

The measured results show that the type of firewood affects emission parameters of the heat source.

Computer modelling is becoming more powerful and developed, therefore gaining in popularity. It is emerging

as an attractive tool to assist the combustion engineer in such areas as new process design, plant scale-up, retrofitting, and pollutant control. Therefore the numerical simulation of particulate matter formation will be done in the future research.

### Conflict of Interests

There is no conflict of interests regarding the publication of this paper.

### Acknowledgments

This work was supported by the VEGA no. 1/1353/12 and sponsored in the frame of program OPV-Podpora kvality vzdelávania a rozvoj ľudských zdrojov v oblasti technického výskumu a vývoja v priestore modernej vedomostnej spoločnosti ITMS 26110230117.

### References

- [1] L. Dzurenda, *Spáovanie Dreva a Kôry, Vydanie I. Zvolen, Vydavateľstvo Tu Vo Zvolene, 2005.*
- [2] M. Carnogurska, M. Prihoda, and T. Brestovic, "Modelling of nitrogen oxides formation applying dimensional analysis," *Chemical and Process Engineering*, vol. 32, no. 3, pp. 175–184, 2011.
- [3] P. Chudíková, M. Taušová, K. Erdélyiová, and P. Tauš, "Potential of dendromass in Slovak Republic and its actual exploitation in thermic economy," *Acta Montanistica Slovaca*, vol. 15, no. 2, pp. 139–145, 2011.
- [4] S. Shurupov, *Some Factors that Govern Particulate Carbon Formation during Pyrolysis of Hydrocarbons*, Combustion Institute, Pittsburgh, Pa, USA, 2000.
- [5] B. S. Haynes and H. Wagner, "Soot formation," *Progress in Energy and Combustion Science*, vol. 7, no. 4, pp. 229–273, 1981.
- [6] L. Naeher, M. Brauerb, M. Lipsettc et al., "Wood smoke health effects," *Inhalation Toxicology*, vol. 19, no. 1, pp. 67–106, 2007.
- [7] J. Cernecký, A. Neupauerová, I. Janosko, and M. Soldan, *Technika Životného Prostredia*, Technická Univerzita vo Zvolene, 2010.
- [8] J. Jandačka, M. Malcho, and M. Mikulík, *Biomasa Ako Zdroj Energie*, Juraj Štefuň—Georg, 2007.
- [9] J. Jandačka, R. Nosek, Š. Papučík et al., *Drevné Pelety a Aditíva*, Juraj Štefuň—Georg, 2011.
- [10] N. Kaliyan and R. V. Morey, "Factors affecting strength and durability of densified biomass products," *Biomass and Bioenergy*, vol. 33, no. 3, pp. 337–359, 2009.
- [11] P. Nemeč, A. Caja, and R. Lenhard, "Visualization of heat transport in heat pipes using thermocamera," *Archives of Thermodynamics*, vol. 31, no. 4, pp. 125–132, 2010.
- [12] P. Mihal'ov and M. Čarnogurská, "Návrh regulácie kotla PK4," *Acta Mechanica Slovaca*, vol. 3, pp. 153–158, 1999.
- [13] R. Nosek, J. Jandačka, and M. Holubčík, *Effect of Additives to Wood Pellets Properties, Power Control And Optimization*, Kuching, Malaysia, 2010.

## Research Article

# Heat Removal from Bipolar Transistor by Loop Heat Pipe with Nickel and Copper Porous Structures

**Patrik Nemeč, Martin Smitka, and Milan Malcho**

*Department of Power Engineering, Faculty of Mechanical Engineering, University of Žilina, Univerzitná 1, 01026 Žilina, Slovakia*

Correspondence should be addressed to Patrik Nemeč; patrik.nemec@fstroj.uniza.sk

Received 28 February 2014; Accepted 19 March 2014; Published 18 May 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Patrik Nemeč et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Loop heat pipes (LHPs) are used in many branches of industry, mainly for cooling of electrical elements and systems. The loop heat pipe is a vapour-liquid phase-change device that transfers heat from evaporator to condenser. One of the most important parts of the LHP is the porous wick structure. The wick structure provides capillary force to circulate the working fluid. To achieve good thermal performance of LHP, capillary wicks with high permeability and porosity and fine pore radius are expected. The aim of this work was to develop porous structures from copper and nickel powder with different grain sizes. For experiment copper powder with grain size of 50 and 100  $\mu\text{m}$  and nickel powder with grain size of 10 and 25  $\mu\text{m}$  were used. Analysis of these porous structures and LHP design are described in the paper. And the measurements' influences of porous structures in LHP on heat removal from the insulated gate bipolar transistor (IGBT) have been made.

## 1. Introduction

The trend development of electronic components is miniaturization of the dimension. It leads to an increase of waste heat. This heat often leads to lower performance and failure of electronic components in case of insufficient cooling. In order to maintain appropriate working conditions, waste heat must be removed. One possibility to remove waste heat is to use loop heat pipe (LHP). LHPs are two-phase heat transfer devices that utilize the evaporation and condensation of a working fluid to remove heat and the capillary forces developed in fine porous wicks to circulate the fluid. LHP consists of an evaporator with wick, a condenser, a compensation chamber, and liquid and vapor line (Figure 1). Only the evaporator and the compensation contain wicks; the rest of the loop is made of smooth wall tubing. The wick in the evaporator is made with fine pores for purpose of developing capillary pressure to circulate fluid around the loop, while the wick in the compensation chamber is made with larger pores for purpose of managing fluid ingress and egress. The operating principle of the LHP is as follows. As heat is applied to the evaporator, liquid is vaporized and the menisci formed at the liquid/vapour interface in the evaporator wick develop

capillary forces to push the vapour through the vapour line to the condenser. Vapour condenses in the condenser and the capillary forces continue to push liquid back to the evaporator. The waste heat from the heat source provides the driving force for the circulation of the working fluid and no external pumping power is required. The two-phase compensation chamber stores excess liquid and controls the operating temperature of the loop [1, 2].

In order for the loop to continue to function, the wick in the evaporator must develop a capillary pressure to overcome the total pressure drop in the loop. One of the advantages of a capillary loop is that the meniscus in the evaporator wick will automatically adjust its radius of curvature such that the resulting capillary pressure is equal to the total system pressure drop. The total pressure drop in the system is the sum of frictional pressure drops in the evaporator grooves, the vapour line, the condenser, the liquid line, and the evaporator wick, plus any static pressure drop due to gravity:

$$\Delta P_{\text{total}} = \Delta P_{\text{groove}} + \Delta P_{\text{vap}} + \Delta P_{\text{con}} + \Delta P_{\text{liq}} + \Delta P_w + \Delta P_g. \quad (1)$$

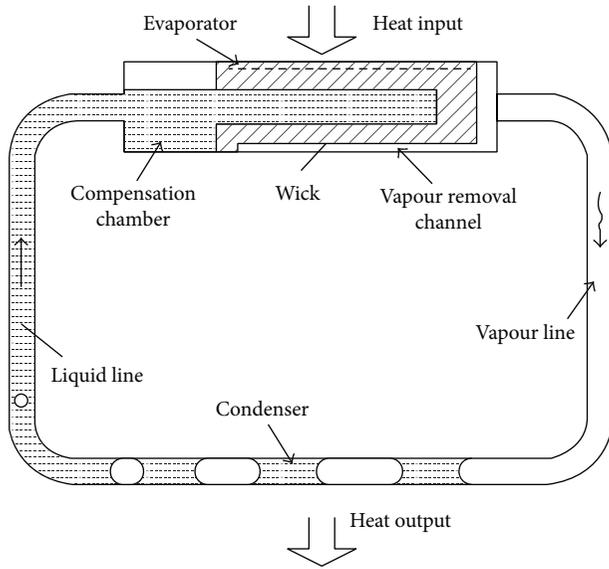


FIGURE 1: Schematic diagram of LHP [3].

The capillary pressure rise that the wick can develop is given by

$$\Delta P_{\text{cap}} = \frac{2\sigma \cdot \cos \theta}{R}, \quad (2)$$

where  $\sigma$  is the surface tension of the working fluid,  $R$  is the radius of curvature of the meniscus in the wick, and  $\theta$  is the contact angle between the liquid and the wick. As the heat load to the evaporator increases, so will the mass flow rate and the total pressure drop in the system. In response, the radius of curvature of the meniscus decreases so as to provide a higher capillary pressure that matches the total system pressure drop. The radius of curvature will continue to decrease with increasing heat loads until it is equal to the pore radius of the wick,  $R_p$ . Under this condition, the wick has reached its maximum capillary pumping capability:

$$\Delta P_{\text{cap,max}} = \frac{2\sigma \cdot \cos \theta}{R_p}. \quad (3)$$

Further increase of the heat load will lead to vapour penetration through the wick and system depressed. Thus, under normal operation, the following condition must be satisfied at all times [4]:

$$\Delta P_{\text{total}} \leq \Delta P_{\text{cap}}. \quad (4)$$

Williams and Harris [5] investigated the in-plane and cross-plane properties of step-graded metal felt wicks for heat pipe applications. Porosity, effective pore radius, and liquid permeability were determined using imbibition, capillary flow porometry, and pressure-flow rate data, respectively. The authors determined that many of the correlations in the literature for pore size and permeability are too general in nature, echoing the conclusions of Bonnefoy et al. [6] in regard to effective thermal conductivity.

Holley and Faghri [7] outlined methods for permeability and effective pore radius measurements based on the rate-of-rise test.

Typically, the rate-of-rise test requires observing the liquid front as it rises in a dry wick partially immersed in a liquid pool. As the precise location of this front can be difficult to detect, the authors devised a method using mass uptake rather than the meniscus front to determine the rate of rise of liquid in the wick. By analyzing the climbing meniscus, the authors developed a series of equations which could be used to numerically reduce the mass uptake data to yield permeability and pore size results.

Several relationships for permeability can be found; the most common is the Blake-Kozeny equation [8, 9], which gives the permeability of a bed of packed spheres as

$$K = \frac{r_p^2 \epsilon^3}{37.5(1 - \epsilon)^2}, \quad (5)$$

where  $K$  is permeability,  $r_p$  is pore radius, and  $\epsilon$  is porosity.

Ren and Wu [10] modeled the effect of wick effective thermal conductivity in LHP evaporators; a two-dimensional axisymmetric model was developed yielding results in agreement with the literature in some respects, namely, the position of the liquid front in relation to a heated fin [11, 12].

Zhao and Liao [12] present temperature profiles indicating decreasing heat leak for increasing heat flux in a bed of packed spheres.

Iverson et al. [13] studied heat and mass transport in sintered copper wick structures. Wick samples were mounted vertically with the lower section immersed in a pool of water. A heater mounted to the back face of the wick applied power to the sample and the resulting temperature gradients were measured along with the mass flow rate of working fluid.

The majority of heat load is used in vaporization on the outer surface of wick [14]. The rest of heat input (called "heat leak") is conducted across the wick and is proportional to the effective thermal conductivity (ETC) of the capillary wicks [15]. Lower thermal conductivity of the porous wick ensures lesser heat conduction to the liquid inside the wick inner surface and maintains the operating temperature and thus the thermal resistance of the whole LHP.

Ku [4] and Furukawa [16] in research works dealt with LHP heat leak models utilize conductance parameter and the temperature difference between the evaporator and compensation chamber. Consider

$$\dot{Q}_{e,cc} = G_{e,cc} (T_e - T_{cc}), \quad (6)$$

where  $Q$  is heat leak,  $G$  is conductance parameter, and  $T$  is temperature of the evaporator and compensation chamber.

In steady state operation, the heat leak to the compensation chamber must be offset by the liquid returning from the condenser; (7) results, where  $\Delta T$  represents the subcooling of the returning fluid

$$\dot{Q}_{e,cc} = \dot{m} c_p \Delta T, \quad (7)$$

where  $m$  is mass flow and  $c_p$  is specific heat.

Chuang [17] developed a steady state LHP model which breaks the overall heat leak into two separate components: axially from the evaporator to the compensation chamber and radially from the heat source to the evaporator core. These two effects are related in that the formation of vapour bubbles in the evaporator core due to radial leak reduces the overall heat flow path back to the compensation chamber, increasing axial leak [4].

Chuang derived the following expressions for the axial and radial heat leak, respectively:

$$\begin{aligned} \dot{Q}_{\text{leak},a} &= k_{\text{eff}}A \left( \frac{T_e - T_{cc}}{L} \right) + (\text{Nu}k_f\pi L) \left( \frac{T_e - T_{cc}}{2} \right), \\ \dot{Q}_{\text{leak},r} &= \frac{2\pi k_{\text{eff}}L\zeta}{(r_o/r_i)^\zeta - 1} \Delta T_w, \end{aligned} \quad (8)$$

where  $Q_{\text{leak}}$  is heat leak power,  $k_{\text{eff}}$  is effective thermal conductivity,  $A$  is area,  $L$  is characteristic length,  $\text{Nu}$  is Nusselt number,  $k_f$  is fluid thermal conductivity, and  $\zeta$  represents a nondimensional ratio of advection and conduction given by

$$\zeta = \frac{\dot{m}c_p}{2\pi k_{\text{eff}}L}. \quad (9)$$

In his analysis and experiment, Chuang assumed this parameter to be zero, that is, pure conduction. For the low power cases studying this assumption was valid and resulted in low error; however, for high power levels or low wick conductivity, this assumption loses validity.

## 2. LHP Wick Structure

To achieve good thermal performance, capillary wicks with high permeability and porosity and fine pore radius are expected. These parameters depend mainly on the manufacturing process. The most frequently used wicks are made of sintered metal, like nickel, copper, titanium, stainless steel, or polymers (polyethylene, polypropylene, and PTFE) [18–20].

Tap powder sintering technique using a graphite matrix is used by Reimbrecta et al. [21] to prepare Ni wicks dedicated for capillary pump applications. Graphite is recommended because it shows a low interaction with nickel at the usual sintering temperatures. Xin et al. [22] used two different methods, the cold-pressing sintering and direct loose sintering for development of Ni and Ni-Cu (90% nickel and 10% copper) wicks for loop heat pipes. The optimal Ni-Cu capillary wick is prepared using direct loose sintering technique, with mean pore radii of  $0.54 \mu\text{m}$ . Huang and Franchi [23] fabricated bimodal wick structure using copper screen mesh and two powder materials (nickel filamentary powder and spherical copper powder). But these wicks may be produced with inherent shortcomings. Samanta et al. [24] developed by metal injection moulding Ni wick structures and perform study of its physical characteristic depending on sintering time (30, 60, and 90 min) and temperature (900, 930, and  $950^\circ\text{C}$ ). Gernert et al. [25] develop fine pore wick structure for LPH. Wu et al. [26] discuss effect of sintering temperature curve in wick structure manufactured for LHP.

According to [14] the main parameters of wick are porosity, pore diameter, and permeability. The optimal porosity of sintered wick is between 30 and 75% regardless of the pore diameter. The sintered material porosity increases when the temperature or the forming pressure decreases. The optimal permeability is between  $10^{-14}$  and  $3 \cdot 10^{-13} \text{ m}^2$ . The pore diameters of these various porous materials are between 1 and  $20 \mu\text{m}$ , except for copper, which has larger pore diameters (between 20 and  $100 \mu\text{m}$ ).

In [22] the optimal capillary wick was found to be sintered at  $650^\circ\text{C}$  for 30 min using direct loose sintering technique, with 90% nickel and 10% copper. The wick reaches the porosity of 70% and a mean pore diameter of  $1.8 \mu\text{m}$ . In [4] biporous nickel wicks were fabricated. Porosity of 77.4% was achieved using cold pressure sintering method, at a temperature of  $700^\circ\text{C}$ , with a pore former content of 30% in volume.

## 3. Characterization of Sintered Structures

According to the above-mentioned experiences with sintered structures for LHP we decide to make wick structures from nickel and copper powder. At first we do analysis of several sintered structures depending on grain size, sintering temperature and sintering time on porosity, pore size, and strength. In electric furnace etalons from copper powders with grain sizes of 50 and  $100 \mu\text{m}$  and nickel powders with grain size of 10 and  $25 \mu\text{m}$  were sintered. The copper powders were sintered at temperature of 800 and  $950^\circ\text{C}$  for time of 30 and 90 minutes and nickel powders were sintered at temperature of  $600^\circ\text{C}$  for time of 30 and 90 minutes.

**3.1. Porosity Measuring.** The porosity of a wick structure describes the fraction of void space in the material, where the void may contain working fluid [27]. For the porosity measuring, the weight method was used. At first, the sample was weighed in dry state. Secondly, the sample was soaked with distilled water ( $\rho = 0.998 \text{ g}\cdot\text{cm}^{-3}$  at  $20^\circ\text{C}$ ). The weight of absorbed water was estimated by the difference between both values and then a deduction of the “empty space” (thus the total pore volume) and the porosity. Consider

$$\varepsilon = \frac{M_{\text{soaked sample}} - M_{\text{dry sample}}}{V_{\text{total}} \cdot \rho_{\text{distilled water}}}, \quad (10)$$

where  $\varepsilon$  is wick structure porosity,  $M$  is weight of porous sample,  $V$  is pore volume of the porous sample, and  $\rho$  is density.

The results of porosity measuring are shown in Tables 1, 2, 3, and 4.

**3.2. Microscopic Analysis of Pore Size.** Investigation of etalons sintered structures by microscopic analysis shown, how influential sintered temperature and time pore size and ratio grain size to pore size of each structure. In Figure 2 are pictures created by 100 times zoom of porous structures sintered from copper powder grain size of 50 and  $100 \mu\text{m}$ . In the first two pictures it is seen that the structures sintered at temperature of

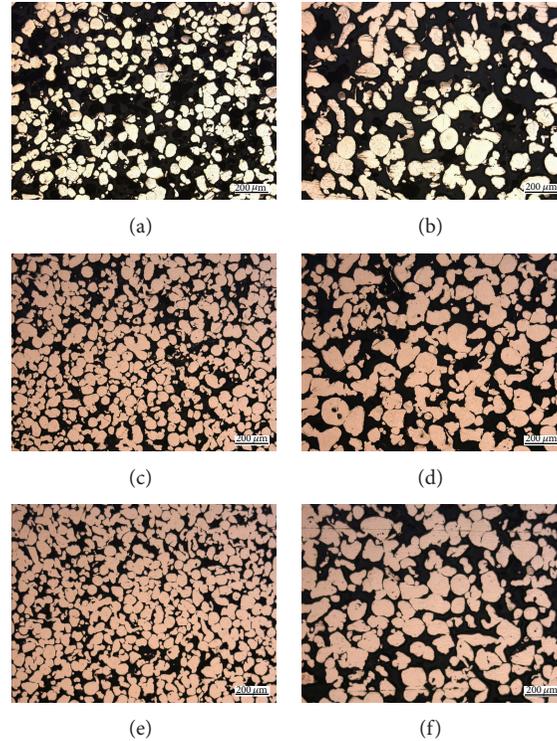


FIGURE 2: Microscopic pictures of the sintered structures from copper powders. (a) Grain size of  $50 \mu\text{m}$ , sintering temperature of  $800^\circ\text{C}$ , and sintering time of 30 min. (b) Grain size of  $100 \mu\text{m}$ , sintering temperature of  $800^\circ\text{C}$ , and sintering time of 30 min. (c) Grain size of  $50 \mu\text{m}$ , sintering temperature of  $950^\circ\text{C}$ , and sintering time of 30 min. (d) Grain size of  $100 \mu\text{m}$ , sintering temperature of  $950^\circ\text{C}$ , and sintering time of 30 min. (e) Grain size of  $50 \mu\text{m}$ , sintering temperature of  $950^\circ\text{C}$ , and sintering time of 90 min. (f) Grain size of  $100 \mu\text{m}$ , sintering temperature of  $950^\circ\text{C}$ , and sintering time of 90 min.

TABLE 1: Porosity of sintered structures from copper powder with grain size of  $50 \mu\text{m}$ .

Grain size ( $\mu\text{m}$ )	50	50	50	50
Sintering temperature ( $^\circ\text{C}$ )	800	800	950	950
Sintering time (min)	30	90	30	90
Porosity (%)	55	54	52	50

TABLE 2: Porosity of sintered structures from copper powder with grain size of  $100 \mu\text{m}$ .

Grain size ( $\mu\text{m}$ )	100	100	100	100
Sintering temperature ( $^\circ\text{C}$ )	800	800	950	950
Sintering time (min)	30	90	30	90
Porosity (%)	58	56	55	52

$800^\circ\text{C}$  have two times bigger pore than powder grain. Comparison of etalons sintered at temperatures 800 and  $950^\circ\text{C}$  shows that the etalons sintered at temperature of  $800^\circ\text{C}$  have so much bigger pore size than at temperature of  $950^\circ\text{C}$ . It is meaning that pore sizes are so much width to create capillary action in structure. Comparison of etalons sintered at the same temperature and various time intervals observed that the time of sintering at the temperature nearest the melting temperature of sintering material is not decisive. And the last

TABLE 3: Porosity of sintered structures from nickel powder with grain size of  $10 \mu\text{m}$ .

Grain size ( $\mu\text{m}$ )	10	10
Sintering temperature ( $^\circ\text{C}$ )	600	600
Sintering time (min)	30	90
Porosity (%)	69	67

TABLE 4: Porosity of sintered structures from nickel powder with grain size of  $25 \mu\text{m}$ .

Grain size ( $\mu\text{m}$ )	25	25
Sintering temperature ( $^\circ\text{C}$ )	600	600
Sintering time (min)	30	90
Porosity (%)	72	70

comparison of etalons at the same sintering temperature and time interval observes that the grain size of sintered material has impact on pore size. According to microscopic analysis of sintered structures, which clarifies their shape and profile, we can conclude that the main influencing factor of pore size is grain size, sintering temperature, and not so much sintering time.

In Figure 3 there are pictures created by 500 times zoom of porous structures sintered from nickel powder grain size

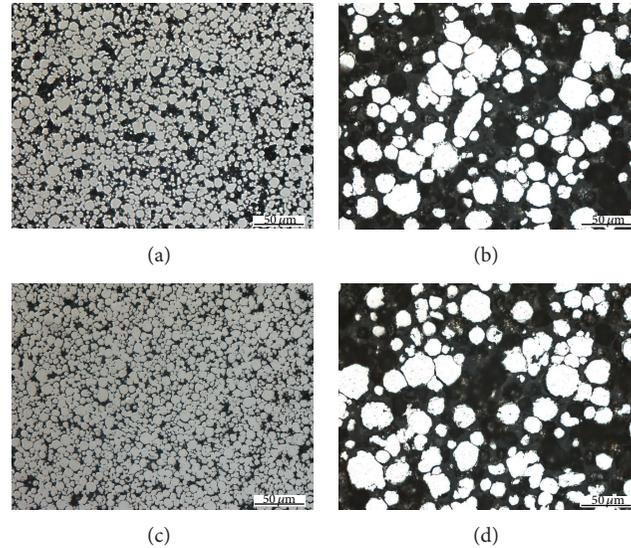


FIGURE 3: Microscopic pictures of the sintered structures from nickel powders. (a) Grain size of  $10\ \mu\text{m}$ , sintering temperature of  $600^\circ\text{C}$ , and sintering time of 30 min. (b) Grain size of  $25\ \mu\text{m}$ , sintering temperature of  $600^\circ\text{C}$ , and sintering time of 30 min. (c) Grain size of  $10\ \mu\text{m}$ , sintering temperature of  $600^\circ\text{C}$ , and sintering time of 90 min. (d) Grain size of  $25\ \mu\text{m}$ , sintering temperature of  $600^\circ\text{C}$ , and sintering time of 90 min.



FIGURE 4: Copper powder sintered wick structure.



FIGURE 5: Nickel powder sintered wick structure.

of 10 and  $25\ \mu\text{m}$ . Comparison of etalons sintered from nickel powder leads to the same conclusion findings as etalons sintered from copper powder. To the formation of pore size sintered structure does not affect sintering time but grain size.

#### 4. Design of LHP

From results of porosity measurement and microscopic analysis for wick structure of LHP two copper etalons and two nickel etalons were chosen. The first structure was made of copper grain size of  $50\ \mu\text{m}$  and sintered at temperature  $950^\circ\text{C}$  for 30 minutes (Figure 4). The second structure was made of copper grain size of  $100\ \mu\text{m}$  and sintered at temperature of  $950^\circ\text{C}$  for 30 minutes. The third structure was made of nickel grain size of  $10\ \mu\text{m}$  and sintered at temperature of  $600^\circ\text{C}$  for 90 minutes (Figure 5). The fourth structure was made of nickel grain size of  $25\ \mu\text{m}$  and sintered at temperature of  $600^\circ\text{C}$  for 90 minutes. The wick structures were sintered in sand form (mold) manufactured according to model of required shape in muffle furnace.

All parts of LHP (evaporator, compensation chamber, vapor, and liquid line) were made from copper pipes. As a working fluid distilled water was used. In the evaporator

sintered wick structure from copper powder was inserted. To avoid heat loss (it is also called heat leak) into the compensation chamber a brass flange with rubber seal was inserted between the evaporator and the compensation chamber. In Figure 6 there is the model of design LHP and the main parameters of LHP design are in Table 5.

#### 5. Measurement and Results

The LHP with sintered wick structure was proposed to test cooling of IGBT. On the evaporator of LHP the aluminum block with fixed insulated gate bipolar transistor (IGBT) was mounted. For better heat transport thermal conductive paste was applied on the connection between IGBT and aluminum block and between aluminum block and the evaporator [28]. The condenser of LHP was made as tube heat exchanger [29]. The cooling circle of heat exchanger was regulated by the thermostat at constant temperature of  $20^\circ\text{C}$  [30, 31]. In Figure 7 there is schema of the measuring unit.

The temperature of the IGBT was measured by thermocouple inserted under IGBT. The maximum permissible

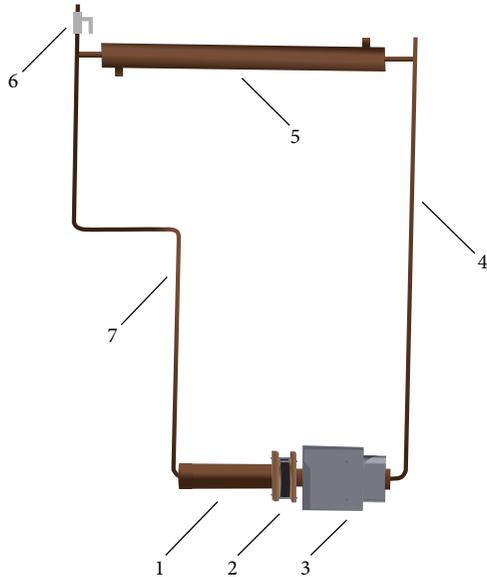


FIGURE 6: Model of design LHP: 1: compensation chamber, 2: rubber seal, 3: evaporator, 4: vapor line, 5: condenser, 6: filling valve, and 7: liquid line.

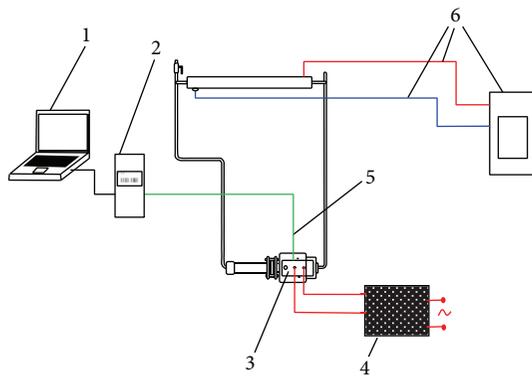


FIGURE 7: Schematic diagram of measuring device: 1: PC, 2: logger, 3: IGBT, 4: power supply voltage and current, 5: thermocouple, and 6: thermostat.

temperature of IGBT is  $100^{\circ}\text{C}$ . Transistor was connected to DC power of source and it was gradually loaded by DC. Like this was performed measurement impact of four kinds wick structures in LHP to heat remove from IGBT. The results from measurement of IGBT cooling by LHP with copper wick structures are shown in Figures 8 and 9, and results from measurement of IGBT cooling by LHP with nickel wick structures are shown in Figures 10 and 11. In Figures 8 and 9 it is seen that on start-up of LHP at input power of 100 W the temperature of evaporator increases and only after time, when the LHP starts to operate, the temperature of evaporator decreases and is stabilized. After first stabilization of the temperature input power gradually increased for 50 W. In Figures 10 and 11 it is seen that decrease of temperature did not occur at start-up of LHP with nickel wick structures.

TABLE 5: Main design parameters of the LHP.

Evaporator	
Total length (mm)	130
Active length (mm)	86
Outer/inner diameter (mm)	28/26
Material	Copper
Saddle	
Size (length/height/width)	118/89/40
Material	Alumina
Sintered copper powder	
Number of vapor grooves	6
Porosity (%)	52–55
Outer/inner diameter (mm)	26/8
Sintered nickel powder	
Number of vapor grooves	6
Porosity (%)	67–70
Outer/inner diameter (mm)	26/8
Compensation chamber	
Outer/inner diameter (mm)	35/33
Length (mm)	110
Charge mass	
Distilled water	60%
Vapor line	
Length (mm)	670
Outer/inner diameter (mm)	6/4
Liquid line	
Length (mm)	820
Outer/inner diameter (mm)	6/4
Condenser	
Length (mm)	420
Outer/inner diameter (mm)	6/4

Comparing results of dependence of temperature on input power of IGBT cooled by LHP with various variants of sintered wick structure, the LHP with nickel wick structure did not show better properties of heat removal than LHP with copper wick structure. Comparing dependencies of temperature course on input power of IGBT cooled by LHP with first and second wick structure it is seen that at load of up to 200 W has both LHP almost the same results. At higher input power than 200 W loaded into IGBT it is seen that the LHP with first structure did not remove heat from IGBT and the temperature of IGBT exceeded  $100^{\circ}\text{C}$ . The LHP with second wick structure is able to cool the IGBT under temperature of  $100^{\circ}\text{C}$  until the IGBT input power of 450 W. Comparing dependencies of temperature course on input power of IGBT cooled by LHP with third and fourth wick structure it is seen that evaporator temperature of LHP with third structure at input power of 100 W gradually increases with time and is stabilized at temperature  $92^{\circ}\text{C}$ . The LHP with fourth wick structure is able to cool the IGBT under temperature of  $100^{\circ}\text{C}$  until the IGBT input power of 250 W.

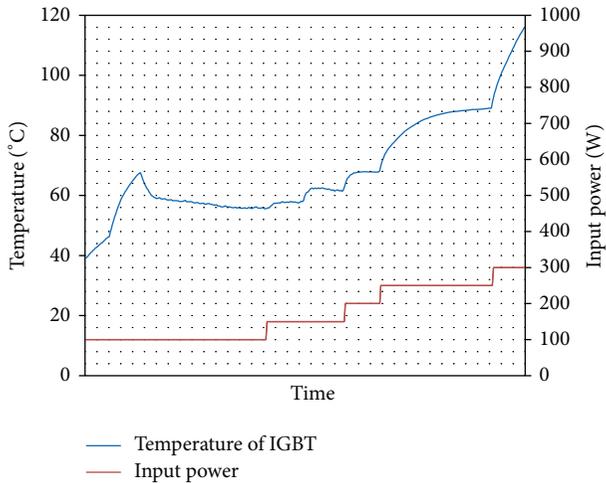


FIGURE 8: Dependence of temperature on input power of IGBT cooled by LHP with first variant of wick structure.

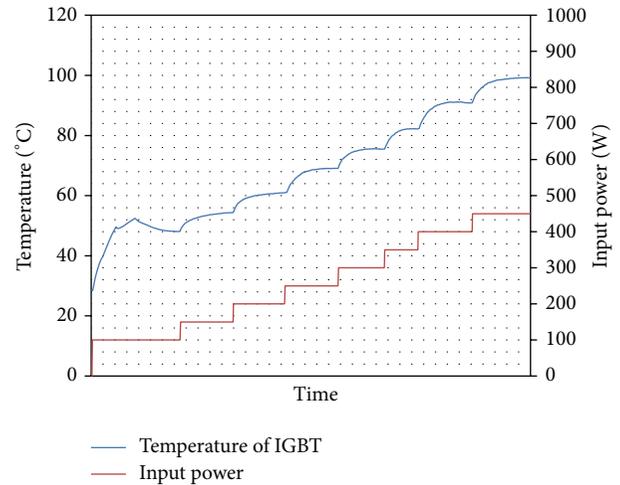


FIGURE 9: Dependence of temperature on input power of IGBT cooled by LHP with second variant of wick structure.

**6. Discussion**

This experiment was performed in framing scientific research of porous structures suitable for LHP and finding ability of heat removal produced by IGBT. We lead off previous research works about LHP in which materials specification suitable for porous structure was preferred. We choose copper and nickel powder with two various granularities. At first etalons were manufactured from each material sintered at various temperatures and times. It was observed that temperature is the main influencing factor on wick structure porosity and pore size is depending on powder grain size. After them for each material one wick structure was manufactured with best characteristics of porosity and pore size and used in LHP for IGBT cooling. The knowledge gained from the IGBT cooling by LHP has given us the information necessary to know how much heat flux is LHP able to remove from heat source. This piece of information will be in the future useful in the design of cooling devices working with the LHP. In the future we would like to focus deeper on analysis of the physical characteristics (e.g., thermal conduction and capillary pressure) of manufactured wick structures and on research of construction design LHP able to remove heat by natural convection to the surroundings.

**7. Conclusion**

According to microscopic analysis of sintered structures, which clarifies their shape and profile, we can conclude that the main influencing factors of pore size are grain size, sintering temperature, and not so much sintering time. The measurement comparison of dependency IGBT temperature from input power cooled by LHP with copper or nickel wick structure can conclude, however, that in both cases the structures had the same porosity and better effect on heat removal from IGBT that had porous structure with bigger pore size. Generally the smallest pore size could cause the

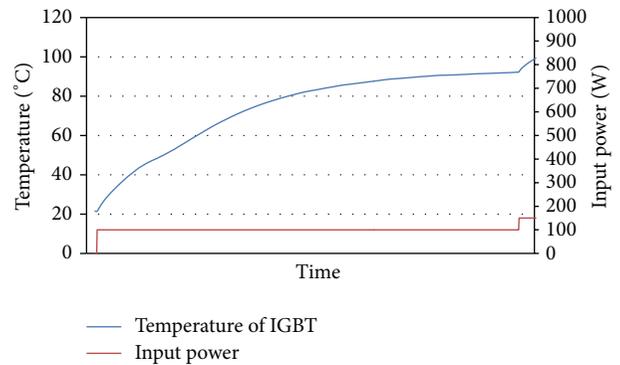


FIGURE 10: Dependence of temperature on input power of IGBT cooled by LHP with third variant of wick structure.

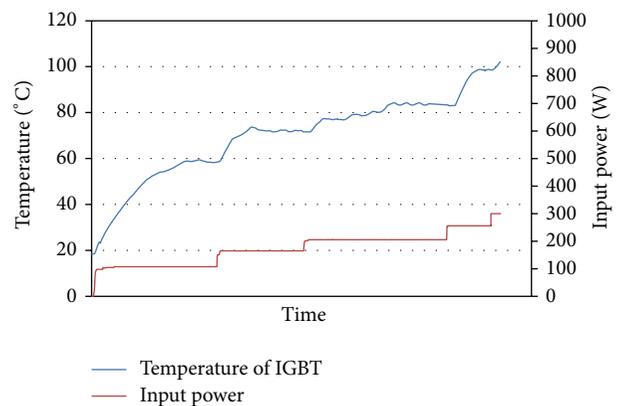


FIGURE 11: Dependence of temperature on input power of IGBT cooled by LHP with fourth variant of wick structure.

low capillary pressure in sintered wick structures against total pressure in whole LHP system.

### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

### Acknowledgments

This paper was created within the solution of project APVV-0577-10 and sponsored in the frame of program OPV-Podpora kvality vzdelávania a rozvoj ľudských zdrojov v oblasti technického výskumu a vývoja v priestore modernej vedomostnej spoločnosti ITMS 26110230117.

### References

- [1] T. Kaya and J. Ku, "Thermal operational characteristics of a small-loop heat pipe," *Journal of Thermophysics and Heat Transfer*, vol. 17, no. 4, pp. 464–470, 2003.
- [2] K. H. Cheung, T. T. Hoang, J. Ku, and T. Kaya, "Thermal performance and operational characteristics of loop heat pipe (NRL LHP)," SAE Technical Paper no. 981813, 1998.
- [3] Y. F. Maydanik, "Loop heat pipes," *Applied Thermal Engineering*, vol. 25, no. 5-6, pp. 635–657, 2005.
- [4] J. Ku, "Operating characteristics of loop heat pipes," in *Proceedings of the 29th International Conference on Environmental System*, Denver, Colo, USA, July 1999.
- [5] R. R. Williams and D. K. Harris, "Cross-plane and in plane porous properties measurements of thin metal felts: applications in heat pipes," *Experimental Thermal and Fluid Science*, vol. 27, no. 3, pp. 227–235, 2003.
- [6] M. Bonnefoy, J. M. Ochterbeck, B. L. Drolen, and M. N. Nikitkin, "Effective thermal conductivity of saturated sintered Nickel loop heat pipe wicks," in *Proceedings of the 37th Thermophysics Conference, American Institute of Aeronautics and Astronautics*, pp. 1–10, Portland, Ore, USA, June 2004.
- [7] B. Holley and A. Faghri, "Permeability and effective pore radius measurements for heat pipe and fuel cell applications," *Applied Thermal Engineering*, vol. 26, no. 4, pp. 448–462, 2006.
- [8] S. W. Chi, *Heat Pipe Theory and Practice: A Sourcebook*, Hemisphere, New York, NY, USA, 1976.
- [9] M. Čarnogurská, M. Příhoda, T. Brestovič, J. Molínek, and R. Pyszko, "Determination of permeability and inertial resistance coefficient of filter inserts used in the cleaning of natural gas," *Journal of Mechanical Science and Technology*, vol. 26, no. 1, pp. 103–111, 2012.
- [10] C. Ren and Q. Wu, "Heat transfer in loop heat pipes capillary wick: effect effective thermal conductivity," *Journal of Thermophysics and Heat Transfer*, vol. 21, no. 1, pp. 134–140, 2007.
- [11] D. Khrustalev and A. Faghri, "Heat transfer in the inverted meniscus type evaporator at high heat fluxes," *International Journal of Heat and Mass Transfer*, vol. 38, no. 16, pp. 3091–3101, 1995.
- [12] T. S. Zhao and Q. Liao, "On capillary-driven flow and phase-change heat transfer in a porous structure heated by a finned surface: measurements and modeling," *International Journal of Heat and Mass Transfer*, vol. 43, no. 7, pp. 1141–1155, 2000.
- [13] B. D. Iverson, T. W. Davis, S. V. Garimella, M. T. North, and S. S. Kang, "Heat and mass transport in heat pipe wick structures," *Journal of Thermophysics and Heat Transfer*, vol. 21, no. 2, pp. 392–404, 2007.
- [14] S. Launay, V. Sartre, and J. Bonjour, "Parametric analysis of loop heat pipe operation: a literature review," *International Journal of Thermal Sciences*, vol. 46, no. 7, pp. 621–636, 2007.
- [15] X. Gongming, C. Kehang, Z. Yong, and C. Lin, "Reduction of effective thermal conductivity for sintered LHP wicks," *International Journal of Heat and Mass Transfer*, vol. 53, no. 13-14, pp. 2932–2934, 2010.
- [16] M. Furukawa, "Model-based method of theoretical design analysis of a loop heat pipe," *Journal of Thermophysics and Heat Transfer*, vol. 20, no. 1, pp. 111–121, 2006.
- [17] P. Y. A. Chuang, *An improved steady-state model of loop heat pipes based on experimental and theoretical analyses [Ph.D. thesis]*, The Pennsylvania State University, State College, Pa, USA, 2003.
- [18] J. Li, Y. Zou, L. Cheng, R. Singh, and A. Akbarzadeh, "Effect of fabricating parameters on properties of sintered porous wicks for loop heat pipe," *Powder Technology*, vol. 204, no. 2-3, pp. 241–248, 2010.
- [19] D. Reay and P. Kew, *Heat Pipes*, Butterworth-Heinemann, Boston, Mass, USA, 5th edition, 2006.
- [20] V. Pastukhov, Y. F. Maidanik, C. Vershinin, and M. A. Korukov, "Miniature loop heat pipes for electronics cooling," *Applied Thermal Engineering*, vol. 23, no. 9, pp. 1125–1135, 2003.
- [21] E. G. Reimbrechta, M. C. Fredel, E. Bazzo, and F. M. Pereira, "Manufacturing and microstructural characterization of sintered nickel wicks for capillary pumps," *Material Research*, vol. 2, no. 3, pp. 225–229, 1999.
- [22] G. Xin, K. Cui, Y. Zou, and L. Cheng, "Development of sintered Ni-Cu wicks for loop heat pipes," *Science in China Series E: Technological Sciences*, vol. 52, no. 6, pp. 1607–1612, 2009.
- [23] X. Huang and G. Franchi, "Design and fabrication of hybrid bi-modal wick structure for heat pipe application," *Journal of Porous Materials*, vol. 15, no. 6, pp. 635–642, 2008.
- [24] S. K. Samanta, B. B. Sharma, P. Das, and A. K. Lohar, "Development of tubular Ni wicks used in LHP for space applications," *Frontiers in Heat Pipes*, vol. 2, Article ID 043004, 2011.
- [25] N. J. Gernert, G. J. Baldassarre, and J. M. Gottschlich, "Fine pore loop heat pipe wick structure development," SAE Technical Paper no. 961319, 1996.
- [26] S. C. Wu, C. J. Huang, W. H. Yang, J. C. Chang, and C. C. Kung, "Effect of sintering temperature curve in wick manufactured for loop heat pipe," *World Academy of Science, Engineering and Technology*, vol. 62, 2012.
- [27] L. Orman, "Possibility of the application of microstructures in heating and ventilation systems," *Structure and Environment*, vol. 2, no. 1, pp. 41–45, 2010.
- [28] P. Nemeč, M. Malcho, M. Smitka, and J. Matušov, "Performance parameters of closed loop thermosyphon," in *Communications: Scientific Letters of the University of Žilina*, vol. 14, pp. 53–57, University of Žilina, 2012.
- [29] R. Lenhard and J. Jandačka, "Two-phase modeling of interphase heat transport from the condensation to evaporation part of the heat-pipe," in *Proceedings of the 11th International Conference of Numerical Analysis and Applied Mathematics*, vol. 1558 of AIP Proceedings, pp. 2138–2141, Rhodes, Greece, 2013.
- [30] J. Jandačka, M. Holubčík, Š. Papučík, and R. Nosek, "Combustion of pellets from wheat straw," *Acta Montanistica Slovaca*, vol. 17, no. 4, pp. 283–289, 2012.

- [31] K. Kaduchová, R. Lenhard, and J. Jandačka, "Optimization of heat exchanger for indirectly heated water heater," *EPJ Web of Conferences*, vol. 25, Article ID 01036, 10 pages, 2012.

## Research Article

# Structural Safety Analysis Based on Seismic Service Conditions for Butterfly Valves in a Nuclear Power Plant

Sang-Uk Han, Dae-Gyun Ahn, Myeong-Gon Lee, Kwon-Hee Lee, and Seung-Ho Han

Department of Mechanical Engineering, Dong-A University, Busan 604-714, Republic of Korea

Correspondence should be addressed to Seung-Ho Han; shhan85@dau.ac.kr

Received 26 February 2014; Accepted 9 March 2014; Published 11 May 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Sang-Uk Han et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The structural integrity of valves that are used to control cooling waters in the primary coolant loop that prevents boiling within the reactor in a nuclear power plant must be capable of withstanding earthquakes or other dangerous situations. In this study, numerical analyses using a finite element method, that is, static and dynamic analyses according to the rigid or flexible characteristics of the dynamic properties of a 200A butterfly valve, were performed according to the KEPIC MFA. An experimental vibration test was also carried out in order to verify the results from the modal analysis, in which a validated finite element model was obtained via a model-updating method that considers changes in the *in situ* experimental data. By using a validated finite element model, the equivalent static load under SSE conditions stipulated by the KEPIC MFA gave a stress of 135 MPa that occurred at the connections of the stem and body. A larger stress of 183 MPa was induced when we used a CQC method with a design response spectrum that uses 2% damping ratio. These values were lower than the allowable strength of the materials used for manufacturing the butterfly valve, and, therefore, its structural safety met the KEPIC MFA requirements.

## 1. Introduction

The availability factor of nuclear power plants has been significantly improved worldwide, whereby nuclear power is becoming more economically competitive with fossil fuels for base-load electricity generation in many countries [1]. Nuclear power accounted for nearly 35% of domestic production electricity in 2012 and is gradually increasing in dependency [2]. Structural integrity includes valves that can withstand earthquakes and other dangerous situations, because they are used to control cooling waters in the primary coolant loop to prevent boiling within the reactor in a nuclear power plant. An accident due to the leakage of radioactive matters, however, can inflict catastrophic damage on the environment nearby. Therefore, with the enhanced awareness of the potential of an earthquake to cause such damage, qualifying the valves has now become standard practice, that is, establishing their ability to withstand a seismic load without damage. Strict safety guidelines should be carried out as defined by the KEPIC MFA [3], which indicates the verification of seismic adequacy with prescribed safety rates for structures and equipment. The verification of the seismic

adequacy consists of environmental qualifications for handling the effects of heat- and radiation-induced degradation and of seismic qualifications that should be carried out either by numerical analysis or by experimental tests using a shake-table or by comparison with past experiences. The seismic qualification for the use of shake-table testing is normally very exorbitant and these facilities may not be available in many places and can show only single isolated structures or equipment without simulating structural connections to the secondary component, which may change the dynamic behavior when compared with the actual as-installed structures [4, 5]. Another venerable method for seismic qualification is purely analytical and uses a finite element method. The reliability of this method totally depends on the finite element model, which generally cannot produce the dynamic behavior of as-installed structures even for structurally simple components. Thus, finite element models are usually verified via an experimental modal test. The test results are assumed to be correct and the finite element model is tuned to closely correlate with the experimental test results. *In situ* modal test data obtained from the modal tests conducted on the as-installed structure were used directly for the seismic response

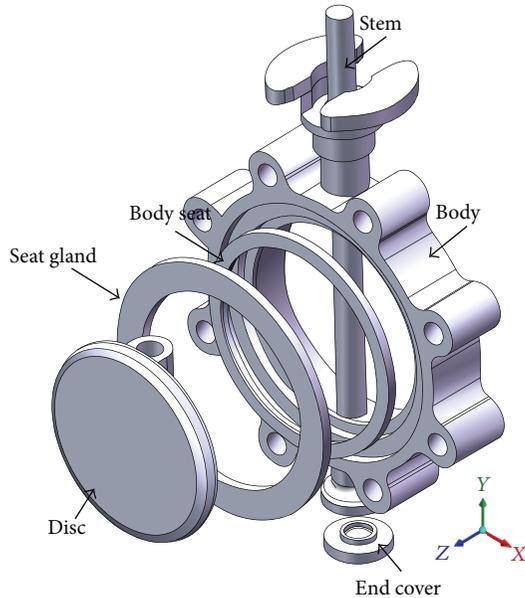


FIGURE 1: Configuration of a 200A butterfly valve.

estimation in order to overcome the limitations when using a finite element model for seismic analysis [6, 7]. However, these methods may not always be practical for many of the structural components in a nuclear power plant due to the difficulties of conducting *in situ* modal tests.

In the present study, the numerical analyses using finite element methods, that is, static and dynamic analyses according to the rigid or flexible characteristics of dynamic properties for a 200A butterfly valve in a nuclear power plant, were performed according to the dictates of the KEPIC MFA [3]. An experimental vibration test was carried out in order to verify the results from the modal analysis, whereby a validated finite element model was obtained via model updating that considered the changes in *in situ* experimental data. By using the validated finite element model, structural safety analysis under seismic service conditions was carried out.

## 2. Seismic Qualification

**2.1. Butterfly Valve.** Figure 1 shows the configuration of the butterfly valve, which is used in a nuclear power plant to control cooling waters in the primary coolant loop preventing boiling within the reactor. Since a radiation leak would have a disastrous effect on the environment, the butterfly valve must not leak and must endure earthquakes and other dangerous situations. The butterfly valve shown in Figure 1 has a 200 mm inner diameter and consists of 6 parts: stem, body, body seat, seat gland, disc, and end cover. These parts are made up of the following materials: Gr. WCB, CF8, T316, T316, A564-630, and T304, respectively.

**2.2. Seismic Qualification Process.** The seismic qualification of the butterfly valve should demonstrate that the valve has the ability to perform its safety functions during and

after being subjected to the forces resulting from a safe shutdown earthquake (SSE) event. The KEPIC MFA [3] describes the approach methods for seismic qualification, which are grouped into 4 general categories: to predict the equipment's performance by numerical analysis, to test the equipment under simulated seismic conditions, to qualify the equipment by a combination of experimental tests and numerical analyses, and to qualify the equipment through the use of experience data. Each of the categories may be adequate to verify the ability of the equipment to meet the seismic qualification requirements. Among these 4 categories, this study focused on predicting the equipment's performance via numerical analysis.

The methods used for numerical analysis include static and dynamic analysis depending on the structure of the equipment and dynamic properties such as the complexity of the equipment and whether the equipment is rigid or flexible [8]. Figure 2 shows the flow chart of a performance assessment based on a seismic qualification using numerical analysis according to the KEPIC MFA [3]. The review stage in the first step takes into account the complexity of the butterfly valve and the adequacy of analytical techniques to properly predict its safe operation during seismic excitation. The butterfly valve should be modeled such that its mass distribution and stiffness characteristics will be adequately represented when using the finite element method. This finite element model can be used to perform a modal analysis in order to determine the rigidity or flexibility. When the natural frequency at the 1st mode calculated by the modal analysis is higher than a cut-off frequency of 33 Hz, that is, the dominant frequency of an earthquake, the butterfly valve is considered sufficiently rigid and may be analyzed statically. But if it is not higher than 33 Hz, a dynamic analysis should be performed because the butterfly valve is considered to be flexible and poses a risk of resonance in the dominant frequency range of an earthquake. In this step, the butterfly valve can be analyzed via response spectrum analysis, in which the responses of stress obtained from each modal response are combined to consider all significant modes. Finally, the structural safety is estimated by comparing the combined stress with the allowable stress of the materials in use.

**2.2.1. Static Analysis.** Static analysis, also known as equivalent static force analysis, is a method that enables calculation of the stresses in each part of the structure to be recreated by a static force, that is, the equivalent of an earthquake. Although static analysis often underestimates structural safety by comparison with dynamic analysis, this simple process prevents the need to perform time-consuming computations. The acceleration responses that are required in order to estimate structural safety during analysis must be determined only according to the maximum peak of the response spectrum using a conservative damping value. The seismic qualification for unit of equipment devices or structural systems can be achieved only by static analysis, because the effect of resonance does not have to be considered. The governing equation of static analysis can be presented as follows:

$$[K] \{U\} = \{F\}, \quad (1)$$

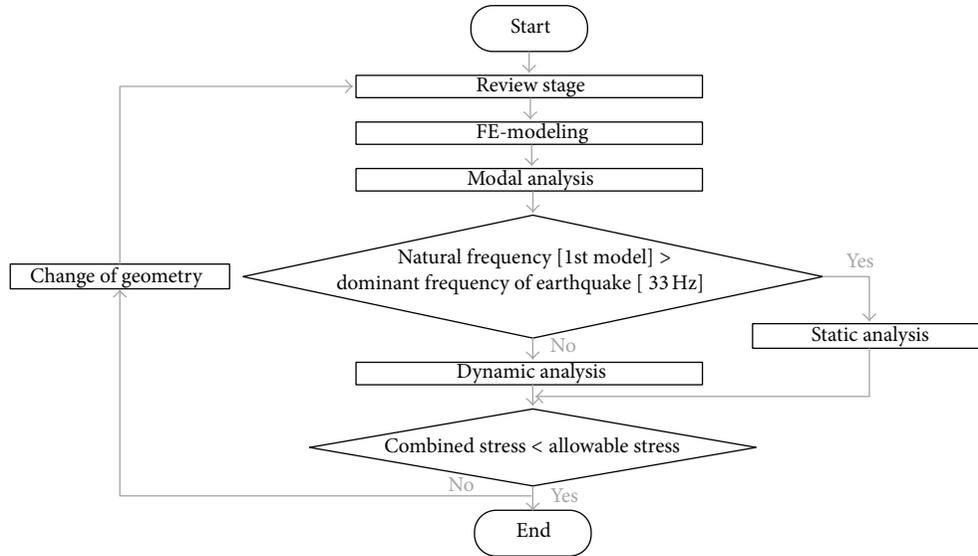


FIGURE 2: Flow chart of performance assessment using numerical analysis.

where  $[K]$ ,  $\{U\}$ , and  $\{F\}$  are designated as the stiffness matrix, the nodal displacement vector, and the external force vector, respectively, caused by the dead load or gross weight. The nodal displacement vector  $\{U\}$  was calculated using the finite element method, which then allowed for the stress distribution at nodal points.

**2.2.2. Dynamic Analysis.** For flexible equipment, where the natural frequency is lower than the dominant frequency of an earthquake, that is, 33 Hz, as shown in Figure 1, dynamic analysis should be carried out based on either the response spectrum method or the time history method. The response spectrum method based on structural dynamics enables the approximate estimation of the dynamic performances of the equipment such as the maximum responses of displacement and stress. The dynamic performances are determined by combining each modal response, which includes all significant modes. This method is commonly used for the dynamic analyses of seismic qualifications. Meanwhile, the time history method can be used to evaluate the time history of dynamic responses due to an earthquake, which displays earthquake-induced motion as a function of time, usually in terms of acceleration. Although the time history method provides relatively accurate dynamic responses, a time-consuming computational work and a complicated procedure are required due to the consideration of a large number of degrees of freedom and to the detailed data from earthquake-induced motion [9, 10]. In the present study, dynamic analysis was conducted using the response spectrum method. The dynamic performances of the equipment can be presented as follows:

$$[M] \{\ddot{U}\} + [C] \{\dot{U}\} + [K] \{U\} = [P(t)], \quad (2)$$

where  $[M]$ ,  $[C]$ ,  $\{\ddot{U}\}$ ,  $\{\dot{U}\}$ , and  $[P(t)]$  are the mass matrix, the damping matrix, the nodal acceleration vector, the nodal velocity vector, and the applied dynamic load vector,

respectively. In the analysis procedure, the nodal velocity vector of the finite element analysis is obtained first, and then the nodal displacement and stress can be calculated. The modal analysis must be carried out before application of the response spectrum analysis, because  $[P(t)]$  is designated as the set of load values induced from the modal responses. After the modal analysis, a response spectrum analysis should be carried out, whereby the responses of stress obtained from each modal response combine all significant modes. There are two rational ways to combine responses from the response spectrum: the SRSS (square root of sum of square) method and the CQC (complete quadratic combination) method [11, 12]. In the SRSS method, the squares of a specific response are summed, and the square root of this sum takes the combined effect into account. The SRSS provides relatively conservative results, except where closely spaced modes apply. In the case of closely spaced modes, the combined response values are often underestimated. Meanwhile, the CQC method combines responses based on the use of cross-modal coefficients, which reflect the duration and frequency content of the seismic events as well as the modal frequencies and damping ratio of the equipment. The present study applied the CQC method. The total mode response,  $R_a$ , obtained by the CQC method can be written as follows:

$$R_a = \left[ \sum_{i=1}^N \sum_{j=1}^N k \varepsilon_{ij} R_i R_j \right]^{1/2}, \quad (3)$$

where  $k$  is 1 when  $i = j$  is valid and is 2 when  $i = j$  is invalid;  $R_i$  and  $R_j$  represent the mode responses at the  $i$ th and  $j$ th modes, respectively.  $\varepsilon_{ij}$  designates the cross-modal coefficient presenting the correlation between the  $i$ th and  $j$ th modes, which can be presented as follows:

$$\varepsilon_{ij} = \frac{8(\xi_i \xi_j)^{1/2} (\xi_i + r \xi_j) r^{3/2}}{(1 - r^2)^2 + 4\xi_i \xi_j r (1 + r^2) + 4(\xi_i^2 + \xi_j^2) r^2}, \quad (4)$$

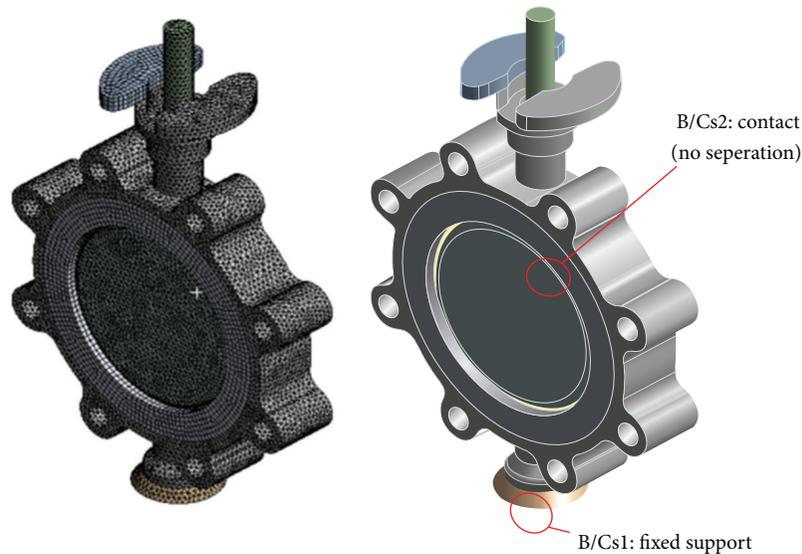


FIGURE 3: Finite element model and boundary conditions for modal analysis based on static analysis.

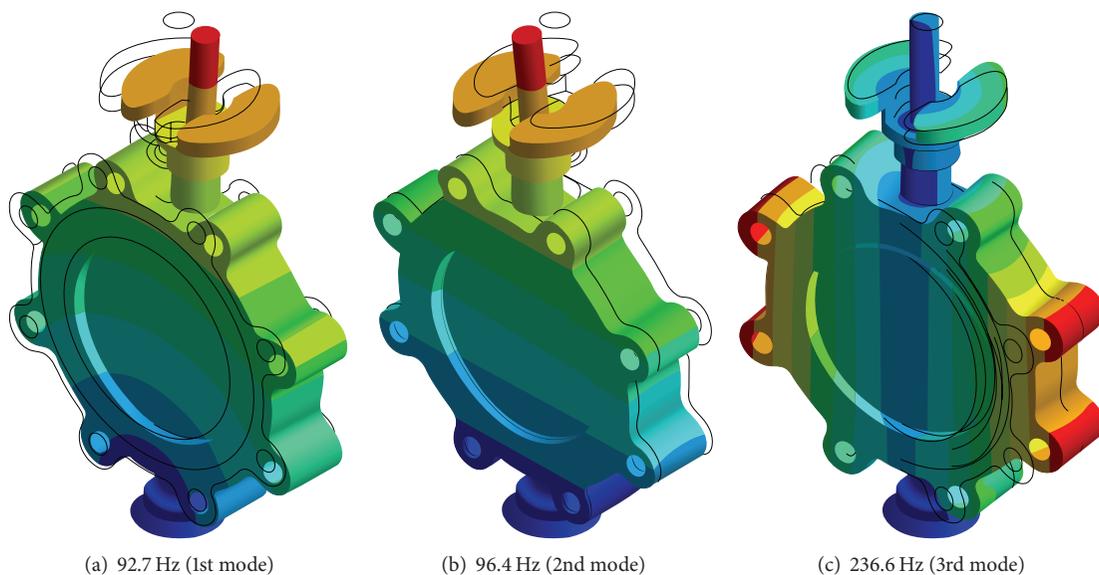


FIGURE 4: Frequencies at each mode obtained by modal analysis.

where  $r$  and  $\xi$  are the ratio of natural frequencies and modal damping, respectively.

### 3. Modal Analysis

**3.1. Modal Analysis Using the Finite Element Method.** The modal analysis for the butterfly valve was performed using the finite element method via the commercial software, ANSYS Workbench [13]. Figure 3 shows the finite element model constructed using the preprocessor option provided in ANSYS Workbench with the boundary conditions applied to the modal analysis. The contact conditions between each part of the finite element model were implemented using special elements, such as CONTA 174 and TARGE 170, and the no-separation contact boundary condition, where a sliding, but

not nonlinear, motion is permitted at the contact surfaces between the disc and body sheet. Furthermore, to investigate the effect of the contact conditions on the frequencies occurring at each mode, other boundary conditions show that the contact surfaces are bonded. Figure 4 shows the natural frequencies obtained by the modal analysis under the no-separation contact boundary conditions, that is, 92.7, 96.4, and 236.6 Hz at the 1st, 2nd, and 3rd modes, respectively. In the case of the bonded-surface boundary conditions, the modal analysis provided natural frequencies of 94.4, 96.3, and 240.4 Hz for each mode, which was a less than 2% error compared with that of the no-separation contact boundary conditions. These results showed that the natural frequencies of the butterfly valve were higher than 33 Hz, which means that the butterfly valve can be considered sufficiently rigid

TABLE 1: Comparison of natural frequencies obtained using initial and modified FE-models.

	1st mode	2nd mode	3rd mode
Natural frequency obtained from the experimental modal test (Hz)	69.3	72.2	219.7
Natural frequency obtained by using the initial FE-model (Hz)	92.7 (25%)	96.4 (25%)	236.6 (7.7%)
Natural frequency obtained by using the modified FE-model (Hz)	68.9 (0.5%)	72.3 (0.1%)	226.6 (3%)

() means % errors (abs.) to natural frequencies obtained from experimental modal test.

and could be analyzed statically according to the process of seismic qualification, as shown in Figure 2.

**3.2. Modification of the Finite Element Model.** The reliability of the modal analysis totally depends on the finite element model. If the mass distribution and boundary conditions of the finite element model cannot be considered equivalent to the as-installed conditions, the dynamic behavior obtained from the finite element model shows a significant difference compared with the as-installed version. Therefore, the finite element model should be verified using the data from the experimental modal test, in which the test results are assumed to be correct, and the finite element model is tuned to closely correlate with the test results. In the present study, we performed a model-updating method that considered changes in the *in situ* experimental modal test data, and a validated finite element model was obtained.

An experimental modal test of the butterfly valve was performed. The end cover of the butterfly valve was welded to a steel plate on the reaction floor. The butterfly valve was instrumented with 5 accelerometers from the B&K Co. [14] with a capacity of 3,000 G, as shown in Figure 5. The location of the accelerometers was selected according to the numerical results from the modal analysis, where 4 accelerometers measured the acceleration at each mode, and the other accelerometer was attached to the steel plate on the reaction floor to compensate for the relative movement of the butterfly valve. Data acquisition was accomplished using NEXUX software from the B&K Co. Singular values of acceleration data occurred by the stroke of an impact hammer were converted into natural frequencies at each mode. Figure 6 shows natural frequencies of 69.3, 72.2, and 219.7 Hz at the 1st, 2nd, and 3rd modes, respectively, and these are listed in Table 1. As the table shows, the experimental values deviated significantly more than 25% from the computed frequencies using the initial finite element model. This might have been caused by a poor reflection of the initial finite element model for the butterfly valve such as simplified and idealized assumptions made while constructing the finite element model. For the

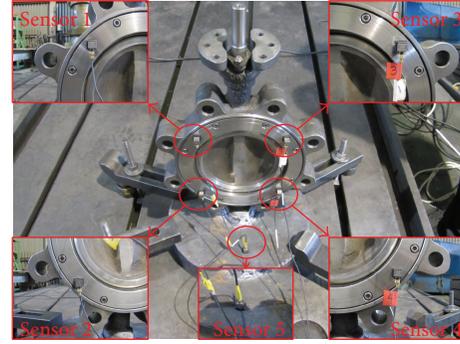


FIGURE 5: Accelerometers attached on the butterfly valve for the experimental modal test.

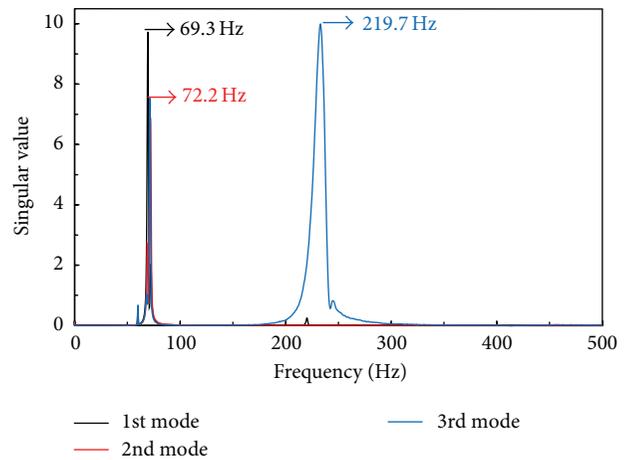


FIGURE 6: Singular values related to the frequencies for each mode.

numerical modal analysis, the butterfly valve was assumed to be rigidly fixed at the steel plate. However, during the experimental modal test, the butterfly valve had either a small degree of rotation or a small deflection at the fixed point. The initial finite element model had to be modified so that it would project confidence for further analysis. To modify the initial finite element model, a trial and error method was used, which is a common part of the model-updating method [15]. Based on the physical understanding of the installation of the butterfly valve, the boundary stiffness at the rigidly fixed steel plate was adjusted, which brought the finite element model predictions close to the experimental results. As shown in Table 1, the computed frequencies from the updated finite element model are almost identical to the measured ones.

## 4. Results of the Structural Safety Analysis

**4.1. Static Analysis.** The structural safety analysis of the butterfly valve was carried out by the seismic qualification based on the static analysis using the validated finite element model, in which the butterfly valve can be considered rigid due to

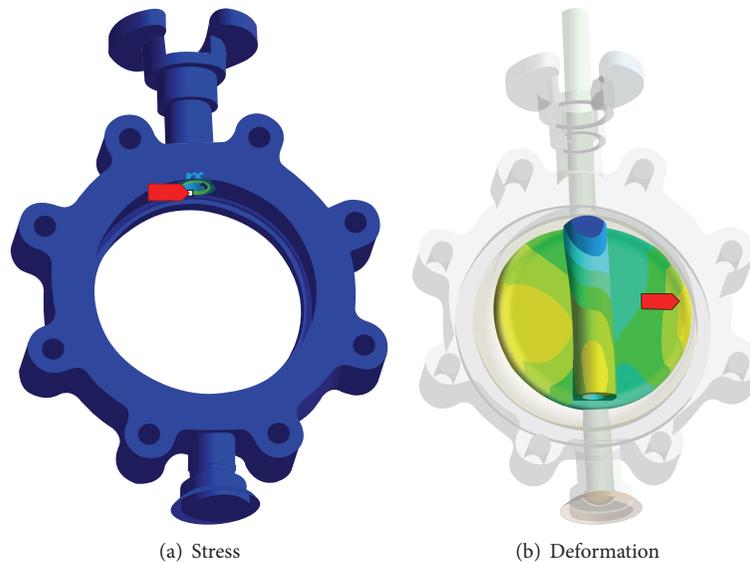


FIGURE 7: Results of structural analysis under reverse pressure according to Grade D of the KEPIC MFA.

TABLE 2: Acceleration values of the SSE load according to the KEPIC MFA.

Horizontal ( $x$ -dir.)	Horizontal ( $y$ -dir.)	Vertical ( $z$ -dir.)
4.5 G	4.5 G	3.0 G

all natural frequencies of higher than 33 Hz, as calculated by the modal analysis. The stresses applied in the butterfly valve were calculated under a combined load, that is, an equivalent static force, defined as Grade D by the KEPIC MFA. In the case of Grade D, the combined load, which was subjected at the center of gravity in the butterfly valve, accounted for dead weight, operation load, and SSE load. The operation load occurred under normal and reverse pressures with respect to flow direction in the butterfly valve, and the SSE load was obtained by acceleration values in 3-dimensional directions, as defined in the KEPIC MFA, as shown in Table 2. Figure 7 shows a contour plot of the applied stresses and deformations, and the results are listed in Table 3. A maximal stress of 135 MPa occurred at the contact area between the top side of the stem and the body under a load combination of the dead weight, the operation load under reverse pressure, and the SSE load. The safety factor that described the structural capacity of the butterfly valve was 1.7 in consideration of the allowable stress of the material that was used for the body, Gr. WCB, with a yield strength of 235 MPa.

#### 4.2. Dynamic Analysis

**4.2.1. The Procedure for Response Spectrum Analysis.** If the butterfly valve is installed in a pipeline system, its natural frequency could be lower than 33 Hz, and a resonance failure might be expected. In this case, dynamic analysis should be

TABLE 3: Results of structural analysis based on static analysis for Grade D of the KEPIC MFA.

	Max. stress (MPa)	Safety factor
Normal pressure	57	4.1
Reverse pressure	135	1.7

performed, so that the pipeline system, including the butterfly valve, could show a flexible dynamic motion. The procedure for response spectrum analysis is provided in the KEPIC END [16]. This procedure is based on dynamic analysis, as mentioned in Section 2.2.2. Figure 8 shows a flow chart for the response spectrum analysis.

**4.2.2. Characteristics of Dynamic Behavior.** For the response spectrum analysis, the natural frequencies of the butterfly valve were obtained by the modal analysis for a full-scale model of a pipeline system, which included this valve. The present study used a simple method to simulate the dynamic behaviors of the full-scale model. This was implemented by modification of the boundary conditions, in which all of the constraints at the valve body were released except for a displacement in the  $z$ -direction. Figure 9 shows a schematic diagram of the boundary conditions for the modal analysis that was based on the dynamic analysis. Figure 10 shows the natural frequencies at each mode. The 1st and 2nd modes did not occur due to its free-body motion. Meanwhile, the natural frequencies at the 3rd, 4th, and 5th modes were calculated as 24.4, 47.9, and 250 Hz, respectively. The natural frequency at the 3rd mode was lower than 33 Hz, so that the pipeline system including the butterfly valve would be considered flexible and could be affected

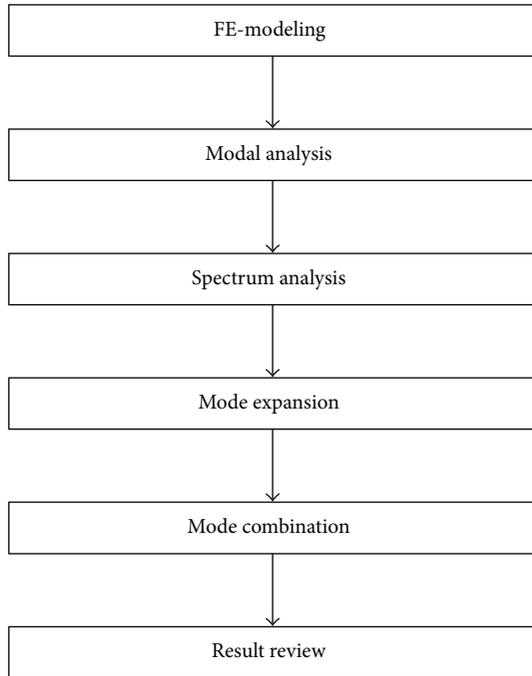


FIGURE 8: Flow chart for response spectrum analysis based on dynamic analysis.

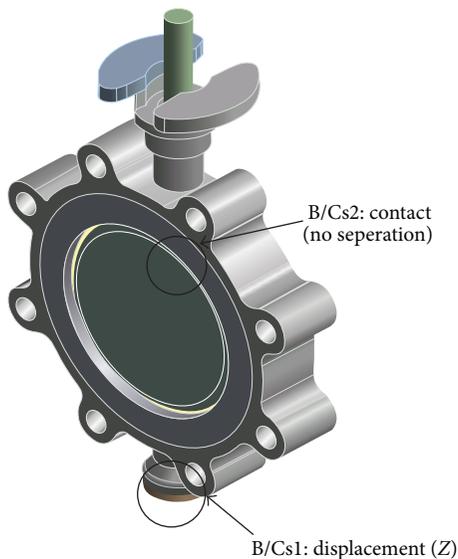


FIGURE 9: Boundary conditions for modal analysis based on dynamic analysis.

by the resonance in the dominant frequency range of an earthquake.

**4.2.3. Structural Safety Assessment by Response Spectrum Analysis.** For the response spectrum analysis, the responses of stress obtained from each modal response at all significant

modes were combined according to the CQC method based on the use of cross-modal coefficients, as shown in (3). The structural analysis that was used to estimate the responses of stress took into account all three load conditions suggested in the KEPIC MFA, that is, RRS (required response spectrum), DRS (design response spectrum), and RIM (required input motion). In general, the DRS is recommended if the RRS and the RIM are not available. In the present study, the load condition regulated in the KEPIC MFA, as shown in Figure 11, was granted to the DRS where a damping ratio of 2% in the horizontal and vertical directions was taken into account in a case where the valves would be applied to a nuclear power plant. Figure 12 shows the results of the structural analysis, in which the maximal stress of 183 MPa occurred at the contact area between the bottom layer of the stem and the body. The characteristics of dynamic behavior were similar to that for the 3rd mode of the modal analysis, as shown in Figure 10. The structural safety factor was 1.3. Although this result is lower than that in the case of static analysis, the structural safety of the butterfly valve met the requirements of the KEPIC MFA.

## 5. Conclusions

In this study, the structural safety analysis of a 200A butterfly valve for use in a nuclear power plant was performed in static and dynamic ways according to the KEPIC MFA. The results are as follows.

- (1) Analytical and experimental modal tests were carried out, and their deviations were taken into account. The initial finite element model was modified to decrease the error range to less than 3%.
- (2) The static analysis provided a maximal stress of 135 MPa at the contact area between the topside of the stem and the body under a load combination of the dead weight, the operation load under reverse pressure, and the SSE load. The safety factor for the structural capacity of the butterfly valve was 1.7.
- (3) In the case of dynamic analysis, the maximal stress was 183 MPa, and the characteristics of dynamic behavior were similar to those for the 3rd mode of the modal analysis. The structural safety factor was 1.3. These values were under the allowable strength for the materials used in the manufacture of the butterfly valve, and, therefore, its structural safety met the requirements of the KEPIC MFA.

The presented findings could be applicable as an index of the structural safety of the butterfly valve based on the seismic qualification in a nuclear power plant. To verify the fatigue requirements according to ASME Section III, the frequency of occurrence based on operating histories should be taken into account, and the fatigue analysis is currently being carried out. These results will be presented elsewhere in the near future.

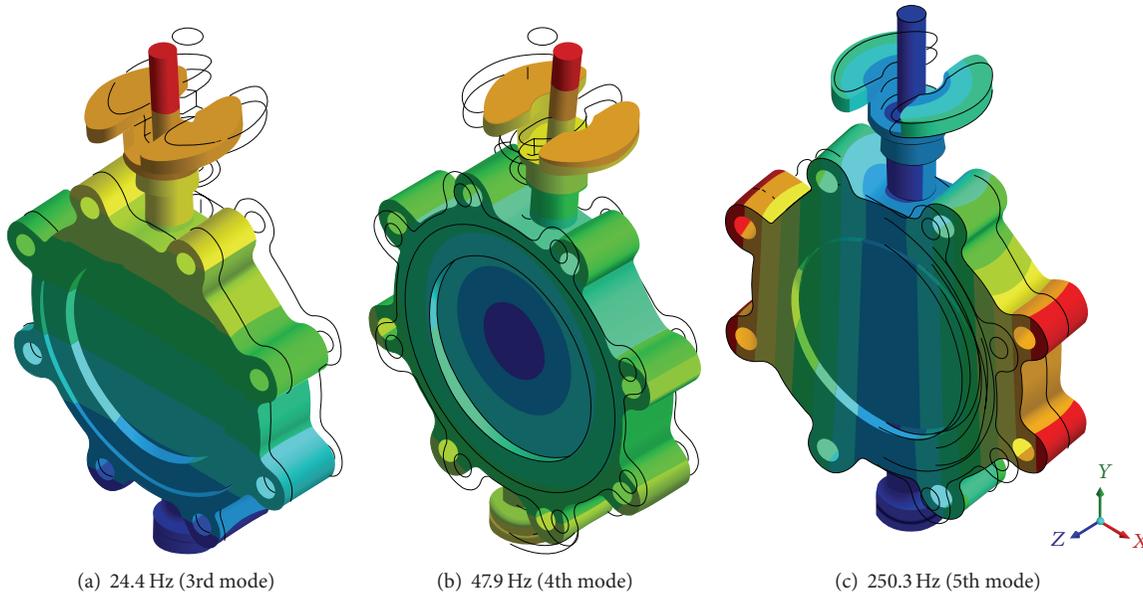


FIGURE 10: Frequencies at each mode obtained by modal analysis based on dynamic analysis.

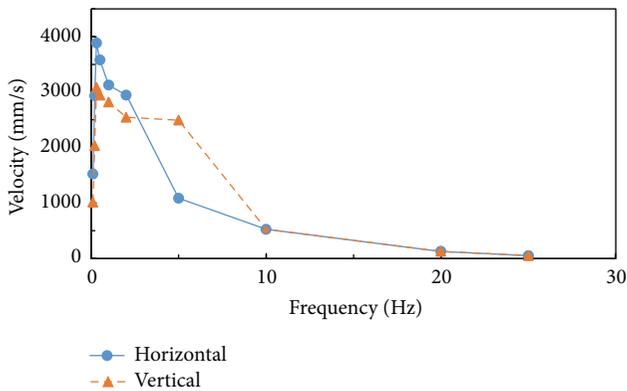


FIGURE 11: Design response spectrum considering a 2% damping ratio in the horizontal and vertical directions, as stipulated in the KEPIC MFA.

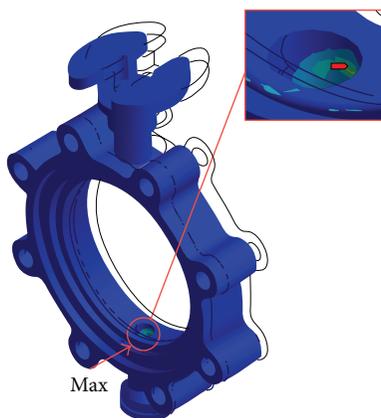


FIGURE 12: Results of structural analysis obtained by dynamic analysis.

**Conflict of Interests**

The authors declare that there is no conflict of interests regarding the publication of this paper.

**Acknowledgments**

This work was supported by the Human Resources Program (No. 20134030200320) of the Korea Institute of Energy Technology Evaluation and Planning (KETEP), granted financial resource from the Ministry of Trade, Industry & Energy, Republic of Korea.

**References**

- [1] International Atomic Energy Agency, *Seismic Design and Qualification for Nuclear Power Plant Safety*, Safety Series no. 50-SG-D15, IAEA, 1992.
- [2] C. S. Kim and S. H. Lee, “The economic effects of nuclear power generation using CGE model,” *Korean Energy Economic Review*, vol. 9, no. 2, pp. 129–152, 2010.
- [3] M-16 KEPIC MFA, *Qualification of Mechanical Equipment—General Requirements*, Korea Electric Power Industry Code, 2010.
- [4] J. K. Sinha, A. R. Rao, and R. K. Sinha, “Realistic seismic qualification using the updated finite element model for in-core components of reactors,” in *Proceedings of the International Symposium on Seismic Evaluation of Nuclear Facilities*, IAEA-CN-106-29, Vienna, Austria, August 2003.
- [5] B. H. Cho, H. H. Jung, S. J. Cho, and C. S. Kim, “The analytical estimation of structural integrity for motor driven auxiliary feedwater pump in nuclear power plants,” in *Proceedings of the 9th International Conference on Fracture & Strength of Solids*, Jeju, Korea, June 2013.
- [6] J. K. Sinha and R. I. K. Moorthy, “Combined experimental and analytical method for a realistic seismic qualification of

- equipment,” *Nuclear Engineering and Design*, vol. 195, no. 3, pp. 331–338, 2000.
- [7] U. Sur, “A test verified model development study for a nuclear water chiller using the seismic qualification analysis and test,” *Nuclear Engineering and Technology*, vol. 43, no. 4, pp. 355–360, 2011.
- [8] IEEE Std 344-1987, *IEEE Recommended Practice for Seismic Qualification of Class 1E Equipment for Nuclear Power Generating Stations*, The Institute of Electrical and Electronics Engineers, New York, NY, USA, 1987.
- [9] R. E. Spears, “Unique method for generating design earthquake time histories,” in *Proceedings of the ASME Pressure Vessels and Piping Conference (PVP '08)*, pp. 3–10, Chicago, Ill, USA, July 2008.
- [10] M. J. Jhung and Y. H. Ryu, “Study on dynamic response of mechanical component to earthquake,” *Journal of Nuclear Science and Technology*, vol. 47, no. 11, pp. 1065–1074, 2010.
- [11] X. Zhou, R. Yu, and L. Dong, “The Complex-Complete-Quadratic-Combination (CCQC) method for seismic responses of non-classically damped linear MDOF system,” in *Proceedings of the 13th World Conference on Earthquake Engineering*, Vancouver, Canada, August 2004.
- [12] Q. Z. Khan, “Evaluation of effects of response spectrum analysis on height of building,” in *Proceedings of the International Conference on Sustainable Built Environment (ICSBE '10)*, Kandy, Sri Lanka, December 2010.
- [13] ANSYS—*Simulation Driven Product Development*, <http://www.ansys.com>.
- [14] Brüel & Kjær *Sound & Vibration Measurement*, <http://www.bksv.com>.
- [15] J. K. Sinha and M. I. Friswell, “The use of model updating for reliable finite element modelling and fault diagnosis of structural components used in nuclear plants,” *Nuclear Engineering and Design*, vol. 223, no. 1, pp. 11–23, 2003.
- [16] E-5 KEPIC END, *Nuclear Electrical and I&C—Equipment Qualification*, Korea Electric Power Industry Code, 2010.

## Research Article

# Distributed Optimal Power and Rate Control in Wireless Sensor Networks

Meiqin Tang,<sup>1</sup> Jianyong Bai,<sup>2</sup> Jing Li,<sup>3</sup> and Yalin Xin<sup>1</sup>

<sup>1</sup> School of Mathematics and Statistics, Ludong University, Yantai 264025, China

<sup>2</sup> Department of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

<sup>3</sup> School of Information Engineering, Nanchang University, Nanchang 330031, China

Correspondence should be addressed to Meiqin Tang; [meiqin@ualberta.ca](mailto:meiqin@ualberta.ca)

Received 23 January 2014; Accepted 20 February 2014; Published 8 May 2014

Academic Editors: N. Barsoum, P. Vasant, and G.-W. Weber

Copyright © 2014 Meiqin Tang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the rapid development of wireless sensor networks, reducing energy consumption is becoming one of the important factors to extend node lifetime, and it is necessary to adjust the launching power of each node because of the limited energy available to the sensor nodes in the networks. This paper proposes a power and rate control model based on the network utility maximization (NUM) framework, where a weighting factor is used to reflect the influence degree of the sending power and transmission rate to the utility function. In real networks, nodes interfere with each other in the procedure of transmitting signal, which may lead to signal transmission failure and may negatively have impacts on networks throughput. Using dual decomposition techniques, the NUM problem is decomposed into two distributed subproblems, and then the conjugate gradient method is applied to solve the optimization problem with the calculation of the Hessian matrix and its inverse in order to guarantee fast convergence of the algorithm. The convergence proof is also provided in this paper. Numerical examples show that the proposed solution achieves significant throughput compared with exiting approaches.

## 1. Introduction

In recent years, with the consistent development of wireless sensor networks, they have been extensively applied in environmental monitoring, volcano monitoring, forest fire prevention, military facilities, and other fields [1]. A wireless sensor network is composed of a large number of microsensors which are capable of sensing, data processing, and transmission. It transmits signals through the self-organization form of networks and the cooperation among nodes. The energy of sensor nodes is very limited, and generally the battery of nodes is not non-renewable or its' updating is costly which results in an energy-constrained problem. Data transmission consumes most energy of sensor nodes. Therefore, how to control transmission power has become a key problem in the research of wireless sensor networks for low-energy consumption in data transmission under the premise of quality of services.

In wireless sensor networks, power control is to continuously adjust the transmission power of network nodes

through a reasonable design. It can ensure the network connectivity and reduce the mutual interference among nodes to extend the lifetime of the whole sensor network. We consider the power and rate control problem based on the network utility framework, which has been introduced in wired [2, 3] and wireless networks [4–11] since the publication of the seminal paper [2] by Kelly et al. in 1998. In the past few years, a great deal of research effort has been made on the energy conservation in wireless sensor networks. One important perspective is to maximize the network lifetime [4, 5] while guaranteeing the required traffic rate. However, since sensor nodes are assumed to have fixed source rates, it is likely that the network cannot sustain the rate for the given system resource constraints. To this end, rate allocation in wireless sensor networks has been studied in recent years [6, 7]. In [6], the use of lexicographical max-min rate allocation was advocated for the nodes and a polynomial-time algorithm was developed for exploiting the parametric analysis technique from linear programming. The rate maximization problem in [7] was formulated as a concave utility maximization and

a subgradient algorithm was proposed to solve it distributively. Game theory was applied to the power control problem based on the network utility maximization (NUM) framework in wireless sensor networks [8, 9]; the appropriate transmission power was selected to improve the network topology, reduce the collisions, increase good-put, and confirm the network connectivity. In [10], two algorithms were presented to compute the transmission power of each node with the objectives of minimizing the total transmission power and the total interference, respectively. But the time-varying wireless environment was not considered in the paper. The energy-constrained nature of nodes limits the operational lifetime of the network since energy is dissipated in both sensing and communicating data across the network. There is an intrinsic tradeoff between network lifetime maximization and rate allocation in wireless sensor networks. In [11], this tradeoff was characterized by considering a cross-layer design problem in a wireless sensor network with orthogonal link transmissions, and then the dual theory was used to solve the optimization problem. A priced-based distributed power and rate control algorithm was proposed in [12]; it can simulate the cooperation of power control and rate adaptation among the nodes. A joint source-channel maximum likelihood (SCML) decoding framework is proposed in wireless sensor networks (WSNs) [13], and prediction likelihood tree (PLT) approach is applied to exploit the spatiotemporal narrowband properties of the sensor data for sequence detection.

Considering the time-varying rate constraint, this paper presents a power and control algorithm based on the NUM framework which is similar to the algorithm proposed in [11, 12]. We adopt the probability to the constraint, which is more suitable for the varying wireless environment. The target function is decomposed into two optimization subproblems using dual decomposition methods to reduce the complexity of the solution for the large-scale network. We solve the optimization problem using the conjugate gradient method without the calculation of the Hessian matrix and its inverse. This can guarantee fast convergence of the algorithm.

The rest of the paper is organized as follows. The system model is described in Section 2. The proposed algorithm is formulated in Section 3. We give the convergence of the proposed algorithm in Section 3, numerical results are provided in Section 4, and Section 5 concludes.

## 2. Formulation of Power Optimization

**2.1. System Description.** Considering a wireless sensor network that consists of a set of  $N$  sensor nodes and a set of  $L$  wireless communication links, we assume that the link exists only when the nodes communicate directly. The transmission rate of node  $s$  is defined as  $x_s$ , where  $s \in N$  and  $x_s \geq 0$  is satisfied. Let  $p_s$  denote the transmission power of node  $s$  with  $0 \leq p_s \leq p_{\max}$ , where  $p_{\max}$  is the maximum power. In the production and design of wireless sensor nodes, the survival time is generally designed to be not less than a constant value  $T_0$ . Assuming that most of the energy is consumed in the process of signal transmission and the other part of the energy

loss is ignored, the energy of wireless sensor network nodes is stored as a constant value  $E$ , and then  $p_{\max}$  is given by

$$p_{\max} = \frac{E}{T_0}. \quad (1)$$

Assume there exists a link  $l$  ( $l \in L$ ) which is connected with the other and can be modeled as an additive white Gaussian noise (AWGN) channel with a noise spectral density  $N_0$ . The channel capacity  $c_l$  of link  $l$  can be given by Shannons theory [14]:

$$c_l(p_l) = w \log_2 \left( 1 + \frac{p_l K d_l^{-\alpha}}{N_0 w} \right), \quad (2)$$

where  $w$  is the fixed bandwidth, the transmission distance between the transmitter and receiver on link  $l$  is  $d_l$ ,  $K$  is a constant that depends on the transmission frequency, and  $\alpha$  is the path-loss exponent, and therefore we have the following restriction condition:

$$\sum_s x_s \leq c_l(p_l). \quad (3)$$

Since the node may produce mutual influence and interference when transmitting the signal, the transmission of signals between adjacent nodes is not always successful.  $q_s$  is assumed to be the probability of the successful signal transmission between adjacent nodes. We can give the following improved constraint condition:

$$\sum_s q_s x_s \leq c_l(p_l). \quad (4)$$

Now we describe the calculating procedure of probability  $q_s$ . The number of the nodes in link  $l$  which may compete with the node  $s$  is  $n_s$  ( $n_s \in N$ ). In the case of competition, node  $s$  is to measure how much probability of  $q_s$  denotes the transmit success probability of node  $s$ . Given that the success transmitting gain is  $u_s$ , which is related to the utilization ratio of energy for the node, the failure transmitting gain is  $u_c$  and the gain that the node does not transmit is  $u_i$ . Obviously, we can get  $u_c < u_i < u_s$  since the utilization ratio of energy for the node is low when the node transmission fails, while the ratio is high when the node transmission succeeds. Consequently, the transmit success probability of node  $s$  is  $(1 - q_s)^{n_s}$ , the transmission failure probability of node  $s$  is  $1 - (1 - q_s)^{n_s}$ , and the transmission signal gain  $u_t$  of node  $s$  can be given by

$$u_t = (1 - q_s)^{n_s} u_s + [1 - (1 - q_s)^{n_s}] u_c. \quad (5)$$

When the node does not transmit the signal, the signal gain  $u_w$  is

$$u_w = u_i. \quad (6)$$

Each node can choose to transmit the signal or not. There are two cases for transmission: success and failure. If one of the nodes successfully transmits the signal, the other nodes are all in transmission failure state or no transmission state. In order to prolong the lifetime of the whole system, the gain

of the nodes needs to be balanced. Suppose  $u_t = u_\omega$ , and we get

$$q_s = 1 - \left( \frac{u_i - u_c}{u_s - u_c} \right)^{1/n_s}. \quad (7)$$

We can find that the success transmission probability is related not only to the node number but also to the gain of the transmission success, transmission failure, and no transmission. The gains are always defined as

$$u_s = \omega_1 p_s \log p_s, \quad u_c = \omega_2 p_s \log p_s, \quad u_i = 0, \quad (8)$$

where  $\omega_1$  and  $\omega_2$  are constants. Submitting (8) into (7), we can get

$$q_s = 1 - \left( \frac{\omega_2}{\omega_2 - \omega_1} \right)^{1/n_s}. \quad (9)$$

The utility function  $U_s$  based on the NUM framework denotes the satisfaction degree of the user, which is continuously differentiable while increasing. Taking into account the transmission rate and transmission power, the total utility function based on the network utility function is defined as

$$\alpha \sum_s U_s^1(x_s) - (1 - \alpha) \sum_l \sum_s U_s^2(p_l), \quad (10)$$

where  $\alpha$  is the weight which can reflect the influence degree of the utility function to the transmission power and rate, and it satisfies  $0.05 \leq \alpha \leq 0.95$ .

We can observe that a given encoding distortion can be guaranteed by controlling both the source rate and the encoding power. When simply adjusting the source rate or the encoding power to a very low or very high level, the encoding distortion will inevitably become large while the total power consumed at the sensor node will increase fast. In this paper, we consider the power and rate control problem with the power and rate constraints. The optimization problem for resource control is formulated as

$$\begin{aligned} \max \quad & \alpha \sum_s U_s^1(x_s) - (1 - \alpha) \sum_l \sum_s U_s^2(p_l) \\ \text{s.t.} \quad & \sum_s q_s x_s \leq c_l(p_l), \quad l = 1, 2, \dots, L \\ & 0 \leq p_l \leq p_{\max} \\ & 0.05 \leq \alpha \leq 0.95, \\ & x_s \geq 0. \end{aligned} \quad (11)$$

**2.2. Dual Decomposition of the Proposed Algorithm.** Two Lagrange multipliers  $\lambda_l$  and  $\mu_l$  are introduced to the Lagrangian dual function of the primal problem (11), and

the corresponding Lagrangian dual function can be expressed as

$$\begin{aligned} L(x_s, p_l, \lambda_l, \mu_l) &= \alpha \sum_s U_s^1(x_s) - (1 - \alpha) \sum_l \sum_s U_s^2(p_l) \\ &+ \sum_l \lambda_l \left[ c_l(p_l) - \sum_s q_s x_s \right] + \sum_l \mu_l (p_{\max} - p_l) \\ &= \sum_s \left[ \alpha U_s^1(x_s) - (1 - \alpha) \sum_l U_s^2(p_l) - q_s x_s \sum_l \lambda_l \right] \\ &+ \sum_l \lambda_l c_l(p_l) + \sum_l \mu_l (p_{\max} - p_l). \end{aligned} \quad (12)$$

The dual problem is then given by

$$\begin{aligned} \min \quad & D(\lambda, \mu) \\ \text{s.t.} \quad & \lambda \geq 0, \quad \mu \geq 0. \end{aligned} \quad (13)$$

We can get the dual function as follows:

$$\begin{aligned} D(\lambda, \mu) &= \max \sum_s \left[ \alpha U_s^1(x_s) - (1 - \alpha) \sum_l U_s^2(p_l) - q_s x_s \sum_l \lambda_l \right] \\ &+ \sum_l [\lambda_l c_l(p_l) + \mu_l (p_{\max} - p_l)]. \end{aligned} \quad (14)$$

The dual function can be decomposed into two subproblems, which are evaluated separately for the nodes, and then the dual function can be rewritten as

$$\begin{aligned} D(\lambda, \mu) &= \max \alpha U_s^1(x_s) - (1 - \alpha) \sum_l U_s^2(p_l) \\ &- q_s x_s \sum_l \lambda_l + \lambda_s c_l(p_s) + \mu_s (p_{s \max} - p_s). \end{aligned} \quad (15)$$

The subproblem of optimization with  $x_s$  as the variable is

$$\begin{aligned} \max \quad & \alpha U_s^1(x_s) - q_s x_s \sum_l \lambda_l \\ \text{s.t.} \quad & x_s \geq 0. \end{aligned} \quad (16)$$

The subproblem of optimization with  $p_l$  as the variable is

$$\begin{aligned} \max \quad & -(1 - \alpha) \sum_l U_s^2(p_l) + \lambda_l c_l(p_l) + \mu_l (p_{\max} - p_l) \\ \text{s.t.} \quad & 0 \leq p_l \leq p_{\max}. \end{aligned} \quad (17)$$

According to the dual theory, we can get the following equation for the rate suboptimization problem from (16):

$$\begin{aligned} \min \quad & q_s x_s \sum_l \lambda_l - \alpha U_s^1(x_s) \\ \text{s.t.} \quad & x_s \geq 0. \end{aligned} \quad (18)$$

And the power suboptimization problem as described in (17) can be rewritten as

$$\min (1 - \alpha) \sum_l U_s^2(p_l) - \lambda_l c_l(p_l) - \mu_l(p_{\max} - p_l) \tag{19}$$

$$\text{s.t. } 0 \leq p_l \leq p_{\max},$$

where  $f(x_s)$  and  $f(p_l)$  are twice continuously differentiable. We used the conjugate gradient method mentioned before to calculate, which has many advantages. It only needs to seek the function of the first order derivative, which not only alleviates the slow convergence characteristic of the steepest descent method, but also avoids the storage and computation of Hessian matrix and its inverse features in Newton's method. Besides, the program compilation is relatively simple and the computational complexity is relatively small. It is the most effective solution of a large number of linear equations and nonlinear unconstrained optimization problems.

For (16), we obtain the guiding function through the derivation of  $\lambda_l, \mu_l$ :

$$\frac{\partial}{\partial \lambda_l} \left( (1 - \alpha) U_s^1(x_s) - q_s x_s \sum_l \lambda_l \right) = -q_s x_s. \tag{20}$$

For (17), we get the guiding function through the derivation of  $\lambda_l, \mu_l$ :

$$\begin{aligned} \frac{\partial}{\partial \lambda_l} \left( (\alpha - 1) \sum_l U_s^2(p_l) + \lambda_l c_l(p_l) + \mu_l(p_{\max} - p_l) \right) \\ = c_l(p_l) \end{aligned} \tag{21}$$

$$\begin{aligned} \frac{\partial}{\partial \mu_l} \left( (\alpha - 1) \sum_l U_s^2(p_l) + \lambda_l c_l(p_l) + \mu_l(p_{\max} - p_l) \right) \\ = p_{\max} - p_l. \end{aligned} \tag{22}$$

According to (20) and (21), the step iterative algorithm of  $\lambda_l$  is

$$\lambda_l(k+1) = [\lambda_l(k) + \gamma(k) [c_l(p_l) - q_s x_s]]^+. \tag{23}$$

According to (22), the step iterative algorithm of  $\mu_l$  is

$$\mu_l(k+1) = [\mu_l(k) + \kappa(k) (p_{\max} - p_l)]^+, \tag{24}$$

where  $\gamma(k)$  and  $\kappa(k)$  are the step length.

The principle and the procedure of power  $p_l$  and rate  $x_s$  optimal algorithm are similar, which are all optimized based on the conjugate gradient method. And we used decomposition theory to decompose the two algorithms into one separately and they are connected by the Lagrange multipliers to be combined to one problem to assure the total utility is to be maximized. The steps of the optimization algorithm based on the conjugate gradient method for  $x_s$  are as follows.

*Step 1.* Given an initial value of  $p_l(0), \mu_l(0)$ , and  $\mu_l(0)$ , where the error  $\varepsilon > 0$ , calculate  $f(0)$  as follows:

$$f(0) = f(p_l(0)), \tag{25}$$

and then set

$$d(0) = -\Delta f(p_l(0)), \quad k = 0. \tag{26}$$

*Step 2.* If

$$\|f(p_l(k))\| \leq \varepsilon \tag{27}$$

is satisfied, stop the calculation; otherwise, go to Step 3.

*Step 3.* Calculate step length factor  $\zeta_k$  through the linear search method, where  $\zeta_k$  satisfies the strong Wolfe linear search criteria:

$$f(p_l(k)) + \zeta_k d_k = \min_{\zeta > 0} f(p_l(k) + \zeta(k) d(k)). \tag{28}$$

Assume

$$p_l(k+1) = p_l(k) + \zeta_k d_k,$$

$$\lambda_l(k+1) = [\lambda_l(k) + \gamma(k) [c_l(p_l) - q_s x_s]]^+, \tag{29}$$

$$\mu_l(k+1) = [\mu_l(k) + \kappa(k) (p_{\max} - p_l)]^+.$$

*Step 4.* Calculate

$$\chi_k = \frac{\|\nabla f(p_l(k+1))\|^2}{\|\nabla f(p_l(k))\|^2}, \tag{30}$$

$$d_{k+1} = -\nabla f(p_l(k+1)) + \chi_k d_k.$$

*Step 5.* Consider

$$k = k + 1; \tag{31}$$

then, go to Step 2.

In the algorithm, we must guarantee the direction is decreased, which requires

$$\begin{aligned} (\nabla f(p_l(k)), d(k)) &= (\nabla f(p_l(k)), -\nabla f(p_l(k))) \\ &= -\|\nabla f(p_l(k))\|^2 < 0. \end{aligned} \tag{32}$$

So we use the strong Wolfe linear search criteria to calculate the step length factor  $\zeta_k$ , and the direction of  $d_k$  is definitely the descent direction.

The steps of the optimization algorithm based on the conjugate gradient method for  $x_s$  are similar to  $p_l$ .

*Step 1.* Given an initial value of  $x_s(0), \mu_l(0)$ , and  $\mu_l(0)$ , where the error  $\varepsilon' > 0$ , calculate  $f'(0)$ :

$$f'(0) = f'(p_l(0)), \tag{33}$$

and then set

$$d'(0) = -\Delta f(p_l(0)), \quad k' = 0. \tag{34}$$

*Step 2.* If

$$\|f'(p_l(k))\| \leq \varepsilon' \tag{35}$$

is satisfied, stop the calculation; otherwise, go to Step 3.

Step 3. Calculate the step length factor  $\zeta'_k$  through the linear search method, where  $\zeta'_k$  satisfies the strong Wolfe linear search criteria:

$$f'(x_s(k)) + \zeta'_k d'_k = \min_{\zeta' > 0} f'(x_s(k) + \zeta'(k) d'(k)). \quad (36)$$

Assume

$$\begin{aligned} x_s(k+1) &= x_s(k) + \zeta'_k d'_k, \\ \lambda_l(k+1) &= [\lambda_l(k) + \gamma(k) [c_l(p_l) - q_s x_s]]^+, \\ \mu_l(k+1) &= [\mu_l(k) + \kappa(k) (p_{\max} - p_l)]^+. \end{aligned} \quad (37)$$

Step 4. Calculate

$$\begin{aligned} \chi'_k &= \frac{\|\nabla f'(x_s(k+1))\|^2}{\|\nabla f'(x_s(k))\|^2}, \\ d'_{k+1} &= -\nabla f'(p_l(k+1)) + \chi'_k d'_k. \end{aligned} \quad (38)$$

Step 5. Consider

$$k = k + 1, \quad (39)$$

and then go to Step 2.

In the algorithm, we must guarantee the direction is decreased, which requires

$$\begin{aligned} (\nabla f'(x_s(k)), d(k)) &= (\nabla f'(x_s(k)), -\nabla f'(x_s(k))) \\ &= -\|\nabla f'(x_s(k))\|^2 < 0. \end{aligned} \quad (40)$$

### 3. Convergence Analyses

For the general function, the conjugate gradient method under certain conditions is convergent, and the convergence speed is generally superior to that of the steepest descent method. The  $p_l$ 's convergence will be proved as follows. It is similar to  $x_s$ .

**Proposition 1.** Assume that  $f(p_l)$  is in the bounded set

$$L = \{p_l \in R \mid f(p_l) \leq f(p_l(0))\}, \quad (41)$$

which is continuously differentiable and has a lower bound; then, the sequence obtained by the conjugate gradient method  $\{p_l\}$  converges to  $p_l^*$ , which is the stagnation point of  $f(p_l)$ .

*Proof.* Given that  $\{p_l(k)\}$  is a finite sequence of number, according to the algorithms termination conditions, the last  $p_l^*$  must meet

$$g(p_l^*) = 0, \quad (42)$$

so  $p_l^*$  is the stagnation point of  $f(p_l)$ .

If  $\{p_l(k)\}$  is an infinite series, then for all  $k$ ,

$$\nabla f(p_l(k)) \neq 0. \quad (43)$$

Since

$$d_k = -\nabla f(p_l(k)) + \chi_k d_{k-1}, \quad (44)$$

we get

$$(\nabla f(p_l(k)), d_k) = -\|\nabla f(p_l(k))\|^2 < 0, \quad (45)$$

where  $d_k$  is the decline direction. Since  $\{f(p_l(k))\}$  is the lower bound of the sequence and is monotonically decreasing,

$$f(p_l(k)) \rightarrow f(p_l^*). \quad (46)$$

The convergence analysis of  $x_s^*$  is similar to  $p_l^*$ .  $\square$

### 4. Numerical Example

We simulate a network which consists of nine sensor nodes and one sink node. The sensor nodes will transmit their sensing data to the sink node, and all the nodes are randomly deployed in an area of 100 m  $\times$  100 m. The utility function is set to be in the log form; for example,  $U_1$  is set to be  $\log 2(x)$  and  $U_2 = -p$ . The fixed bandwidth  $w$  is set to 5 MHz and the maximum of  $p$  is set to be 1 mW for all links. The path-loss exponent  $\sigma$  is set to 2. We will show the network performance with different  $\alpha$  values.

First, we show the convergence figures for the optimal rates and power. The maximum iteration of the algorithm is set to be 200. Figures 1 and 2 show the optimal video rates for sensor nodes when  $\alpha$  is set to 0.05 and 0.95, respectively, from which we can find that the proposed algorithm can converge within 20 iteration steps. We also find the located rates of some nodes are large, while the others are small. This is because these nodes transmit the data to the sink node directly, which will not relay the data and will transmit the data more efficiently. In wireless sensor networks, the available energy is limited, and Figures 3 and 4 show the consumed video power of the nodes, from which we can also indicate the proposed algorithm can get to the optimum efficiently and quickly. From these four figures, we can find that the proposed algorithm can get higher throughput and uses less energy.

In Figure 5, we find that when we set  $\alpha$  to 0.95, the utility is large, but the value of lifetime is small; when the  $\alpha$  is set to 0.05, the value of lifetime is large, while the network utility is small. For this reason, we can find that there is apparently a tradeoff between network utility and network lifetime in energy-limited wireless networks. Different values can be set according to actual need.

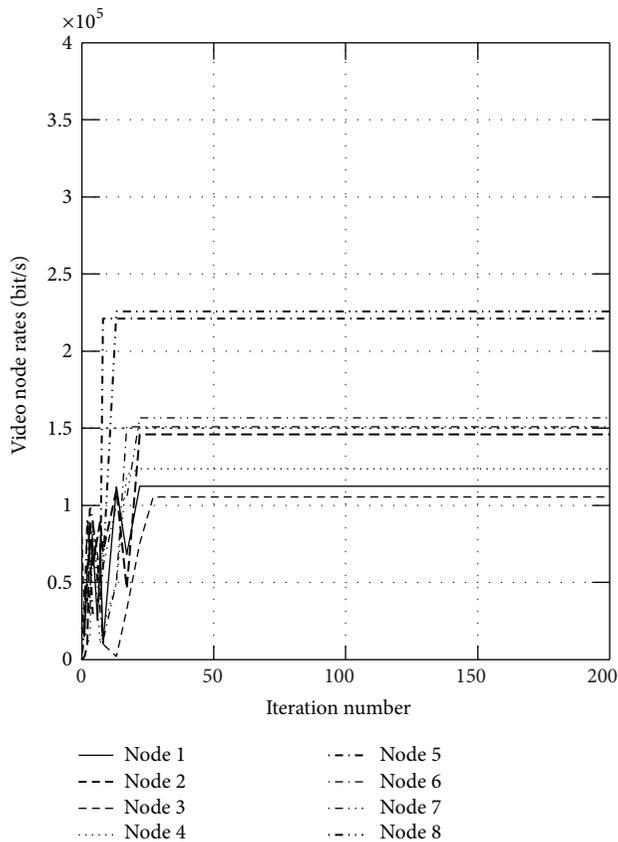
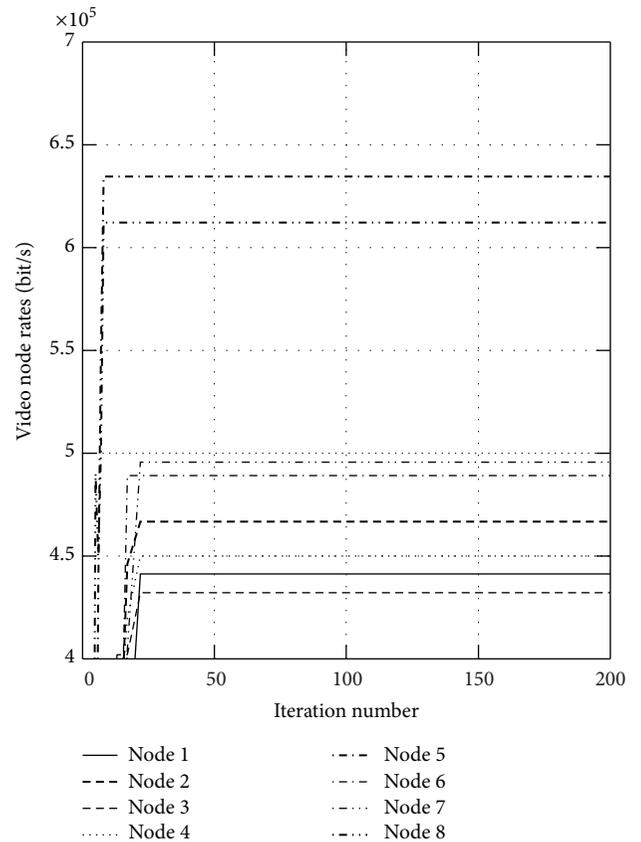
In order to verify the performance of the proposed algorithm, we compared it with previous works in [11, 12], and the network setup is the same. Tables 1 and 2 give the data rate comparison with different values. From these two tables we can find that the proposed algorithm can get higher utility than the algorithm in [12] and minUtility [11] and less power consumption than the algorithm in [12] and MaxUtility [11], because minUtility only seeks minimum energy consumption and while MaxUtility only seeks the maximum utility. Through adding the probability for the rate constraint, our algorithm can get higher data rate.

TABLE 1: The data rate comparison with different methods for  $\alpha = 0.05$ .

Method	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7	Node 8
The proposed algorithm	$1.1256e + 05$	$1.4621e + 05$	$1.0543e + 05$	$1.2390e + 05$	$2.2111e + 05$	$1.5098e + 05$	$1.5678e + 05$	$2.2567e + 05$
The algorithm in [12]	$1.1021e + 05$	$1.3609e + 05$	$0.9763e + 05$	$1.1907e + 05$	$1.8798e + 05$	$1.3142e + 05$	$1.3988e + 05$	$2.0995e + 05$
MaxUtility [11]	$1.26e + 05$	$1.4569e + 05$	$1.1314e + 05$	$1.3086e + 05$	$1.9209e + 05$	$1.7142e + 05$	$1.6678e + 05$	$1.8765e + 05$
MinEnergy [11]	92	89.21	98.74	99.2356	101.48	99.672	96.32	102.54

TABLE 2: The data rate comparison with different methods for  $\alpha = 0.95$ .

Method	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7	Node 8
The proposed algorithm	$4.4123e + 05$	$4.6678e + 05$	$4.321e + 05$	$4.5009e + 05$	$6.3457e + 05$	$4.8914e + 05$	$4.9562e + 05$	$6.1209e + 05$
The algorithm in [12]	$4.2265e + 05$	$4.4324e + 03e + 05$	$4.5616e + 05$	$4.3190e + 05$	$4.4999e + 05$	$6.2008e + 05$	$4.81998e + 05$	$4.785e + 05$
MaxUtility [11]	$4.5532e + 05$	$4.7121e + 03e + 05$	$4.35143e + 05$	$4.590e + 05$	$4.8654e + 05$	$4.2672e + 05$	$4.996e + 05$	$4.9867e + 05$
MinEnergy [11]	104.56	105.32	102.78	103.45	128.9	107.8	103.81	127.65

FIGURE 1: The optimal video rates for sensor nodes when  $\alpha$  is set to 0.05.FIGURE 2: The optimal video rates for sensor nodes when  $\alpha$  is set to 0.95.

Tables 3 and 4 show the power of the different algorithms, for which we can see that the proposed algorithm get less power consumption than the algorithm proposed in [12] and MaxUtility [11], which shows the proposed algorithm is very efficiently.

## 5. Conclusion

This paper proposes a new power control method based on network utility maximization framework for wireless sensor networks, and a trade-off parameter for the utility and

TABLE 3: The data power comparison with different methods for  $\alpha = 0.05$ .

Method	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7	Node 8
The proposed algorithm	$0.6321e - 6$	$0.6102e - 6$	$0.6198e - 6$	$0.6218e - 6$	$0.5187e - 6$	$0.6018e - 6$	$0.5989e - 6$	$0.5266e - 6$
The algorithm in [12]	$0.7112e - 6$	$0.7426e + 03e - 6$	$0.7002e - 6$	$0.7221e - 6$	$0.6789e - 6$	$0.7108e - 6$	$0.7685e - 6$	$0.6351e - 6$
MaxUtility [11]	$0.9168e - 05$	$0.9041e - 05$	$0.9908e - 05$	$0.9127e - 05$	$0.8873e - 05$	$0.9162e - 05$	$0.9092e - 05$	$0.8904e - 05$
MinEnergy [11]	$0.8465e - 09$	$0.8169e - 09$	$0.83287e - 09$	$0.8658e - 09$	$0.1099e - 10$	$0.8365e - 09$	$0.8102e - 09$	$0.1214e - 10$

TABLE 4: The data rate comparison with different methods for  $\alpha = 0.95$ .

Method	Node 1	Node 2	Node 3	Node 4	Node 5	Node 6	Node 7	Node 8
The proposed algorithm	$0.3214e - 6$	$0.3189e - 6$	$0.3102e - 6$	$0.3336e - 6$	$0.1765e - 6$	$0.3087e - 6$	$0.3096e - 6$	$0.1924e - 6$
The algorithm in [12]	$0.4219e - 6$	$0.4468e + 03e - 6$	$0.4097e - 6$	$0.4354e - 6$	$0.2176e - 6$	$0.4677e - 6$	$0.4781e - 6$	$0.2674e - 6$
MaxUtility [11]	$0.5532e - 05$	$0.5096e - 05$	$0.5989e - 05$	$0.5213e - 05$	$0.3786e - 05$	$0.5011e - 05$	$0.5124e - 05$	$0.3989e - 05$
MinEnergy [11]	$0.4107e - 09$	$0.4231e - 09$	$0.4358e - 09$	$0.4691e - 09$	$0.0519e - 10$	$0.4218e - 09$	$0.449e - 09$	$0.0627e - 10$

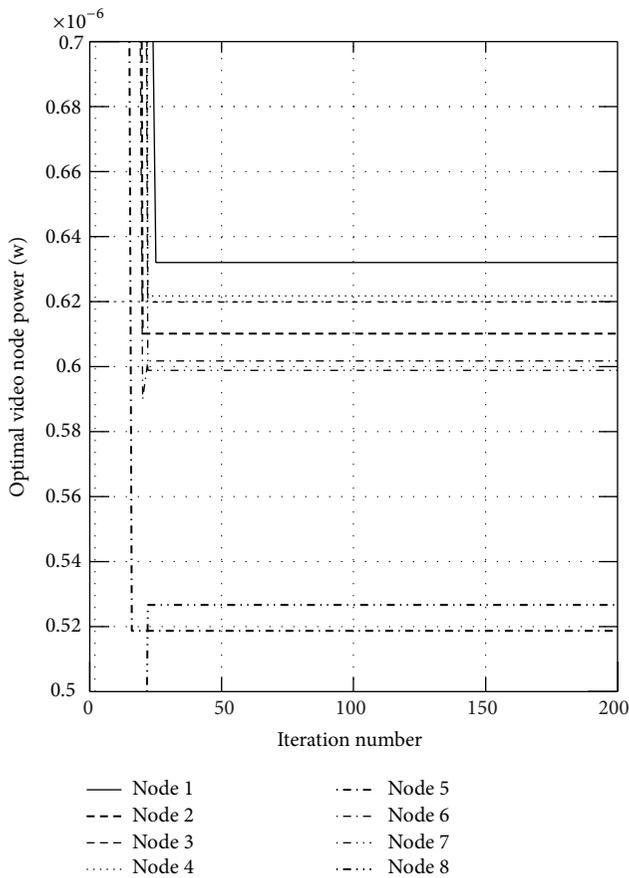


FIGURE 3: The consumed video power of the nodes when  $\alpha$  is set to 0.05.

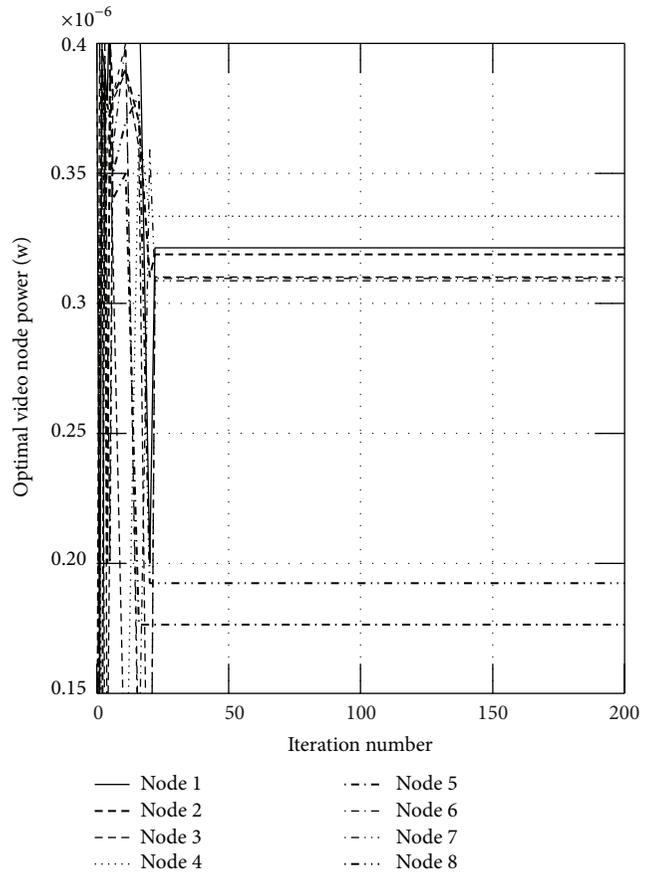


FIGURE 4: The consumed video power of the nodes when  $\alpha$  is set to 0.95.

lifetime is introduced to the system. Considering the time varying environment, the probability is added to the rate constraint, which is natural in the context of various applications.

Moreover, we put forward a new priced-based distributed algorithm using a gradient method. The algorithm is designed to keep acceptable throughput. Simulation studies show that

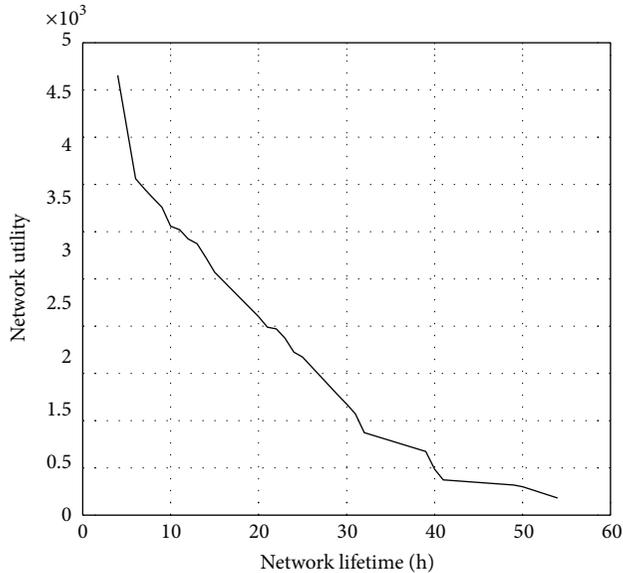


FIGURE 5: The tradeoff between network utility and network lifetime.

the proposed algorithms are effective to solve the optimization problem and outperform the existing approaches in terms of throughput and energy efficiency since desired variables converge to the optimal point very quickly.

### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

### Acknowledgment

This work was supported by the National Science Foundation of China Grants 61174097.

### References

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, 2002.
- [2] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, no. 3, pp. 237–252, 1998.
- [3] S. H. Low and D. E. Lapsley, "Optimization flow control—I: basic algorithm and convergence," *IEEE/ACM Transactions on Networking*, vol. 7, no. 6, pp. 861–874, 1999.
- [4] V. Srinivasan, C.-F. Chiasserini, P. S. Nuggehalli, and R. R. Rao, "Optimal rate allocation for energy-efficient multipath routing in wireless ad hoc networks," *IEEE Transactions on Wireless Communications*, vol. 3, no. 3, pp. 891–899, 2004.
- [5] R. Madan and S. Lall, "Distributed algorithms for maximum lifetime routing in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 5, no. 8, pp. 2185–2193, 2006.
- [6] Y. T. Hou, Y. Shi, and H. D. Sherali, "Rate allocation in wireless sensor networks with network lifetime requirement," in

*Proceedings of the 5th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MoBiHoc '04)*, pp. 67–77, May 2004.

- [7] K.-L. Hung, B. Bensaou, J. Zhu, and F. Nait-Abdesselam, "Energy-aware fair routing in wireless sensor networks with maximum data collection," in *Proceedings of the IEEE International Conference on Communications (ICC '06)*, pp. 3438–3443, July 2006.
- [8] S. Sengupta, M. Chatterjee, and K. A. Kwiat, "A game theoretic framework for power control in wireless sensor networks," *IEEE Transactions on Computers*, vol. 59, no. 2, pp. 231–242, 2010.
- [9] G. Yang and G. Zhang, "A power control algorithm based on non-cooperative game for wireless sensor networks," in *Proceedings of the International Conference on Electronic and Mechanical Engineering and Information Technology (EMEIT '11)*, pp. 687–690, August 2011.
- [10] M. X. Cheng, X. Gong, L. Cai, and X. Jia, "Cross-layer throughput optimization with power control in sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 7, pp. 3300–3308, 2011.
- [11] H. Nama, M. Chiang, and N. Mandayam, "Optimal utility-lifetime trade-off in self-regulating wireless sensor networks: a distributed approach," in *Proceedings of the 40th Annual Conference on Information Sciences and Systems (CISS '06)*, pp. 789–794, March 2006.
- [12] Z. Yang, S. Liao, and W. Cheng, "Joint power control and rate adaptation in wireless sensor networks," *Ad Hoc Networks*, vol. 7, no. 2, pp. 401–410, 2009.
- [13] C. Manoj and A. K. Jagannatham, "Optimal prediction likelihood tree based source-channel ML decoder for wireless sensor networks," *IEEE Signal Processing Letters*, vol. 21, no. 2, pp. 135–139, 2014.
- [14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, New York, NY, USA, 1991.

## Research Article

# Optimal Robust Motion Controller Design Using Multiobjective Genetic Algorithm

Andrej Sarjaš,<sup>1</sup> Rajko Svečko,<sup>1</sup> and Amor Chowdhury<sup>1,2</sup>

<sup>1</sup> Faculty of Electrical Engineering and Computer Science, University of Maribor, Smetanova 17, 2000 Maribor, Slovenia

<sup>2</sup> Margento B.V., Crystal Tower, Orlyplein 10, 1043 DP Amsterdam, The Netherlands

Correspondence should be addressed to Andrej Sarjaš; [andrej.sarjas@um.si](mailto:andrej.sarjas@um.si)

Received 28 February 2014; Accepted 18 March 2014; Published 8 May 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Andrej Sarjaš et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper describes the use of a multiobjective genetic algorithm for robust motion controller design. Motion controller structure is based on a disturbance observer in an RIC framework. The RIC approach is presented in the form with internal and external feedback loops, in which an internal disturbance rejection controller and an external performance controller must be synthesised. This paper involves novel objectives for robustness and performance assessments for such an approach. Objective functions for the robustness property of RIC are based on simple even polynomials with nonnegativity conditions. Regional pole placement method is presented with the aims of controllers' structures simplification and their additional arbitrary selection. Regional pole placement involves arbitrary selection of central polynomials for both loops, with additional admissible region of the optimized pole location. Polynomial deviation between selected and optimized polynomials is measured with derived performance objective functions. A multiobjective function is composed of different unrelated criteria such as robust stability, controllers' stability, and time-performance indexes of closed loops. The design of controllers and multiobjective optimization procedure involve a set of the objectives, which are optimized simultaneously with a genetic algorithm—differential evolution.

## 1. Introduction

Modern positioning systems require high performance and complex operation, which demand highly efficient and complex controller structures. Complex controller algorithms and their designs are mostly related to very complex design procedures and optimisation techniques, with which designers try to achieve the desired performance specification. To date, many advanced design methods have been developed, based on different structures of closed-loop systems and paradigms. The most widely used designs are based on disturbance observers [1–3], internal model control [4], and model based disturbance attenuation [5]. The disturbance observer designs have been mostly used in industrial environments because of its simple transparent structure and capability to ensure proper tradeoff between robustness and performance properties. Tradeoffs between criteria can be handled with heuristic optimization approaches or mathematical programming optimization techniques, abbreviated as LP, QP, SDP, NP, and so forth [6, 7]. Optimization-based methods in

computer aided design (CAD) have proved to be a valuable tool in engineering practice, with which one can achieve proper performance and suitable controllers for closed-loop systems. In general, the optimization procedure can be divided into convex or nonconvex optimization problems.

Mathematical programming approaches, such as LP, QP, and SDP, are very efficient for convex problems, which cover many modern robust controller designs in state space approach [8–10]. However, there are also many control problems, where the evaluation of the objective is not strictly convex or where the desired criteria cannot be combined into a set of convex objective functions. For example, polynomial synthesis with fixed order controller in a transfer function form and controller structure optimization is a basically nonconvex problem [11, 12]. One way to overcome this problem is by combining heuristic optimization methods with new or conventional controller designs [13–15]. This combination provides a set of efficient tools to address complex multivariable problems with performance constraints [13]. There is much evidence which indicates high

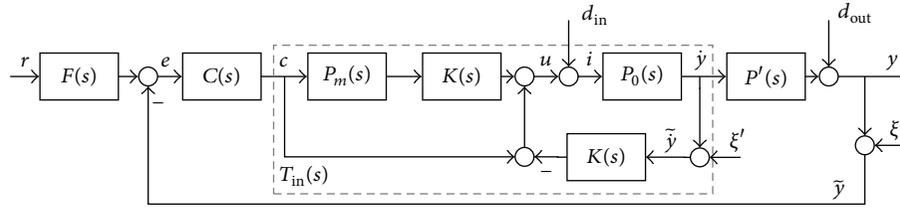


FIGURE 1: Disturbance observer in an RIC framework.

efficiency of the heuristic optimization technique, especially with genetic algorithms (GAs). The GA optimization technique with well-formulated objective functions can preserve satisfactory results of controlled systems while overcoming some problems and limitations of conventional designs [13, 16–18].

This paper considers a robust disturbance observer (DOB) in a robust internal compensator framework (RIC) with multiobjective optimization of different robustness and performance criteria. The DOB structure has already been studied in detail and has several different structural approaches [1, 19]. The more convenient and straightforward methods of DOB are based on Q-filter and inverse nominal models within the internal feedback branch [1, 20, 21]. The more advanced approaches are based on internal reference models in RIC framework [2, 20]. Both methods, Q-filter and RIC, deal well with disturbance attention and have similar approaches and structures [2, 20, 22]. However, the RIC approach is more transparent and reliable in comparison with the Q-filter design. The main difference between both approaches is in the design of their internal loops [5, 22]. The significant disadvantage of the DOB with Q-filter design is the usage of an inverse nominal model within the internal feedback loop. Most mechanical systems are presented as strict proper functions, which prevents direct usage of the model inverse. The second disadvantage is the internal-loop structure with a low-pass Q-filter within a partially positive feedback loop. The property of the Q-filter lowers the closed-loop's stability and the selection is strictly empirical with some vague guidance for the first-, second-, and third-order filters [5, 21, 23, 24]. Each selection of approximated inverse function and Q-filter requires additional assessment of the closed-loop's performance and robustness. The RIC on the other hand has better structural transparency and can be much easily incorporated into the optimization procedure. Internal and external controllers can be acquired from the optimization procedure, so that the robustness of the RIC system can be ensured [25, 26].

This paper describes the design of an RIC disturbance observer with optimization of robustness and performance criteria with multiobjective differential evolution (DE) [27, 28]. The main contribution of the presented paper is a multiobjective optimization approach with simultaneous optimization of a set of criteria for inner and outer loops of the RIC structure. The used robustness and performance criteria are derived directly from the property of the norm  $H_\infty$  and uncertainty models [29, 30]. The criteria have the form of even polynomials, where simple quasi-convex control

problems arise. Robust stability of the system can be quite straightforwardly assessed, if the nonnegativity condition of the polynomial is provided. Even polynomials can be directly used in the optimization technique within genetic algorithm, without any other additional transformation. The second contribution of this paper is the design of a simple parameterized controller structure with selected central characteristic polynomial. The central characteristic polynomial has an admissible region in the stable half plain prescribed so that it ensures the expected performance and stability of the RIC system. The basic idea comes from the pole-colouring technique and regional pole assignment method [31, 32]. In comparison to similar methods, the presented approach uses objective functions of the regional pole placement technique and can be used for nonparametric uncertainties. The presented objectives are optimized with a genetic algorithm, differential evolution (DE). DE has evident advantages over other similar techniques: simple structure, fast convergence, lower space complexity, adaptive parameter settings, and so forth [33]. All design criteria within the DE algorithm are evaluated over the nonnegativity property of robustness criteria and roots calculations of the closed-loop characteristic polynomial for regional pole placement.

This paper is divided into seven sections. In Section 2, we describe the RIC disturbance observer for a positioning system. Section 3 describes the regional pole placement for internal and external loops of the RIC system and applied parameterization of the controller structure. Thereafter, in Section 4, we describe performance and robustness properties with the metric  $H_\infty$ , based on nonnegativity assessments of the even polynomial. Section 5 describes a multiobjective optimization approach with DE and a composite multiobjective function. After Section 5, a design example of an RIC system with multiobjective optimization results is provided in Section 6. The conclusions of the paper are summarized in Section 7.

## 2. RIC Disturbance Observer for Positioning Systems

A disturbance observer in the RIC framework is composed of internal and external loops [34–36]. The design of the internal loop is based on input disturbance attention, where most often appearing disturbances are reaction and load torque, friction, and unmodeled dynamics [20, 34]. The external loop is designed so that it ensures the overall performance of the closed-loop system. The RIC structure is shown in Figure 1, where  $K(s)$ ,  $P_0(s)$ ,  $P'(s)$ ,  $P_m(s)$ ,  $C(s)$ , and  $F(s)$  are the internal

controller, plant, reference model, and 2 DOF structures with an external controller and a prefilter, respectively. The internal loop with  $P_0(s)$  covers the angular velocity in relation to input torques and disturbance rejection  $d_{in}$ , where  $P'(s)$  is the transfer function of the rotary encoder. The transfer function of the rotary encoder is mostly treated as a system with pure integral behavior. The RIC transfer functions are

$$\begin{aligned}
 P_0(s) &= \frac{B_0(s)}{A_0(s)}, & P'(s) &= \frac{B'(s)}{A'(s)}, & K(s) &= \frac{L_K(s)}{R_K(s)}, \\
 P_m(s) &= \frac{B_m(s)}{A_m(s)}, & C(s) &= \frac{L_C(s)}{R_C(s)}, & F(s) &= \frac{L_F(s)}{R_F(s)}.
 \end{aligned}
 \tag{1}$$

The matrix form of the RIC system in Figure 1 is

$$\begin{aligned}
 &\begin{bmatrix} F^{-1} & 0 & 0 & F^{-1} & 0 \\ -C(P_m K + 1) & 1 & 0 & 0 & K \\ 0 & -P_0 P' & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & -P_0 & 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} e \\ i \\ y \\ \tilde{y} \\ \hat{y} \end{pmatrix} \\
 &= \begin{pmatrix} r \\ d_{in} \\ d_{out} \\ \xi \\ \xi' \end{pmatrix},
 \end{aligned}
 \tag{2}$$

where  $r$ ,  $d_{in}$ ,  $d_{out}$ ,  $\xi$ , and  $\xi'$  are reference signal, input disturbance, output disturbance, internal noise, and external noise, respectively.

The closed-loop transfer function is

$$\begin{aligned}
 &\begin{pmatrix} e \\ i \\ y \\ \tilde{y} \\ \hat{y} \end{pmatrix} = \frac{1}{1 + KP_0 + CP_0 P' + KCP_m P_0 P'} \\
 &\times \begin{bmatrix} F(KP_0 + 1) & -P_0 P' & -(1 + KP_0) & -(1 + KP_0) & P_0 P' K \\ FC(P_m K + 1) & 1 & -C(P_m K + 1) & -C(P_m K + 1) & -K \\ P_0 P' FC(P_m K + 1) & P_0 P' & (1 + KP_0) & -P_0 P' C(P_m K + 1) & -P_0 P' F \\ P_0 P' FC(P_m K + 1) & P_0 P' & (1 + KP_0) & (1 + KP_0) & -P_0 P' F \\ P_0 FC(P_m K + 1) & P_0 & -CP_0(P_m K + 1) & -P_0 C(P_m K + 1) & (P_0 P' C + P_0 P' CP_m K + 1) \end{bmatrix} \\
 &\times (r \ d_{in} \ d_{out} \ \xi \ \xi')^T.
 \end{aligned}
 \tag{3}$$

The RIC system in Figure 1 is slightly a modified classic structure. The internal loop in the modified structure embraces only the angular velocity, while the classical internal loop uses for the feedback branch the same output as for the external loop [21, 22]. The modified structure provides indirect decoupling of the internal and external loops, where the poles of the internal loop do not directly influence the zeroes of the external loop. Most RIC designs include standard controller structures in internal loops [20, 24]. Standard controllers, like PID or PI, with integral behaviour improve disturbance rejection for low-frequency disturbance  $d_{in}$ . On the other hand, internal integral behaviour causes a double-integrator effect on the external loop, which significantly lowers the stability domain and generates undesired responses on the output disturbances and reference signals. This paper will consider an RIC structure, shown in Figure 1, where the internal integral behaviour of the controller  $K(s)$  does not directly influence the external loop. For the sake of the RIC structure simplification, we assume that the prefilter function is constant and is equal to  $F(s) = 1$ .

### 3. Regional Pole Placement of the RIC System

The design procedure of the RIC system can be addressed in two steps: the first step covers the internal loop controller design and the second step the external loop controller design. Both controllers are designed with the regional pole placement technique based on the selected central polynomial in the expected stable area. The expected area is freely chosen by the designer and exhibits closed-loop dynamic performance. The parameterization of controllers is applied to ensure good disturbance rejection and robustness properties.

Let us consider the internal loop of the system in Figure 1. The closed-loop system is

$$\begin{aligned}
 &\begin{pmatrix} \dot{y} \\ i \end{pmatrix} = \frac{1}{1 + KP_0} \begin{bmatrix} P_0(1 + KP_m) & P_0 \\ 1 + KP_m & 1 \end{bmatrix} \begin{pmatrix} c \\ d_{in} \end{pmatrix}, \\
 &\begin{pmatrix} \dot{y} \\ i \end{pmatrix} = S_{in} \begin{bmatrix} P_0(1 + KP_m) & P_0 \\ 1 + KP_m & 1 \end{bmatrix} \begin{pmatrix} c \\ d_{in} \end{pmatrix},
 \end{aligned}
 \tag{4}$$

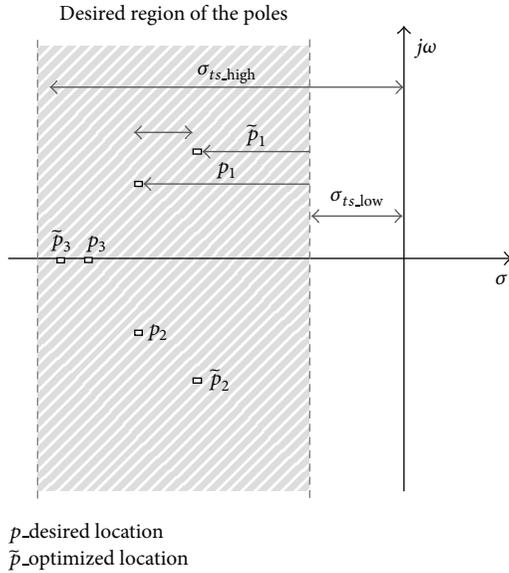


FIGURE 2: Settling-time objective function with settling-time boundaries  $\sigma_{ts,low}$  and  $\sigma_{ts,high}$ .

where  $S_{in}$  is the sensitivity function of the internal closed-loop. The characteristic polynomial is defined as

$$A_0(s)R_K(s) + B_0(s)L_K(s) = C_{in}(s). \quad (5)$$

Equation (5) is the standard starting point of pole placement design. The solution of the equation is provided under strict polynomial degree conditions [37]. The presented approach uses only the conditions when an exact solution does not exist,  $\deg R_k \leq \deg A_0 - 2$ . In this case, the controller

structure does not depend on plant order, like in the classic pole placement design, and is arbitrarily selected by the designer. The only way to satisfy expression (5) is regional pole placement with a prescribed region and deviation assessment, as described in [32]. Deviation is assessed between the given polynomial  $C_{in}(s)$  and the candidate polynomial obtained with optimization. The term prescribed region of the closed polynomial is similar to the term domain stability [29, 38]. The prescribed region with the selected central polynomial offers more leeway for further multiobjective optimization of robustness and performance criteria. The polynomial equation for regional pole placement is

$$A_0(s)R_K(s) + B_0(s)L_K(s) = \tilde{C}_{in}(s), \quad \tilde{C}_{in}(s) \approx C_{in}(s), \quad (6)$$

where  $\tilde{C}_{in}(s)$  is the candidate polynomial obtained with the optimization with controller coefficient  $K(s)$ . The prescribed region of the desired pole location can be derived from the properties of the Laplace stability domain,  $D_{in} = \{s = \sigma + j\omega \in \mathbb{C} \mid \sigma \in -\Re, j\omega \in \Im\}$ , where  $\tilde{C}_{in} \in D_{in}$  holds true. The possible objective functions of the prescribed region for controllers  $K(s)$ ,  $C(s)$  are described in the following subsections.

**3.1. Settling-Time Objective Function.** Roughly speaking, the settling time of the closed-loop system is inversely proportional to the real part of dominant poles in the complex domain  $D_{in}$ . The objective function can be composed so that all closed-loop poles lie in a vertical stripe, bounded by parameters  $\sigma_{ts,high}$  and  $\sigma_{ts,low}$ . The vertical stripe belongs to domain  $D_{in}$ . Parameters  $\sigma_{ts,high}$  and  $\sigma_{ts,low}$  specify the objective function criteria of the closed-loop settling time (see Figure 2).

The proposed objective function is

$$J_1 = \min_K \left( \begin{array}{c} 1 - \sum_n \frac{|\Re\{\tilde{C}_{in}\} - |\sigma_{ts,low}||}{|\Re\{\tilde{C}_{in}\}|}, \forall \sigma_{ts,high} < \Re\{\tilde{C}_{in}\} < \sigma_{ts,low} \wedge \\ \sum_n \frac{|\Re\{C_{in}\} - |\Re\{\tilde{C}_{in}\}|}{\max(|\Re\{C_{in}\}|, |\Re\{\tilde{C}_{in}\}|)} \end{array} \right), \quad \tilde{C}_{in} \in D_{in}, \quad (7)$$

where  $\sigma_{ts,high}$  and  $\sigma_{ts,low}$  indicate the location of the stripe boundary and  $n$  indicates the number of closed-loop poles. The right boundary  $\sigma_{ts,high}$  of the stripe prevents the possibility of obtaining poles with large negative real part values. Large negative real part pole values can cause a too high dynamic of the closed-loop system. A higher dynamic of the closed-loop can render proper real-time implementation of the discrete controller on the embedded system or cause improper behaviour of controller outputs. Improper higher dynamic of the controller output mostly manifests itself as tempestuous responses, which indicate nonoptimal energy behaviour of the closed-loop system.

The first term of the  $J_1$  indicates a normalized value of the objective if all poles lie on the left of the vertical line  $\sigma_{ts,low}$ . The second term assesses the deviation of the real part pole value between the desired  $p$  and the optimized  $\tilde{p}$  pole location, where  $p$  is  $p \in C_{in}$ ,  $\tilde{p} \in \tilde{C}_{in}$ .

**3.2. Rise-Time Objective Function.** The rise time of the closed-loop system is roughly proportional to the inverse of the dominant poles imaginary values and may be upper-bounded by the parameter  $\omega_{tr}$ . The objective function can be composed so that all closed-loop poles lie on the horizontally bounded stripe with bounds  $\pm j\omega_{tr}$  (see Figure 3). The selected stripe belongs to domain  $D_{in}$ .

The proposed rise-time objective function is

$$J_2 = \min_K \left( \begin{array}{l} 1 - \sum_n \frac{||\Im\{\tilde{C}_{in}\}| - |\omega_{tr}||}{|\omega_{tr}|}, \forall |\Im\{\tilde{C}_{in}\}| < |\omega_{tr}| \\ \sum_n \frac{||\Im\{C_{in}\}| - |\Im\{\tilde{C}_{in}\}||}{\max(|\Im\{C_{in}\}|, |\Im\{\tilde{C}_{in}\}|)} \end{array} \right), \quad \tilde{C}_{in} \in D_{in}. \quad (8)$$

The objective function  $J_2$  describes the desired stripe in the complex plain, where the first term indicates the position inside the horizontal stripe and the second the imaginary deviation between the desired  $p$  and the optimized  $\tilde{p}$  poles.

**3.3. Damping-Ratio Objective Function.** Damping ratio is also an important design criterion of the closed-loop dynamic performance and is related to the angle of the vectors between the negative real and the imaginary axis of the dominant poles [32]. The damping ratio is determined with the angle boundary  $\pm\varphi_c$  (see Figure 4).

The proposed damping-ratio objective function is

$$J_3 = \min_K \left( 1 - \sum_n \frac{||\varphi\{\tilde{C}_{in}\}| - |\varphi_c||}{|\varphi_c|}, \forall |\varphi\{\tilde{C}_{in}\}| < |\varphi_c| \right), \quad \tilde{C}_{in} \in D_{in}. \quad (9)$$

The proposed objective function  $J_3$  represents the cone region of the desired poles. All objective functions  $J_1$ - $J_3$  represent min-optimization procedure, where all objectives are normalized on the interval [0-1].

**3.4. The Composite Objective Function of the Dynamic Criteria.** The composite objective function represents the intersection of objectives  $J_1$ - $J_3$ . A graphical presentation of the composite objective function is shown in Figure 5.

The composite objective function can be presented as multiobjective criteria for DE, where the expected solution area is equal to

$$\begin{aligned} \Psi_{in} &= (J_1 \cap J_2 \cap J_3) \wedge \min(\text{dev}(p_{in}, \tilde{p}_{in})) \\ &\quad \forall \Psi_{in} \in D_{in}, \\ \forall p_{in} &= \{p_{in} \in C_{in} \mid C_{in} \in D_{in}\}, \\ \forall \tilde{p}_{in} &= \{\tilde{p}_{in} \in \tilde{C}_{in} \mid \tilde{C}_{in} \in D_{in}\}. \end{aligned} \quad (10)$$

The multiobjective optimization approach will be discussed in the following sections.

**3.5. Internal Controller Parameterization  $K(s)$ .** Controller parameterization is an applied technique, which ensures satisfied properties of the closed-loop system. The internal

controllers  $K$  and  $C$  can be parameterized so as to ensure good input disturbance rejection and robustness properties. In many applications, the input disturbance  $d_{in}$  in positioning systems represents the load torque and Coulomb's friction with a typical low-frequency characteristic [20]. The internal system can provide good elimination of low-frequency disturbances  $d_{in}$  and reference tracking  $c$  if the sensitivity function  $|S_{in}| \ll 1$  has a proper damping effect on the frequency span  $B = [\omega_l, \omega_h]$ ,  $c(B) \wedge d_{in}(B) \gg 1$ . Under this assumption, the controller  $K$  can be parameterized using different polynomial structures, with a known effect on the sensitivity function  $|S_{in}|$ . Controller parameterization with integral action  $R_K(s) = R'_K(s)s$  has a well-known influence on the sensitivity function at low-frequency characteristics, such as the known simple structures PI and PID. For the sake of operational safety, the integral action can in many cases be approximated with the stable pole  $(s + \delta)$ ,  $\{\delta \in \mathbb{R} \mid 0 < \delta \ll 1\}$ , where  $K(s)$  denominator polynomial is equal to  $R'_K(s)(s + \delta)$ . In this case, the damping of the sensitivity function  $S_{in}$  is lowered by parameter  $\delta$ . The lowered damping value of  $S_{in}$  is not noticeable in real-time operation, especially if the absolute damping value is lower than the absolute measurement accuracy.

A minimized tracking error  $e$  and good low-frequency input disturbance rejection  $d_{in}$  can be achieved if the following holds true:

$$\lim_{\omega \rightarrow B} |K(B)| \gg 1, \quad B = \{B \in \mathbb{R} \mid 0 < B \leq \omega_{low}\}, \quad (11)$$

where  $B$  is the low-frequency span of the sensitivity function. Sensitivity function  $S_{in}$  for span  $B$  with controller parameterization  $(s + \delta)$  is

$$\begin{aligned} \lim_{\omega \rightarrow B} |S_{in}(\delta, \omega)| &= \lim_{\omega \rightarrow B} \left( |S_{in}(\omega)| \sqrt{\omega^2 + \delta^2} \right), \\ \lim_{\omega \rightarrow B} |S(B)| \approx \delta &\implies |S(\delta, B)| \approx 0. \end{aligned} \quad (12)$$

Parameter  $\delta$  determines the closed-loop tracking accuracy and the capability of disturbance rejection with characteristic polynomial  $\tilde{C}_{in}(s)$ . Controller structure  $K$  is parameterized as

$$R_K(s) = \begin{cases} R'_K(s) \prod_{k=1}^{\tilde{p}} (s + \delta_k), \\ R'_K(s) R_{Kpar}(s). \end{cases} \quad (13)$$

Parameter  $\delta_k$  represents an additional parameter for further optimization of performance and robustness criteria. The polynomial equation with parametric solutions is

$$A_0(s) R'_K(s) \prod_{k=1}^{\tilde{p}} (s + \delta_k) + B_0(s) L_K(s) = \tilde{C}_{in}(s), \quad (14)$$

$$A_0(s) R_K(s, \delta) + B_0(s) L_K(s) = \tilde{C}_{in}(s).$$

The solution of the optimization procedure is a proper set of polynomial coefficients, where the internal polynomial  $\tilde{C}_{in}(s)$  belongs to the  $\Psi_{in}$  and a strong proximity condition  $\tilde{C}_{in}(s) \approx C_{in}(s)$  holds true.

3.6. *The Design of the External Controller C(s).* Controller  $C(s)$  design technique is the same as for the internal controller  $K(s)$ , where the external characteristic polynomial  $\tilde{C}_{out}$  is

$$\begin{aligned} & A_0(s) A'(s) A_m(s) R_K(s) R_C(s) \\ & + B_0(s) L_K(s) A'(s) A_m(s) R_C(s) \\ & + B_0(s) B'(s) L_C(s) A_m(s) R_K(s) \\ & + B_0(s) B'(s) B_m(s) L_K(s) L_C(s) = \tilde{C}_{out}(s). \end{aligned} \quad (15)$$

The possible controller parameterization  $C(s)$  is

$$\Gamma(s) R'_C(s) \prod_{k=1}^{\tilde{p}} (s + \eta_k) + Y(s) L_C(s) = \tilde{C}_{out}(s), \quad (16)$$

$$\Gamma(s) R_C(s, \eta) + Y(s) L_C(s) = \tilde{C}_{out}(s),$$

where polynomials  $\Gamma(s)$  and  $Y(s)$  are

$$\begin{aligned} \Gamma(s) &= A'(s) A_m(s) (A_0(s) R_K(s) + B_0(s) L_K(s)), \\ Y(s) &= B_0(s) B'(s) (A_m(s) R_K(s) + B_m(s) L_K(s)). \end{aligned} \quad (17)$$

Coefficient  $\eta$  can also be used as an additional parameter for criteria optimization in the same sense as parameter  $\delta$ .

The controller  $C(s)$  is the proper solution of the optimization problem for external loop if the following conditions are satisfied:

$$\begin{aligned} \Psi_{out} &= (J_{out1} \cap J_{out2} \cap J_{out3}) \wedge \min(\text{dev}(p_{out}, \tilde{p}_{out})) \\ & \quad \forall \Psi_{out} \in D_{out}, \\ \forall p_{out} &= \{p_{out} \in C_{out} \mid C_{out} \in D_{out}\}, \\ \forall \tilde{p}_{out} &= \{\tilde{p}_{out} \in \tilde{C}_{in} \mid \tilde{C}_{in} \in D_{out}\}. \end{aligned} \quad (18)$$

Objective functions  $J_{out1}-J_{out3}$  have the same properties and meanings as objective functions  $J_1-J_3$ .

#### 4. Nonparametric Uncertainty and Robustness Criterion of RIC

The robustness of the RIC system is assessed using uncertainty models  $\Delta P$  and a stable proper input weight  $V'_{noise}$ .

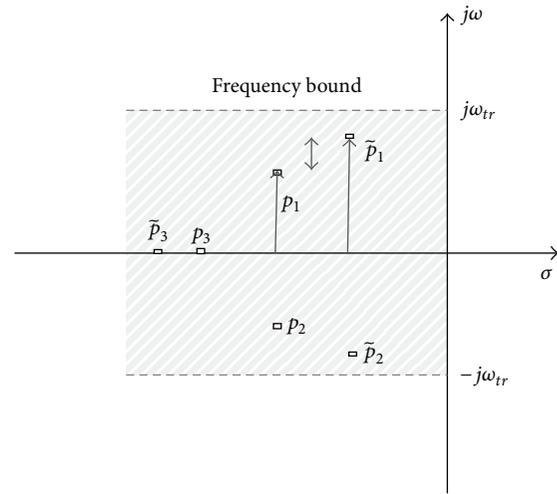


FIGURE 3: Rise-time objective function with a horizontal boundary  $\pm j\omega_{tr}$ .

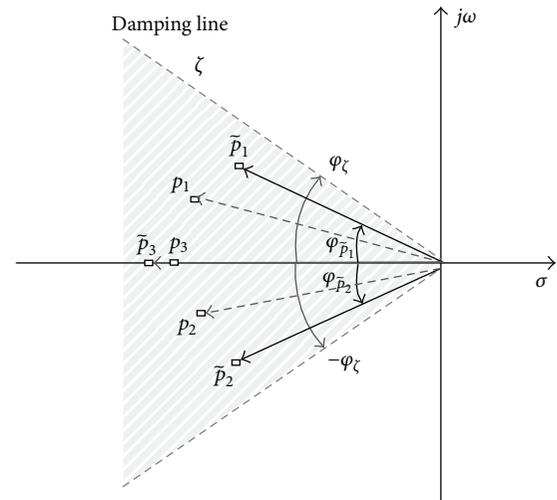


FIGURE 4: Damping-ratio objective function with parameter  $\pm\phi_\zeta$ .

The uncertainty models describe model deviation within the given frequency space and are represented with a nominal model and stable uncertainty weights  $\Delta W$  [30, 39]. Weights  $V'_{noise}$  represent measured noise spectrum of the signal  $w'_1$ . For RIC design, we assume that the reference model is  $P_m \approx P_0 \in PH_\infty$ . The internal loop of the RIC with uncertainty models is shown in Figure 6.

The robustness assessment of the internal loop should be considered with multiplicative and inverse uncertainty. The multiplicative uncertainty represents uncertainties by lower frequencies, while the inverse uncertainty represents uncertainties by higher frequencies [39].

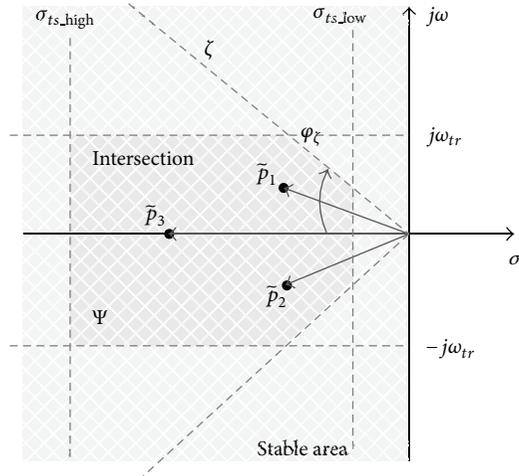


FIGURE 5: Intersection of objectives  $J_1$ - $J_3$ .

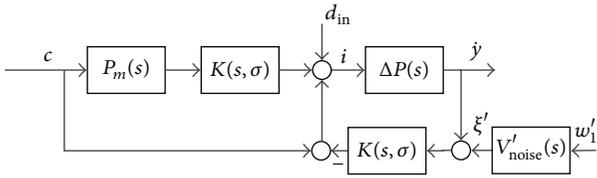


FIGURE 6: Internal loop with uncertainty models  $\Delta P$ .

Let us consider the multiplicative uncertainty model  $\Delta P_M = P_0(1 + \Delta W_M)$  with nominal plant  $P_0$ . The closed-loop characteristic using the multiplicative uncertainty model  $\Delta P_M$  is

$$\begin{pmatrix} \tilde{y} \\ \tilde{i} \end{pmatrix} = (1 + KP_0(1 + \Delta W_M))^{-1} \times \begin{bmatrix} P_0(1 + \Delta W_M)(1 + KP_m) & P_0(1 + \Delta W_M) \\ 1 + KP_m & 1 \end{bmatrix} \begin{pmatrix} c \\ d_{in} \end{pmatrix}. \quad (19)$$

Robust stability for the multiplicative uncertainty is preserved if the following holds true:

$$\begin{aligned} \left\| \Delta W_M \frac{KP_0}{1 + KP_0} \right\|_{\infty} < 1 \\ \forall |\Delta(\omega)| < 1 \quad \omega = \{\omega \in \mathbb{R} \mid 0 \leq \omega < \infty\}, \quad (20) \\ \left\| \Delta W_M T_{in} \right\|_{\infty} < 1, \end{aligned}$$

where  $T_{in}$  is a complementary sensitivity function of the internal loop.

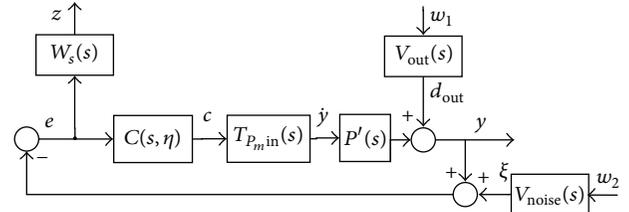


FIGURE 7: External loop optimization structure.

The internal loop with inverse uncertainty  $\Delta P_I = P_0(1 + \Delta W_I)^{-1}$  is defined as

$$\begin{aligned} \begin{pmatrix} \tilde{y} \\ \tilde{i} \end{pmatrix} &= (1 + KP_0(1 + \Delta W_I)^{-1})^{-1} \\ &\times \begin{bmatrix} P_0(1 + \Delta W_I)^{-1}(1 + KP_m) & P_0(1 + \Delta W_I)^{-1} \\ 1 + KP_m & 1 \end{bmatrix} \quad (21) \\ &\times \begin{pmatrix} c \\ d_{in} \end{pmatrix}. \end{aligned}$$

The robustness for inverse models is satisfied if the following holds true:

$$\begin{aligned} \left\| \Delta W_I \frac{1}{1 + KP_0} \right\|_{\infty} < 1 \\ \forall |\Delta(\omega)| < 1 \quad \omega = \{\omega \in \mathbb{R} \mid 0 \leq \omega < \infty\}, \quad (22) \end{aligned}$$

where  $S_{in}$  is the sensitivity function of the internal loop.

The noise suppression  $\xi'$  of the internal loop can be assessed with the following criterion:

$$T_{w'_1} = (1 + K(\delta)\Delta P)^{-1}V'_{noise}. \quad (23)$$

The robustness criterion of noise suppression with the uncertainty model is

$$T_{w'_1} = \min_K \left\| \begin{bmatrix} W_M & W_I \end{bmatrix} \begin{bmatrix} T_{in} & 0 \\ 0 & S_{in} \end{bmatrix} V'_{noise} \right\|. \quad (24)$$

After derivation of robust stability conditions of the internal loop, robust stability conditions for the external loop can be presented. The purpose of external controller design is to ensure overall performance characteristics like proper reference tracking  $r$ , output disturbance rejection  $d_{out}$ , and good measurement noise cancellation  $\xi$ . The optimization structure of the external loop is shown in Figure 7.

Selected weights  $W_s$ ,  $V_{out}$ ,  $V_{noise}$  belong to the  $PH_{\infty}$  domain. Input weights  $V_{out}$ ,  $V_{noise}$  represent the input spectral characteristics of disturbance  $d_{out}$  and noise  $\xi$ , where the performance weight  $W_s$  is selected to ensure a smooth frequency characteristic of the external-loop sensitivity function  $S_{out}(s)$ . A smooth characteristic of the sensitivity function is also an additional indicator of the closed-loop time-performance characteristics.

The closed-loop characteristic of the external loop with weights  $W_s, V_{out}, V_{noise}$  is

$$T_{zw} = W_s^{-1} \left[ (1 + C(\eta) T_{in} P')^{-1} (1 + C(\eta) T_{in} P')^{-1} (1 + C(\eta) T_{in} P')^{-1} \right] \begin{bmatrix} 1 \\ V_{out} \\ V_{noise} \end{bmatrix} \tag{25}$$

$$= W_s^{-1} \begin{bmatrix} S_{out} & S_{out} & S_{out} \end{bmatrix} \begin{bmatrix} 1 \\ V_{out} \\ V_{noise} \end{bmatrix}.$$

The optimal solution of the optimization problem is

$$\min_C \|T_{zw}\|_{\infty} = \gamma_{min}. \tag{26}$$

The main goal of expression (26) is to minimize the influences of inputs  $w_1, w_2$  on the sensitivity function  $S_{out}$  and to ensure final feedback performance with  $S_{out} \approx W_s \wedge |S_{out}| < |W_s|$ .

*4.1. Robustness Criteria Assessment via Nonnegativity of the Even Polynomial.* Let us consider the property of the norm  $\|\cdot\|_{\infty}$  related to robust stability with uncertainty models [30, 39]. The robust stability criterion is ensured with condition  $\|\cdot\|_{\infty} < 1$ , where the criterion is derived for a given transfer function  $P_0(s) = B(s)A^{-1}(s)$ . The condition  $\|P_0\|_{\infty} < 1$  holds if the even polynomial  $\pi(\omega^2)$  is a strictly nonnegative function,

$$A(\omega^2) - B(\omega^2) = \pi(\omega^2) > 0. \tag{27}$$

The property of the function  $\pi(\omega^2)$  is derived from the function  $\Phi(j\omega)$ , defined as in [30]:

$$\Phi(j\omega) = \gamma^2 I - \frac{B(j\omega) B(-j\omega)}{A(j\omega) A(-j\omega)}. \tag{28}$$

Function  $\Phi(j\omega)$  is a strictly continuous function for all  $\omega \in \mathbb{R} \cup \{\infty\}$  and has no imaginary zeroes. The norm of the function equals  $\|P_0\|_{\infty} \leq \gamma$  only if  $\Phi(j\omega) > 0$  for all  $\omega \in \mathbb{R}$ . The robustness property for nonparametric uncertainty can be assessed with even polynomial  $\pi(\omega^2) > 0$  (27), where we assume that  $\gamma = 1$  and  $A(j\omega)A(-j\omega) = A(\omega^2)$ ,  $B(j\omega)B(-j\omega) = B(\omega^2)$  holds true.

**Theorem 1.** *The norm  $\|\cdot\|_{\infty}$  of the system with polynomials  $A, B$  is  $\|B/A\|_{\infty} < 1$  only if the corresponding polynomial  $\pi(\omega^2)$ , for all  $\omega$ , is a strictly nonnegative function and  $B(s)A(s)^{-1}$  belongs to  $PH_{\infty}$ .*

*Proof.* The simple explanation of the norm  $\|\cdot\|_{\infty}$  is  $\|B/A\|_{\infty} := \sup_{\omega} |B(j\omega)A^{-1}(j\omega)|$ , where we assume that  $B(s)A(s)^{-1}$  belongs to  $PH_{\infty}$ . The norm of the transfer function is  $\|B/A\|_{\infty} < 1$  only if the ratio of polynomials  $|B(j\omega)|/|A(j\omega)|^{-1} < 1$ . It is obvious from the above-mentioned that the difference (27),  $\pi(\omega^2)$ , must be a strictly

nonnegative function and that  $\pi(\omega^2)$  has no real zeroes. The positivity condition of the  $\pi(\omega^2)$  is an objective of robustness assessment.  $\square$

Based on condition (27), the controllers' robustness property can be achieved with the assessment of the nonnegativity of the even polynomial. Each single robustness criterion ((20), (22), (24), and (26)) can be presented in the form of an even polynomial and a nonnegativity condition.

*4.2. Robust Stability Assessment with a Nonnegativity Condition for the Internal Loop.* The internal-loop robustness conditions are presented with expressions (20), (22), and (24). Based on condition (27), the even polynomial  $\pi_M(\omega^2, \delta)$  for multiplicative uncertainty (20) is

$$\begin{aligned} \pi_M(s, \delta) &= \widetilde{C}_{in}(s) \widetilde{C}_{in}(-s) w_{aM}(s) w_{aM}(-s) \\ &\quad - B_0(s) B_0(-s) L_K(s, \delta) L_K(-s, \delta) w_{bM}(s) \\ &\quad \times w_{bM}(-s), \\ \pi_M(\omega^2, \delta) &= \widetilde{C}_{in} w_{aM}(\omega^2) - B_0 L_K w_{bM}(\omega^2, \delta) > 0, \end{aligned} \tag{29}$$

where the weight  $W_M(s)$  is  $W_M(s) = w_{bM}(s)w_{aM}^{-1}(s)$ .

Even polynomial  $\pi_I(\omega^2, \delta)$  for robust stability with inverse uncertainty model (22) with stable weight  $W_I(s) = w_{bI}(s)w_{aI}^{-1}(s)$  is

$$\begin{aligned} \pi_I(s, \delta) &= \widetilde{C}_{in}(s) \widetilde{C}_{in}(-s) w_{aI}(s) w_{aI}(-s) \\ &\quad - A_0(s) A_0(-s) R_K(s, \delta) R_K(-s, \delta) w_{bI}(s) \\ &\quad \times w_{bI}(-s), \\ \pi_I(\omega^2, \delta) &= \widetilde{C}_{in} w_{aI}(\omega^2) - A_0 R_K w_{bI}(\omega^2, \delta) > 0. \end{aligned} \tag{30}$$

Even polynomial  $\pi_{T_{w'_1}}(\omega^2, \delta)$  for condition (24) with stability weights  $W_M(s)$  and  $W_I(s)$  and input weight  $V'_{\text{noise}}(s) = v'_{\text{bnoise}}(s)v'^{-1}_{\text{anoise}}(s)$  is

$$\begin{aligned} \pi_{T_{w'_{1M}}}(\omega^2, \delta) &= w_{aM}\tilde{C}_{\text{in}}v'_{\text{anoise}}(\omega^2) \\ &\quad - w_{bM}A_0A_mR_Kv'_{\text{bnoise}}(\omega^2, \delta), \\ \pi_{T_{w'_{1S}}}(\omega^2, \delta) &= w_{aS}\tilde{C}_{\text{in}}v'_{\text{aout}}(\omega^2) \\ &\quad - w_{bS}A_0A_mR_Kv'_{\text{bout}}(\omega^2, \delta), \\ \pi_{T_{w'_1}}(\omega^2, \delta) &= 0.5 \left( \pi_{T_{w'_{1M}}}(\omega^2, \delta) + \pi_{T_{w'_{1S}}}(\omega^2, \delta) \right) > 0. \end{aligned} \tag{31}$$

To simplify the optimization procedure in comparison to the classic approach, the composed condition (24) is treated similarly as the criterion in augments plant transformation in a classic  $H_\infty$  design, wherein the condition  $\pi_{T_{w'_1}}(\omega^2, \delta)$  (31) is a simple even polynomial function. In the multi-objective optimization approach, the condition  $\pi_{T_{w'_1}}(\omega^2, \delta)$  (31) can be also considered as two distinct functions:  $\pi_{T_{w'_{1M}}}(\omega^2, \delta)$  and  $\pi_{T_{w'_{1S}}}(\omega^2, \delta)$ .

**4.3. Robust Stability Assessment with a Nonnegativity Condition for the External Loop.** A polynomial function for robustness and performance condition (26) for the external loop in Figure 7 can be formulated the same way as conditions (24) and (31). The even polynomial of condition (26) with performance weights  $W_S(s) = w_{bS}(s)w_{aS}^{-1}(s)$ ,  $V_{\text{out}}(s) = v_{\text{bout}}(s)v_{\text{aout}}^{-1}(s)$ , and  $V_{\text{noise}}(s) = v_{\text{bnoise}}(s)v_{\text{anoise}}^{-1}(s)$  is

$$\begin{aligned} \pi_{w_1z}(\omega^2, \eta) &= w_{aS}C_{\text{out}}(\omega^2) - w_{bS}A_0A_mA'R_KR_C(\omega^2, \eta), \\ \pi_{w_2z}(\omega^2, \eta) &= w_{aS}C_{\text{out}}v_{\text{aout}}(\omega^2) \\ &\quad - w_{bS}A_0A_mA'R_KR_Cv_{\text{bout}}(\omega^2, \eta), \\ \pi_{zw}(\omega^2, \eta) &= \left( \pi_{w_1z}(\omega^2, \eta)^2 + \pi_{w_2z}(\omega^2, \eta)^2 \right) > 0. \end{aligned} \tag{32}$$

As with condition (31), we can use composite objective function  $\pi_{zw}(\omega^2, \eta)$  to simplify the optimization procedure.

Additional criteria are imposed to achieve strong stability of the RIC, where the real-time operational safety for the controlled system must be preserved in case some parts of the feedback system fail, for example, sensors' failure, electronic driver failure, and faulty motor. Controllers' stability can be assessed over strict positive realness (SPR) criteria [40]. Controllers  $K(s)$ ,  $C(s)$  are stable transfer functions if SPR's conditions hold true:

$$\begin{aligned} \Re [K(j\omega) + K(-j\omega)] &> 0, \\ \Re [C(j\omega) + C(-j\omega)] &> 0, \end{aligned} \tag{33}$$

where  $\Re[K(j\omega)]$  and  $\Re[C(j\omega)]$  are even functions. The even polynomials of stability conditions (33) are

$$\begin{aligned} \pi_{K_{\text{SPR}}}(\omega^2, \sigma) &= 2(R_K(-j\omega, \sigma)L_K(j\omega) + R_K(j\omega, \sigma)L_K(-j\omega)) > 0, \\ \pi_{C_{\text{SPR}}}(\omega^2, \eta) &= 2(R_C(-j\omega, \eta)L_C(j\omega) + R_C(j\omega, \eta)L_C(-j\omega)) > 0. \end{aligned} \tag{34}$$

The robust conditions and the stability property of controllers (29)–(32), (34) can be directly used in a multiobjective optimization algorithm with DE.

### 5. Multiobjective Optimization with DE for RIC Design

The paper presents a multiobjective optimization procedure with DE. Genetic algorithm DE is known as a very efficient and powerful stochastic optimization procedure. DE includes similar operation steps as other GAs, but, in comparison with other classic GAs, DE deviates the current population with scaled differences between two randomly selected members [28]. There are many reasons for using the DE algorithm for solving various optimization problems in different scientific disciplines. Compared to other similar algorithms, DE has a simpler structure and is easier to implement on real problems. Because of its simple structure, it is suitable for large scale optimization problems. The control parameters of DE have been well studied, and their influence on the optimization procedure is well known.

This paper presents the design of a robust, efficient, simple, and structured disturbance observer, based on simple objective functions for performance and robustness assessment. The main problem in optimal control design is formulating proper efficient objective functions, which can be further used in a corresponding optimization algorithm. Objective functions with their properties must ensure optimal or suboptimal solutions of the given problem. The given optimization problem in RIC structure with regional pole placement technique does not provide a convex set of objective functions. There also do not exist general straightforward procedures to convert such objectives to convex problem, which are mostly in conflict with each other and unrelated. For example, unrelated criteria are strong stability and robustness, closed-loop pole location and controller stability, controller structure and robustness performance, and so forth. In such cases where engineering simplifications are needed, it is very appropriate to use a metaheuristic optimization tool, like DE. For this reason, this paper does not deal with the efficiency of multiobjective DE algorithms and their variants but only considers the efficiency of the used approach in an optimal control problem in an RIC framework with formed objective functions for dynamic properties (7)–(9) and robustness (29)–(32) and (34).

Multiobjective optimization problems with DE involve multiple objectives, which must be optimized simultaneously,

- (1) Select central polynomials  $C_{in}(s), C_{out}(s)$  and weights  $W(s), V(s)$ .
- (2) Select controllers' structure  $K, C$  and parameters  $\sigma, \eta$ .  
 $\deg K = \deg C_{in} - \deg P_0$ ,  
 $\deg R_K = \deg L_K$ ,  
 $\deg C = \deg C_{out} - (\deg C_{in} + \deg P_m + \deg P')$ ,  
 $\deg R_C = \deg L_C$ .
- (3) Parameters' selection of the DE-optimization algorithm.  
*(Number of parents-NP, differential weight-F, crossover probability-CR, mutation strategy)*
- (4) **while** (min  $f(\omega^2, X)$ )
- (5) DE selects value of parameters  $\sigma, \eta$ .
- (6) Solve polynomial equations (14), (16).
- (7) Derive convex polynomials  $\pi(\omega^2, X)$ .
- (8) Find current Pareto-front of the  $f(\omega^2, X)$ .
- (9) **end while**

ALGORITHM 1: Multiobjective DE algorithm.

and a set of possible solutions must be obtained. The set of possible solutions is evaluated based on the concept of dominance and Pareto optimality [14, 17, 33]. The presented RIC design uses DEMO algorithm presented by Robič and Filipič [42]. The DEMO combines the advantages of DE with the mechanisms of Pareto-based ranking and crowding distance sorting. The advantage of the DEMO algorithm is that it ensures convergence to the Pareto-front and a uniform spread of individuals along the front [42].

A multiobjective function is composed of derived objective functions (7)–(9), (29)–(32), and (34) and is equal to

$$\min_{\omega^2 \vee X} f(\omega^2, X) = \begin{bmatrix} J_1(X) \wedge J_{out1}(X) \\ J_2(X) \wedge J_{out2}(X) \\ J_3(X) \wedge J_{out3}(X) \\ \pi_M(\omega^2, X) \\ \pi_I(\omega^2, X) \\ \pi_{T_{w'_1}}(\omega^2, X) \\ \pi_{zw}(\omega^2, X) \\ \pi_{K_{SPR}}(\omega^2, X) \\ \pi_{C_{SPR}}(\omega^2, X) \end{bmatrix}, \quad (35)$$

$$\text{s.t. } f(\omega^2, X) > 0,$$

$$X \in P,$$

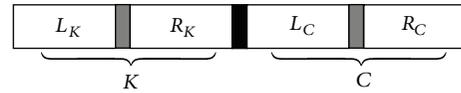
where  $P$  is parameter space and  $X$  is decision vector. Decision vector  $X$  contains the coefficients of polynomials  $L_K, R_K, L_C, R_C$  (1) with possible parameterization coefficients  $\delta$  and  $\eta$ . Figure 8 shows the structure of the decision variable  $X$ .

*Note.* The solution  $X \in P$  is Pareto-optimal if and only if there does not exist  $\tilde{X} \in P$  that satisfies  $f(\omega^2, X) < f(\omega^2, \tilde{X})$ .

The optimization algorithm starts with the preselected central polynomials  $C_{in}, C_{out}$ , where the controllers' degrees are prescribed in the same way as in the classic polynomial design (6), (16) [37]. The degree of controller  $K$  is equal to the condition  $\deg K = \deg C_{in} - \deg P_0$ , where  $\deg C_{in} > \deg P_0$  holds true. The degree of controller  $C$  can be determined

TABLE 1: Parameters of an electromechanical system.

Motor parameters	
$J$	$6.2 \cdot 10^{-3} \text{ kgm}^2$
$B$	$4.2 \cdot 10^{-3} \text{ Nms}$
$k_p$	$0.081 \text{ Nm/A}$
$R$	$7.9 \Omega$
$L$	$5.7 \text{ mH}$

FIGURE 8: Structure of decision variable  $X$ .

the same way as for controller  $K$ , where  $\deg C = \deg C_{out} - (\deg C_{in} + \deg P_m + \deg P')$  and  $\deg C_{out} \geq (\deg C_{in} + \deg P_m + \deg P')$  holds true. The optimization procedure is shown in Algorithm 1.

## 6. The Design Procedure for RIC Controllers

The proposed robust motion controller design is demonstrated by an electromechanical positioning system with the parameters given in Table 1.

The parameters  $J, B, k_p, R,$  and  $L$  are rotor inertia, viscous friction, torque factor, terminal resistance, and rotor induction, respectively. To ensure simple controller structures of  $K(s)$  and  $C(s)$ , we use a simplified model of the mechanical system  $P_0(s)$ . The simplification is often applicable if the plant has more strongly expressed dominant stable poles in comparison to other stable poles. The given model is

$$P_0(s) = \frac{\omega(s)}{i(s)} = \frac{k_p}{Js + B} = \frac{130.6}{s + 0.667}, \quad (36)$$

$$P'(s) = \frac{\varphi(s)}{\omega(s)} = \frac{1}{s}.$$

TABLE 2: Uncertainty parameters of the positioning system.

Uncertainty parameters	
$\Delta J$	$[2.4 \div 9.8] \cdot 10^{-3} \text{ kgm}^2$
$\Delta B$	$[3.3 \div 6.5] \cdot 10^{-3} \text{ Nms}$
$\Delta k_p$	$[0.6 \div 1.3] \text{ Nm/A}$

TABLE 3: RIC control requirements.

Feedback loop RIC requirements	
Settling time	$t_s < 1.4 \text{ s}$
Nominal plant overshooting	$M_{pn} < 3\%$
Worst-case plant overshooting	$M_{\text{worst}} < 20\%$
Tracking accuracy	$e_{\text{error}} <  0.0015  \text{ rad}$
Tracking reference frequency	$[0 \div 1] \text{ rad/s}$
Input disturbance rejection	$[0.05] \text{ Nm}, [0 \div 0.2] \text{ rad/s}$
Output disturbance rejection	$[0.5] \text{ rad}, [0 \div 0.5] \text{ rad/s}$
Input current limit	$\pm 1.5 \text{ A}$
Voltage limit	$\pm 14.8 \text{ V}$
Robust stabilization	
Stable and low-order controllers	$K, C$

The output of the controller  $K(s)$  is an armature current  $i[A]$  through a magnetic coil, which is proportional to the electrical torque. To avoid additional nonlinearities, it is sensible to consider terminal voltage  $|Ri + k_e \omega| < |U_{\text{limit}}|$ , especially if the system is driven conventional by an H-bridge and PWM signal. The coefficient  $k_e$  is an electromotive force constant.

The uncertainty plant is given by

$$\Delta P(s) = \frac{\Delta k_p}{\Delta J s + \Delta B}, \quad (37)$$

where the parameters vary in intervals according to changed operation points, friction, load, gear box, and so forth. The uncertainty parameter intervals are shown in Table 2.

The estimated uncertainty weights for the robustness criteria (20), (22), and (24) are

$$\begin{aligned} W_M(s) &= \frac{1.34 \times s^3 + 1.156 \times s^2 + 0.32 \times s + 0.062}{s^3 + 1.57 \times s^2 + 0.48 \times s + 0.096}, \\ W_I(s) &= \frac{0.65 \times s^2 + 0.39 \times s + 0.083}{s^2 + 0.82 \times s + 0.17}, \\ V'_{\text{noise}}(s) &= \frac{0.0023 \times s + 1.2 \cdot 10^{-6}}{s + 980.3}. \end{aligned} \quad (38)$$

The desired requirements of the RIC systems are given in Table 3.

The reference model  $P_m(s)$  is selected so as to ensure proper internal dynamic of the RIC structure, where holds  $P_m(s) \approx P_0(s)$ . The selected  $P_m(s)$  is

$$P_m(s) = \frac{289.2}{s + 1.5}. \quad (39)$$

TABLE 4: Allowed region for optimized internal-loop poles  $\tilde{C}_{in}(s)$  in a complex plain.

Allowed poles region $D_{in}$ for $\tilde{C}_{in}$	
$\sigma_{ts,low}$	-0.7
$\sigma_{ts,high}$	-4
$\pm j\omega_{tr}$	$\pm j10$
$\pm \varphi_\zeta$	$\pm 30^\circ$

TABLE 5: Allowed region for optimized internal-loop poles  $\tilde{C}_{out}(s)$  in a complex plain.

Allowed poles region $D_{out}$ for $\tilde{C}_{out}$	
$\sigma_{ts,low}$	-0.35
$\sigma_{ts,high}$	-9
$\pm j\omega_{tr}$	$\pm j14$
$\pm \varphi_\zeta$	$\pm 74^\circ$

According to the RIC dynamic requirement for input and output disturbance rejection and tracking property, the additional performance weights are selected as follows:

$$\begin{aligned} W_S(s) &= \frac{2.1 \times s^2 + 0.46 \times s + 0.001}{s^2 + 3.98 \times s + 0.99}, \\ V_{out}(s) &= \frac{0.002 \times s + 0.0012}{s + 0.0014}, \\ V_{noise}(s) &= \frac{2.9 \cdot 10^{-3} \times s + 1.2 \cdot 10^{-3}}{s + 1.001 \cdot 10^3}. \end{aligned} \quad (40)$$

The controller transfer function  $K(s)$  is selected so as to ensure simple structure and maintain desirable closed-loop performance. The selected low-order structure is

$$K(s, \tilde{r}_{k0}) = \frac{l_{k1}s + l_{k0}}{r_{k1}s + \delta}, \quad (41)$$

where the parameter  $\delta$  represents controller parameterization (14) and the allowable desired value is given on the interval  $[10^{-4} - 10^{-2}]$ . The value of  $\delta$  ensures proper input disturbance rejection for low-frequency signals and stable approximate integral behaviour.

Accordingly, the following internal central polynomial  $C_{in}(s)$  is chosen on the selected controller structure  $K(s)$  on the condition  $\deg C_{in} = \deg K + \deg P_0$ . The internal central polynomial is

$$C_{in}(s) = s^2 + 2 \times s + 1. \quad (42)$$

The polynomial  $C_{in}(s)$  ensures proper dynamic and stability of the internal system. The selected allowed region  $D_{in}$  of the optimized internal loop poles  $\tilde{C}_{in}(s)$  is presented in Table 4.

The selected controller structure  $C(s)$  is

$$C(s, \tilde{r}_{c0}) = \frac{l_{c2}s^2 + l_{c1}s + l_{c0}}{r_{c2}s^2 + r_{c1}s + \eta}. \quad (43)$$

TABLE 6: Parameters of DE-optimization algorithm.

Optimization algorithm parameters	
Number of parents NP	50
Differential weight F	0.85
Crossover probability CR	0.92
Mutation strategy	DE/rand/1/bin

TABLE 7: Polynomials  $C_{in}$ ,  $\tilde{C}_{in}$  roots comparisons.

Polynomials	$C_{in}(s)$	$\tilde{C}_{in}(s)$
	-1	$-0.944 - j0.04$
	-1	$-0.944 + j0.04$

TABLE 8: Polynomials  $C_{out}$ ,  $\tilde{C}_{out}$  roots comparisons.

Polynomials	$C_{out}(s)$	$\tilde{C}_{out}(s)$
	$-3.4 - j12.3$	$-3.09 - j10.2$
	$-3.4 + j12.3$	$-3.09 + j10.2$
	$-2.5 - j3$	$-2.61 - j2.76$
	$-2.5 + j3$	$-2.61 + j2.76$
	-1.2	-1.09
	-0.5	-0.48

The parameter  $\eta$  represents  $C(s)$  controller parameterization in such a way that the double-integrator effect in the external loop is avoided. The admissible value is  $\eta > 50$ . According to controller structures  $K(s)$  and  $C(s)$  and the reference model  $P_m(s)$ , the external central polynomial is selected as

$$C_{out}(s) = s^6 + 13.5 \times s^5 + 234 \times s^4 + 1286 \times s^3 + 4171 \times s^2 + 4773 \times s + 1490, \quad (44)$$

where the following condition holds true:  $\deg C_{out} = \deg C + (\deg C_{in} + \deg P_m + \deg P')$ . The selected allowed region  $D_{out}$  of the optimized external-loop poles  $\tilde{C}_{out}(s)$  is presented in Table 5.

The structure of the decision variable  $X$  for the given example is shown in Figure 9.

The selected parameters of DE algorithm are presented in Table 6.

### 7. Results

Optimized controllers  $K(s)$  and  $C(s)$  after optimization with DE and objective function (35) are as follows:

$$K(s) = \frac{13.56 \cdot 10^{-3} \times s + 8.4 \cdot 10^{-3}}{s + 0.18 \cdot 10^{-3}}, \quad (45)$$

$$C(s) = \frac{0.27 \times s^2 + 2.3 \times s + 2.29}{s^2 + 9.324 \times s + 102.1}.$$

The optimization results are presented below. The difference between the selected internal central polynomial  $C_{in}(s)$  and the optimized polynomial  $\tilde{C}_{in}(s)$  is presented in Table 7.

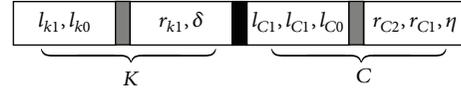


FIGURE 9: The structure of the decision variable  $X$  with 10 parameters.

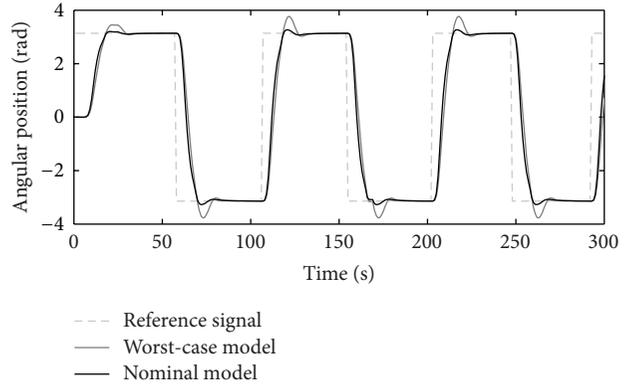


FIGURE 10: Step-reference tracking.

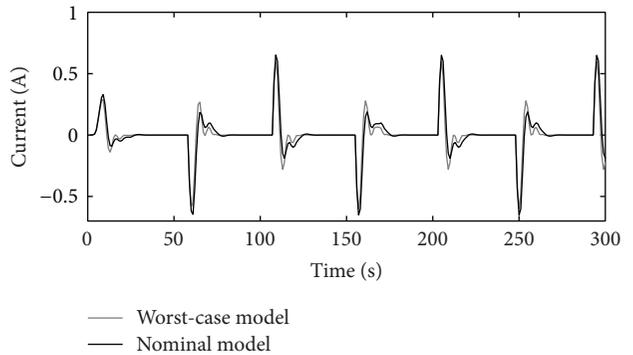


FIGURE 11: Output of the controller  $K(s)$ .

A comparison of polynomials  $C_{out}(s)$  and  $\tilde{C}_{out}(s)$  is presented in Table 8.

From the results of polynomial comparisons in Tables 7 and 8, it is evident that the optimization procedure with Pareto-optimal solution ensures a satisfactory fitting of pole positions in the dominant region. The obtained controllers  $K(s)$  and  $C(s)$  are stable, and all poles of internal and external loops lie in the prescribed region.

The final values of robust criteria (20), (22), (24), and (26) are presented in Table 9.

The tracking RIC capabilities on step-reference signals for the nominal and worst-case systems are shown in Figure 10. The reference signal presents a rotation of the RIC system for half a turn to the left and to the right. The worst-case model is selected accordingly, as a possible real operational case with values  $\max(\Delta J)$ ,  $\max(\Delta B)$ , and  $\min(\Delta k_p)$ , chosen from Table 2. Controllers'  $K$  and  $C$  outputs are shown in Figures 11, 12, and 13, respectively.

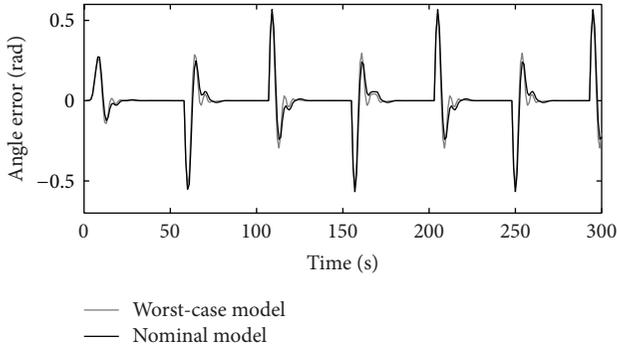


FIGURE 12: Output of the controller  $C(s)$ .

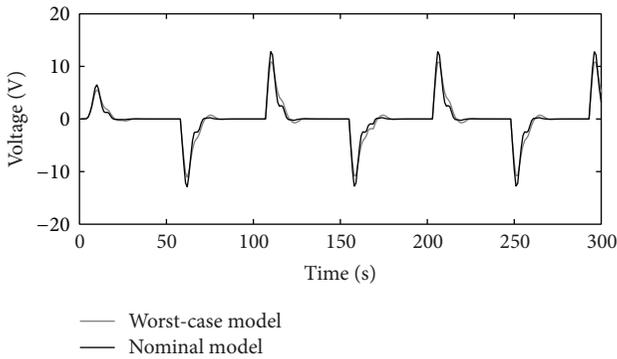


FIGURE 13: Output-voltage of the RIC system.

The disturbance rejection capability of the positioning system is shown in Figures 14, 15, 16, and 17.

Figures 10–17 show that the robust motion controller presented in the RIC framework satisfies all control design criteria. Table 9 and Figures 10–17 provide evidence that the stability and performance conditions are preserved. The system does not exceed the limit values for the operation voltage and current in the given operation interval, so that the system does not exhibit additional nonlinear or oscillating behaviour (Figures 11 and 13). The influence of the measured noise is also minimized with conditions  $\|T_{w'}\|_{\infty}$  and  $\|T_{wz}\|_{\infty}$ , Table 9. The system has good reference tracking and disturbance rejection within the prescribed area of system uncertainty (Figures 14–17) and simple low-order structures of the internal and external controllers.

### 8. Conclusion

This paper presented the design of a robust RIC structure for a positioning system. The proposed approach shows the capability of robustness and performance optimization over nonnegativity of an even polynomial, where regional pole placement is used. The even polynomial can be also formulated for other types of uncertainty and performance criteria. Controllers  $K$  and  $C$  can be parameterized approximately by using known characteristics, which allows the possibility of preserving strong stability of the controlled system. Optimization with the multicriterion algorithm, such

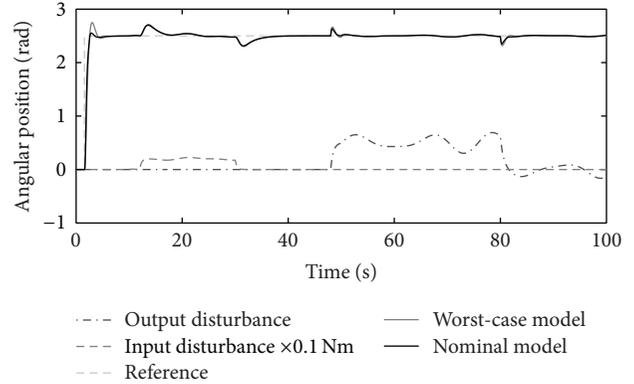


FIGURE 14: Input-output disturbance rejection.

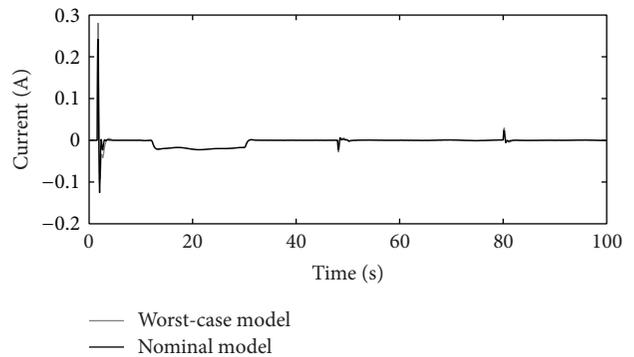


FIGURE 15: Output of the controller  $K(s)$  with disturbance attenuation.

TABLE 9: Value of robust criteria.

Criteria	$\ \cdot\ _{\infty}$	Value
	$\ T_{in}W_M\ _{\infty}$	0.82
	$\ S_{in}W_I\ _{\infty}$	0.62
	$\ T_{w'}\ _{\infty}$	0.23
	$\ S_{out}W_s^{-1}\ _{\infty}$	0.73
	$\ W_s^{-1}S_{out}W_{out}\ _{\infty}$	0.123
	$\ W_s^{-1}S_{out}W_{noise}\ _{\infty}$	0.47

as DE, offers the possibility of including many criteria and an arbitrary number of free parameters, as shown in the presented example. The criteria can include system knowledge, uncertainties, and perturbation characteristics, as well as criteria related to frequency and time domain characteristics of a closed-loop system. The presented results in the design example confirmed the validity of the proposed approach. The DE-optimization procedure with different robustness and performance criteria can be used in a wide range of different controller and feedback structures designs. The presented approach can be straightforward extended to the  $H_2$  or  $H_{\infty}/H_2$  controller design. Further work will be focused on the pole placement robust state space controller design for MIMO system, where the polynomial equation introduces a set of parametric solutions. In MIMO case, the exact solution of the polynomial equation is limited to the number

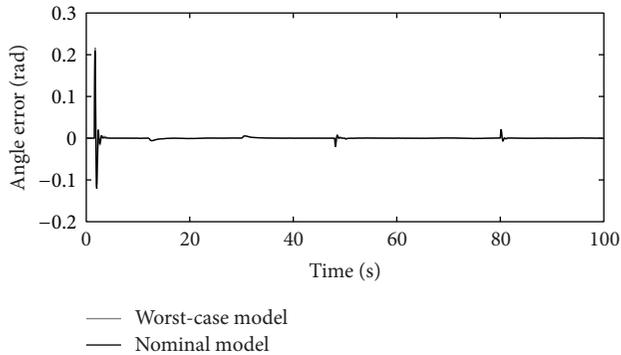


FIGURE 16: Output of the controller  $C(s)$  with disturbance attenuation.

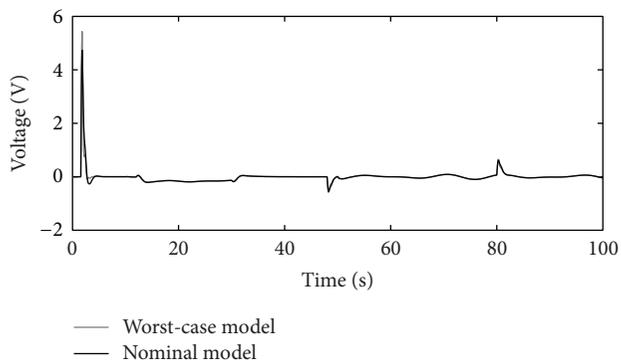


FIGURE 17: Output voltage of the RIC system with disturbance attention.

of the inputs and outputs of the system. The parametric solutions can be used as optimization parameters, similar to the presented approach.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

- [1] Y. Choi, K. Yang, W. K. Chung, H. R. Kim, and I. H. Suh, "On the robustness and performance of disturbance observers for second-order systems," *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 315–320, 2003.
- [2] M. R. Mataušek and A. I. Ribić, "Design and robust tuning of control scheme based on the PD controller plus disturbance observer and low-order integrating first-order plus dead-time model," *ISA Transactions*, vol. 48, no. 4, pp. 410–416, 2009.
- [3] B. Yao, M. Al-Majed, and M. Tomizuka, "High-performance robust motion control of machine tools: an adaptive robust control approach and comparative experiments," *IEEE/ASME Transactions on Mechatronics*, vol. 2, no. 2, pp. 63–76, 1997.
- [4] Y. Wang, Z. H. Xiong, and H. Ding, "Robust internal model control with feedforward controller for a high-speed motion platform," in *Proceedings of the IEEE IRS/RSJ International Conference on Intelligent Robots and Systems (IROS '05)*, pp. 187–192, August 2005.
- [5] B. K. Kim and W. K. Chung, "Unified analysis and design of robust disturbance attenuation algorithms using inherent structural equivalence," in *Proceedings of the American Control Conference*, pp. 4046–4051, June 2001.
- [6] P. B. Stephen, C. A. Hax, and T. L. Magnanti, *Applied Mathematical Programming*, Addison-Wesley, 1997.
- [7] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.
- [8] H. Khatibi, A. Karimi, and R. Longchamp, "Fixed-order controller design for polytopic systems using LMIs," *IEEE Transactions on Automatic Control*, vol. 53, no. 1, pp. 428–434, 2008.
- [9] T. Bakka and H. R. Karimi, "Robust  $H_\infty$  dynamic output feedback control synthesis with pole placement constraints for offshore wind turbine systems," *Mathematical Problems in Engineering*, vol. 2012, Article ID 616507, 18 pages, 2012.
- [10] P. Gahinet and P. Apkarian, "Linear matrix inequality approach to  $H_\infty$  control," *International Journal of Robust and Nonlinear Control*, vol. 4, no. 4, pp. 421–448, 1994.
- [11] L. Jin and Y. C. Kim, "Fixed, low-order controller design with time response specifications using non-convex optimization," *ISA Transactions*, vol. 47, no. 4, pp. 429–438, 2008.
- [12] K. J. Hunt, "Polynomial LQG and  $H_\infty$  infinity controller synthesis: a genetic algorithm solution," in *Proceedings of the 31st IEEE Conference on Decision and Control*, vol. 4, pp. 3604–3609, 1992.
- [13] A. Shenfield and P. J. Fleming, "Multi-objective evolutionary design of robust controllers on the Grid," *Engineering Applications of Artificial Intelligence*, vol. 27, pp. 17–27, 2014.
- [14] A. A. Hussein and R. Sarker, "The pareto differential evolution algorithm," *International Journal on Artificial Intelligence Tools*, vol. 11, no. 4, pp. 531–555, 2002.
- [15] L. Wang and L.-P. Li, "Fixed-structure  $H_\infty$  controller synthesis based on differential evolution with level comparison," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 1, pp. 120–129, 2011.
- [16] K. J. Hunt, "Polynomial LQG and  $H_\infty$  infinity controller synthesis: a genetic algorithm solution," in *Proceedings of the 31st IEEE Conference on Decision and Control*, vol. 4, pp. 3604–3609, 1992.
- [17] A. Chipperfield and P. Fleming, "Gas turbine engine controller design using multiobjective genetic algorithms," in *Proceedings of the 1st IEE/IEEE International Conference on Genetic Algorithms in Engineering Systems: Innovations and Applications (GALESIA '95)*, pp. 214–219, September 1995.
- [18] G. Avanzini and E. A. Minisci, "Evolutionary design of a full-envelope full-authority flight control system for an unstable high-performance aircraft," *Proceedings of the Institution of Mechanical Engineers G—Journal of Ae*, vol. 225, no. 10, pp. 1065–1080, 2011.
- [19] H. Kobayashi, S. Katsura, and K. Ohnishi, "An analysis of parameter variations of disturbance observer for motion control," *IEEE Transactions on Industrial Electronics*, vol. 54, no. 6, pp. 3413–3421, 2007.
- [20] H. S. Lee and M. Tomizuka, "Robust motion controller design for high-accuracy positioning systems," *IEEE Transactions on Industrial Electronics*, vol. 43, no. 1, pp. 48–55, 1996.
- [21] T. Umeno and Y. Hori, "Robust speed control of DC servomotors using modern two degrees-of-freedom controller design," *IEEE Transactions on Industrial Electronics*, vol. 38, no. 5, pp. 363–368, 1991.

- [22] B. K. Kim and W. K. Chung, "Advanced disturbance observer design for mechanical positioning systems," *IEEE Transactions on Industrial Electronics*, vol. 50, no. 6, pp. 1207–1216, 2003.
- [23] K. Kong and M. Tomizuka, "Nominal model manipulation for enhancement of stability robustness for disturbance observer," *International Journal of Control, Automation, and Systems*, vol. 11, no. 1, pp. 12–20, 2013.
- [24] K. Kong, J. Bae, and M. Tomizuka, "Control of rotary series elastic actuator for ideal force-mode actuation in human-robot interaction applications," *IEEE/ASME Transactions on Mechatronics*, vol. 14, no. 1, pp. 105–118, 2009.
- [25] A. Sarja, R. Sveko, and A. Chowdhury, "Strong stabilization servo controller with optimization of performance criteria," *ISA Transactions*, vol. 50, no. 3, pp. 419–431, 2011.
- [26] A. Chowdhury, A. Sarjaš, P. Cafuta, and R. Svečko, "Robust controller synthesis with consideration of performance criteria," *Optimal Control Applications and Methods*, vol. 32, no. 6, pp. 700–719, 2011.
- [27] R. Storn and K. Price, "Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.
- [28] S. Das and P. N. Suganthan, "Differential evolution: a survey of the state-of-the-art," *IEEE Transactions on Evolutionary Computation*, vol. 15, no. 1, pp. 4–31, 2011.
- [29] D. Henrion, M. Šebek, and V. Kučera, "Positive polynomials and robust stabilization with fixed-order controllers," *IEEE Transactions on Automatic Control*, vol. 48, no. 7, pp. 1178–1186, 2003.
- [30] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, Prentice-Hall, New Jersey, NJ, USA, 1997.
- [31] F. Yang, M. Gani, and D. Henrion, "Fixed-order robust  $H_\infty$  controller design with regional pole assignment," *IEEE Transactions on Automatic Control*, vol. 52, no. 10, pp. 1959–1963, 2007.
- [32] M. T. Söylemez and N. Munro, "A note on pole assignment in uncertain system," *International Journal of Control*, vol. 66, no. 4, pp. 487–497, 1997.
- [33] T. Tušar and B. Filipič, "Differential evolution versus genetic algorithm in multiobjective optimization," in *Evolutionary Multi-Criterion Optimization*, vol. 4403 of *Lecture Notes in Computer Science*, pp. 257–271, 2007.
- [34] A. Tesfaye, H. S. Lee, and M. Tomizuka, "A sensitivity optimization approach to design of a disturbance observer in digital motion control systems," *IEEE/ASME Transactions on Mechatronics*, vol. 5, no. 1, pp. 32–38, 2000.
- [35] B. K. Kim, W. K. Chung, H.-T. Choi, I. H. Suh, and Y. H. Chang, "Robust internal loop compensator design for motion control of precision linear motor," in *Proceedings of the IEEE International Symposium on Industrial Electronics (ISIE '99)*, pp. 1045–1050, July 1999.
- [36] M. S. Nasser, J. Poshtan, M. S. Sadegh, and F. Tafazzoli, "Novel system identification method and multi-objective-optimal multivariable disturbance observer for electric wheelchair," *ISA Transactions*, vol. 52, no. 1, pp. 129–139, 2013.
- [37] V. Kučera, "The pole placement equation—a survey," *Kybernetika*, vol. 30, no. 6, pp. 578–584, 1994.
- [38] D. D. Šiljak and D. M. Stipanović, "Robust D-stability via positivity," *Automatica*, vol. 35, no. 8, pp. 1477–1484, 1999.
- [39] J. Doyle, B. A. Francis, and A. Tannenbaum, *Feedback Control Theory*, Macmillan Publishing, New York, NY, USA, 1990.
- [40] H. Khatibi, A. Karimi, and R. Longchamp, "Fixed-order controller design for polytopic systems using LMIs," *IEEE Transactions on Automatic Control*, vol. 53, no. 1, pp. 428–434, 2008.
- [41] R. Storn and K. Price, "Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.
- [42] T. Robič and B. Filipič, "DEMO: differential evolution algorithm for multi-objective optimization," in *Proceeding of 3rd International Conference of Evolution Multi-Criterion Optimization LNCS*, vol. 3410, pp. 520–533, 2005.

## Research Article

# Optimal Location through Distributed Algorithm to Avoid Energy Hole in Mobile Sink WSNs

Li Qing-hua,<sup>1,2</sup> Gui Wei-hua,<sup>1</sup> and Chen Zhi-gang<sup>1</sup>

<sup>1</sup> School of Information Science and Engineering, Central South University, Changsha 410083, China

<sup>2</sup> Institute of Technology, Lishui University, Lishui, Zhejiang 323000, China

Correspondence should be addressed to Li Qing-hua; [netow@163.com](mailto:netow@163.com)

Received 2 January 2014; Accepted 20 February 2014; Published 7 May 2014

Academic Editors: N. Barsoum, V. N. Dieu, and P. Vasant

Copyright © 2014 Li Qing-hua et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In multihop data collection sensor network, nodes near the sink need to relay on remote data and, thus, have much faster energy dissipation rate and suffer from premature death. This phenomenon causes energy hole near the sink, seriously damaging the network performance. In this paper, we first compute energy consumption of each node when sink is set at any point in the network through theoretical analysis; then we propose an online distributed algorithm, which can adjust sink position based on the actual energy consumption of each node adaptively to get the actual maximum lifetime. Theoretical analysis and experimental results show that the proposed algorithms significantly improve the lifetime of wireless sensor network. It lowers the network residual energy by more than 30% when it is dead. Moreover, the cost for moving the sink is relatively smaller.

## 1. Introduction

Wireless sensor networks are a self-organizing network composed of random distribution of intensive, low-cost, integrated sensors with the function units of sensing module, data-processing module, and short-range wireless communication module. By the measurement the surrounding environment such as humidity, temperature, heat, infrared, sonar, radar, and seismic waves with a variety of built-in sensors to sensor network has achieved a full range of monitoring and controlling the physical world. It is an important integral part of long-term planning of next generation Internet and has already been widely used in the past few years [1].

Sensor nodes usually cannot be replaced or reallocated energy in wireless sensor network, and most applications need to ensure long-term monitoring of certain areas (most applications have prespecified lifetime requirements); for example, the application mentioned in [2] requires that the effective monitoring time for the network should be greater than 9 months to extend the life of sensor network and, thus, is of great significance.

However, research to improve the network life is of great challenges. There is a sensor network-specific “energy hole”

phenomenon, which refers to premature death of those nodes in the hotspot. In multihop data collection sensor network, nodes near the sink have to suffer more routing load [3], so the energy consumption level is higher than nodes in other regions. This is known as the hotspot. Study shows that because of the impact of the energy hole, the network residual energy is as high as 90% [4] when the network is out of function.

Different from the general network with static sink, intelligent mobile robots can act as a mobile sink in the network to collect data. When the residual energy near the sink becomes small, sink repeatedly moves to the location with more abundant remaining energy so as to achieve a balanced energy consumption rate among the entire network, avoiding the energy hole and obtaining longer network lifetime.

Despite a lot of research on the mobile sink, different from previous studies, the main contribution of this paper is as follows.

- (1) Many researchers recognize the existence of hotspots near the sink, so mobile sink strategy is to move sink to the areas with the highest remaining energy in order to avoid energy hole near the sink and

achieve the purpose of balance energy consumption. According to our study, when the sink is seated at different locations in the network, power consumption near sink is of different sizes. Energy consumption in different directions is also not the same. In many cases (especially when the sink is at the edge of the network) the power consumption in network-centric side of sink is much higher than the other side near the edge of the network. Meanwhile, on the side near network-centric, even energy consumption in regions far from the sink is higher than that of the near-edge side. Consequently, simple conclusion is that “areas near sink” which have higher energy consumption cannot effectively guide the design of sink mobility, and they are too general and difficult to be implemented in practice.

In this paper, through differential analysis, precise energy consumption of each node, for the first time, is given for arbitrary sink location in the network. This lays the foundation for mobile sink and energy balancing strategy.

- (2) We propose distributed sink mobile strategy. Based on theoretical analysis of network energy consumption, it can estimate the power consumption of the network before sink moves to the new anchor and predict the optimal next location for sink according to the actual energy consumption of current network. In addition, previous studies fail to present a good solution to the cost of the sink migration. The distributed strategies proposed in this paper are different from previous studies. As the sequence of sink anchors does not affect the ultimate energy consumption and network lifetime, we usually do not move sink according to the calculated result. Instead, we compute a couple of anchors prior to moving and then traverse all the locations according to the minimum moving cost in each step in order to minimize the total cost.

The organization of this paper is as follows. Section 2 introduces relevant research. The network model and the problem statement are presented in Section 3. Section 4 introduces characteristics of data forwarding and energy consumption. It is the basis of theoretical research of our paper. Section 5 describes distributed sink mobile strategy. Section 6 discusses performance and experimental comparison. Section 7 is a summary of the whole paper.

## 2. Related Work

There are many existing researches handling energy hole problem. They can be divided into two categories based on the sink mobility: static sink network (for short, static sink) and mobile sink network (for short, mobile sink). Although this paper focuses on mobile sink, solutions to avoid energy hole in static sink can often be used here. Therefore, we first introduce related research with static sink.

- (a) Analysis and evaluation model. Li and Mohapatra [5] first propose a mathematical model to analyze

the energy hole problem. They assume that the network is uniformly distributed in a circle network and then analyze and compare performance of several existing strategies to avoid energy hole from the perspective of network traffic [6]. They point out that hierarchical structure and data compression mechanism are effective for alleviating the energy hole problem, increasing the rate of data collection, in turn, exacerbating the energy hole [5] and effectiveness of increasing number of the nodes is not obvious.

- (b) Node density control strategy. The principle is that the higher energy consumption near the sink causes the energy hole problem, and therefore it is effective to deploy more nodes near the sink in order to reduce the data load of every node, alleviating the impact of energy hole. This is the so-called nonuniform distribution strategy. Another manifestation of node density control strategy is to deploy more nodes in the place where hotspots are possible to appear which is also effective. Such studies can be found in the literature [7–9].
- (c) Adjustable sensing and communication range. The main principle of this type of strategy is that the node transmission power is adjustable; for example, Berkeley Motes node has 100 transmission power levels [10]. Node energy consumption is directly proportional to  $\alpha$  power of transmission distance and therefore uses smaller transmission range in hotspots near sink and larger transmission range in regions far away from the sink which can achieve balanced energy consumption. This strategy can be combined with node density control strategy. For example, Hossain et al. [11] set the interval between nodes near the sink smaller so that the required energy to transmit data is small, set the interval between nodes far away from the sink larger, so the required energy to transmit data is larger. In this way, it can achieve balanced energy consumption among all of the nodes. Such studies can be found in the literature [12, 13].

The above discussion is under the situation where sink is stationary after deployment. However, with the development of intelligent robot, the research of mobile sink has attracted more and more attention [14]. The research in mobile sink can be summarized into the following categories.

- (a) Relay nodes: such method is to use relay node in hotspot to avoid energy hole. Relay nodes can be both stationary and mobile. The role of mobile relay nodes is essentially similar with that of mobile sink. Related research can be found in the literature [15, 16].
- (b) Single mobile sink: in this kind of networks there is only one sink, and it is movable. Luo and Hubaux [17] propose an early solution for mobile sink to solve the unbalanced energy consumption. As the sink can move, the nodes around it keep on changing over times, and thus it can avoid energy hole around the sink. The author proves that in a circular sensor network, putting sink in the center of a circle is

the best way to save energy. Also, he claims that sink moving along the edge of the network can achieve minimum energy consumption.

Luo also puts forward a strategy that mobile sink moves along the anchor (anchor points) to collect data in [18]. The main idea is that when the sink stays in an anchor it collects data and gets the situation of energy consumption over the whole network in order to determine the interval to stay in every anchor.

As the energy consumption rate near the sink is relatively higher, [19] analogizes sink as “lawn” (mower) and the region with higher remaining energy as “grass” where relatively higher. So the strategy of moving the sink is the “lawn” (mower) moving to the region where “grass” is relatively higher repeatedly.

Reference [20] presents a mobile sink trajectory optimization algorithm and the main idea is as follows. At first, the mobile sink moves along a straight line and collects information about network data and energy consumption information. Mobile sink then adjust the trajectory using the latest information which collected in the process of data collection so that the mobile sink moves near the nodes in order to reduce the cost of data communication and thus to form an optimal trajectory of sink. The paper discusses random movement, forecast movement, and the network performance of different nodes of data collection patterns (passive, multihop, and limited multihop). Reference [21] sums up the target of mobile sink strategy into three categories: (1) minimizing average energy consumption; (2) minimizing largest energy consumption; (3) minimizing relative energy consumption and proposes corresponding mobile algorithms. More research on single mobile sink can be found in [22–24].

- (c) Multiple mobile sinks: compared with single mobile sink, multiple mobile sinks will increase the cost of the network, but the network performance (network lifetime, network delays) can be greatly improved and, therefore, is subject to a wide range of research. However, mobile sinks require mutual cooperation and mutual coordination of movement between several sinks, and thus the study is more complicated than research of single mobile sink.

The latest research about multiple mobile sinks can be found in [24] which studies the optimization strategy of multiple mobile sinks and compares the situations of a number of stationary sinks and sinks moving along the hexagon surrounding. They find that sink mobility can significantly increase the network life, and the more points sink stays along the hexagon surrounding, the better balanced energy consumption the network will achieve. Mirela Marta, finally, proposes the distributed mobile optimization algorithm to maintain connectivity between the mobile sinks. Related research about multiple mobile sinks can also be found in [23, 24].

In addition, whether the research handles single or multiple mobile sink, they may adopt different strategies for data collection. It can be summed up in two circumstances.

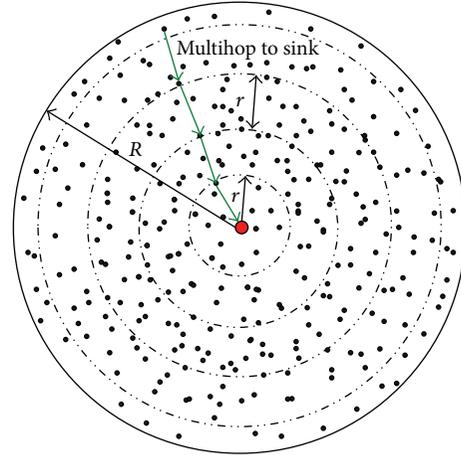


FIGURE 1: Network architecture model.

One is that mobile sink and static sink adopt the same type of data collection, using a static network sophisticated data collection strategies, such as single-hop, multihop, and subcluster. The other is to use different type data collection strategies from static sink, such as sensor nodes cache data so that the data is sent to sink when it arrives at vicinity of sensing, or the data is sent to a fixed regional cache, and then sink moves along the fixed region to collect data.

### 3. Network Model and Problem Description

Network architecture model: we apply the module similar with [17, 19], a typical wireless sensor network for cyclical data collection and a circle with radius of  $R$ ; see Figure 1. In this network, there are  $n$  nodes and,  $\{N_0, N_1, N_2, N_3, \dots, N_n\}$ ,  $N_0$  stands for sink and it can move throughout the network; others represent work nodes and cannot move after initially deployed. Communication range of nodes is noted with  $r$ ; the difference from general sensor networks is that the transmission range is adjustable, and nodes automatically adjust its communication range based on the distance between two nodes; for example, Berkeley Motes node has 100 transmission levels [7, 19]. Each work node will sense data in each cycle. We use the mature shortest path protocol for collecting data [8] and sending them to sink with multihop [8, 9].

Energy consumption model: we use typical energy consumption model; the cost of moving mobile sink is calculated according to formula (1), cost for sending data is calculated according to formula (2), cost for receiving data is calculated according to formula (3), and specific details can be found in the literature [14]:

$$E_{\text{sink}}(s) = sE_e, \tag{1}$$

$$E_{\text{member}} = \begin{cases} 1E_{\text{elec}} + l\epsilon_{fs}d^2 & \text{if } d < d_0, \\ 1E_{\text{elec}} + l\epsilon_{\text{amp}}d^4 & \text{if } d > d_0, \end{cases} \tag{2}$$

$$E_{\text{Rx}}(l) = lE_{\text{elec}}, \tag{3}$$

where  $E_{\text{elec}}$  stands for the energy loss of firing circuit. If the transmission distance is less than the threshold  $d_0$ ,

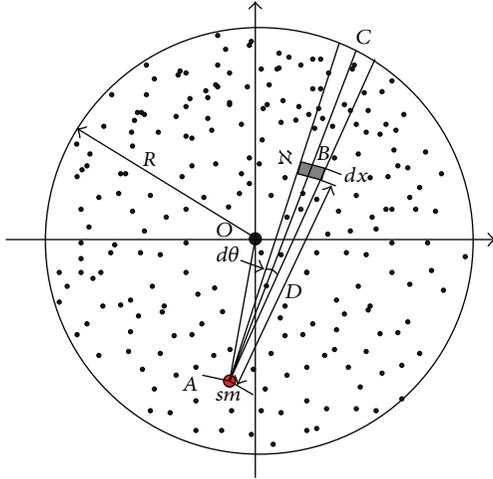


FIGURE 2: Network computing model.

power amplifier loss is based on free-space model; when the transmission distance is greater than or equal to the threshold value, it uses multipath attenuation model.  $\epsilon_{fs}$ ,  $\epsilon_{amp}$  represent the power for these two models' amplification, respectively. Energy for receiving  $l$  bit of data refers to formula (3). In this paper, the above specific parameters come from the literature [14].

**Problem Description.** For a given mobile sensor network shown in Figure 1, the problem can be described as follows: how to move and choose the anchors of mobile sink to maximize the network lifetime? Here we term the rounds of data collection till the first node dies as the network lifetime [7, 14].

#### 4. Analysis Load of Node and the Tactor That Affects the Network Lifetime

**4.1. Analysis Load of Node.** When the sink moves to an arbitrary location such as  $(x_0, y_0)$ , if it is able to calculate the data load of each node, it then will be easy to calculate the energy consumption of each node based on formulas (2) and (3) so as to learn energy consumption of the entire network. Therefore, this paper will compute data load for each sensor node when sink is located at arbitrary  $(x_0, y_0)$ . To the best of our knowledge, this paper gives derivation of data load in the network. It is also the basis for sink strategy in this paper.

**Theorem 1.** Suppose the center of network be  $O(0,0)$ , the sink moved to  $A(x_0, y_0)$ , an optional sensor node  $B$  at  $(x_b, y_b)$ , and the intersection point of  $AB$  extension with the network border is  $(x_c, y_c)$ , then the data load for  $B$  node is as follows:

$$D_r^x = \frac{\{(a-1-i)c + ((a-1-i)(a+i)r/2)\}}{(ir+c)}$$

if  $D=ir+c$ ,  $i \in \{0 \cdots a\}$ ,  $c \in \{b \cdots r\}$  //receive data,

$$D_t^x = \frac{\{(a-1-i)c + ((a-1-i)(a+i)r/2)\}}{(ir+c)+1}$$

if  $D=ir+c$ ,  $i \in \{0 \cdots a\}$ ,  $c \in \{b \cdots r\}$  //send data,

$$D_r^x = \frac{\{(a-i)c + ((a+1+i)(a-i)r/2)\}}{(ir+c)}$$

if  $D=ir+c$ ,  $i \in \{0 \cdots a\}$ ,  $c \in \{0 \cdots b\}$  //receive data,

$$D_t^x = \frac{\{(a-i)c + ((a+1+i)(a-i)r/2)\}}{(ir+c)+1}$$

if  $D=ir+c$ ,  $i \in \{0 \cdots a\}$ ,  $c \in \{0 \cdots b\}$  //send data,

where  $R_1 = |AC|$ ,  $\alpha = \left\lfloor \frac{R_1}{r} \right\rfloor$ ,  $R_1 = ar + b \mid b \leq r$ ,

$$D = |AB| = ir + c \mid i \in \{0 \cdots a\},$$

$$i = \left\lfloor \frac{D}{r} \right\rfloor, \quad c = D - ir, \quad c \in \{0 \cdots r\},$$

$$|AC| = \sqrt{(x_c - x_0)^2 + (y_c - y_0)^2},$$

$$|AB| = \sqrt{(x_b - x_0)^2 + (y_b - y_0)^2}.$$

(4)

*Proof.* This paper applies the shortest path routing protocol to transmit data to sink through multihop. For an arbitrarily node  $B(x_b, y_b)$ , see Figure 2;  $C$  represents intersection point of  $AB$  extension with the network border, and the data load for  $B$  is the amount of data whose distance from  $B$  is integer multiple of  $r$  on line  $BC$ . First, we calculate the coordinates of  $C(x_c, y_c)$ .

Equation of line  $AB$ :

$$y = \frac{y_b - y_0}{x_b - x_0} (x - x_0) + y_0. \quad (5)$$

Equation of the circle:

$$x^2 + y^2 = R^2. \quad (6)$$

Formula (5) can be simplified as

$$y = \frac{y_b - y_0}{x_b - x_0} (x - x_0) + y_0 = \frac{y_b - y_0}{x_b - x_0} x - \frac{y_b - y_0}{x_b - x_0} x_0 + y_0. \quad (7)$$

Let  $g_1 = (y_b - y_0)/(x_b - x_0)$ ,  $g_2 = -(y_b - y_0)/(x_b - x_0)x_0 + y_0 = y_0 - g_1x_0$ ,  $y = g_1x + g_2$ ; we can work out  $(x_c, y_c)$  by substituting it in formula (6):

$$(1 + g_1^2)x^2 + 2g_1g_2x + g_2^2 - R^2 = 0. \quad (8)$$

Solving the coordinates of  $C$  can be divided into several situations as follows.

First. When  $x(i) \neq x_0$

coordinates of  $C$  are as follows:

$$x_c = \frac{-2g_1g_2 \pm \sqrt{(2g_1g_2)^2 - 4(1+g_1^2)(g_2^2 - R^2)}}{2(1+g_1^2)},$$

$$y_c = g_1 \frac{-2g_1g_2 \pm \sqrt{(2g_1g_2)^2 - 4(1+g_1^2)(g_2^2 - R^2)}}{2(1+g_1^2)} + g_2,$$

where  $g_1 = \frac{y_b - y_0}{x_b - x_0}$ ,  $g_2 = -\frac{y_b - y_0}{x_b - x_0}x_0 + y_0;$

(9)

if  $x_b < x_0$  then

$$x_c = \frac{-2g_1g_2 - \sqrt{(2g_1g_2)^2 - 4(1+g_1^2)(g_2^2 - R^2)}}{2(1+g_1^2)},$$

$$y_c = g_1 \frac{-2g_1g_2 - \sqrt{(2g_1g_2)^2 - 4(1+g_1^2)(g_2^2 - R^2)}}{2(1+g_1^2)} + g_2;$$

(10)

if  $x_b > x_0$  then

$$x_c = \frac{-2g_1g_2 + \sqrt{(2g_1g_2)^2 - 4(1+g_1^2)(g_2^2 - R^2)}}{2(1+g_1^2)},$$

$$y_c = g_1 \frac{-2g_1g_2 + \sqrt{(2g_1g_2)^2 - 4(1+g_1^2)(g_2^2 - R^2)}}{2(1+g_1^2)} + g_2.$$

(11)

Second. When  $x_b = x_0$

if  $y_b = y_0$  then this is the sink itself; no data needs to be sent;

if  $y_b \neq y_0$  then  $x_c = x_0$   $x_c^2 + y_c^2 = R^2;$

if  $y_b > y_0$  then  $y_c = \sqrt{R^2 - x_c^2};$

if  $y_b < y_0$  then  $y_c = -\sqrt{R^2 - x_c^2}.$

According to coordinate of  $C$ , the length of line  $AC$  is

$$|AC| = \sqrt{(x_c - x_0)^2 + (y_c - y_0)^2}. \tag{12}$$

The length of line  $AB$  is

$$|AB| = \sqrt{(x_b - x_0)^2 + (y_b - y_0)^2}. \tag{13}$$

Let  $R_1 = |AC|$ ,  $\alpha = [R_1/r]$ ,  $R_1 = \alpha r + b \mid b \leq r:$

$$D = |AB| = ir + c \mid i \in \{0 \cdots \alpha\},$$

$$i = \left\lfloor \frac{D}{r} \right\rfloor, \quad c = D - ir, \quad c \in \{0 \cdots r\}. \tag{14}$$

Data load of  $B$  is calculated as follows. Its distance from sink is  $D = |AB| = ir + c \mid i \in \{0 \cdots a\}$ ,  $x \in \{0 \cdots b\}$ .

Then check sector area  $\aleph$  with angle of  $d\theta$  and width of  $dx$  (see Figure 2). The dimensions of this area are approximately  $\aleph_s = D d\theta dx$ . The number of nodes in this ring is  $\rho D d\theta dx$ . If it is located in the  $\{ir \cdots ir + b\} \mid i \in \{0 \cdots a\}$ th ring, that is to say, the location is  $D = ir + c \mid i \in \{0 \cdots a\}$ ,  $c \in \{0 \cdots b\}$ , then data load of  $\aleph$  is as follows.

It is responsible of forwarding all the remote data in sector area whose width is  $dx$  and is integer multiple of  $r$  away from  $\aleph$ . The dimension of these areas can be computed as

$$d\theta ((i+1)r + c) dx + d\theta ((i+2)r + c) dx$$

$$+ d\theta ((i+3)r + c) dx + \cdots d\theta (ar + c) dx \tag{15}$$

$$= d\theta dx \left( (a-i)c + \left( \frac{(i+1+a)(a-i)r}{2} \right) \right).$$

This is the dimension of area  $\aleph$  which is responsible of forwarding data. Then data load of  $\aleph$  is

$$d\theta dx \left( (a-i)c + \left( \frac{(i+1+a)(a-i)r}{2} \right) \right) \rho. \tag{16}$$

Data sent is

$$\left\{ d\theta dx \left( (a-i)c + \left( \frac{(i+1+a)(a-i)r}{2} \right) \right) \right.$$

$$\left. + d\theta (ir + c) dx \right\} \rho. \tag{17}$$

It can be assumed that the data load is uniformly shared by each node in a very small region. Then data load of each node is

$$\frac{d\theta dx ((a-i)c + ((i+1+a)(a-i)r/2)) \rho}{d\theta (ir + c)}, \tag{18}$$

$$dx \rho = \frac{((a-i)c + ((i+1+a)(a-i)r/2))}{(ir + c)}.$$

Data sent is  $\{d\theta dx(ac + ((1+a)ar/2)) + d\theta(ir+c)dx\} \rho / d\theta(ir+c)dx \rho = 1 + ((a-i)c + ((i+1+a)(a-i)r/2)) / (ir+c)$ .

If  $D = ir + c \mid i \in \{0 \cdots a\}$ ,  $c \in \{b \cdots r\}$  is located in the  $\{ir+b, ir+r\}$ th ring data load of  $\aleph$  can be computed as follows

It is responsible of forwarding all the remote data in sector area whose width is  $dx$  and is integer multiple of  $r$  away from  $\aleph$ . The dimension of these areas can be computed as

$$d\theta ((i+1)r + c) dx + d\theta ((i+2)r + c) dx$$

$$+ d\theta ((i+3)r + c) dx + \cdots d\theta ((a-1)r + c) dx \tag{19}$$

$$= d\theta dx \left( (a-i-1)c + \left( \frac{(a-i-1)(a+i)r}{2} \right) \right).$$

Then data received by  $\aleph$  is

$$d\theta dx \left( (a-i-1)c + \left( \frac{(a-i-1)(a+i)r}{2} \right) \right) \rho. \tag{20}$$

Data sent is

$$\left\{ d\theta dx \left( (a-i-1)c + \left( \frac{(a-i-1)(a+i)r}{2} \right) \right) + d\theta c dx \right\} \rho. \quad (21)$$

It can be assumed that the data load is uniformly shared by each node in a very small region. Then received data of each node is

$$\frac{\{(a-1-i)c + ((a-i-1)(a+i)r/2)\}}{(ir+c)}. \quad (22)$$

Data sent is

$$1 + \frac{\{(a-1-i)c + ((a-i-1)(a+i)r/2)\}}{(ir+c)}. \quad (23)$$

□

Based on Theorem 1, Figure 3 shows the energy consumption map when sink is at different locations. As can be seen from the chart, energy consumption of the network is different when sink is in different locations; nodes near sink do not necessarily suffer higher energy cost. For example, when sink is located near the circumference, power consumption in network-centric side of sink is much higher than the other side near the edge of the network. In order to show more clearly the data load for sink in different orientations and positions, Figure 4 selects three straight lines from the map and Figure 8 shows the statistical chart of the data load on these three lines.

As can be seen from Theorem 1, data load of node is directly proportional to distance ( $BC$  length) from this node to circumference. The longer  $BC$  is, the more data node  $B$  has to bear. It can be seen from Figure 5 that  $SA'$  line is the longest, and therefore  $SA'$  line has to undertake the most amount of data, but data load of  $SA$  line on the other side is the smallest, although  $SA$  is the closest to sink. The situation of  $BB'$  and  $CC'$  lines is also not the same. This indicates generally the idea that nodes near sink suffer higher volume of data which is wrong. Besides, energy consumption will change with sink location and has close relations to the transmission range. Figure 6 shows data load of  $CC'$  in Figure 4 under different communication range  $r$ . The general rule is that the bigger  $r$  is, the smaller the amount of data the node has to bear. The extreme case is when  $r = R$  every node sends data directly to sink. Each node at this time only has to bear one unit of data, which is the least.

**Corollary 2.** Note the transmission range  $r$  with  $f_r^i(x)$ ; sink has moved to  $A(x_0, y_0)$ , an arbitrary node  $B(x_b, y_b)$ ; then the energy consumption of node  $B$  is

$$f_r^i(x) = \begin{cases} D_r E_{elec} + D_t E_{elec} + D_t \varepsilon_{fs} x^2 & \text{if } x < d_0, i = 0 \\ D_r E_{elec} + D_t E_{elec} + D_t \varepsilon_{amp} x^4 & \text{if } x \geq d_0, i = 0 \\ D_r E_{elec} + D_t E_{elec} + D_t \varepsilon_{fs} r^2 & \text{if } x < d_0, i \neq 0 \\ D_r E_{elec} + D_t E_{elec} + D_t \varepsilon_{amp} x^4 & \text{if } x \geq d_0, i \neq 0. \end{cases} \quad (24)$$

*Proof.* According to Theorem 1, the amount of received data of nodes  $D = ir + x$  away from the sink is  $D_r$  and the amount of sent data is  $D_t = D_r + 1$ . Substituting them into energy formulas (1) and (2) will lead to Corollary 2. □

Based on Corollary 2, Figure 7 shows the energy consumption under different sink locations and different  $r$ . As can be seen from the figure, the energy consumption of mobile sink is very complex. So it requires careful planning for moving sink.

In order to examine the energy consumption of the network in detail, Figure 8 shows energy consumption map of the three lines selected from Figure 4. Figure 9 shows energy consumption of nodes located at different distance from the sink under various transmission ranges computed through Corollary 2. Different from data load, when node is located at the first  $r$  distance ring from sink, the transmission range is actually the distance to sink, so it is less than  $r$ . When distance from node to sink is larger than  $r$ , transmission range is all the same as  $r$ . Therefore, the energy consumption does not decline as the distance grows. Its law is that in the first ring where  $r$  is small, node close to sink has to undertake much higher data load; although the communication range is smaller than  $r$ , energy consumption rate is still very high. With the increase of  $r$ , data load declines very fast (see Figure 6), and thus energy consumption of nodes nearest from sink also drops very quickly. Although data load in the edge of first ring is smaller than that within the first ring, the transmission range is larger and hence energy consumption increases. Consequently, there is concave within  $r$  from sink in the power consumption chart (see Figure 9 energy consumption chart  $r = 125$ ). From the second ring, the communication range is all  $r$ , but data load is smaller as they are farther from sink, energy consumptions also a decreasing function of distance from sink (see Figure 9). This paper reveals details of network energy consumption for the first time.

**4.2. The Factor That Affects the Network Lifetime.** Data collection strategy for mobile sink is as follows: sink moves to a new location to carry out  $k$  rounds of data collection and then moves to a new anchor. It goes on like this model until the death of the network. Suppose the total time sink can pause during the survival of the entire network which is  $\Gamma$ . Then goal for sink mobility is to find suitable anchors to achieve the longest life span of the network. And the essence is to maximize the number of rounds of data collection, that is,  $\text{Max}(k \times \Gamma)$ . As we define the lifetime as an interval till the death of the first node. Therefore, in order to achieve the longest life expectancy, it is necessary to minimize the largest energy consumption among all the nodes.

**4.2.1. The Impact of the Number of Anchors.** After migrating to the purpose location, sink will conduct data collection. The purpose location is termed as the anchor. The number of locations that sink can pause is termed as anchor number. In fact, the case of static sink can be summed up as the situation where there is only one anchor. From above we know that the center is the best sink location now. In addition, we are

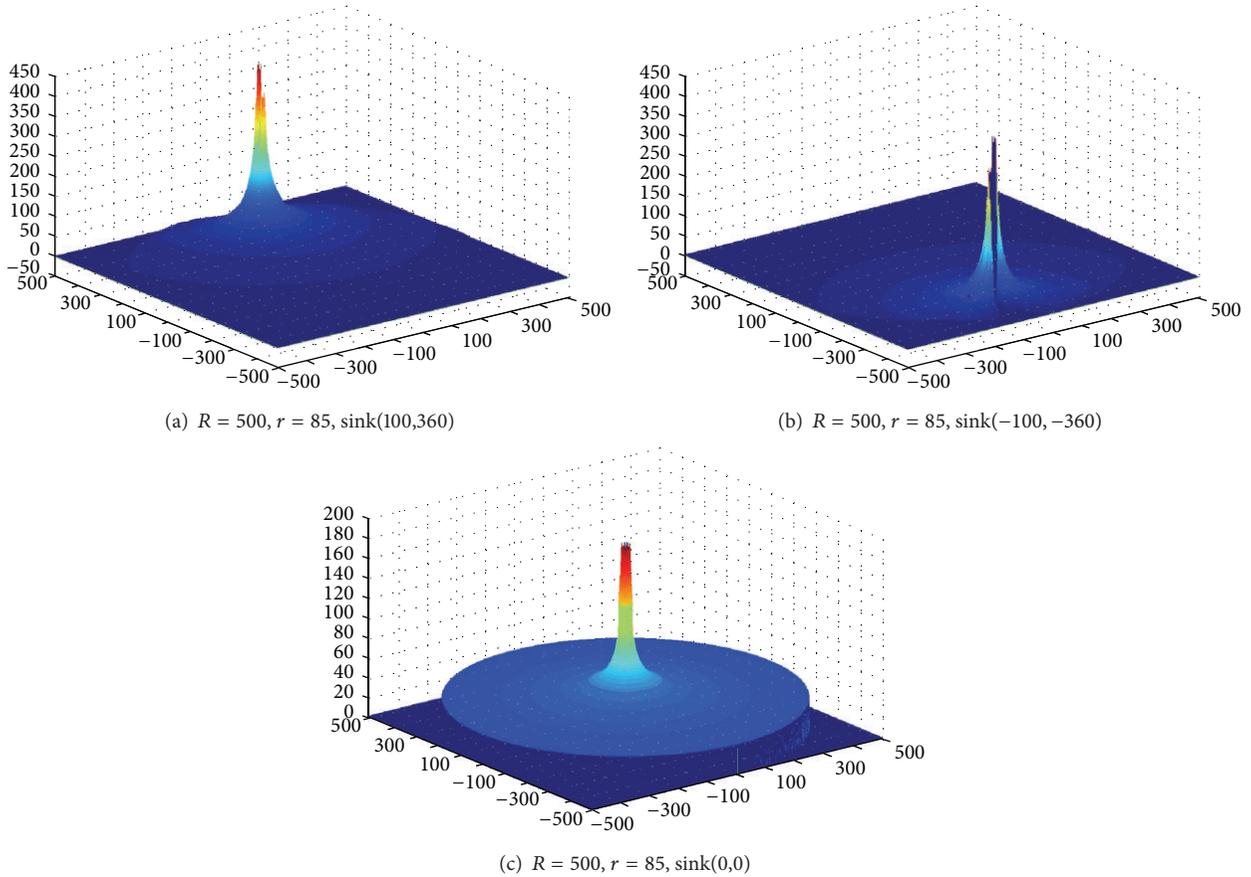


FIGURE 3: Data load of network.

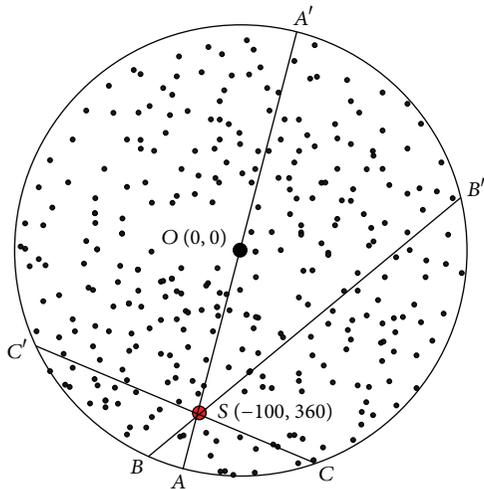


FIGURE 4: Selected  $AA', BB',$  and  $CC'$  positions.

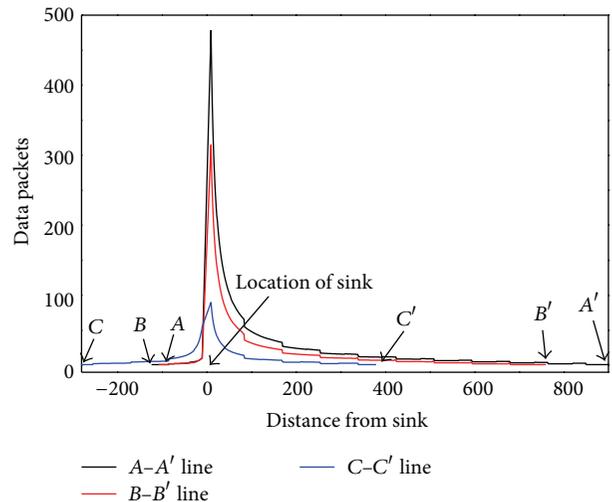


FIGURE 5: Data load in different distances from sink in different directions.

able to get the largest network lifetime (equal to static sink). However, when the number of anchor numbers increases to be more than one, the analysis of network life expectancy has become more complex. We assume that sink can move across the whole network, and the anchor number is infinite in theory. But the actual anchor number is limited as follows.

- (1) Based on the principle of symmetry of a circular network with Proposition 2 (claim 2) in Luo and Hubaux's paper [17] have proven that the track of

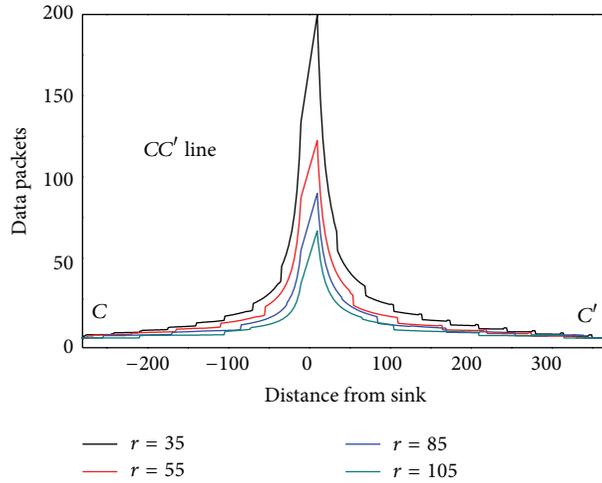


FIGURE 6: Data load under different communication ranges.

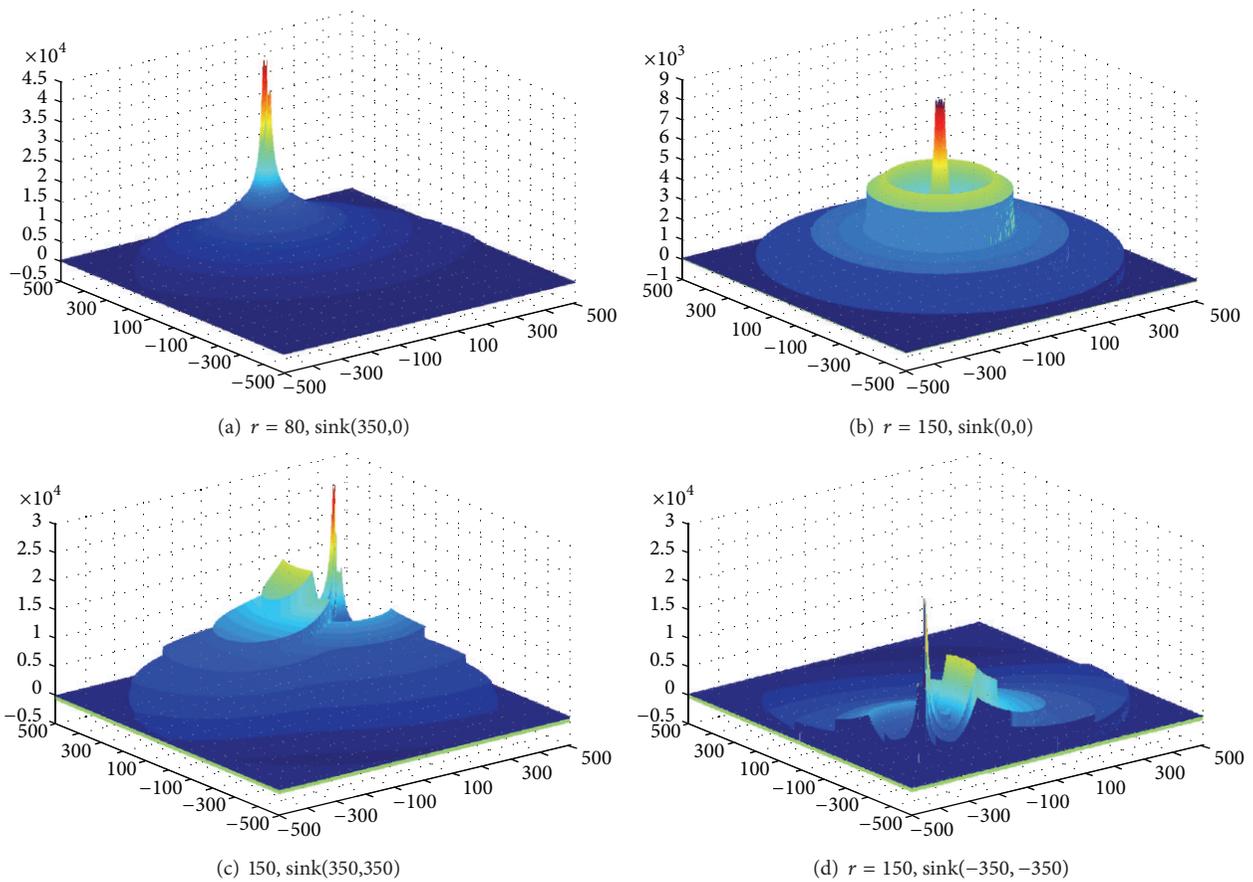


FIGURE 7: Energy consumption of network ( $R = 500$ ).

mobile sink is a ring. That is, trajectory of sink is limited to a circular track within the circle.

- (2) Because of the limited energy resource of sensor node, the life span of the network is limited as  $n$  rounds. The sink conducts at least one round of data collection at each anchor; then anchor number is restricted to the  $n$  points on the circular track. Also, based on the

principle of symmetry, these  $n$  anchor points should be uniformly distributed on the circular trajectory to achieve optimal network performance.

- (3) The actual physical characteristics of the network, such as sink mobility, network topography, and geography, can limit the anchor sink's stay.

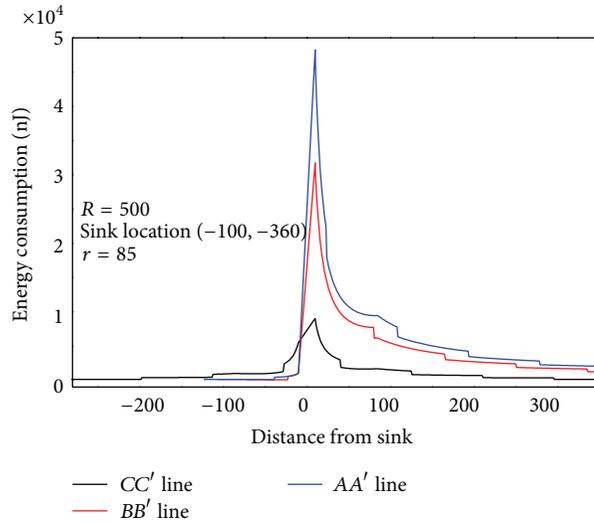


FIGURE 8: Energy consumption of nodes different distances from the sink in different directions.

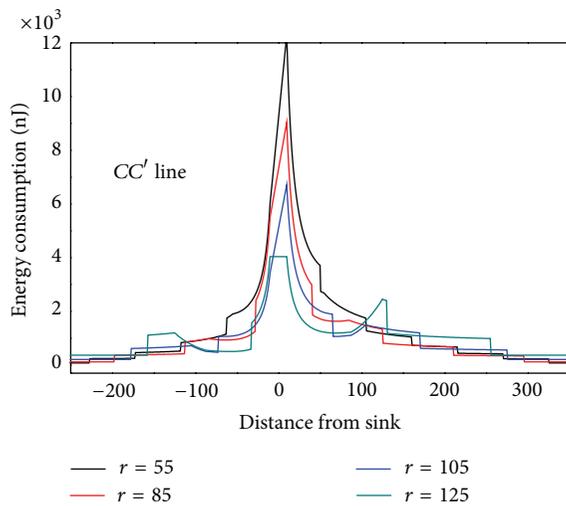


FIGURE 9: Energy consumption under different communication ranges.

When the sink is located on a track (the track is a ring whose center is at origin; when the radius of ring is  $R_m$ , it is termed as trajectory or orbit  $R_m$ ) and the anchor location is determined, the algorithm below can compute the energy consumption for each network grid.

Based on Theorem 1, Figure 10(a) shows energy situation of five rounds of data collection for each anchor, and then move to the next location; there is a total of 50 anchors and 250 rounds of data collection. And in Figure 10(b) there is a total of 250 anchors and 250 rounds of data collection. It can be seen that under different anchor numbers, even if the two situations conduct the same rounds of data collection (the same network lifetime), the largest energy consumption in network whose anchor number is bigger is less than half of that in the former network (less anchor number). It indicates that anchor numbers have a greater impact on the network

life expectancy. If we define quotient obtained by dividing the maximum energy consumption among all the nodes by the number of rounds of data collection as the network energy efficiency, then the energy efficiency of the network is higher, the higher life expectancy will be under the same initial energy and vice versa.

We also find that network energy efficiency increases with the growth of the anchor. But the efficiency will stay stable after anchor increases to a certain threshold.

The anchor number is mainly affected by the initial energy of node. If the initial energy is small, the life span of the network  $n$  will be short. However the anchor number cannot exceed  $n$ , so it will affect the network efficiency when  $n$  is small. Similarly, the larger the network is, the more energy for one round data collection will be, thus decrease the number of anchor.

Anchor number will not affect the energy efficiency of the network after it rises over certain degree, and the degree is related with the scale of the network. Let the radius of the sink track be  $R_m$ ; study in this paper shows that the degree will be greater when  $R_m$  increases and vice versa. The most special situation is that if  $R_m = 0$ , then degree of sink is  $n = 0$ ; that is, the anchor number does not affect the network energy efficiency. From the above analysis we can see that anchor number of sink has great impact on network life expectancy. Therefore, previous studies which do not consider anchor number usually simplify the problem.

**4.2.2. The Impact of Radius of Mobile Trajectory.** Figure 11 shows the energy situation when  $R_m$  is, respectively, 2000 m, 1200 m, 1000 m, and 800 m in network with  $R = 2000$  m and 50 sink anchors. Largest energy in  $R_m = 2000$  m is not the minimum; instead,  $R_m = 1200$  m has the minimum energy consumption. The above description indicates that the optimal trajectory of mobile sink needs to be carefully optimized and there will be a better route for the longest network life.

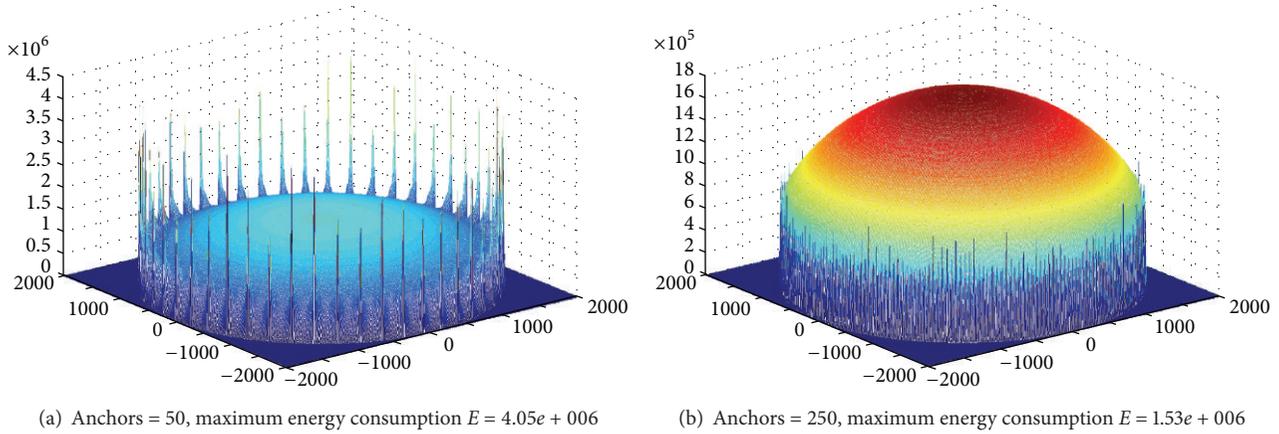


FIGURE 10: Energy consumption map under different anchor number ( $R = 2000$ ,  $R_m = 2000$  m, and  $r = 85$ ).

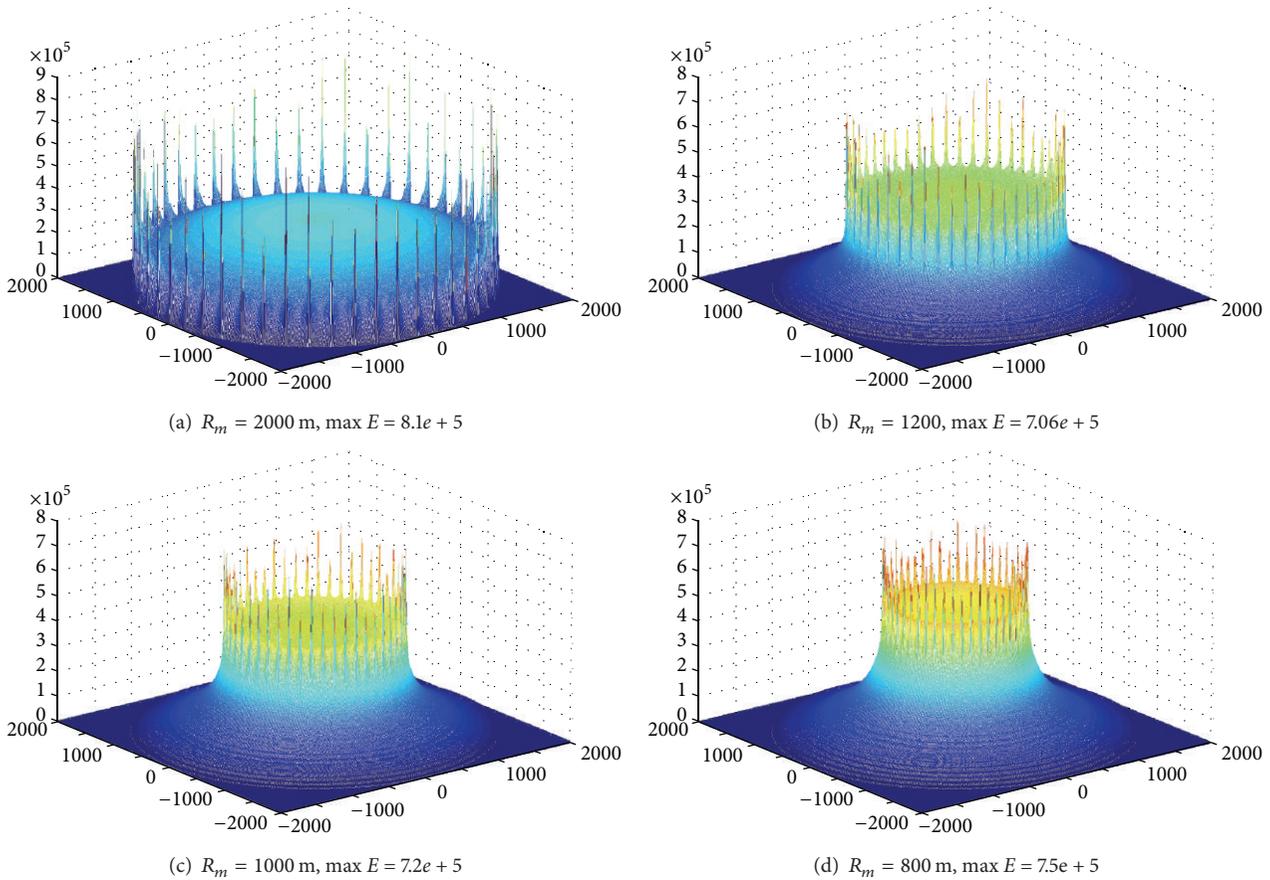


FIGURE 11: Energy consumption for different values of  $R_m$  ( $R = 2000$ , rounds = 50, anchors = 50,  $r = 85$ ).

4.2.3. *The Impact of Communication Range.* The communication range  $r$  has a great impact on the network lifetime. Figure 12 shows the energy situation of the line which is 620 m from sink to the circumference (similar to  $|SA'| = 620$  m in Figure 4). It can be seen that energy situation changes with different  $r$ . The rule is as follows: with the increase of  $r$ , energy consumption of nodes near sink decreases, and energy consumption of nodes far from sink grows. The right chart in

Figure 12 shows energy consumption under different length of  $|SA'| = L$ . The rule is that the greater  $L$  is, the higher energy consumption will become. Accordingly, the largest lifetime of static sink network is as follows: when  $L = R$ , the value of  $r$  which can minimize the maximum energy consumption on  $|SA'|$  is the value that can achieve the largest network lifetime. The energy consumption should be energy cost after one round of data collection. The quotient obtained by dividing

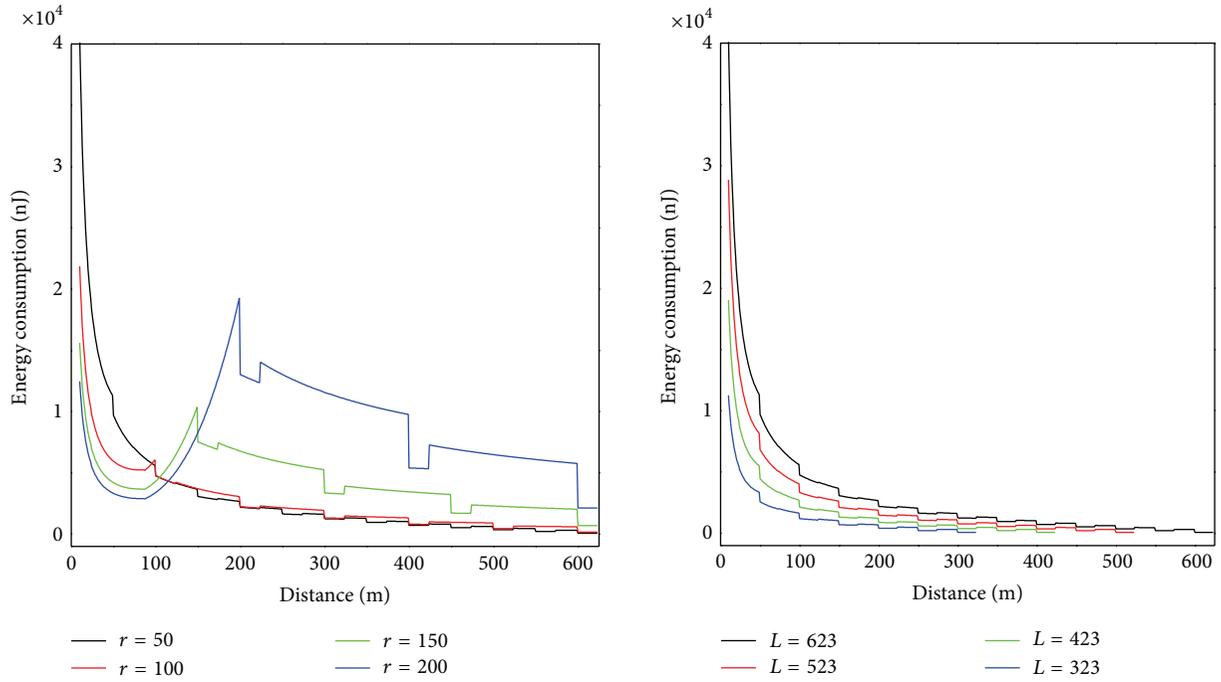


FIGURE 12: The relationship between energy consumption and  $r$ .

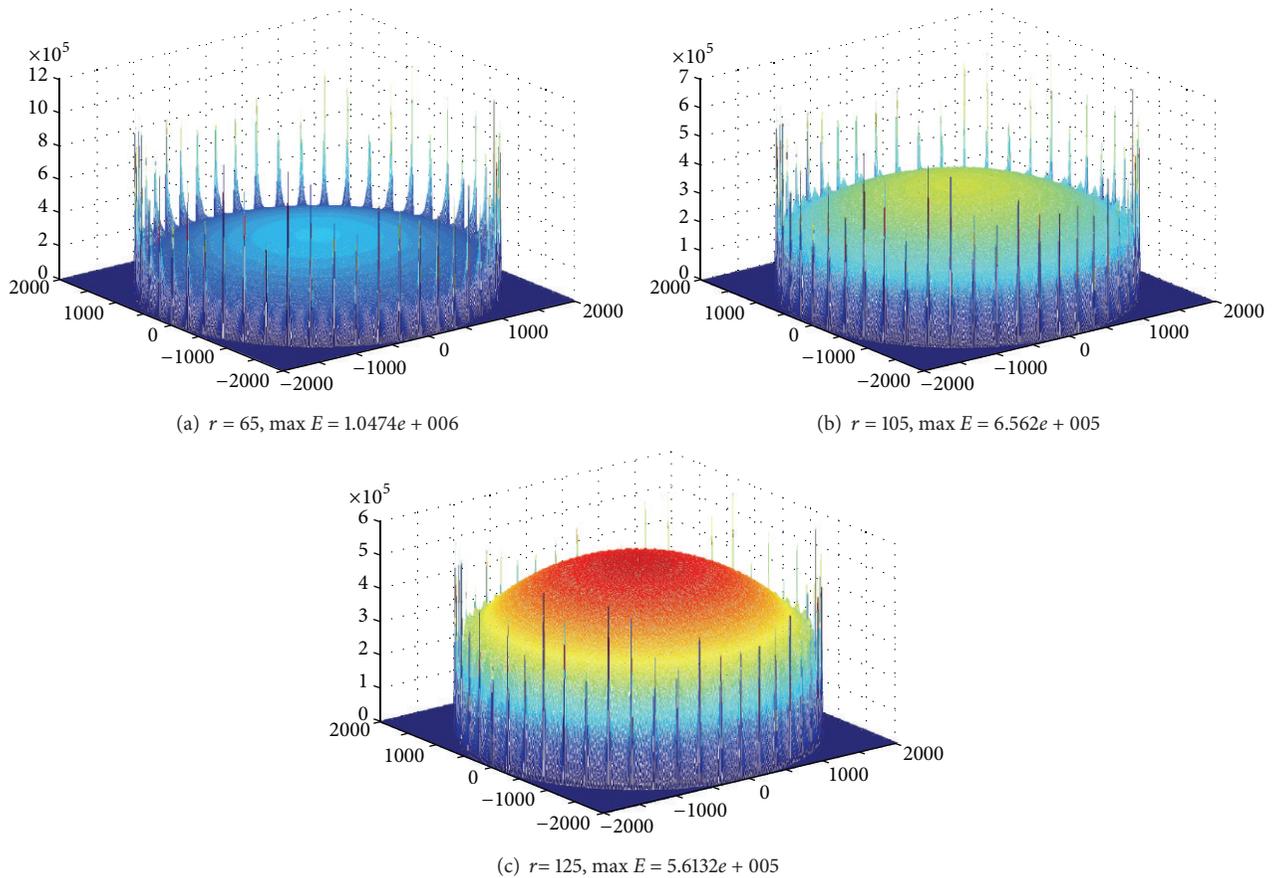


FIGURE 13: The largest energy consumption of mobile sink network under different  $r$  ( $R = 2000$  m,  $R_m = 2000$  m, rounds = 50, anchors = 50).

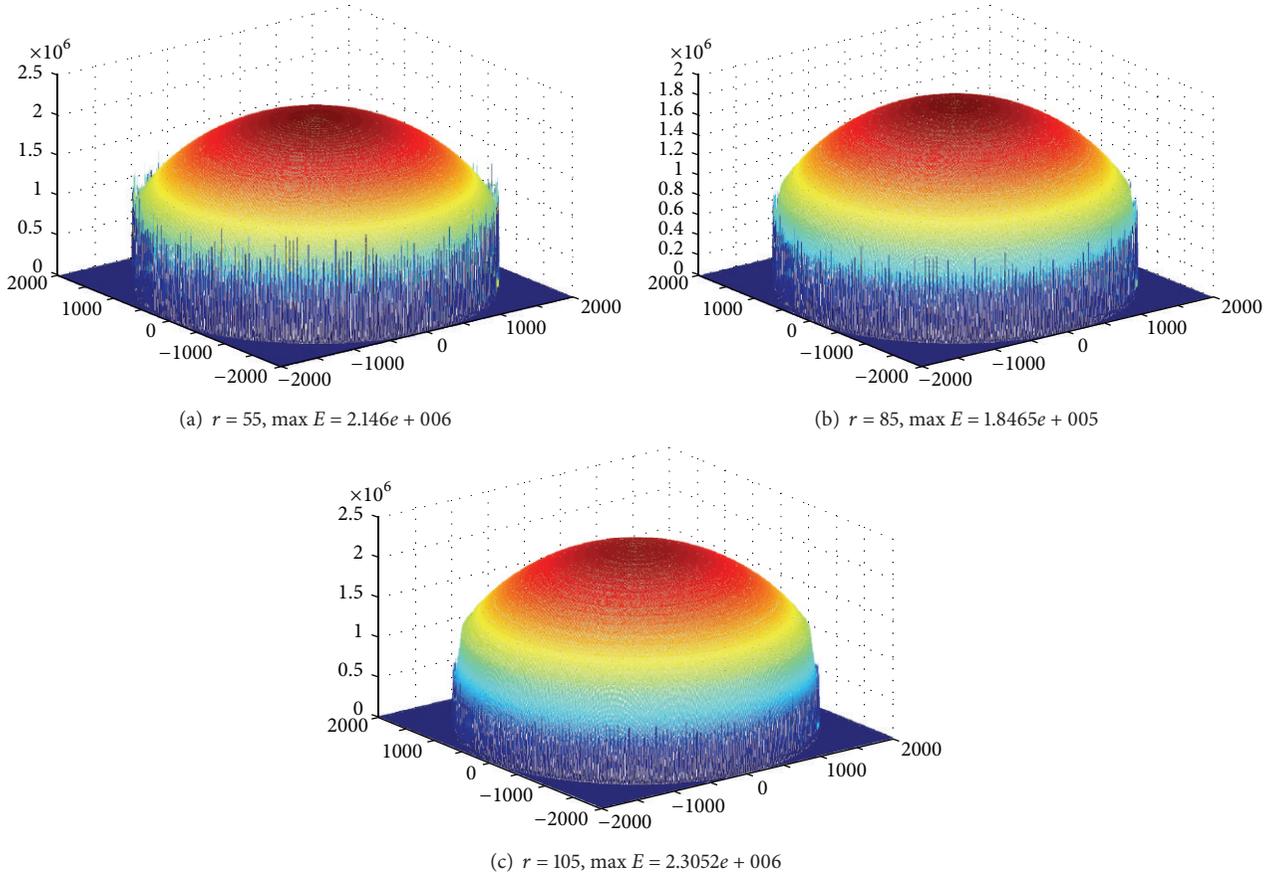


FIGURE 14: The largest energy consumption of mobile sink network under different  $r$  ( $R = 2000$  m,  $R_m = 2000$  m, rounds = 300, anchors = 300).

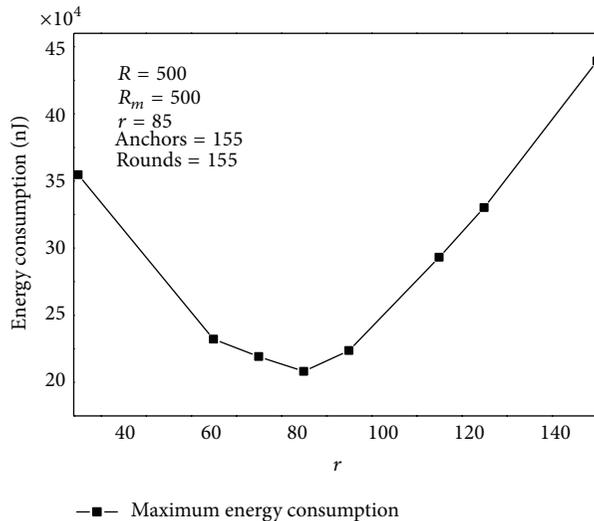


FIGURE 15: Maximum energy consumption under different  $r$ .

node initial energy by the largest energy consumption is the largest network lifetime with static sink.

Communication range  $r$  also has a great impact on the lifetime of mobile sink network. We can prove this point through experiment. Below is in the network with

$R = 2000$ , sink moving along the circumference, and sink pausing anchor should be only 50. Energy consumption in the trajectory is low; energy consumption on the trajectory is high. Therefore, when  $r$  increases, energy consumption in the trajectory will increase, and energy consumption on the

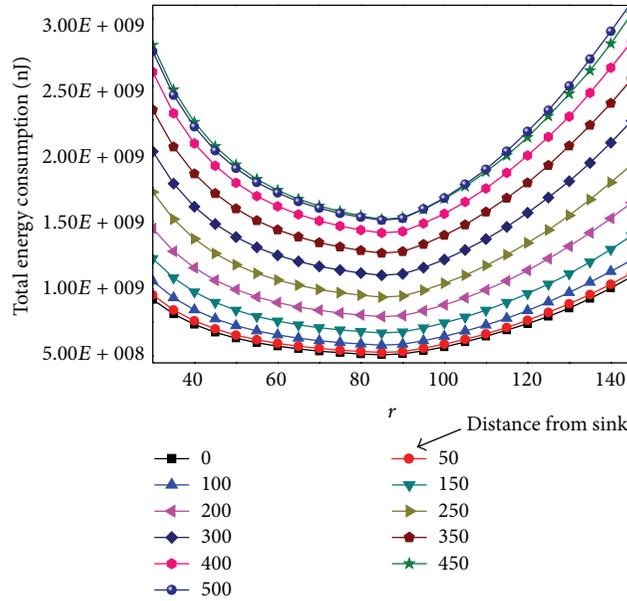


FIGURE 16: Overall energy consumption under different  $r$ .

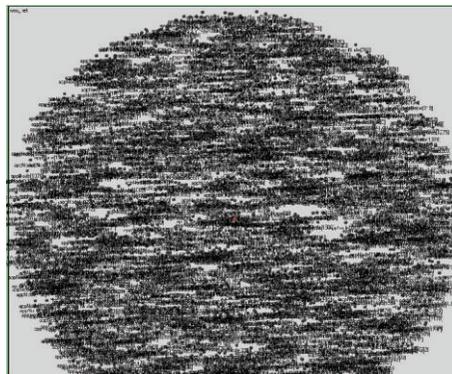


FIGURE 17: Randomly deployed network.

trajectory will reduce. Figures 13(a), 13(b), and 13(c), respectively, show different maximum energy consumption under different  $r$ . We can see that the difference between network largest energy is nearly doubled when setting different  $r$  values.

Below is in the network with  $R = 2000$ , sink moving along the circumference, and sink pausing anchor increasing to 300. We can set different  $r$  values to achieve the smallest largest energy consumption, that is, the maximum network lifetime. Figures 14(a), 14(b), and 14(c) show different maximum energy consumption under different  $r$ . We can see that network lifetime is also not the same.

Figure 15 shows the energy situation after 155 rounds of data collection under different based on Theorem 1. There is a best to obtain the smallest total energy consumption and the largest network lifetime.

Also  $r$  is related with the network overall energy consumption. Figure 16 shows the total energy consumption under different  $r$  after one round of data collection. The

relationship between network total energy consumption and  $r$  is a concave curve; the concave point minimizes total energy consumption, which is also the best  $r$  value. The best  $r$  is not affected by the location of sink. As can be seen from the chart the amount of energy is the least when  $r = 85$  m.

## 5. A Distributed Algorithm for Mobile Sink

**5.1. Distributed Algorithm for Mobile Sink.** In our theoretical calculation nodes are uniformly distributed, and the external environment is homogeneous. But in practice, the deployment of network is always not uniform, and the external environment is not necessarily homogeneous (affected by physical barriers such as restrictions on the terrain). We find in experiments that the deployment of nodes has a great impact on the network lifetime, such as Figures 17 and 18. 3000 nodes are deployed in network with radius of 500 m. In Figure 18 nodes are randomly deployed; in fact, the deployment is uneven. Sparse region will lead to particularly

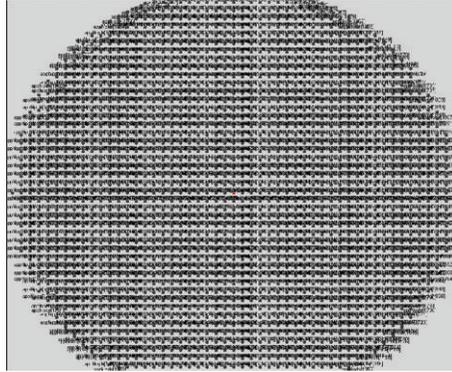


FIGURE 18: Uniformly deployed network.

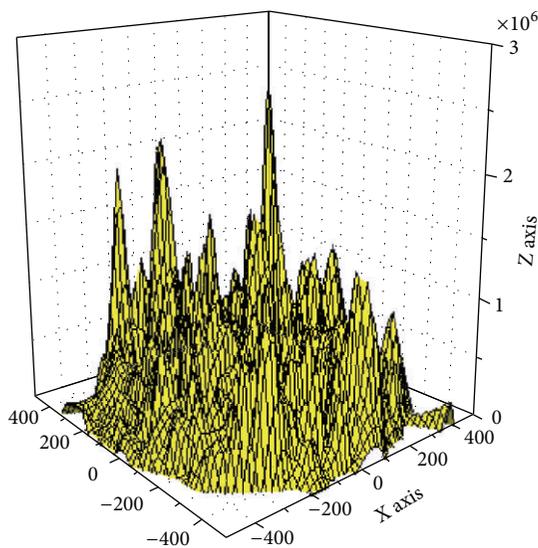


FIGURE 19: Energy consumption of randomly deployed network.

high energy consumption in individual node; these nodes are located at  $(-104, -387)$   $(76, -20)$   $(-30, 279)$ . The energy consumption map is shown in Figure 19. Figure 20 is an energy consumption diagram of strictly uniform deployment. From observation of experimental results, the trend shown in Figure 19 is in line with the theoretical results. However, it is not as optimal as the results of Figure 20. It can be seen that offline methods cannot adjust sink location according to real world situation. Online distributed algorithm adaptability optimizes the choice of the sink anchors according to the actual energy consumption of the network, so its efficiency is relatively high and adaptable. Therefore, based on the previous theoretical analysis, we propose distributed mobile sink algorithm to avoid the energy hole.

The idea of Algorithm 1 is first to mesh network, then move sink to a grid  $(X_0, Y_0)$ , calculate the total energy of all the nodes by adding the actual energy consumed and the precalculated theoretical result, and then choose the largest energy  $EE$  as the *compare energy* for sink at  $(X_0, Y_0)$  and repeatedly we switch sink to next grid and calculate its *compare energy*. From all the *compare energies*, we choose the

minimum one and set the correspondent sink coordinate as the optimal next pause anchor.

The complexity of Algorithm 1 is  $m * n * n1 * |r|$ ;  $m * n$  represents the number of grid,  $n1$  represents the number of nodes, and  $|r|$  represents the transmission level of nodes.

**5.2. Analysis of Mobile Strategy.** Algorithm 1 only considers maximizing the life span of the network; sink pausing path is shown in Figure 21(a). However, it fails to discuss the mobility of sink. In fact, it may not be feasible to purely act according to Algorithm 1 in real application. Because the speed of the mobile sink should be slow, sink cannot move around the network without limitations in speed as described in Algorithm 1. In addition, sink mobility will bring extra recourse dissipation and cost; thus the shorter sink moves the better. For these reasons, we recognize that Algorithm 1 can be improved according to the fact that if the total pause of anchors for sink is  $N$ , then the network life is irrelevant with the sequence of pausing. Therefore, we can precalculate  $k$ -anchors in Algorithm 1, rather than one single anchor. As long as there is at least one of them located within the scope of current sink migration capability, then ideal moving path of sink is to form a ring around the network. Since the algorithm for calculating  $k$ -anchors is similar with Algorithm 1, due to space limitations, we omitted it here.

## 6. Performance Analysis and Experimental Results Comparison

This paper applies OMNET++ to carry out experiments; OMNET++ is an open network simulation platform which is open source, component-based and modular for large network and has been widely recognized by the academic community [10]. Experimental parameters are shown in Table 1 from the literature [14], if there is no special note.

### 6.1. Mobile Sink Network Performance Analysis and Experimental Comparison

**6.1.1. Impact of the Number of Anchor.** Figures 22 and 23 show, respectively, the theoretical calculations and experimental results of energy consumption under different numbers of anchors. The charts show that theoretical results

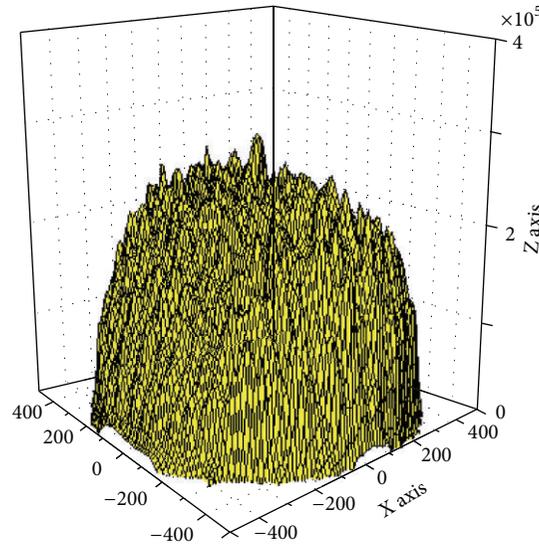


FIGURE 20: Energy consumption of uniformly deployed network.

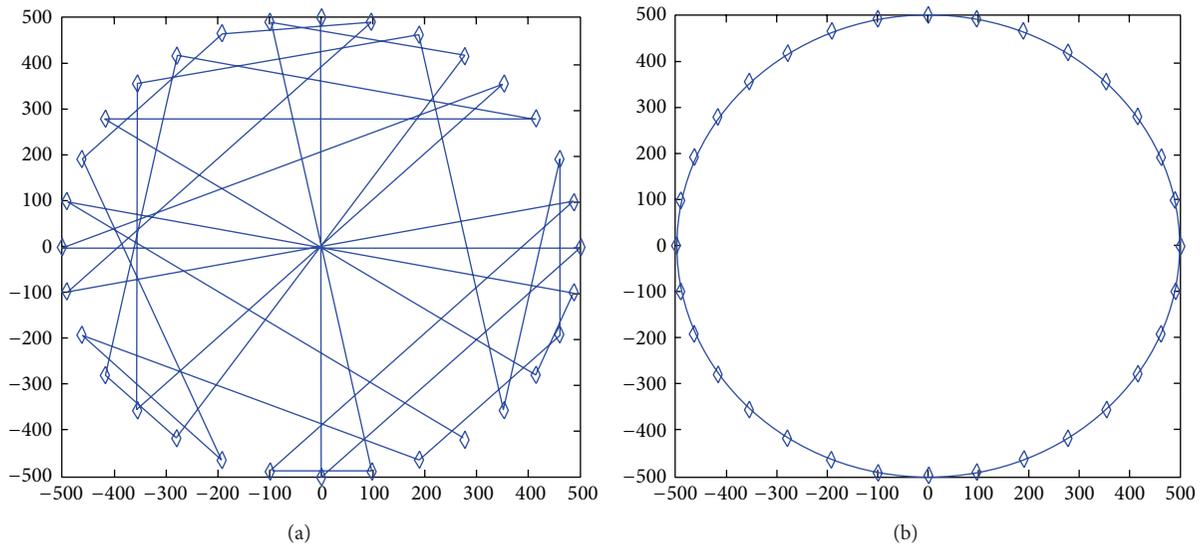


FIGURE 21: Path for mobile sink.

match the experimental outcome. Figures 24 and 25 compare the maximum energy consumption under different numbers of anchors. The maximum energy consumption is different. We can obtain the *energy per round* by dividing the largest energy consumption by the number of rounds. It is clear that if the *energy per round* is small, then the efficiency of network is high. Figures 26 and 27, respectively, show the *energy per round*. From the chart we can find that the more the anchors are, the better the network efficiency will be. However, after the number of anchor exceeds a certain degree, the network efficiency is stable.

6.1.2. *Impact of Transmission Range.* Figures 28 and 29 show the comparison of network energy dissipation between theoretical analysis and experimental under different transmission ranges and mobile trajectory. We can see a comprehensive impact of  $r$  and trajectory radius on the

network energy consumption and the trend experimental outcomes is very close to the theoretical results.

Figures 30 and 31 show the maximum energy consumption when sink is on different route. If sink is located in the same trajectory energy consumption can be different for various communication radiuses. If sink is located in distinct trajectory discrepancy between network energy consumption is great. Therefore, optimization of the mobile program requires comprehensive consideration of variety of factors.

Figures 32 and 33 show the energy consumption of nodes on the diameter after 20 rounds of data capture under different  $r$  and  $R_m$ . The tendency in experimental result is in line with the theory.

6.2. *Performance Analysis of the Distributed Algorithm.* The experimental scene is as follows: network radius  $R = 500$ , the numbers of sensor node equal to 3000. Similar with the

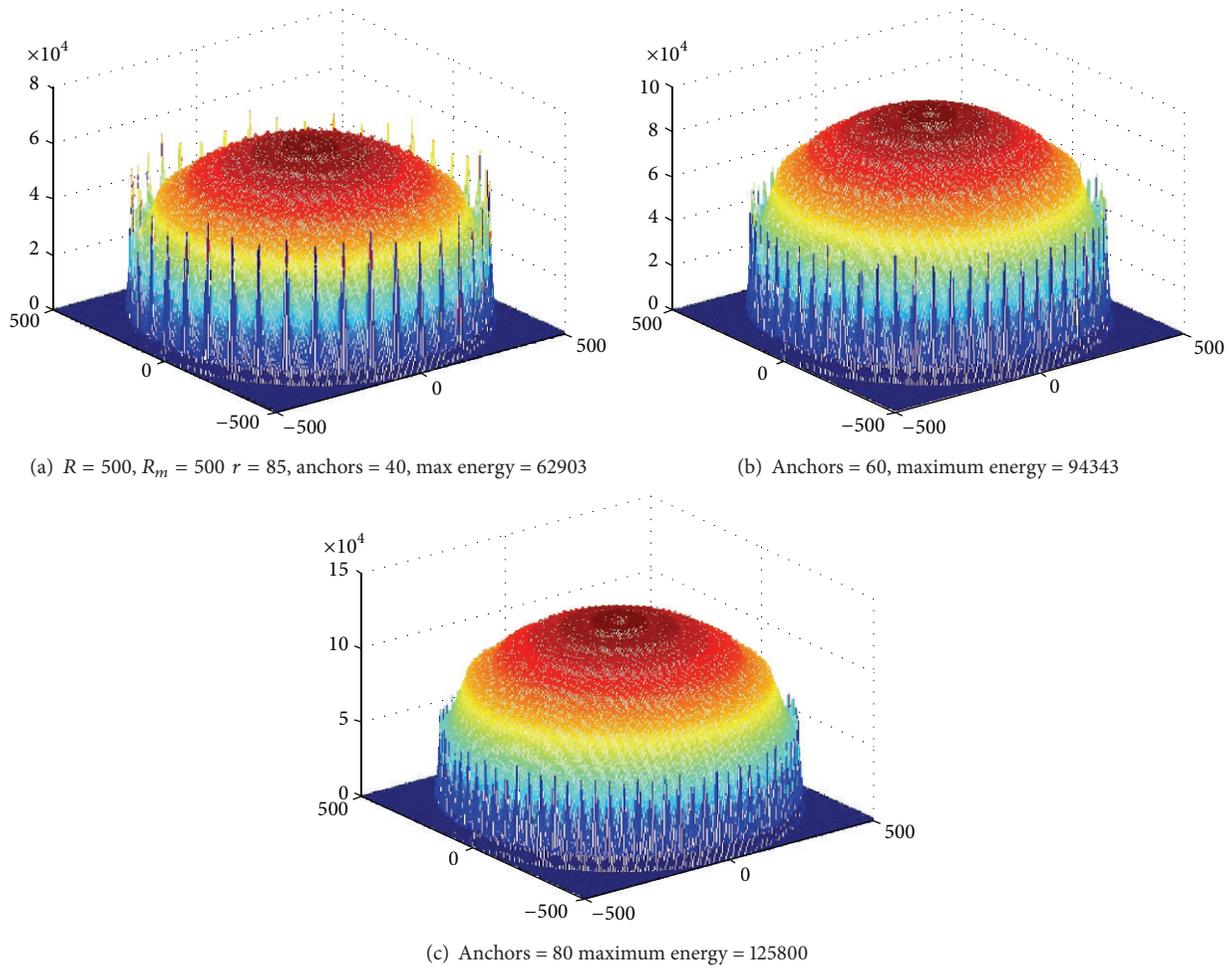


FIGURE 22: Energy consumption under different numbers of anchor (theoretical value).

scene shown in Figure 18, sink collects data for 150 rounds. Based on previous experiments, we know that centralized algorithm in this paper is the same as LUO [17]. We only need to contrast the centralized algorithm, static sink, and distributed algorithm.

As theoretical analysis and comparison have been discussed previously, we only compare experimental results here. Figure 34 shows the energy consumption of these three algorithms after 150 data gathering.

From the experimental results in Figure 34, the energy consumption of static sink is more than the mobile sink approach by nearly an order. The maximum energy cost in distributed algorithm is the smallest. It shows that the proposed algorithm of distributed mobile sink can further improve network lifetime.

Figure 35 shows the comparison of maximum energy consumption between centralized algorithm and distributed algorithm in the process of data collection. Figure 36 shows the trajectory of sink computed by distributed algorithm. We can see that the actual optimized track of mobile network is not complete the same as calculated routine. Distributed algorithm has extended the footprint of sink to inside of the network, rather than only moving on the circle.

Distributed algorithm proposed in this paper is superior to all other algorithms as distributed algorithm can adaptively adjust the next location of sink based on actual energy consumption of the entire network. There are similar endeavors in previous studies. This paper can first calculate theoretical energy consumption of the entire network, effectively guiding the sink migration. In contrast with the algorithm which considers energy of the nodes only within one hop of current sink, the algorithm here takes into account the overall energy consumption. Our algorithm is also superior to those algorithms which take area near the sink certainly as hotspot and each time move sink from hotspot to the region with the highest residual energy. Because we can neither simply believe that nodes near sink spend the highest energy consumption, nor that moving sink to the region which remains the highest energy will minimize the maximum energy consumption.

**6.3. Network Remaining Energy Comparison.** Figure 37 shows the residual energy ratio between static sink and mobile sink after 150 rounds of data gathering in network of  $R = 500$  m. The remaining energy ratio of static network remains greater than 70 percent all through the run, while

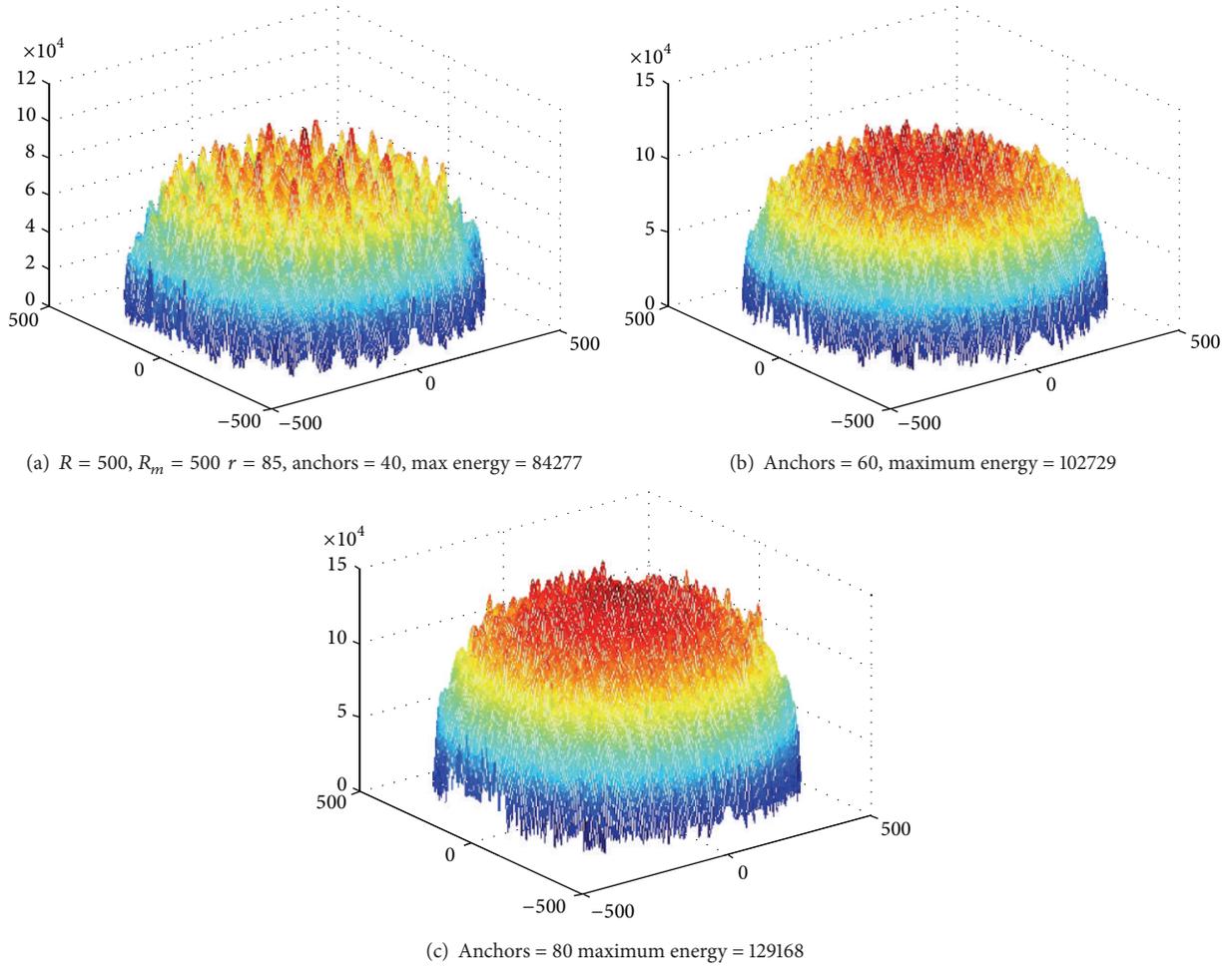


FIGURE 23: Energy consumption under different numbers of anchor (experimental value).

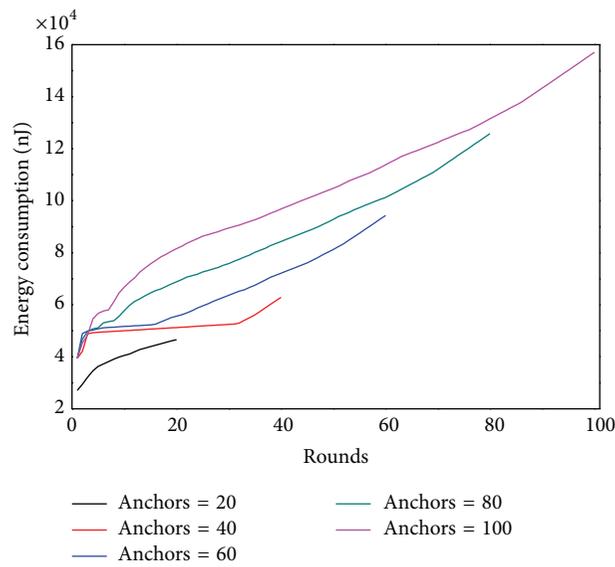


FIGURE 24: Maximum energy consumption under different numbers of anchor (theoretical value).

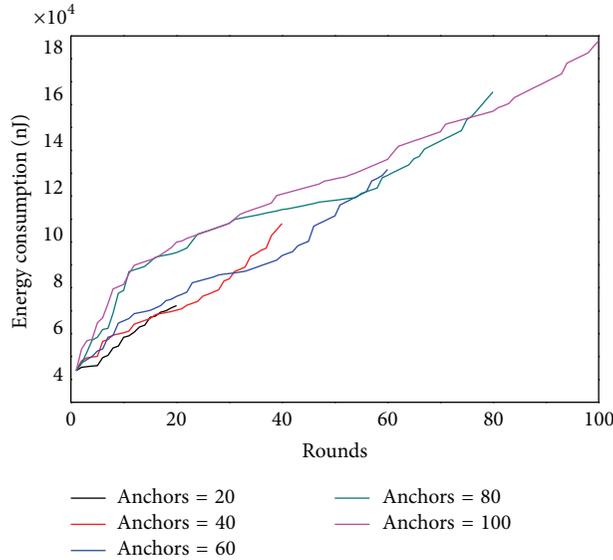


FIGURE 25: Maximum energy consumption under different numbers of anchor (experimental value).

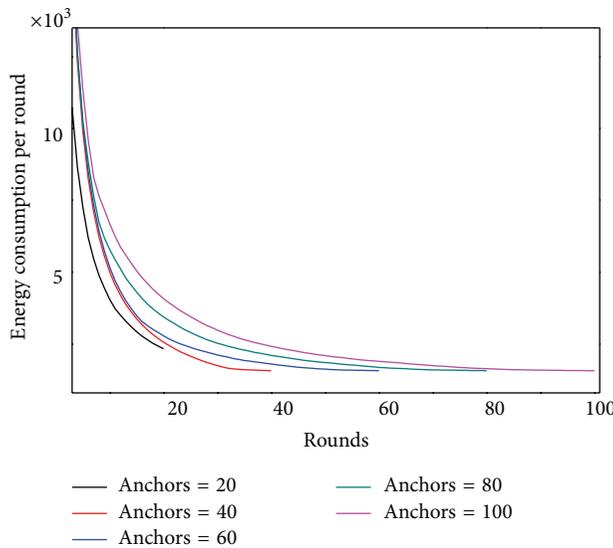


FIGURE 26: Energy per round under different numbers of anchor (theoretical value).

the remaining energy ratio of mobile sink strategy drops with the rounds of data collection and finally stays unchanged. Sink mobility has greatly enhanced the energy utilization rate (by 30%). Figure 38 shows the network remaining energy ratio under different transmission radius and it indicates remaining energy ratios are not the same as communication range changes (residual energy ratio =  $1 - (\text{total energy consumption of all the nodes}) / (\text{number of nodes} \times \text{maximum energy consumption})$ ).

We also analyzed the impact of network parameters on the network performance. The node density, from previous study, has little impact on network performance. However, uniformity of the deployment has impact on the network life span. The more uniformly the nodes are deployed, the higher network life expectancy will be. In addition, other

factors include the initial nodes energy and the network size. The greater the initial node energy is, the more round it can carry out data collection and vice versa. While increasing the network scale will decrease the network lifetime. As the number of rounds of data collection is equal to the minimum docking anchor of sink node, if the network lifetime is higher, docking anchors will become more and thus energy efficiency increases. Finally, we also experiment on the impact of parameters on the network used in Table 1. Generally, they influence the network performance, however, not the trend.

### 7. Conclusion and Future Works

The main contribution of this paper is as follows: (1) presents a method to accurately calculate energy expenditure of the

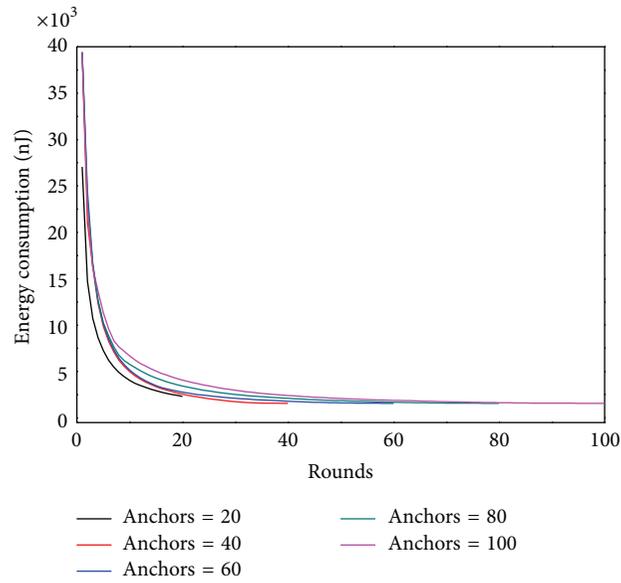


FIGURE 27: Energy per round under different numbers of anchor (experimental value).

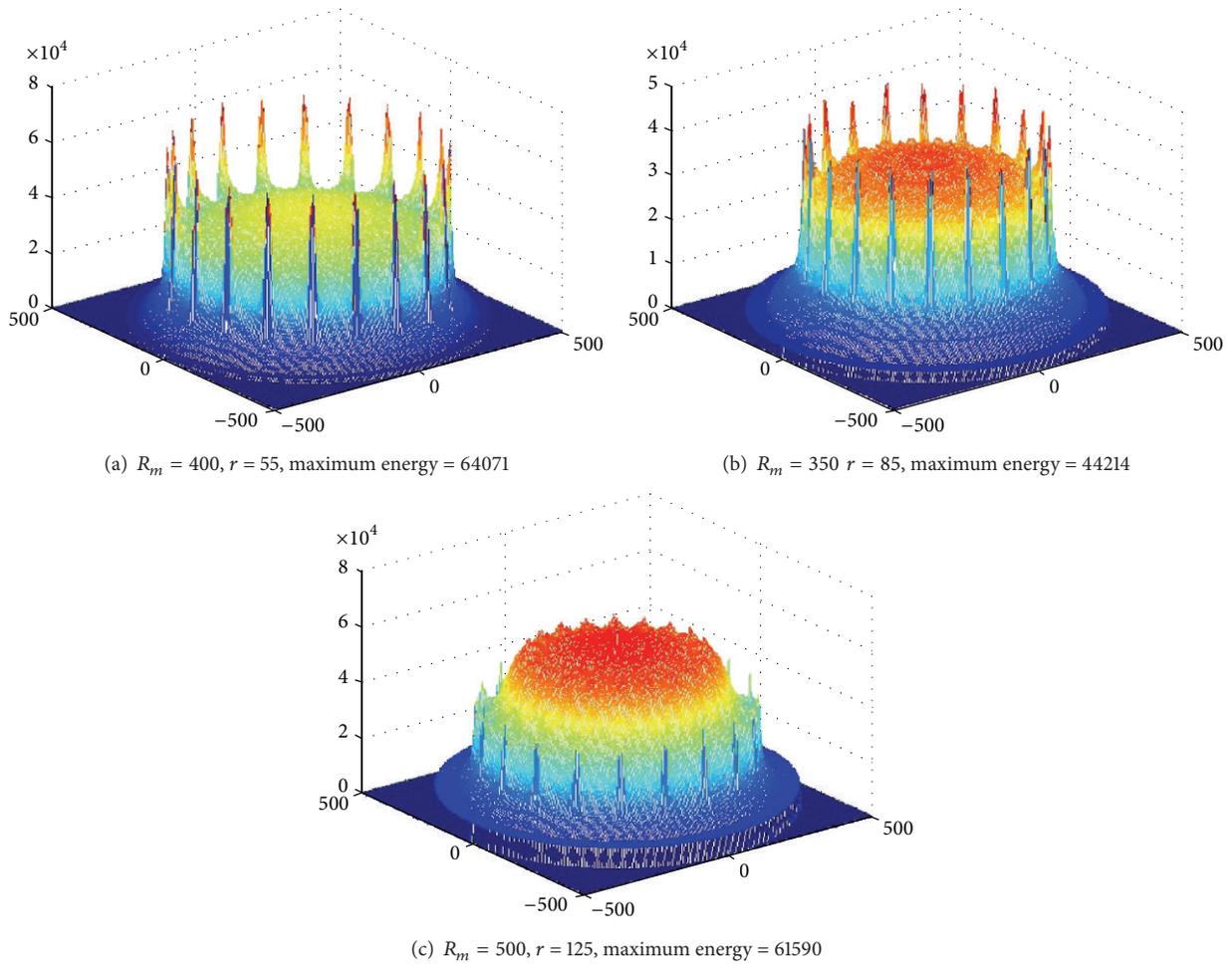


FIGURE 28: Energy consumption under different transmission range (theoretical value,  $R = 500$ , rounds = 20, anchors = 20).

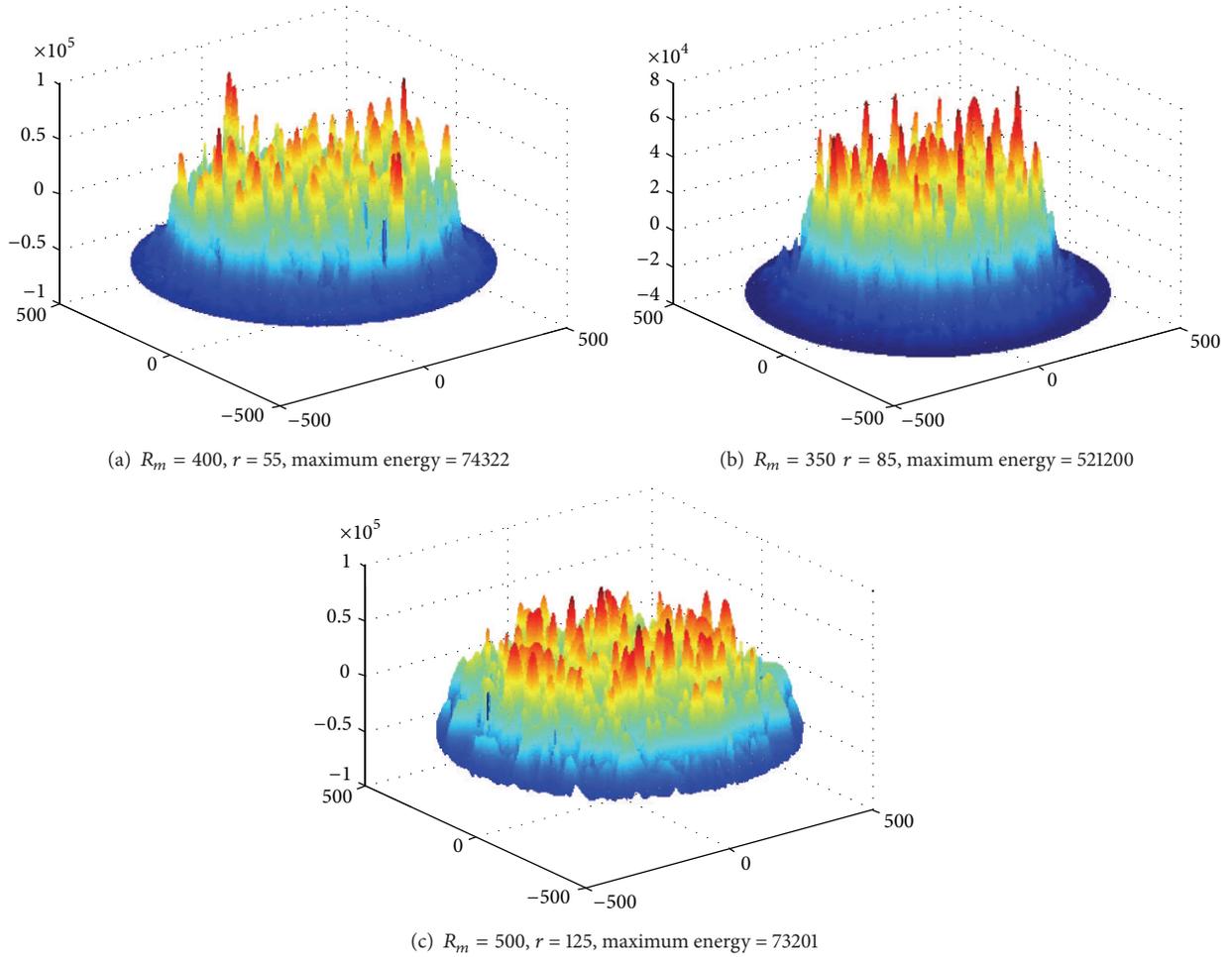


FIGURE 29: Energy consumption under different transmission range (experimental value,  $R = 500$ , rounds = 20, anchors = 20).

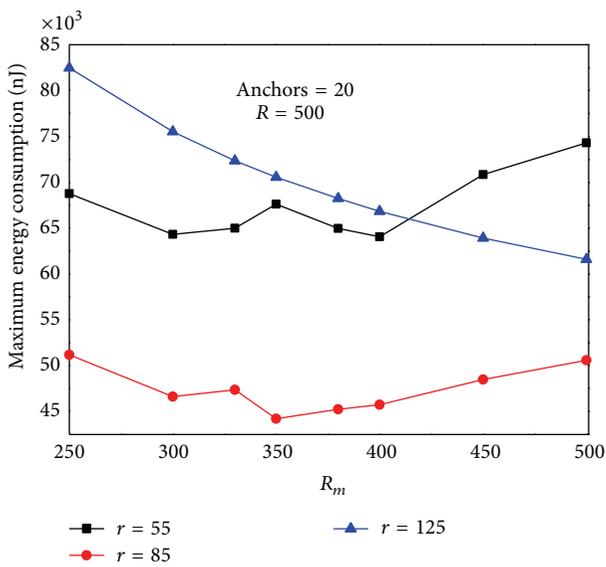


FIGURE 30: Maximum energy consumption under different transmission range and  $R_m$  (theoretical value).

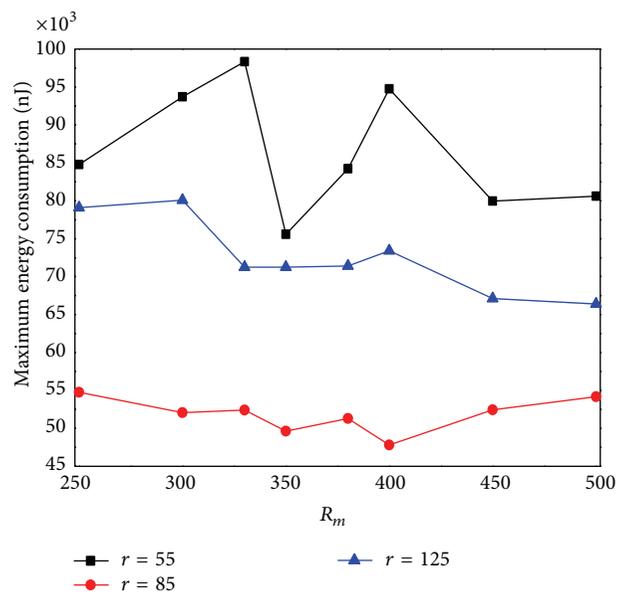


FIGURE 31: Maximum energy consumption under different transmission range and  $R_m$  (experimental value).

```

//To find the optimal sink anchor
(1) Maxenergy = 0;
(2) min_global_energy = ∞; //current value that minimize the maximum energy consumption
(3) For  $x_0 = -R : R$ 
(5) For  $y_0 = R^2 - x_0^2$ 
//sink can pause at any point of the network
(6) Maxenergy = 0;
//set the initial maximum energy consumption as 0
(7) for  $i = 1 : n$  //number of nodes
(8) for each  $r$  do
(9) computer energy  $E(i)$  consumption use Corollary 2
//compute energy cost of  $i$  node when sink is at  $(x_0, y_0)$ 
(10)  $Ea(i) = e(i) + E(i)$ 
//  $e(i)$  represent the actual energy consumption of node  $i$ 
(11) If  $Maxenergy < Ea(i)$  //when the energy
// consumption of  $i$  node is larger than current maximum
(12)  $Maxenergy = Ea(i)$  //compare energy of
//sink( $x_0, y_0$ ) is the energy consumption of node  $i$ 
(13) Endif
(14) End for
(15) End for;
(16) If  $min\_global\_energy > Maxenergy$  // Maxenergy
//represent the maximum energy consumption of sink( $x_0, y_0$ )
(17)  $min\_global\_energy = Maxenergy$ 
//minimize the comprehensive energy
(18)  $x_{best} = x_0; y_{best} = y_0;$ 
//next sink anchor to minimize the comprehensive energy
(19) Endif
(20) End for
(21) End for
(22) Return( $x_{best}, y_{best}$ );
//return the optimal next sink anchor

```

ALGORITHM 1: Find best location( $e(1 : n)$ ).

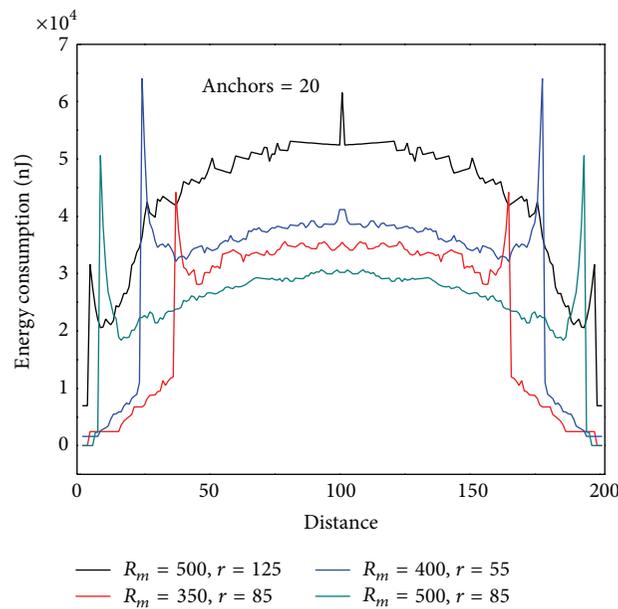


FIGURE 32: Energy consumption after combining  $r$  and  $R_m$  (theoretical value).

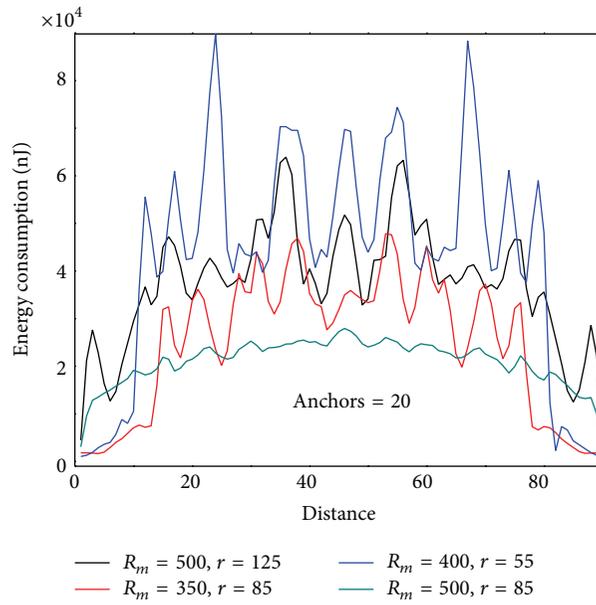


FIGURE 33: Energy consumption after combining  $r$  and  $R_m$  (experimental value).

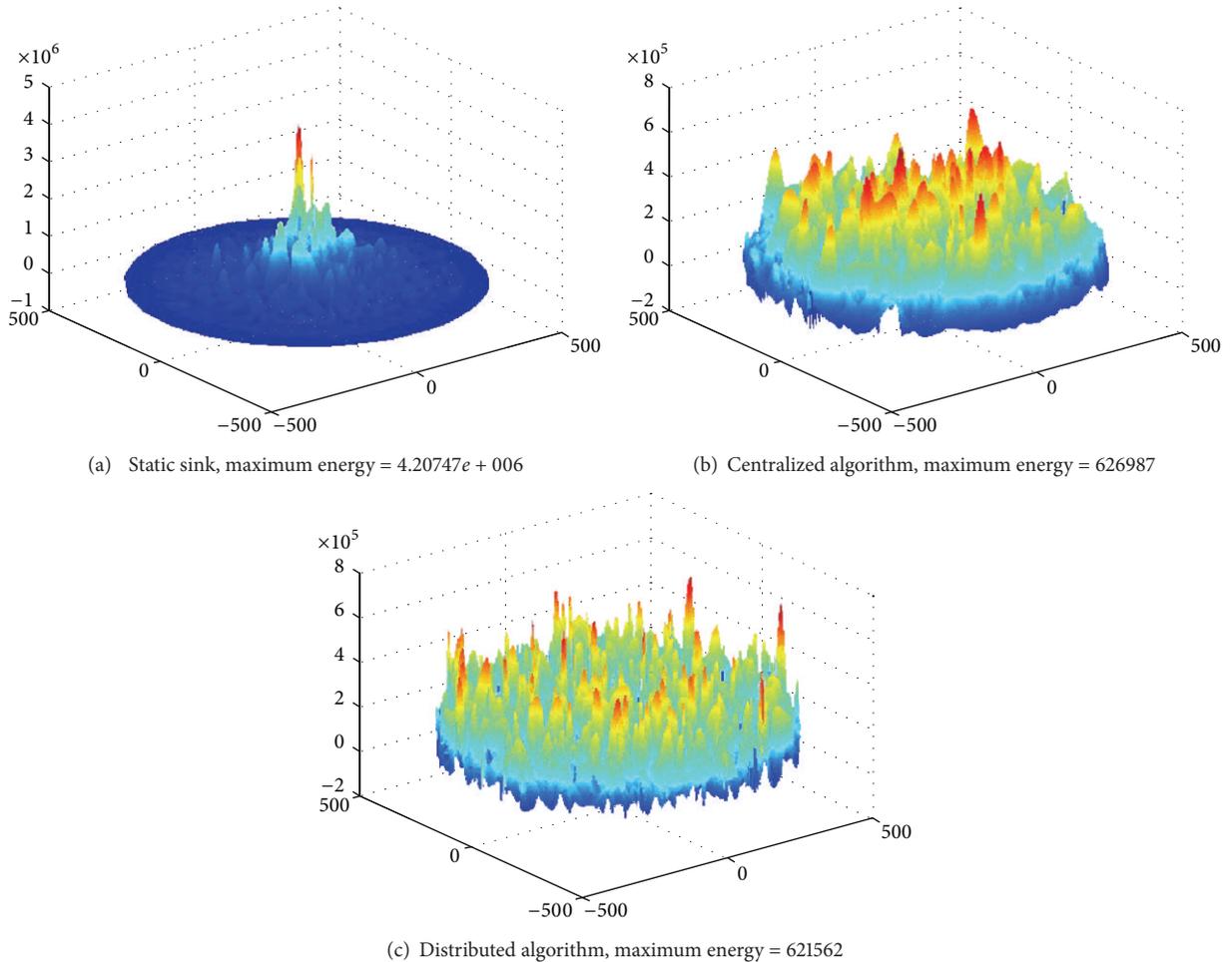


FIGURE 34: Experimental results comparison of static sink, centralized algorithm, and distributed algorithm.

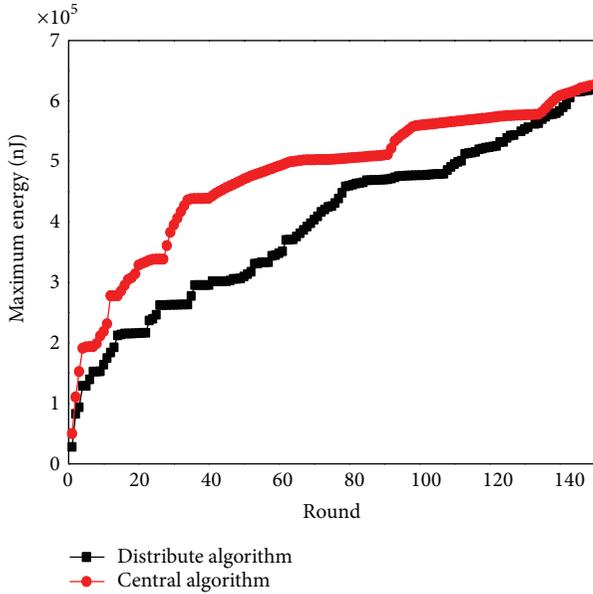


FIGURE 35: Maximum energy comparisons between centralized algorithm and distributed algorithm.

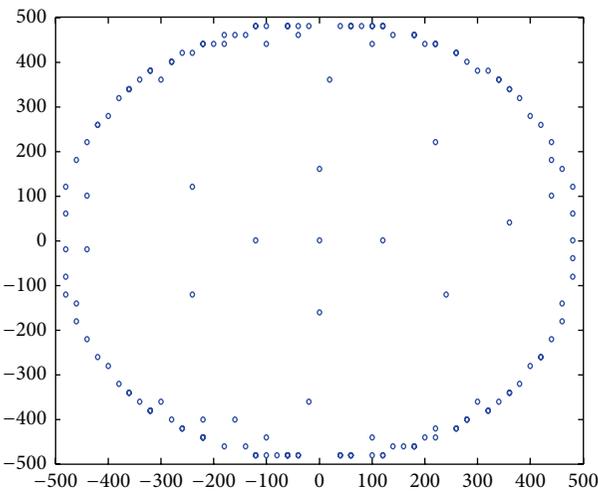


FIGURE 36: Mobile sink positions in distributed algorithm.

network when sink is located anywhere; (2) proposes an online distributed algorithm combined with practical sensor networks to achieve better practical significance. In theory analysis and experiment section, we analyze the factors that affect network life in detail. The conclusion is of more general significance.

Although this paper can present a more precise calculation of the energy consumption of the network, the complexity of distributed mobile algorithm is still relatively large. There are several future research works; although there are many theoretical computing for regular network, most of the networks are irregular. Their optimization of mobile sink has more far-reaching applications. Therefore, the next step we plan to extend the energy consumption calculation

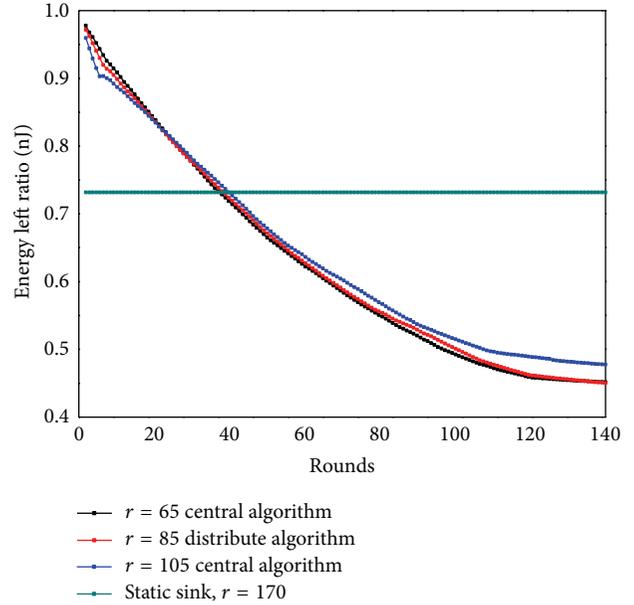


FIGURE 37: Residual energy ratios under static sink and mobile sink.

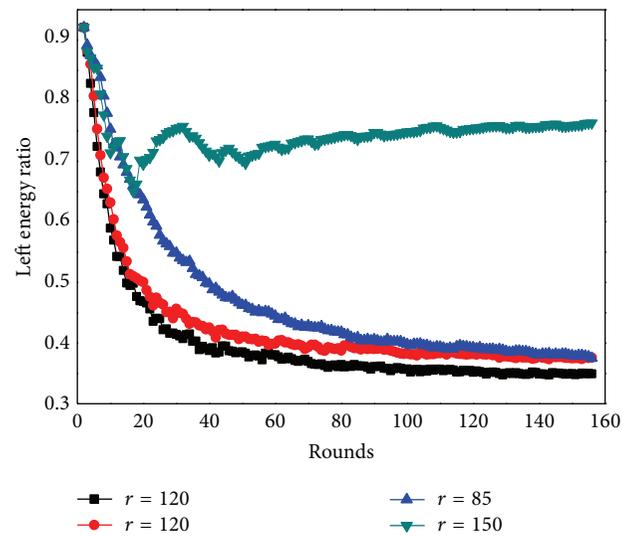


FIGURE 38: Residual energy ratios under different transmission radius.

to general irregular network. Secondly, this paper handles the situation of single mobile sink and multiple mobile sinks, because of its complexity, has so far failed to reach a satisfying method. Based on the conclusion and theoretical calculations in this paper, we will pursue preferable solution for network with multiple mobile sinks [24].

**Conflict of Interests**

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

This work was supported by China Postdoctoral Science Foundation (2012M511756).

## References

- [1] W. Wang, V. Srinivasan, and K.-C. Chua, "Extending the lifetime of wireless sensor networks through mobile relays," *IEEE/ACM Transactions on Networking*, vol. 16, no. 5, pp. 1108–1120, 2008.
- [2] I. Chatzigiannakis, A. Kinalis, and S. Nikolettseas, "Efficient data propagation strategies in wireless sensor networks using a single mobile sink," *Computer Communications*, vol. 31, no. 2, pp. 896–914, 2008.
- [3] A. Liu, X. Jin, G. Cui, and Z. Chen, "Deployment guidelines for achieving maximal lifetime and avoiding energy holes in sensor network," *Information Sciences*, vol. 230, pp. 197–226, 2013.
- [4] S. Olariu and I. Stojmenović, "Design guidelines for maximizing lifetime and avoiding energy holes in sensor networks with uniform distribution and uniform reporting," in *Proceedings of the 25th IEEE International Conference on Computer Communications (INFOCOM '06)*, Barcelona, Spain, April 2006.
- [5] J. Li and P. Mohapatra, "Analytical modeling and mitigation techniques for the energy hole problem in sensor networks," *Pervasive and Mobile Computing*, vol. 3, no. 3, pp. 233–254, 2007.
- [6] X. Tang and J. Xu, "Adaptive data collection strategies for lifetime-constrained wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 19, no. 6, pp. 721–734, 2008.
- [7] X. Wu, G. Chen, and S. K. Das, "Avoiding energy holes in wireless sensor networks with nonuniform node distribution," *IEEE Transactions on Parallel and Distributed Systems*, vol. 19, no. 5, pp. 710–720, 2008.
- [8] M.-H. Chiang and G. T. Byrd, "Neighborhood-aware density control in wireless sensor networks," in *Proceedings of the IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC '08)*, pp. 122–129, June 2008.
- [9] J. Lian, K. Naik, and G. B. Agnew, "Data capacity improvement of wireless sensor networks using non-uniform sensor distribution," *International Journal of Distributed Sensor Networks*, vol. 2, no. 2, pp. 121–145, 2006.
- [10] A. Varga, "The OMNET++ discrete event simulation system," version 3.3, <http://www.omnetpp.org>.
- [11] A. Hossain, T. Radhika, S. Chakrabarti, and P. K. Biswas, "An approach to increase the lifetime of a linear array of wireless sensor nodes," *International Journal of Wireless Information Networks*, vol. 15, no. 5, pp. 72–81, 2008.
- [12] M. Cardei, J. Wu, and M. Lu, "Improving network lifetime using sensors with adjustable sensing ranges," *International Journal of Sensor Networks*, vol. 1, no. 1, pp. 41–49, 2006.
- [13] P. Chen, B. O'Dea, and E. Callaway, "Energy efficient system design with optimum transmission range for wireless ad hoc networks," in *Proceedings of the International Conference on Communications (ICC '02)*, vol. 2, pp. 945–952, May 2002.
- [14] A. A. Somasundara, A. Kansal, D. D. Jea, D. Estrin, and M. B. Srivastava, "Controllably mobile infrastructure for low energy embedded networks," *IEEE Transactions on Mobile Computing*, vol. 5, no. 8, pp. 958–972, 2006.
- [15] Y. T. Hou, Y. Shi, H. D. Sherali, and S. F. Midkiff, "On energy provisioning and relay node placement for wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2579–2590, 2005.
- [16] J. Luo, J. Panchard, M. Piorkowski, M. Grossglauser, and J. P. Hubaux, "MobiRoute: routing towards a mobile sink for improving lifetime in sensor networks," in *Proceedings of the IEEE International Conference on Distributed Computing in Sensor Networks (DCOSS '06)*, vol. 4026, pp. 480–497, 2006.
- [17] J. Luo and J.-P. Hubaux, "Joint mobility and routing for lifetime elongation in wireless sensor networks," in *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '05)*, pp. 1735–1746, Miami, Fla, USA, March 2005.
- [18] M. Ma and Y. Yang, "SenCar: an energy-efficient data gathering mechanism for large-scale multihop sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 18, no. 10, pp. 1476–1488, 2007.
- [19] Y. Bi, L. Sun, J. Ma, N. Li, I. A. Khan, and C. Chen, "HUMS: an autonomous moving strategy for mobile sinks in data-gathering sensor networks," *Eurasip Journal on Wireless Communications and Networking*, vol. 2007, Article ID 64574, 15 pages, 2007.
- [20] D. Vass and A. Vidács, "Positioning mobile base station to prolong wireless sensor network lifetime," in *Proceedings of the 1st International Conference on Emerging Networking Experiments and Technologies (CoNEXT '05)*, pp. 300–301, Toulouse, France, October 2005.
- [21] W. Liang, P. Schweitzer, and Z. Xu, "Approximation algorithms for capacitated minimum forest problems in wireless sensor networks with a mobile sink," *IEEE Transactions on Computers*, vol. 62, no. 10, pp. 1932–1944, 2013.
- [22] E. Ekici, Y. Gu, and D. Bozdog, "Mobility-based communication in wireless sensor networks," *IEEE Communications Magazine*, vol. 44, no. 7, pp. 56–62, 2006.
- [23] G. Shi, M. Liao, M. Ma, and Y. Shu, "Exploiting sink movement for energy-efficient load-balancing in wireless sensor networks," in *Proceedings of the 1st ACM International Workshop on Foundations of Wireless Ad Hoc and Sensor Networking and Computing (FOWANC '08)*, vol. 5, pp. 39–44, May 2008.
- [24] J. Luo and J. P. Hubaux, "Joint sink mobility and routing to maximize the lifetime of wireless sensor networks: the case of constrained mobility," *IEEE/ACM Transactions on Networking*, vol. 18, no. 3, pp. 871–884, 2010.

## Research Article

# Global Optimization Ensemble Model for Classification Methods

**Hina Anwar, Usman Qamar, and Abdul Wahab Muzaffar Qureshi**

*Department of Computer Engineering, College of Electrical & Mechanical Engineering (E&ME), National University of Sciences and Technology (NUST), H-12, Islamabad 46000, Pakistan*

Correspondence should be addressed to Usman Qamar; [usmanq@ceme.nust.edu.pk](mailto:usmanq@ceme.nust.edu.pk)

Received 24 February 2014; Accepted 19 March 2014; Published 27 April 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Hina Anwar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supervised learning is the process of data mining for deducing rules from training datasets. A broad array of supervised learning algorithms exists, every one of them with its own advantages and drawbacks. There are some basic issues that affect the accuracy of classifier while solving a supervised learning problem, like bias-variance tradeoff, dimensionality of input space, and noise in the input data space. All these problems affect the accuracy of classifier and are the reason that there is no global optimal method for classification. There is not any generalized improvement method that can increase the accuracy of any classifier while addressing all the problems stated above. This paper proposes a global optimization ensemble model for classification methods (GMC) that can improve the overall accuracy for supervised learning problems. The experimental results on various public datasets showed that the proposed model improved the accuracy of the classification models from 1% to 30% depending upon the algorithm complexity.

## 1. Introduction

According to Han and Kamber [1], “Data Mining is known to be a part of knowledge discovery (KDD) process in which data is analysed and summarized from different perspectives and converted into useful information. It helps in extracting the hidden and valid data which has the potential of being transformed into useful information.” It is similar to machine learning process and can also be termed as supervised learning process. Supervised learning is the process of data mining for deducing rules from training datasets. A broad array of supervised learning algorithms exists, every one of them with its own advantages and drawbacks.

In classification the first step is to divide the data in two portions known as training set and testing set [2]. In these datasets, one attribute must be necessarily defined as class. According to Han et al. [2], the two steps of the classification task are model construction and model usage. In this task, the model is built with the help of trained dataset and then this trained model is used to allocate the unseen records as precisely as possible. While training dataset is used to build and train the model, the testing dataset is used to validate and test the model accuracy [3], which bring us to some of the basic issues that affect the accuracy of a classifier while solving a supervised learning problem. For instance, the bias-variance

tradeoff, the dimensionality curse, or the noise in the dataset all contribute towards a decreasing accuracy. Bias arises when the classifier cannot represent the true function; that is, the classifier under fits the data; that is, when it is training on any data set than for a specific input value it is methodically inaccurate when predicting the right outcome for that input value. In contrast to this, variance occurs when the algorithm over fits the data and for a specific input value in a dataset it gives a different outcome every time the training dataset is changed. Another problem that can affect the accuracy of a classifier is the dimensionality or the number of attributes or features in a dataset. If we input a large number of attributes in a classification algorithm even for problems where decision depends on subset of all those attributes, then performance of the classifier will be clouded by high variance due to high dimension of dataset. Therefore if a dataset with high dimension is being used the classifier must make a tradeoff between high bias and low variance. The classification results are also altered by the noise in data, that is, redundant records, incorrect records, missing records, outliers, and so forth. All these problems affect the accuracy of a classifier. Usually the improvements done in a classifier or ensemble model are limited to a very narrow spectrum and they cannot be applied to another classifier under the same conditions.

TABLE I: Parameter values in GA.

Parameter	Value
Selection scheme	Roulette wheel
Crossover type	shuffle
Probability of crossover	0.5
Probability for initial population	0.5
Probability of mutation	1/number of attribute
Maximal fitness	Infinity

Classification accuracy is normally improved through ensemble models like bagging (which averages the prediction of a number of classification models), boosting (it uses the voting scheme over a number of classification models), or a combination of classifiers from different or same families as discussed in Section 2.

Therefore, in this paper we propose a global optimization using the idea of ensemble models for classification methods and prove through experimental results that our model improves the classification accuracy of various classifiers on various different public datasets. Section 2 of the paper gives an insight into the previously related work. Design of the proposed model is given in Section 3. Section 4 explains the implementation. Section 5 gives the result and analysis. Section 6 contains the conclusion and future work.

## 2. Literature Review

As mentioned earlier that so far no global optimization ensemble model is present which can help in improving the classification and prediction accuracy for supervised learning problems which are generally affected by a spectrum of issues like dimensionality, accuracy rate, data quality, and so forth. Although, no global solution exists for these problems but some other efforts have been made to resolve these issues and all of them are either algorithm specific or data specific. Every approach has tackled the problem of classification accuracy rate from a different angle and perspective. One such work is [16] where Dash et al. have proved through comparison of various classification techniques like support vector machine (SVM) with polynomial kernel, support vector machine with RBF kernel, radial basis function network (RBFN), and multilayer perceptron network (MLP) with and without feature extraction. It was found that for construction of high performance classification model for microarray dataset, partial least square (PLS) regression method is the suitable feature selection method instead of hybrid dimensionality reduction scheme and feature selection combined with various classification techniques can yield better results. Lin and Chen in [17] combined PSO- (particle swarm optimization-)based approach with commonly used classification technique LDA (linear discriminant analysis). This research also emphasizes the importance of feature selection and its positive effect on classification accuracy. Authors of this study have compared the performance of this combined model called PSOLDA with many other feature selection techniques like forward selection, back propagation selection, and so forth and shown through experimental results that for many public datasets

the proposed combined model (PSOLDA) has higher classification accuracy rate. Bryll et al. [18] developed a new wrapper method AB (attribute bagging) to improve the classification accuracy implementing a two-stage method in which first a suitable size was provided for training data and then randomly a subset of attributes was selected for voting scheme. This method was compared with bagging which was used with some decision tree algorithms and some rule induction algorithms, and it was found that the AB performs better in terms of accuracy and constancy. And authors conclude that attribute partitioning is better than data partitioning for improving the accuracy in an ensemble method. Abbott [19] compared boosting with an ensemble of models across the algorithm families. These combined models used voting as the selection scheme and authors report that boosting performs better because it focuses on complicated cases in data and takes into account the confidence value of a particular classification decision. Sohn and Lee [20] tried to improve the classification accuracy of algorithms like neural network and decision trees by applying different approaches including bagging, boosting, and clustering. However for the particular problem of road traffic accident classification clustering leading to classification was found to be more effective. Smith and Martinez [21] suggested that outliers and noise should be eliminated from the dataset as it will yield better results in terms of classification accuracy. Because by removing or filtering these instances the dataset becomes clean of all the cases that could be misclassified. As there is no general definition or guide available as to what noise is and what an outlier is, therefore the identification of these two elements in any dataset is difficult. Furthermore PRISM was found to be one of the best algorithms for finding cases that could be outliers. Dimensionality reduction problem has been an interesting topic for researchers in a diverse spectrum of fields like image detection, voice detection, microarrays, neural network patterns, and so forth. As discussed by Zamalloayz et al. [22], Liu et al. [23], and Raymer et al. [24] genetic algorithm (GA) is quite a popular method under research and is found to be quite effective for feature selection and classification accuracy improvement. All these researches related to GA are data specific or algorithm specific. In [22] the performance of GA is compared with other feature reduction and extraction techniques like liner discriminant analysis (LDA); principle component analysis (PCA) for one dataset GA was found to perform better while for the other dataset LDA and PCA showed promising results. In [23] the genetic algorithm is combined with the boosting technique in order to improve accuracy of classification. The improved version assigns higher weight to the misclassified instances in order to shift the focus on them in the next iteration. This process tends to achieve higher accuracy with less number of evaluations than the original GA. In [24] genetic algorithm is implemented in combination with K-nearest neighbor classifier and feature extraction; reduction and classifier training are all done simultaneously and results are compared with other industry standard feature extraction and reduction technique like liner discriminant analysis and sequential floating forward feature selection.

Despite all this extensive work on ensemble methods and feature reduction problem and various classification

TABLE 2: (a) Parameter optimization using grid search. (b) Parameter configuration for classifier.

(a)

Parameter	Operator	Grid range	Combination	Scale	Optimal value
Population size	GA-layer 1	2–100	2, 3, 6, 11, 18, 27, 37, 50, 65, 81, 100	Quadratic	6
Maximum number of generation	GA-layer 1	1–50	1, 6, 11, 16, 21, 26, 30, 35, 40, 45, 50	Linear	16
Number of iterations	CV-layer 2	2–50	2, 4, 6, 10, 14, 19, 26, 33, 41, 50	Quadratic	10
Sampling Size	Bagging-layer 3	0–1.0	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1	Linear	0.6
Number of iterations	Bagging-layer 3	1–100	1, 2, 5, 10, 17, 26, 37, 50, 64, 81, 100	Quadratic	10

(b)

Operator name	Parameter configuration
ID 3	Criterion: information_gain
	Minimal size of split: 4
	Minimal leaf size: 2
	Minimal gain: 0.1
Decision tree	Criterion: Information_gain
	Minimal size for split: 4
	Minimal leaf size: 2
	Minimal gain: 0.1
	Maximal depth: 20
Random forest	Confidence: 0.5
	Number of trees: 10
	Criterion: Information_gain
	Minimal leaf size: 2
	Minimal gain: 0.1
Rule induction	Maximum depth: 20
	Confidence: 0.5
	Criterion: information_gain
	Sample ratio: 0.7
K-NN	Pureness: 0.6
	Minimal prune benefit: 0.6
	K nearest neighbors: 11
	Weighted vote: true
Naïve bayes	Measure type: nominal measures
	Nominal measure: Dice similarity
W-AODE	Laplace correction: true
W-PART	Frequency for super parents: 1.0
	Confidence threshold: 0.5
W-J48	Minimum objects per leaf: 2.0
	Confidence threshold: 0.5
	Minimum objects per leaf: 2.0

algorithms for improving the accuracy rate in classification, there is no global optimization ensemble model suggested so far that can improve the accuracy of classification methods with any dataset. Therefore in this paper we design and implement such a global optimization model.

### 3. Design

The idea was to implement the concept of ensemble model in order to create an improved global model. Figure 1 shows the overall design of the ensemble model.

Layer 1 was providing antidote for dimensionality curse. As discussed in the literature review the dimensionality reduction or feature reduction is necessary in order to improve the classification accuracy. Therefore in our model the first layer contains the data set, preprocessing operator, and a feature reduction operator. For feature selection, genetic algorithm (GA) is used as it has shown to produce better results than other feature reduction techniques [23–25].

Maximal fitness of GA is set to infinity as there is no absolute maxima for the fitness function which means the GA will keep on selecting the best of best until the stop criteria are

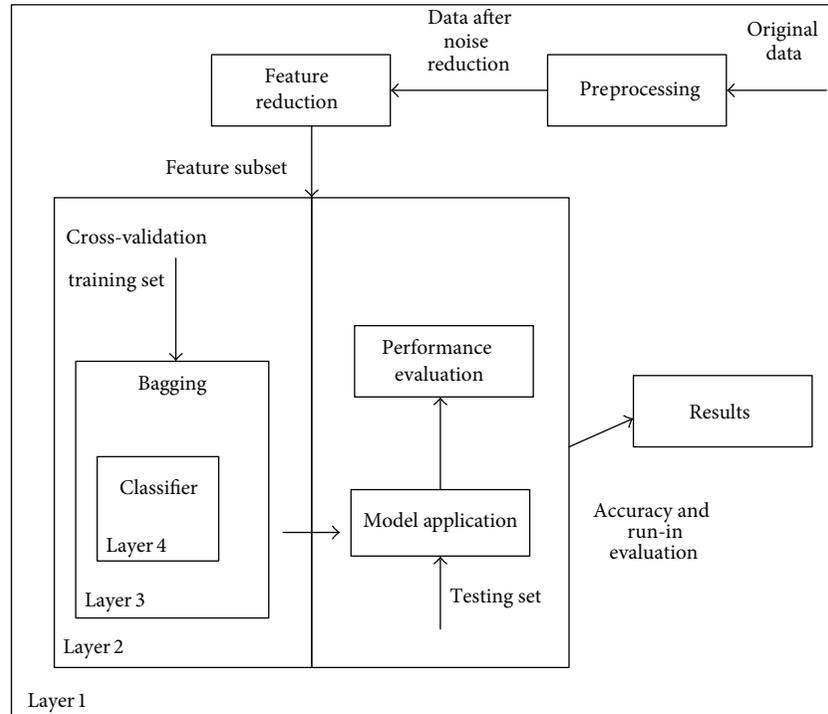


FIGURE 1: Design of global optimization ensemble model for classification methods (GMC).

met which in this case is the maximum number of generation. Roulette wheel selection scheme was used for selecting individuals because it has the obvious advantage that it does not ignore or discard any individuals and each individual is given a chance of being chosen as even the weakest of individuals might be hiding valuable information. And as we are striving for a global solution, therefore a selection method that preserves diversity and is fast to converge sounds good. Crossover type was set to shuffle because shuffle crossover is related to uniform crossover. A single crossover position (as in single-point crossover) is selected. But before the variables are exchanged, they are randomly shuffled in both parents. After recombination, the variables in the offspring are unshuffled. This removes positional bias as the variables are randomly reassigned each time crossover is performed. The parameter values chosen for GA are shown in Table 1.

Parameter optimization for the operators in each layer was done by implementing global optimization operator using grid search. This methodology involves setting up of grids in the decision space and evaluating the values of the objective function at each grid point. The point which corresponds to the best value of the objective function is considered to be the optimum solution. For all the layers, a total of 5 parameters were optimized using grid search optimizations. From each attribute 11 combinations were proposed; this means for optimizing these 5 parameters total 161051 combinations were tested. Table 2(a) shows all the parameter and there optimized values.

In layer 2 partition of training and testing dataset was done using X-fold crossvalidation. The data set is divided into  $n$  subsets, and the holdout method is repeated  $k$  times. Each

time, one of the  $n$  subsets is used as the test set and the other  $n - 1$  subsets are put together to form a training set. Then the average error across all  $n$  trials is computed. The advantage of this method was that it matters less how the data gets divided. Every data point gets to be in a test set exactly once and gets to be in a training set  $n - 1$  times. Besides, the variance of the resulting estimate is also reduced as  $n$  is increased. Stratified sampling scheme was used in CV with number of iteration set to 10 as shown in Table 2(a). In stratified sampling the random subsets are created but the distribution of class in those subsets is the same as the whole dataset. Thus this type of sampling reduces variance. For example we have a data set of 180 employees and we want a sample set of 40 employees. The first step is to calculate the percentage of male female in each group, that is,

- (i) percentage of male members in full-time category =  $90/180 = 50\%$ ,
- (ii) percentage of male members in part-time category =  $18/180 = 10\%$ ,
- (iii) percentage of female members in full-time category =  $9/180 = 5\%$ ,
- (iv) percentage of female members in part-time category =  $63/180 = 35\%$ .

This calculation tell us that of our desired sample of 40 employees, 50 percent should be male (full time), 10 percent should be male (part-time), 5 percent should be female (full-time), and 35 percent should be female (part-time). This means that we need to calculate the 50% of 40 which is 20. Similarly 10% of 40 is 4, 5% of 40 is 2, and 35% of 40 is 14. This

is the final ratio of records in each category in our sample of 40 employees.

Layer 3 did an optimal bias-variance tradeoff. Accuracy improvement is done by implementing bootstrap aggregation (bagging). Bagging is a machine learning ensemble meta-algorithm which reduces both bias and variance in order to help avoid overfitting. Although it is usually applied to decision tree models, it can be used with any type of model. Bagging is a special case of the model averaging approach. Parameter setting for bagging is shown in Table 2(a). We are using bagging instead of boosting because error = noise error + bias + variance bagging can reduce both bias and variance but mostly it reduces just variance and it hardly ever increases error. For high-bias classifiers, it can reduce bias and for high-variance classifiers, it can reduce variance, while boosting in the early iterations is primarily a bias-reducing method. In later iterations, it appears to be primarily a variance-reducing method. It may increase error and margins and is not good with data with noise. That is the reason why we chose bagging instead of boosting for bias and variance tradeoff.

Classifiers were placed in layer 4 with parameters configuration done according to the dataset. All classifier parameters were set to obtain the optimal model in order to reduce the bias. The setting used for each classifier is shown in Table 2(b).

#### 4. Implementation

Implementation and testing are done using core i3 processor with 4 GB RAM, while coding is done using XML. Preprocessing is performed on every dataset according to requirements of the classifier used in order to remove noise from data and do type conversions. The model is implemented and tested in RapidMiner5 [15].

*Step 1* (algorithm selection). As we are optimizing the model for supervised learning problems, therefore the following linear and nonlinear classifiers were selected, implemented, and tested.

*Step 2* (data set selection). Datasets from various different fields are selected such as banking, medicine, and census data. Selection of the datasets was based on

- (i) relevant to classifier,
- (ii) frequency of citations,
- (iii) a wide range of values in terms of attributes classes and attribute characteristic.

In total 7 datasets from various fields are used for experimentation. The classifiers used in the implementation and the details of the datasets used are given in Tables 3 and 4 respectively.

Suitable classifier for each dataset is selected and indicated as shown in Table 5.

*Step 3* (simple classification using validation technique). First each dataset is classified using the classifier mentioned for each dataset and the results are validated using the X-fold

TABLE 3: Classifiers.

Classifiers
KNN [4, 5]
Decision tree [6, 7]
ID3 [8, 9]
Random forest [10, 11]
Logistic regression [12, 13],
Rule induction [14]
W-AODE, W-PART, W-Prism, W-J48 [15]

cross-validation technique where  $x = 10$  for all classifiers. Sampling technique used for validation is “stratified sampling.”

Results consisting of classification accuracy and classification error are recorded for each classifier.

*Step 4* (classification using global optimization ensemble model for classification methods (GMC)). All the classifiers are now encapsulated in the proposed generic optimization ensemble model and executed for results. Parameters of all the classifiers are the same as in Step 3. Now the improved results consisting of optimized classification accuracy are recorded for every classifier and compared with the previous result in order to calculate the improvement percentage.

#### 5. Results

The results for each data set and the corresponding accuracy comparison between simple classification and GMC model are given in this section.

Table 6 shows that using the GMC model for optimizing the classification accuracy for cancer dataset has improved from 1.13% to 29.76% depending on the classifier and the bias-variance tradeoff its inner complexity offers.

Table 7 shows that using the GMC model for optimizing the classification accuracy for heart disease dataset has improved from 2.4% to 14.54% depending on the classifier and the bias-variance tradeoff its inner complexity offers.

Table 8 shows that using the GMC model for optimizing the classification accuracy for wine dataset has improved from 3.92% to 19.67% depending on the classifier and the bias-variance tradeoff its inner complexity offers.

As shown in Table 9, using the GMC model for optimizing the classification accuracy for adult income dataset has improved from 1% to 6.5% depending on the classifier and the bias-variance tradeoff its inner complexity offers.

As shown in Table 10, using the GMC model for optimizing the classification accuracy for sonar dataset has improved from 4.82% to 15.36% depending on the classifier and the bias-variance tradeoff its inner complexity offers.

As shown in Table 11, using the GMC model for optimizing the classification accuracy for educational dataset has improved from 8% to 26% depending on the classifier and the bias-variance tradeoff its inner complexity offers.

TABLE 4: Data set details.

Data set	Number of cases	Number of attribute	Number of classes	Attribute characteristics	Missing values
Cancer dataset	699	9	2	Numeric	Yes
Diabetes dataset	768	9	2	Integer and real	No
Heart disease dataset	303	14	2	Categorical, integer, and real	Yes
Adult income dataset	1000	15	2	Integer and nominal	No
Wine dataset	178	13	3	Real and integer	No
Sonar dataset	208	61	2	Real and nominal	yes
Educational progress dataset	50	9	3	Nominal	No

TABLE 5: Data set and suitable classifiers.

Dataset	Classifier	Capabilities
All datasets	K-NN	Polynomial, numerical, binomial attributes, and labels can handle missing values
All datasets	Decision tree	Polynomial, numerical, and binomial attributes cannot handle numeric labels and can handle missing values
Heart, wine, and educational and sonar dataset	Rule induction	
Cancer, heart, adult income dataset	ID3	Can only handle binomial and polynomial labels and attributes and cannot handle missing values
All datasets	W-AODE	
All datasets	W-Prism	
Educational progress and sonar and adult income dataset	Random forest	Polynomial, numerical, and binomial attributes cannot handle numeric labels and cannot handle missing values
All datasets	W-PART	
All datasets	W-J48	
Sonar, diabetes, cancer, and adult income dataset	Logistic regression	Numerical attributes and binomial labels cannot handle missing values

TABLE 6: Results for cancer dataset: comparison of optimized classification accuracy using GMC model with simple classification using different classifiers.

Algorithm	Classification accuracy	Optimized classification accuracy	Improvement %
K-NN	66.81%	96.57%	<b>29.76%</b>
Decision tree	94.42%	96.71%	<b>2.29%</b>
ID3	66.52%	85.27%	<b>18.52%</b>
W-PART	94.71%	97.28%	<b>2.57%</b>
W-Prism	90.13%	96.28%	<b>6.15%</b>
W-J48	94.71%	96.71%	<b>2%</b>
W-AODE	97.00%	100%	<b>3%</b>
Logistic regression	95.01%	96.14%	<b>1.13%</b>

TABLE 7: Results for heart disease dataset: comparison of optimized classification accuracy using GMC model with simple classification using different classifiers.

Algorithm	Classification accuracy	Optimized classification accuracy	Improvement %
K-NN	50.82%	59.75%	8.93%
Decision tree	44.89%	59.43%	14.54%
ID3	47.52%	55.48%	8.24%
W-PART	50.52%	60.08%	9.56%
W-Prism	47.51%	56.09%	8.58%
W-AODE	55.47%	61.13%	5.66%
W-J48	49.87%	61.05%	11.18%
Rule induction	57.72%	59.76%	2.4%

TABLE 8: Results of wine dataset: comparison of optimized classification accuracy using GMC model with simple classification using different classifiers.

Algorithm	Classification accuracy	Optimized classification accuracy	Improvement %
K-NN	70.75%	90.42%	<b>19.67%</b>
Decision tree	91.57%	95.49%	<b>3.92%</b>
W-PART	90.42%	96.67%	<b>6.25%</b>
W-Prism	52.32%	61.27%	<b>8.95%</b>
W-AODE	71.34%	75.26%	<b>3.92%</b>
W-J48	90.46%	96.63%	<b>6.17%</b>
Rule induction	86.37%	93.27%	<b>6.9%</b>

TABLE 9: Results of adult income dataset: comparison of optimized classification accuracy using GMC model with simple classification using different classifiers.

Algorithm	Classification accuracy	Optimized classification accuracy	Improvement %
K-NN	76.70%	83.20%	<b>6.5%</b>
Decision tree	80.00%	82.20%	<b>2.20%</b>
ID3	75.60%	78.60%	<b>3%</b>
W-PART	81.00%	83.50%	<b>2.4%</b>
W-Prism	81.10%	82.20%	<b>1.1%</b>
W-AODE	80.80%	82.60%	<b>1.8%</b>
W-J48	81.50%	83.00%	<b>1.5%</b>
Random forest	76.10%	77.30%	<b>1.2%</b>
Logistic regression	79.00%	80.00%	<b>1%</b>

TABLE 10: Results of sonar dataset: comparison of optimized classification accuracy using GMC model with simple classification using different classifiers.

Algorithm	Classification accuracy	Optimized classification accuracy	Improvement %
K-NN	69.71%	74.57%	<b>4.86%</b>
Decision Tree	73.57%	83.67%	<b>10.1%</b>
W-PART	75.48%	83.17%	<b>7.69%</b>
W-Prism	48.02%	63.38%	<b>15.36%</b>
W-J48	70.24%	82.21%	<b>11.97%</b>
Rule induction	71.66%	76.48%	<b>4.82%</b>
Random forest	68.26%	75.36%	<b>7.1%</b>
Logistic regression	74.55%	80.29%	<b>5.74%</b>

TABLE 11: Results of educational dataset: comparison of optimized classification accuracy using GMC model with simple classification using different classifiers.

Algorithm	Classification accuracy	Optimized classification accuracy	Improvement %
K-NN	46%	54%	<b>8%</b>
Decision Tree	42%	56%	<b>14%</b>
ID3	20%	44%	<b>24%</b>
W-PART	32%	54%	<b>22%</b>
W-Prism	24%	50%	<b>26%</b>
W-J48	44%	58%	<b>14%</b>
W-AODE	46%	56%	<b>10%</b>
SVM	60%	76%	<b>16%</b>
Random forest	48%	58%	<b>12%</b>
Rule induction	44%	54%	<b>10%</b>

TABLE 12: Results of diabetes dataset: comparison of optimized classification accuracy using GMC model with simple classification using different classifiers.

Algorithm	Classification accuracy	Optimized classification accuracy	Improvement %
K-NN	73.70%	77.48%	4%
Decision tree	74.0%	75.39%	1.39%
W-PART	73.83%	77.34%	3.51%
W-Prism	57.42%	67.97%	10.55%
W-J48	74.08%	77.22%	3.14%
W-AODE	66.54%	69.14%	2.6%
Logistic regression	76.00%	77.95%	1.95%

As shown in Table 12, using the GMC model for optimizing the classification accuracy for diabetes dataset has improved from 1.39% to 10.55% depending on the classifier and the bias-variance tradeoff its inner complexity offers.

It can be seen that for K-NN the improvement in accuracy is as high as 30%. This is a significant increase in accuracy and shows the effectiveness of the GMC model. For other algorithms such as decision tree or logistic regression the increase in classification accuracy is varying between 1% and 3%. Although this may indicate the shortcoming of GMC model, but this is not the case in reality. During the experimentation it was noted that for some supervised techniques such as decision tree, the accuracy of the classifier was already very high (90% or more); therefore the possibility of further improving the classifier was rather limited. Thus in this case GMC model could only increase the accuracy by a small amount. For example in the case of decision tree, the GMC model increased the average accuracy from 91.1% to 93.2%. However for other supervised learning algorithms there existed a large gap within which the classifier accuracy may further be increased. This explains the significant increase classifier accuracy in algorithms such as K-NN. Thus although GMC model is dependent upon the algorithm in terms of how much classifier accuracy may be improved, yet it has in all cases increased the accuracy of the classifier.

## 6. Conclusion and Further Work

In order to solve the basic issues of supervised learning problems like dimensionality reduction, bias-variance tradeoff, and noise, we used the concept of ensemble models to design an optimized global ensemble model for classification methods (GMC). The model was designed in layers with each layer solving one of the basic issues of supervised learning. We proved through experimentation that if classifiers are enclosed in our model there accuracy improves from 1% to 30% depending upon the algorithm complexity and its capability of handling bias and variance. Our model yielded better results than when the classifiers were used alone or in combination.

The model can be further optimized for extremely large data set in real time. In that case the optimization will focus on the reduction of execution time as well as further improvement in accuracy. Parallel processing can be introduced into the model for minimizing time. There are a lot of

optimization techniques available and a separate research and comparison can be carried out between all those techniques and the effect of those techniques on the global model. Furthermore, research can be carried out on this model for unsupervised learning problems with data sets related to more diverse fields.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

- [1] J. Han and M. Kamber, *Data Mining: Concepts & Techniques*, 2nd edition, 2006.
- [2] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, 2nd edition, 2005.
- [3] S. Kotsiantis, D. Kanellopoulos, and P. Pintelas, "Data preprocessing for supervised learning," *International Journal of Computer Science*, vol. 1, no. 2, pp. 111–117, 2006.
- [4] G. Toussaint, "Geometric proximity graphs for improving nearest neighbor methods in instance-based learning and data mining," *International Journal of Computational Geometry and Applications*, vol. 15, no. 2, pp. 101–150, 2005.
- [5] O. Söder, 2008, [http://www.fon.hum.uva.nl/praat/manual/kNN\\_classifiers.1...What\\_is\\_a\\_kNN\\_classifier\\_.html](http://www.fon.hum.uva.nl/praat/manual/kNN_classifiers.1...What_is_a_kNN_classifier_.html).
- [6] S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 21, no. 3, pp. 660–674, 1991.
- [7] J. Li, Department of Statistics, The Pennsylvania State University, <http://sites.stat.psu.edu/~jiali/course/stat597e/notes2/trees.pdf>.
- [8] H. Wang, L. Wang, and L. Yi, "Maximum entropy framework used in text classification," in *Proceedings of the IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS '10)*, vol. 2, pp. 828–833, October 2010.
- [9] T. M. Mitchell, *Machine Learning*, McGraw-Hill, 1997.
- [10] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [11] T. K. Ho, "Random decision forest," in *Proceedings of the 3rd International Conference on Document Analysis and Recognition*, pp. 278–282, Montreal, QC, Canada, August 1995.
- [12] D. W. Hosmer and S. Lemeshow, *Applied Logistic Regression*, John Wiley & Sons, 2nd edition, 2000.

- [13] J. Cohen, P. Cohen, S. G. West, and L. S. Aiken, *Applied Multiple Regression/Correlation Analysis For the Behavioral Sciences*, Routledge, 3rd edition, 2002.
- [14] J. R. Quinlan, "Generating production rules from decision trees," in *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, J. McDermott, Ed., pp. 304–307, 1987.
- [15] Rapid miner Tutorial, October 2009, <http://www.rapid-i.com/downloads/tutorial/rapidminer-4.6-tutorial.pdf>.
- [16] S. Dash, B. Patra, and B. K. Tripathy, "A hybrid data mining technique for improving the classification accuracy of microarray data set," *International Journal of Information Engineering and Electronic Business*, vol. 2, pp. 43–50, 2012.
- [17] S.-W. Lin and S.-C. Chen, "PSOLDA: a particle swarm optimization approach for enhancing classification accuracy rate of linear discriminant analysis," *Applied Soft Computing Journal*, vol. 9, no. 3, pp. 1008–1015, 2009.
- [18] R. Bryll, R. Gutierrez-Osuna, and F. Quek, "Attribute bagging: improving accuracy of classifier ensembles by using random feature subsets," *Pattern Recognition*, vol. 36, no. 6, pp. 1291–1302, 2003.
- [19] D. W. Abbott, "Combining models to improve classifier accuracy and robustness," in *Proceedings of the 2nd International Conference on Information Fusion (Fusion '99)*, vol. I, pp. 289–295, 1999.
- [20] S. Y. Sohn and S. H. Lee, "Data fusion, ensemble and clustering to improve the classification accuracy for the severity of road traffic accidents in Korea," *Safety Science*, vol. 41, no. 1, pp. 1–14, 2003.
- [21] M. R. Smith and T. Martinez, "Improving classification accuracy by identifying and removing instances that should be misclassified," in *Proceedings of the International Joint Conference on Neural Network (IJCNN '11)*, pp. 2690–2697, August 2011.
- [22] M. Zamalloayz, L. J. Rodriguez-Fuentesy, M. Penagarikanoy, G. Bordely, and J. P. Uribez, "Feature dimensionality reduction through genetic algorithms for faster speaker recognition," in *Proceedings of the 16th European Signal Processing Conference (EUSIPCO '08)*, 2008.
- [23] B. Liu, B. McKay, and H. A. Abbass, "Improving genetic classifiers with a boosting algorithm," in *Proceedings of the Congress on Evolutionary Computation (CEC '03)*, vol. 4, pp. 2596–2602, December 2003.
- [24] M. L. Raymer, W. F. Punch, E. D. Goodman, L. A. Kuhn, and A. K. Jain, "Dimensionality reduction using genetic algorithms," *IEEE Transactions on Evolutionary Computation*, vol. 4, no. 2, pp. 164–171, 2000.
- [25] D. N. Kumar, "Optimization Methods: Advanced Topics in Optimization—Direct and Indirect Search Methods," <http://www.gninagpur.in/nptel/courses/Webcourse-contents/IISc-BANG/OPTIMIZATION%20METHODS/pdf/Module.8/M8L4.LN.pdf>.

## Research Article

# Absolute Stability Criteria for Large-Scale Lurie Direct Control Systems with Time-Varying Coefficients

Fucheng Liao<sup>1</sup> and Di Wang<sup>1,2</sup>

<sup>1</sup> School of Mathematics and Physics, University of Science and Technology Beijing, Beijing 100083, China

<sup>2</sup> School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

Correspondence should be addressed to Fucheng Liao; [fc\\_liao@ustb.edu.cn](mailto:fc_liao@ustb.edu.cn)

Received 18 December 2013; Accepted 9 February 2014; Published 27 April 2014

Academic Editors: N. Barsoum, V. N. Dieu, and P. Vasant

Copyright © 2014 F. Liao and D. Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The absolute stability of large-scale Lurie direct control systems with time-varying coefficients is investigated. Based on the decomposition method for large-scale systems and technique of the nonsingular  $M$ -matrix, a suitable scalar Lyapunov function as a weighted sum is constructed. By estimating its total time derivative, some absolute stability criteria and practical corollaries are derived. Furthermore, the results are extended to multiple nonlinearities. The salient feature of this paper is that the criteria which we propose allow for the situation that the norms of time-varying coefficients are unbounded. The main idea of the methodology is that even if the coefficients are norm-unbounded, by restricting their relative magnitude, the problem of negative definiteness for the derivative can also be changed into the problem of stability for a constant matrix. Finally, some numerical examples are included to illustrate the effectiveness of the proposed criteria.

## 1. Introduction

The absolute stability problem has an important position in the analysis and design of control systems. In fact, as a typical class of nonlinear system, the problem of the absolute stability of the Lurie control system has been studied for almost 70 years [1–3] and has proved a fruitful area of research [4–8]. Recently, the problem of Lurie control systems has been extended. Among the studies, [9, 10] studied the absolute stability of large-scale Lurie systems, [11–13] considered the robust stability of uncertain Lurie systems, and [14, 15] discussed Lurie systems with time delays. In addition, there are many monographs on these topics [16, 17].

Nevertheless, most of the papers on Lurie control systems are confined to the norm-bounded coefficients. The absolute stability of Lurie control systems with time-varying and norm-unbounded coefficients has received little attention. In [18], a research method was introduced for the stability of large-scale systems with time-varying coefficients. The core result was that if the order of infinities for the interconnected elements is far less than the order of isolated subsystems, then the large-scale system is still asymptotically stable.

Reference [19] promoted the results and considered the problem of robust exponential control for a class of large-scale systems with uncertainties and unbounded coefficients. On the other hand, the approach is also an effective way to investigate the Lurie control systems with time-varying coefficients. In particular, in the case of the Lurie indirect control systems, it is relatively easier to study them using this method. In [20, 21], this approach was applied to the Lurie indirect control systems with norm-unbounded coefficients, and some absolute stability criteria for this kind of system were obtained. Reference [22] subsequently extended the criteria to systems with multiple nonlinearities and large-scale Lurie indirect systems. However, because of the linear relationship between  $\sigma$  ( $\sigma$  is not an independent variable) and the other state variables in Lurie direct control systems, this became the main roadblock in judging the total time derivative of the Lyapunov function. Reference [23] overcame this difficulty and derived the sufficient condition of absolute stability for the Lurie direct control systems with norm-unbounded coefficients.

The problem proposed in this paper is more general than that described above: it focuses on large-scale Lurie

direct control systems with time-varying coefficients and systems with multiple nonlinearities. The study of stability for large-scale systems is not a new one; it began in the 1960s [24], and the research method is basically the decomposition method for large-scale systems [25, 26]. From the viewpoint of cybernetics, this is known as decentralized control [27, 28]. The basic idea is decomposing the large-scale system into a certain number of lower-order isolated subsystems and constructing the Lyapunov function of the large-scale system through the isolated subsystems. Our objective in the following analysis is to extend this approach and study large-scale Lurie direct control systems with norm-unbounded coefficients. For brevity, we will not involve the concept of the isolated subsystem but will directly give the Lyapunov function by employing the isolated subsystem and estimate the upper bound of its total derivative.

*Notation.* Throughout this paper,  $\|x\|$  denotes the Euclidean norm  $\|x\| = \sqrt{x^T x}$ , where  $x = (x_1 \ x_2 \ \dots \ x_m)^T$  is a column vector and  $T$  denotes the transposition. For vectors  $x = (x_1 \ x_2 \ \dots \ x_m)^T, y = (y_1 \ y_2 \ \dots \ y_m)^T, x \leq y$  ( $x < y$ ) means that  $x_i \leq y_i$  ( $x_i < y_i$ ) ( $i = 1, 2, \dots, m$ ).  $\|A\|$  represents the Euclidean norm of a matrix  $A$  which is induced by the Euclidean vector norm  $\|x\|$ ; that is,  $\|A\| = \max_{\|x\|=1} \|Ax\|$ .  $\lambda(A)$  denotes an arbitrary eigenvalue of matrix  $A$ , and  $A > 0$  ( $A < 0$ ) means that  $A$  is positive definite (negative definite). The symbol  $I$  stands for the time interval  $(\tau, +\infty)$ , where  $\tau \in R$  or  $\tau = -\infty$ .  $\lim_{t \rightarrow +\infty} f(t)$  represents the upper limit of the function  $f(t)$ ; that is,  $\lim_{t \rightarrow +\infty} f(t) = \lim_{u \rightarrow +\infty} \sup_{t \geq u} (f(t))$ . When the limit of function exists, the upper limit must exist and be equal to its limit, so if we change the upper limit involved in this paper into the limit, the conclusions still hold. The reason why we employ an upper limit to describe the theorems and corollaries is that the upper limit makes the stability conditions less conservative.

## 2. Absolute Stability of Large-Scale Lurie Systems with Single Nonlinearity

Consider the following large-scale Lurie direct control system with time-varying coefficients and single nonlinearity:

$$\begin{aligned} \dot{x}_i &= \sum_{j=1}^r A_{ij}(t) x_j + b_i(t) f(\sigma), \quad (i = 1, 2, \dots, r), \\ \sigma &= \sum_{j=1}^r c_j^T(t) x_j, \end{aligned} \tag{1}$$

where  $x_i \in R^{n_i}$  ( $i = 1, 2, \dots, r$ ) are the state,  $b_i(t), c_i(t) \in R^{n_i}$  ( $i = 1, 2, \dots, r$ ) are the vector function,  $b_i(t)$  continuous and  $c_i(t)$  ( $i = 1, 2, \dots, r$ ) have derivative in time interval  $I$ ,  $\sum_{i=1}^r n_i = n$ , and  $A_{ij}(t)$  ( $i, j = 1, 2, \dots, r$ ) are  $n_i \times n_j$  matrix functions and are continuous in  $I$ . The nonlinearity  $f(\sigma)$  is a continuous function and satisfies

$$\begin{aligned} f(\cdot) &\in K[0, +\infty) \\ &= \{f(\cdot) \mid f(0) = 0, 0 < \sigma f(\sigma) < +\infty, \sigma \neq 0\}. \end{aligned} \tag{2}$$

System (1) is said to be absolutely stable if its zero solution is globally asymptotically stable for any nonlinearity  $f(\sigma) \in K[0, +\infty)$  [16, 17].

Basic assumptions about system (1) are proposed in the following.

(A1) We assume that there exist positive definite symmetrical constant matrices  $P_1, P_2, \dots, P_r$  such that

$$\begin{aligned} \lambda [A_{ii}^T(t) P_i + P_i A_{ii}(t)] &\leq -s_i(t) \leq -s_i \leq -s, \\ (i = 1, 2, \dots, r), \quad \forall t > T, \end{aligned} \tag{3}$$

where  $T \in I, P_i \in R^{n_i \times n_i}, s_i(t) > 0, s_i > 0$  ( $i = 1, 2, \dots, r$ ) are known functions and constants, respectively, and  $s = \min(s_1, s_2, \dots, s_r)$ .

*Remark 1.* Compared with [23], condition A1 just guarantees the global asymptotic stability of  $\dot{x}_i = A_{ii}(t)x_i$  ( $i = 1, 2, \dots, r$ ), not all the linear part of the first equation of (1). So it is exceedingly weak.

(A2) We assume that

$$\sum_{i=1}^r c_i^T(t) b_i(t) \leq -g(t), \quad \forall t > T, \tag{4}$$

where  $g(t) > 0$  is a known function.

*Remark 2.* In [23, 29], we know that  $c_i^T b_i < 0$  is the necessary condition for absolute stability in the case of Lurie direct control systems with constant coefficients.

(A3) We assume that

$$\begin{aligned} \frac{2 \|P_i b_i(t)\|}{\sqrt{s_i(t) g(t)}} &\leq \alpha_i, \quad \frac{\|\sum_{j=1}^r c_j^T(t) A_{ji}(t) + c_i^T(t)\|}{\sqrt{s_i(t) g(t)}} \leq \beta_i, \\ (i = 1, 2, \dots, r), \quad \forall t > T, \end{aligned} \tag{5}$$

where  $\alpha_i, \beta_i$  ( $i = 1, 2, \dots, r$ ) are constant.

(A4) We assume that

$$\frac{2 \|P_i A_{ij}(t)\|}{\sqrt{s_i(t) s_j(t)}} \leq \gamma_{ij}, \tag{6}$$

$$(i, j = 1, 2, \dots, r; i \neq j), \quad \forall t > T,$$

where  $\gamma_{ij}$  ( $i, j = 1, 2, \dots, r; i \neq j$ ) are constant.

As we know, the norm-unbounded and time-varying coefficients in system (1) are the main roadblocks in estimating the total time derivative of the Lyapunov function. But, by using the  $s_i(t), g(t)$  ( $i = 1, 2, \dots, r$ ) in conditions A1 and A2 and placing them in the denominator, the relative magnitude of the norm-unbounded coefficients can be restricted in conditions A3 and A4. That is, the “infinite” nature can be expressed by the “finite” form, which makes the study of the absolute stability of large-scale systems (1) feasible. Therefore, we have the following results.

**Theorem 3.** Under A1, A2, A3, and A4, the system described by (1) is absolutely stable if the matrix

$$G = \begin{pmatrix} -1 & \gamma_{12} & \cdots & \gamma_{1r} & \alpha_1 \\ \gamma_{21} & -1 & \cdots & \gamma_{2r} & \alpha_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_{r1} & \gamma_{r2} & \cdots & -1 & \alpha_r \\ \beta_1 & \beta_2 & \cdots & \beta_r & -1 \end{pmatrix} \quad (7)$$

is stable.

*Proof.* Choose a candidate Lyapunov function for system (1) as

$$V = \sum_{i=1}^r d_i x_i^T P_i x_i + d_{r+1} \int_0^\sigma f(\sigma) d\sigma, \quad (8)$$

where  $d_1, \dots, d_r, d_{r+1}$  are positive numbers that will be determined later. From condition A1, we get that  $\sum_{i=1}^r d_i x_i^T P_i x_i$  is a positive definite quadratic form, and, by the properties of  $f(\sigma)$ , we know that  $V$  in (8) is radially unbounded, is positive definite, and has an infinitesimal upper bound.

*Remark 3.* This is very different from the Lurie indirect control system. For the Lurie indirect control system,  $\sigma$  is an independent component of the state vector. Because of this, in order to guarantee that  $V$  in (8) is positive definite in  $R^{n+1}$  and is radially unbounded, we need to also assume  $\int_0^{\pm\infty} f(\sigma) d\sigma = +\infty$  in [20–22].

Based on the decomposition theory of large-scale systems, and employing a similar modus operandi, we let

$$\begin{aligned} V_i &= x_i^T P_i x_i \quad (i = 1, 2, \dots, r); \\ V_{r+1} &= \int_0^\sigma f(\sigma) d\sigma. \end{aligned} \quad (9)$$

Then, (8) can be written as

$$V = \sum_{i=1}^{r+1} d_i V_i = (d_1, \dots, d_r, d_{r+1}) \begin{pmatrix} V_1 \\ \vdots \\ V_r \\ V_{r+1} \end{pmatrix}. \quad (10)$$

Now, we calculate the time derivatives of  $V_1, V_2, \dots, V_{r+1}$  along the trajectories of (1), respectively, and then combine them to get the total time derivative of  $V$ . The time derivative of each  $V_i$  ( $i = 1, 2, \dots, r$ ) along the trajectories of (1) can be processed as

$$\begin{aligned} \dot{V}_i|_{(1)} &= 2x_i^T P_i \dot{x}_i \\ &= 2x_i^T P_i \left( \sum_{j=1}^r A_{ij}(t) x_j + b_i(t) f(\sigma) \right) \\ &= x_i^T \left( A_{ii}^T(t) P_i + P_i A_{ii}(t) \right) x_i + 2 \sum_{\substack{j=1 \\ j \neq i}}^r x_i^T P_i A_{ij}(t) x_j \\ &\quad + 2x_i^T P_i b_i(t) f(\sigma). \end{aligned} \quad (11)$$

By taking the property of norm and using A1, A3, and A4, we obtain

$$\begin{aligned} \dot{V}_i|_{(1)} &\leq -s_i(t) \|x_i\|^2 + 2 \sum_{\substack{j=1 \\ j \neq i}}^r \|P_i A_{ij}(t)\| \|x_i\| \|x_j\| \\ &\quad + 2 \|P_i b_i(t)\| \|x_i\| |f(\sigma)| \\ &= \sqrt{s_i(t)} \|x_i\| \left( -\sqrt{s_i(t)} \|x_i\| \right. \\ &\quad \left. + \sum_{\substack{j=1 \\ j \neq i}}^r \frac{2 \|P_i A_{ij}(t)\|}{\sqrt{s_i(t) s_j(t)}} \sqrt{s_j(t)} \|x_j\| \right. \\ &\quad \left. + \frac{2 \|P_i b_i(t)\|}{\sqrt{s_i(t) g(t)}} \sqrt{g(t)} |f(\sigma)| \right) \quad (12) \\ &\leq \sqrt{s_i(t)} \|x_i\| \\ &\quad \times \left( -\sqrt{s_i(t)} \|x_i\| + \sum_{\substack{j=1 \\ j \neq i}}^r \gamma_{ij} \sqrt{s_j(t)} \|x_j\| \right. \\ &\quad \left. + \alpha_i \sqrt{g(t)} |f(\sigma)| \right), \quad \forall t > T. \end{aligned}$$

On the other hand, the time derivative of  $V_{r+1}$  along the trajectories of (1) is given by

$$\begin{aligned} \dot{V}_{r+1}|_{(1)} &= f(\sigma) \dot{\sigma}|_{(1)} \\ &= f(\sigma) \left( \sum_{j=1}^r \dot{c}_j^T(t) x_j + \sum_{j=1}^r c_j^T(t) \dot{x}_j \right) \\ &= f(\sigma) \\ &\quad \times \left( \sum_{j=1}^r \dot{c}_j^T(t) x_j \right. \\ &\quad \left. + \sum_{j=1}^r c_j^T(t) \left( \sum_{i=1}^r A_{ji}(t) x_i + b_j(t) f(\sigma) \right) \right) \\ &= f(\sigma) \left( \sum_{i=1}^r \left( \sum_{j=1}^r \dot{c}_j^T(t) A_{ji}(t) + \dot{c}_i^T(t) \right) x_i \right. \\ &\quad \left. + \sum_{j=1}^r c_j^T(t) b_j(t) f(\sigma) \right). \end{aligned} \quad (13)$$

From A2 and A3, we have

$$\begin{aligned} \dot{V}_{r+1}|_{(1)} &\leq \sqrt{g(t)} |f(\sigma)| \sum_{i=1}^r \frac{\|\sum_{j=1}^r c_j^T(t) A_{ji}(t) + \dot{c}_i^T(t)\|}{\sqrt{s_i(t)g(t)}} \\ &\quad \times \sqrt{s_i(t)} \|x_i\| - g(t) f^2(\sigma) \\ &\leq \sqrt{g(t)} |f(\sigma)| \\ &\quad \times \left( \sum_{i=1}^r \beta_i \sqrt{s_i(t)} \|x_i\| - \sqrt{g(t)} |f(\sigma)| \right), \end{aligned} \tag{14}$$

$\forall t > T.$

Combining (12) with (14), we derive

$$\begin{aligned} &\begin{pmatrix} \dot{V}_1 \\ \vdots \\ \dot{V}_r \\ \dot{V}_{r+1} \end{pmatrix}_{(1)} \\ &\leq \begin{pmatrix} \sqrt{s_1(t)} \|x_1\| & & & \\ & \ddots & & \\ & & \sqrt{s_r(t)} \|x_r\| & \\ & & & \sqrt{g(t)} |f(\sigma)| \end{pmatrix} \\ &\quad \times G \begin{pmatrix} \sqrt{s_1(t)} \|x_1\| \\ \vdots \\ \sqrt{s_r(t)} \|x_r\| \\ \sqrt{g(t)} |f(\sigma)| \end{pmatrix}, \end{aligned} \tag{15}$$

$\forall t > T.$

Here, we define the following diagonal matrix:

$$D = \text{diag}(d_1, \dots, d_r, d_{r+1}). \tag{16}$$

Then, employing inequality (15), we get

$$\begin{aligned} \dot{V}|_{(1)} &= (d_1, \dots, d_r, d_{r+1}) \begin{pmatrix} \dot{V}_1 \\ \vdots \\ \dot{V}_r \\ \dot{V}_{r+1} \end{pmatrix}_{(1)} \\ &\leq (d_1, \dots, d_r, d_{r+1}) \\ &\quad \times \begin{pmatrix} \sqrt{s_1(t)} \|x_1\| & & & \\ & \ddots & & \\ & & \sqrt{s_r(t)} \|x_r\| & \\ & & & \sqrt{g(t)} |f(\sigma)| \end{pmatrix} \\ &\quad \times G \begin{pmatrix} \sqrt{s_1(t)} \|x_1\| \\ \vdots \\ \sqrt{s_r(t)} \|x_r\| \\ \sqrt{g(t)} |f(\sigma)| \end{pmatrix} \end{aligned}$$

$$\begin{aligned} &= (\sqrt{s_1(t)} \|x_1\| \cdots \sqrt{s_r(t)} \|x_r\| \sqrt{g(t)} |f(\sigma)|) \\ &\quad \times \frac{G^T D + DG}{2} \begin{pmatrix} \sqrt{s_1(t)} \|x_1\| \\ \vdots \\ \sqrt{s_r(t)} \|x_r\| \\ \sqrt{g(t)} |f(\sigma)| \end{pmatrix}, \end{aligned} \tag{17}$$

$\forall t > T.$

Since  $G$  is stable, this implies that the real part of each eigenvalue of  $-G$  is positive. From the equivalent propositions of the nonsingular  $M$ -matrix in [23, 30], we know that there exists a positive diagonal matrix  $D = \text{diag}(d_1, \dots, d_r, d_{r+1})$  ( $d_i > 0, i = 1, \dots, r, r + 1$ ) such that

$$\frac{(-G)^T D + D(-G)}{2} \tag{18}$$

is positive definite; that is,  $((G^T D + DG)/2)$  is negative definite. Here, we choose the above  $d_1, \dots, d_r, d_{r+1}$  in (8) and let  $-\omega$  be the biggest eigenvalue of  $((G^T D + DG)/2)$  (clearly  $-\omega < 0$ ). So, according to (17), we have

$$\begin{aligned} \dot{V}|_{(1)} &\leq -\omega \left( \sum_{i=1}^r s_i(t) \|x_i\|^2 + g(t) f^2(\sigma) \right) \leq -s\omega \sum_{i=1}^r \|x_i\|^2, \\ &\quad \forall t > T. \end{aligned} \tag{19}$$

This implies that, as to all  $f(\sigma) \in K[0, +\infty)$ ,  $\dot{V}|_{(1)}$  is negative definite. Thus, according to the Lyapunov theorems, system (1) is absolutely stable. The proof is completed.  $\square$

It should be noted that A3 and A4 can be weakened by establishing upper limits. Therefore, the following corollaries are obtained.

(A3') Assume that

$$\begin{aligned} \overline{\lim}_{t \rightarrow +\infty} \frac{2 \|P_i b_i(t)\|}{\sqrt{s_i(t)g(t)}} &= \bar{\alpha}_i, \quad (i = 1, 2, \dots, r), \\ \overline{\lim}_{t \rightarrow +\infty} \frac{\|\sum_{j=1}^r c_j^T(t) A_{ji}(t) + \dot{c}_i^T(t)\|}{\sqrt{s_i(t)g(t)}} &= \bar{\beta}_i, \quad (i = 1, 2, \dots, r), \end{aligned} \tag{20}$$

where  $\bar{\alpha}_i, \bar{\beta}_i$  ( $i = 1, 2, \dots, r$ ) are constant.

(A4') Assume that

$$\overline{\lim}_{t \rightarrow +\infty} \frac{2 \|P_i A_{ij}(t)\|}{\sqrt{s_i(t)s_j(t)}} = \bar{\gamma}_{ij}, \quad (i, j = 1, 2, \dots, r; i \neq j), \tag{21}$$

where  $\bar{\gamma}_{ij}$  ( $i, j = 1, 2, \dots, r; i \neq j$ ) are constant.

**Corollary 4.** Under A1, A2, A3', and A4', the system described by (1) is absolutely stable if the matrix

$$\bar{G} = \begin{pmatrix} -1 & \bar{\gamma}_{12} & \cdots & \bar{\gamma}_{1r} & \bar{\alpha}_1 \\ \bar{\gamma}_{21} & -1 & \cdots & \bar{\gamma}_{2r} & \bar{\alpha}_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \bar{\gamma}_{r1} & \bar{\gamma}_{r2} & \cdots & -1 & \bar{\alpha}_r \\ \bar{\beta}_1 & \bar{\beta}_2 & \cdots & \bar{\beta}_r & -1 \end{pmatrix} \tag{22}$$

is stable.

*Proof.* Let

$$G = \begin{pmatrix} -1 & \bar{\gamma}_{12} + \varepsilon & \cdots & \bar{\gamma}_{1r} + \varepsilon & \bar{\alpha}_1 + \varepsilon \\ \bar{\gamma}_{21} + \varepsilon & -1 & \cdots & \bar{\gamma}_{2r} + \varepsilon & \bar{\alpha}_2 + \varepsilon \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \bar{\gamma}_{r1} + \varepsilon & \bar{\gamma}_{r2} + \varepsilon & \cdots & -1 & \bar{\alpha}_r + \varepsilon \\ \bar{\beta}_1 + \varepsilon & \bar{\beta}_2 + \varepsilon & \cdots & \bar{\beta}_r + \varepsilon & -1 \end{pmatrix}. \quad (23)$$

Namely,  $G$  is a matrix induced by  $\bar{G}$  whose every off-diagonal entry is increased by  $\varepsilon$ .

According to the properties of the matrix, if  $\bar{G}$  is stable, then there exists a sufficiently small scalar  $\varepsilon > 0$  such that  $G$  is also stable. We choose an  $\varepsilon > 0$  that can allow for the stability of  $G$ .

From the assumptions here and the definition of upper limit, as to the above  $\varepsilon$ , there exists a number  $T(\geq \tau)$ , which satisfies that if  $t > T$ , then

$$\frac{2 \|P_i b_i(t)\|}{\sqrt{s_i(t) g(t)}} \leq \sup_{t>T} \left( \frac{2 \|P_i b_i(t)\|}{\sqrt{s_i(t) g(t)}} \right) \leq \bar{\alpha}_i + \varepsilon, \quad (i = 1, 2, \dots, r)$$

$$\begin{aligned} & \frac{\|\sum_{j=1}^r c_j^T(t) A_{ji}(t) + \dot{c}_i^T(t)\|}{\sqrt{s_i(t) g(t)}} \\ & \leq \sup_{t>T} \left( \frac{\|\sum_{j=1}^r c_j^T(t) A_{ji}(t) + \dot{c}_i^T(t)\|}{\sqrt{s_i(t) g(t)}} \right) \leq \bar{\beta}_i + \varepsilon, \quad (24) \\ & \quad (i = 1, 2, \dots, r), \end{aligned}$$

$$\frac{2 \|P_i A_{ij}(t)\|}{\sqrt{s_i(t) s_j(t)}} \leq \sup_{t>T} \left( \frac{2 \|P_i A_{ij}(t)\|}{\sqrt{s_i(t) s_j(t)}} \right) \leq \bar{\gamma}_{ij} + \varepsilon, \quad (i, j = 1, 2, \dots, r; i \neq j).$$

Thus, according to Theorem 3, system (1) is absolutely stable.

If  $\bar{\gamma}_{ij} = 0$  ( $i, j = 1, 2, \dots, r; i \neq j$ ), then the following corollary is proposed.

(A4'') We assume that

$$\lim_{t \rightarrow +\infty} \frac{2 \|P_i A_{ij}(t)\|}{\sqrt{s_i(t) s_j(t)}} = 0, \quad (i, j = 1, 2, \dots, r; i \neq j). \quad (25)$$

Namely,  $\bar{\gamma}_{ij} = 0$  ( $i, j = 1, 2, \dots, r; i \neq j$ ) in condition A4'.  $\square$

**Corollary 5.** Under A1, A2, A3', and A4'', the system described by (1) is absolutely stable if the inequality  $\sum_{i=1}^r \bar{\alpha}_i \bar{\beta}_i < 1$  holds.

*Proof.* According to A4'', the eigenpolynomial of the matrix  $\bar{G}$  can be obtained as

$$\begin{aligned} |\lambda I - \bar{G}| &= \begin{vmatrix} \lambda + 1 & 0 & \cdots & 0 & -\bar{\alpha}_1 \\ 0 & \lambda + 1 & \cdots & 0 & -\bar{\alpha}_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \lambda + 1 & -\bar{\alpha}_r \\ -\bar{\beta}_1 & -\bar{\beta}_2 & \cdots & -\bar{\beta}_r & \lambda + 1 \end{vmatrix} \\ &= \left( \lambda^2 + 2\lambda + \left( 1 - \sum_{i=1}^r \bar{\alpha}_i \bar{\beta}_i \right) \right) (\lambda + 1)^{r-1}, \end{aligned} \quad (26)$$

where  $\lambda = -1$  is an eigenvalue with multiplicity  $(r - 1)$  and the other two eigenvalues satisfy

$$\lambda^2 + 2\lambda + \left( 1 - \sum_{i=1}^r \bar{\alpha}_i \bar{\beta}_i \right) = 0. \quad (27)$$

It is easy to see that all roots of the above equation have a real part if and only if  $\sum_{i=1}^r \bar{\alpha}_i \bar{\beta}_i < 1$ . So  $\bar{G}$  is stable if  $\sum_{i=1}^r \bar{\alpha}_i \bar{\beta}_i < 1$ . This implies that system (1) is absolutely stable by Corollary 4.  $\square$

Consider a more specific case; we have the following corollary.

**Corollary 6.** Under A1, A2, A3', and A4', the system described by (1) is absolutely stable if

$$\bar{\alpha}_i = \bar{\gamma}_{ij} = 0, \quad (i, j = 1, 2, \dots, r; i \neq j) \quad (28)$$

or

$$\bar{\beta}_i = \bar{\gamma}_{ij} = 0, \quad (i, j = 1, 2, \dots, r; i \neq j) \quad (29)$$

holds.

*Proof.* From  $\sum_{i=1}^r \bar{\alpha}_i \bar{\beta}_i = 0 < 1$ , we know that system (1) is absolutely stable by Corollary 5.  $\square$

**Corollary 7.** Under A1, A2, A3', and A4', the system described by (1) is absolutely stable if the following inequalities hold:

$$\begin{aligned} & \sum_{i=1}^r (\bar{\alpha}_i + \bar{\beta}_i) < 2, \\ & \bar{\alpha}_j + \bar{\beta}_j + \sum_{\substack{i=1 \\ i \neq j}}^r (\bar{\gamma}_{ji} + \bar{\gamma}_{ij}) < 2, \quad (j = 1, 2, \dots, r). \end{aligned} \quad (30)$$

*Proof.* Choosing the undetermined coefficients of the Lyapunov function in Theorem 3 as  $d_1 = \cdots = d_r = d_{r+1} = 1$  and combining them with Corollary 4, we just need to prove that

$$\bar{G} + \bar{G}^T = \begin{pmatrix} -2 & \bar{\gamma}_{12} + \bar{\gamma}_{21} & \cdots & \bar{\gamma}_{1r} + \bar{\gamma}_{r1} & \bar{\alpha}_1 + \bar{\beta}_1 \\ \bar{\gamma}_{12} + \bar{\gamma}_{21} & -2 & \cdots & \bar{\gamma}_{2r} + \bar{\gamma}_{r2} & \bar{\alpha}_2 + \bar{\beta}_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \bar{\gamma}_{1r} + \bar{\gamma}_{r1} & \bar{\gamma}_{2r} + \bar{\gamma}_{r2} & \cdots & -2 & \bar{\alpha}_r + \bar{\beta}_r \\ \bar{\alpha}_1 + \bar{\beta}_1 & \bar{\alpha}_2 + \bar{\beta}_2 & \cdots & \bar{\alpha}_r + \bar{\beta}_r & -2 \end{pmatrix} \quad (31)$$

is negative definite. Note that each eigenvalue of  $\bar{G} + \bar{G}^T$  is a real number (since  $\bar{G} + \bar{G}^T$  is a real symmetric matrix) and

every diagonal entry is equal to 2, so, from the Gershgorin circle theorem in [31], we get

$$|\lambda + 2| \leq \sum_{i=1}^r (\bar{\alpha}_i + \bar{\beta}_i),$$

$$|\lambda + 2| \leq \bar{\alpha}_j + \bar{\beta}_j + \sum_{\substack{i=1 \\ i \neq j}}^r (\bar{\gamma}_{ji} + \bar{\gamma}_{ij}), \quad (j = 1, 2, \dots, r).$$
(32)

Namely,

$$\lambda \leq -2 + \sum_{i=1}^r (\bar{\alpha}_i + \bar{\beta}_i),$$

$$\lambda \leq -2 + \bar{\alpha}_j + \bar{\beta}_j + \sum_{\substack{i=1 \\ i \neq j}}^r (\bar{\gamma}_{ji} + \bar{\gamma}_{ij}), \quad (j = 1, 2, \dots, r).$$
(33)

If inequalities (30) hold, then each eigenvalue  $\lambda$  of  $\bar{G} + \bar{G}^T$  satisfies  $\lambda < 0$ , and this implies that  $\bar{G} + \bar{G}^T$  is negative definite. The proof is completed.  $\square$

Moreover, the conditions of Corollary 7 can be weakened as follows.

**Corollary 8.** Under  $A1, A2, A3'$ , and  $A4'$ , the system described by (1) is absolutely stable if the following inequalities

$$\sum_{i=1}^r \bar{\alpha}_i < 1,$$

$$\bar{\beta}_j + \sum_{\substack{i=1 \\ i \neq j}}^r \bar{\gamma}_{ij} < 1, \quad (j = 1, 2, \dots, r)$$
(34)

or

$$\sum_{i=1}^r \bar{\beta}_i < 1,$$

$$\bar{\alpha}_j + \sum_{\substack{i=1 \\ i \neq j}}^r \bar{\gamma}_{ji} < 1, \quad (j = 1, 2, \dots, r)$$
(35)

hold.

*Proof.* From the Gershgorin circle theorem, we know that each eigenvalue  $\lambda$  of  $\bar{G}$  satisfies

$$|\lambda + 1| \leq \sum_{i=1}^r \bar{\alpha}_i,$$

$$|\lambda + 1| \leq \bar{\beta}_j + \sum_{\substack{i=1 \\ i \neq j}}^r \bar{\gamma}_{ij}, \quad (j = 1, 2, \dots, r).$$
(36)

Then, we have

$$|\operatorname{Re}(\lambda) + 1| \leq |\lambda + 1| \leq \sum_{i=1}^r \bar{\alpha}_i,$$

$$|\operatorname{Re}(\lambda) + 1| \leq |\lambda + 1| \leq \bar{\beta}_j + \sum_{\substack{i=1 \\ i \neq j}}^r \bar{\gamma}_{ij}, \quad (j = 1, 2, \dots, r).$$
(37)

Namely,

$$\operatorname{Re}(\lambda) \leq -1 + \sum_{i=1}^r \bar{\alpha}_i,$$

$$\operatorname{Re}(\lambda) \leq -1 + \bar{\beta}_j + \sum_{\substack{i=1 \\ i \neq j}}^r \bar{\gamma}_{ij}, \quad (j = 1, 2, \dots, r).$$
(38)

If inequalities (34) hold, then  $\operatorname{Re}(\lambda) < 0$ , and this implies that  $\bar{G}$  is stable. Similarly, if inequalities (35) hold, we can prove that  $\bar{G}^T$  is stable; that is,  $\bar{G}$  is stable. With the conditions of Corollary 4 being satisfied, we conclude that system (1) is absolutely stable.  $\square$

*Remark 9.* If inequalities (34) and (35) hold together, inequalities (30) hold. Therefore, Corollary 8 is less conservative.

### 3. Absolute Stability of Large-Scale Lurie Systems with Multiple Nonlinearities

Consider the following large-scale Lurie direct control systems with time-varying coefficients and multiple nonlinearities:

$$\dot{x}_i = \sum_{j=1}^r A_{ij}(t) x_j + \sum_{k=1}^m b_{ik}(t) f_k(\sigma_k),$$

$$\sigma_k = \sum_{j=1}^r c_{kj}^T(t) x_j,$$

$$(i, j = 1, 2, \dots, r; k = 1, 2, \dots, m),$$
(39)

where  $x_i \in R^{n_i}$  ( $i = 1, 2, \dots, r$ ) are the state,  $b_{ik}(t) \in R^{n_i}$  ( $i = 1, 2, \dots, r; k = 1, 2, \dots, m$ ),  $c_{kj}(t) \in R^{n_i}$  ( $k = 1, 2, \dots, m; j = 1, 2, \dots, r$ ) are vector functions,  $b_{ik}(t)$  continuous and  $c_{kj}(t)$  have derivatives in time interval  $I$ ,  $\sum_{i=1}^r n_i = n$ , and  $A_{ij}(t)$  ( $i, j = 1, 2, \dots, r$ ) are  $n_i \times n_j$  matrix functions and are continuous in  $I$ . The nonlinearities  $f_k(\cdot)$  ( $k = 1, 2, \dots, m$ ) are continuous functions, and they satisfy

$$f_k(\cdot) \in K[0, +\infty)$$

$$= \{f_k(\cdot) \mid f_k(0) = 0, 0 < \sigma_k f_k(\sigma_k) < +\infty, \sigma_k \neq 0\}.$$
(40)

System (39) is said to be absolutely stable if its zero solution is globally asymptotically stable for any nonlinearity  $f_k(\sigma_k) \in K[0, +\infty)$  [16, 17].

Basic assumptions about system (39) are proposed.

(A5) We assume that there exist positive definite symmetrical constant matrices  $P_1, P_2, \dots, P_r$  such that

$$\lambda [A_{ii}^T(t) P_i + P_i A_{ii}(t)] \leq -s_i(t) \leq -s_i \leq -s,$$

$$(i = 1, 2, \dots, r), \quad \forall t > T,$$
(41)

where  $T \in I$ ,  $s_i(t) > 0$ ,  $s_i > 0$  ( $i = 1, 2, \dots, r$ ) are known functions and constants, respectively, and  $s = \min(s_1, s_2, \dots, s_r)$ .

(A6) We assume that

$$\sum_{j=1}^r c_{kj}^T(t) b_{jk}(t) \leq -g_k(t), \tag{42}$$

$$(k = 1, 2, \dots, m), \quad \forall t > T,$$

where  $g_k(t) > 0$  ( $k = 1, 2, \dots, m$ ) are known functions.

(A7) We assume that

$$\frac{2 \|P_i b_{ik}(t)\|}{\sqrt{s_i(t) g_k(t)}} \leq \alpha_{ik}, \frac{\|\sum_{j=1}^r c_{kj}^T(t) A_{ji}(t) + \dot{c}_{ki}^T(t)\|}{\sqrt{s_i(t) g_k(t)}} \leq \beta_{ki}, \tag{43}$$

$$(i = 1, 2, \dots, r; k = 1, 2, \dots, m), \quad \forall t > T,$$

where  $\alpha_{ik}, \beta_{ki}$  ( $i = 1, 2, \dots, r; k = 1, 2, \dots, m$ ) are constants.

(A8) We assume that

$$\frac{2 \|P_i A_{ij}(t)\|}{\sqrt{s_i(t) s_j(t)}} \leq \gamma_{ij}, \tag{44}$$

$$(i, j = 1, 2, \dots, r; i \neq j), \quad \forall t > T,$$

where  $\gamma_{ij}$  ( $i, j = 1, 2, \dots, r; i \neq j$ ) are constants.

(A9) We assume that

$$\frac{|\sum_{j=1}^r c_{kj}^T(t) b_{jl}(t)|}{\sqrt{g_k(t) g_l(t)}} \leq \mu_{kl}, \quad (k, l = 1, 2, \dots, m; k \neq l), \quad \forall t > T, \tag{45}$$

where  $\mu_{kl}$  ( $k, l = 1, 2, \dots, m; k \neq l$ ) are constants.

In addition, we define the following matrices:

$$G = \begin{pmatrix} -1 & \gamma_{12} & \dots & \gamma_{1r} \\ \gamma_{21} & -1 & \dots & \gamma_{2r} \\ \dots & \dots & \dots & \dots \\ \gamma_{r1} & \gamma_{r2} & \dots & -1 \end{pmatrix},$$

$$F = \begin{pmatrix} -1 & \mu_{12} & \dots & \mu_{1m} \\ \mu_{21} & -1 & \dots & \mu_{2m} \\ \dots & \dots & \dots & \dots \\ \mu_{m1} & \mu_{m2} & \dots & -1 \end{pmatrix}, \tag{46}$$

$$R = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2m} \\ \dots & \dots & \dots & \dots \\ \alpha_{r1} & \alpha_{r2} & \dots & \alpha_{rm} \end{pmatrix},$$

$$L = \begin{pmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1r} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2r} \\ \dots & \dots & \dots & \dots \\ \beta_{m1} & \beta_{m2} & \dots & \beta_{mr} \end{pmatrix}.$$

**Theorem 10.** Under A5, A6, A7, A8, and A9, the system described by (39) is absolutely stable if the matrix

$$Q = \begin{pmatrix} G & R \\ L & F \end{pmatrix} \tag{47}$$

is stable.

*Proof.* Choose a candidate Lyapunov function for system (39) as

$$V = \sum_{i=1}^r d_i x_i^T P_i x_i + d_{r+1} \int_0^\sigma f(\sigma) d\sigma, \tag{48}$$

where  $d_1, \dots, d_r, d_{r+1}, \dots, d_{r+m}$  are positive numbers that will be chosen later. From the properties of  $f_k(\sigma_k)$  and condition A5, we know that  $V$  in (48) is radially unbounded, is positive definite, and has infinitesimal upper bound.

Let

$$V_i = x_i^T P_i x_i \quad (i = 1, 2, \dots, r),$$

$$V_{r+k} = \int_0^{\sigma_k} f_k(\sigma) d\sigma \quad (k = 1, 2, \dots, m). \tag{49}$$

Then, (48) can be written as

$$V = \sum_{i=1}^{r+m} d_i V_i = (d_1 \ \dots \ d_r \ d_{r+1} \ \dots \ d_{r+m}) \begin{pmatrix} V_1 \\ \vdots \\ V_r \\ V_{r+1} \\ \vdots \\ V_{r+m} \end{pmatrix}. \tag{50}$$

First, we calculate the time derivative of each  $V_i$  ( $i = 1, 2, \dots, r$ ) along the trajectories of system (39). Consider

$$\begin{aligned} \dot{V}_i|_{(39)} &= 2x_i^T P_i \dot{x}_i \\ &= 2x_i^T P_i \left( \sum_{j=1}^r A_{ij}(t) x_j + \sum_{k=1}^m b_{ik}(t) f_k(\sigma_k) \right) \\ &= x_i^T (A_{ii}^T(t) P_i + P_i A_{ii}(t)) x_i + 2 \sum_{\substack{j=1 \\ j \neq i}}^r x_i^T P_i A_{ij}(t) x_j \\ &\quad + 2 \sum_{k=1}^m x_i^T P_i b_{ik}(t) f_k(\sigma_k). \end{aligned} \tag{51}$$

From A5, A7, and A8, we have

$$\begin{aligned} \dot{V}_i|_{(39)} &\leq -s_i(t) \|x_i\|^2 + 2 \sum_{\substack{j=1 \\ j \neq i}}^r \|P_i A_{ij}(t)\| \|x_i\| \|x_j\| \\ &\quad + 2 \sum_{k=1}^m \|P_i b_{ik}(t)\| \|x_i\| |f_k(\sigma_k)| \\ &= \sqrt{s_i(t)} \|x_i\| \\ &\quad \times \left( -\sqrt{s_i(t)} \|x_i\| + \sum_{\substack{j=1 \\ j \neq i}}^r \frac{2 \|P_i A_{ij}(t)\|}{\sqrt{s_i(t) s_j(t)}} \sqrt{s_j(t)} \|x_j\| \right. \\ &\quad \left. + \sum_{k=1}^m \frac{2 \|P_i b_{ik}(t)\|}{\sqrt{s_i(t) g_k(t)}} \sqrt{g_k(t)} |f_k(\sigma_k)| \right) \end{aligned}$$



and let  $-\omega$  be the biggest eigenvalue of  $(1/2)(Q^T D + DQ)$  (clearly  $-\omega < 0$ ). Hence, according to (57), we have

$$\begin{aligned} \dot{V}|_{(39)} &\leq -\omega \|U\|^2 = -\omega \left( \sum_{i=1}^r s_i(t) \|x_i\|^2 + \sum_{k=1}^m g_k(t) f_k^2(\sigma_k) \right) \\ &\leq -s\omega \sum_{i=1}^r \|x_i\|^2, \quad \forall t > T. \end{aligned} \tag{58}$$

This implies that, as to all  $f_k(\sigma_k) \in K[0, +\infty)$ ,  $\dot{V}|_{(39)}$  is negative definite. So, system (39) is absolutely stable by the Lyapunov theorems. The proof is completed.  $\square$

At the same time, we can get the following corollaries. The proof for corollaries is similar to that in large-scale Lurie systems with single nonlinearity and thus is omitted.

(A7') We assume that

$$\begin{aligned} \overline{\lim}_{x \rightarrow +\infty} \frac{2 \|P_i b_{ik}(t)\|}{\sqrt{s_i(t) g_k(t)}} &= \bar{\alpha}_{ik}, \\ (i = 1, 2, \dots, r; k = 1, 2, \dots, m), \\ \overline{\lim}_{x \rightarrow +\infty} \frac{\| \sum_{j=1}^r c_{kj}^T(t) A_{ji}(t) + \dot{c}_{ki}^T(t) \|}{\sqrt{s_i(t) g_k(t)}} &= \bar{\beta}_{ki}, \\ (i = 1, 2, \dots, r; k = 1, 2, \dots, m), \end{aligned} \tag{59}$$

where  $\bar{\alpha}_{ik}, \bar{\beta}_{ki}$  ( $i = 1, 2, \dots, r; k = 1, 2, \dots, m$ ) are constant.

(A8') We assume that

$$\overline{\lim}_{x \rightarrow +\infty} \frac{2 \|P_i A_{ij}(t)\|}{\sqrt{s_i(t) s_j(t)}} = \bar{\gamma}_{ij}, \quad (i, j = 1, 2, \dots, r; i \neq j), \tag{60}$$

where  $\bar{\gamma}_{ij}$  ( $i, j = 1, 2, \dots, r; i \neq j$ ) are constant.

(A9') We assume that

$$\lim_{t \rightarrow +\infty} \frac{|\sum_{j=1}^r c_{kj}^T(t) b_{jl}(t)|}{\sqrt{g_k(t) g_l(t)}} = \bar{\mu}_{kl}, \quad (k, l = 1, 2, \dots, m; k \neq l), \tag{61}$$

where  $\bar{\mu}_{kl}$  ( $k, l = 1, 2, \dots, m; k \neq l$ ) are constant.

**Corollary 11.** Under A5, A6, A7', A8', and A9', the system described by (39) is absolutely stable if the matrix

$$\bar{Q} = \begin{pmatrix} \bar{G} & \bar{R} \\ \bar{L} & \bar{F} \end{pmatrix} \tag{62}$$

is stable, where

$$\begin{aligned} \bar{G} &= \begin{pmatrix} -1 & \bar{\gamma}_{12} & \dots & \bar{\gamma}_{1r} \\ \bar{\gamma}_{21} & -1 & \dots & \bar{\gamma}_{2r} \\ \dots & \dots & \dots & \dots \\ \bar{\gamma}_{r1} & \bar{\gamma}_{r2} & \dots & -1 \end{pmatrix}, \\ \bar{F} &= \begin{pmatrix} -1 & \bar{\mu}_{12} & \dots & \bar{\mu}_{1m} \\ \bar{\mu}_{21} & -1 & \dots & \bar{\mu}_{2m} \\ \dots & \dots & \dots & \dots \\ \bar{\mu}_{m1} & \bar{\mu}_{m2} & \dots & -1 \end{pmatrix}, \\ \bar{R} &= \begin{pmatrix} \bar{\alpha}_{11} & \bar{\alpha}_{12} & \dots & \bar{\alpha}_{1m} \\ \bar{\alpha}_{21} & \bar{\alpha}_{22} & \dots & \bar{\alpha}_{2m} \\ \dots & \dots & \dots & \dots \\ \bar{\alpha}_{r1} & \bar{\alpha}_{r2} & \dots & \bar{\alpha}_{rm} \end{pmatrix}, \\ \bar{L} &= \begin{pmatrix} \bar{\beta}_{11} & \bar{\beta}_{12} & \dots & \bar{\beta}_{1r} \\ \bar{\beta}_{21} & \bar{\beta}_{22} & \dots & \bar{\beta}_{2r} \\ \dots & \dots & \dots & \dots \\ \bar{\beta}_{m1} & \bar{\beta}_{m2} & \dots & \bar{\beta}_{mr} \end{pmatrix}. \end{aligned} \tag{63}$$

**Corollary 12.** Under A5, A6, A8', and A9' and with matrices  $\bar{G}, \bar{F}$  being stable, the system described by (1) is absolutely stable if

$$\begin{aligned} \lim_{t \rightarrow +\infty} \frac{2 \|P_i b_{ik}(t)\|}{\sqrt{s_i(t) g_k(t)}} &= 0, \\ \frac{\| \sum_{j=1}^r c_{kj}^T(t) A_{ji}(t) + \dot{c}_{ki}^T(t) \|}{\sqrt{s_i(t) g_k(t)}} &\leq M_{ki}, \\ (i = 1, 2, \dots, r; k = 1, 2, \dots, m) \end{aligned} \tag{64}$$

or

$$\begin{aligned} \frac{2 \|P_i b_{ik}(t)\|}{\sqrt{s_i(t) g_k(t)}} &\leq K_{ik}, \\ \lim_{t \rightarrow +\infty} \frac{\| \sum_{j=1}^r c_{kj}^T(t) A_{ji}(t) + \dot{c}_{ki}^T(t) \|}{\sqrt{s_i(t) g_k(t)}} &= 0, \\ (i = 1, 2, \dots, r; k = 1, 2, \dots, m) \end{aligned} \tag{65}$$

holds, where  $M_{ki}, K_{ik}$  ( $i = 1, 2, \dots, r; k = 1, 2, \dots, m$ ) are constant.

**Corollary 13.** Under A5, A6, A7', A8', and A9', the system described by (39) is absolutely stable if the following inequalities

$$\begin{aligned} \sum_{\substack{j=1 \\ i \neq j}}^r \bar{\gamma}_{ij} + \sum_{j=1}^m \bar{\alpha}_{ij} &< 1, \quad (i = 1, 2, \dots, r), \\ \sum_{\substack{j=1 \\ k \neq j}}^m \bar{\mu}_{kj} + \sum_{j=1}^r \bar{\beta}_{kj} &< 1, \quad (k = 1, 2, \dots, m), \end{aligned} \tag{66}$$

or

$$\sum_{\substack{i=1 \\ i \neq j}}^r \bar{\gamma}_{ij} + \sum_{i=1}^m \bar{\beta}_{ij} < 1, \quad (j = 1, 2, \dots, r),$$

$$\sum_{\substack{i=1 \\ i \neq l}}^m \bar{\mu}_{il} + \sum_{i=1}^r \bar{\alpha}_{il} < 1, \quad (l = 1, 2, \dots, m),$$
(67)

hold.

### 4. Numerical Examples

In this section, two simple numerical examples are introduced to demonstrate the effectiveness of our criteria.

*Example 1.* Consider system (1) with

$$A_{11}(t) = \begin{pmatrix} -2t & 1 \\ t & -3t \end{pmatrix}, \quad A_{12}(t) = \begin{pmatrix} 1 \\ t^2 \end{pmatrix},$$

$$A_{21}(t) = (t^3 \quad -2t), \quad A_{22}(t) = -5t^5,$$

$$b_1 = \begin{pmatrix} -e^{t^2} \\ 2t \end{pmatrix}, \quad b_2 = -6t^3;$$

$$c_1 = (-e^{t^2} \quad 2t), \quad c_2 = -6t^3.$$
(68)

Note that the norms of coefficient matrices for the above system are unbounded; we choose

$$T = 1, \quad P_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad P_2 = 1. \quad (69)$$

Therefore,

$$s_1(t) = 2t, \quad s_1 = 2; \quad s_2(t) = 10t^5, \quad s_2 = 10;$$

$$g(t) = e^{2t^2} - 2t^3. \quad (70)$$

This means that assumptions A1 and A2 are satisfied. From

$$\overline{\lim}_{t \rightarrow +\infty} \frac{2 \|P_1 b_1(t)\|}{\sqrt{s_1(t)} g(t)} = \bar{\alpha}_1 = 0,$$

$$\overline{\lim}_{t \rightarrow +\infty} \frac{2 \|P_2 b_2(t)\|}{\sqrt{s_2(t)} g(t)} = \bar{\alpha}_2 = 0,$$

$$\overline{\lim}_{t \rightarrow +\infty} \frac{\|\sum_{j=1}^2 c_j^T(t) A_{j1}(t) + \dot{c}_1^T(t)\|}{\sqrt{s_1(t)} g(t)} = \bar{\beta}_1 = 0,$$

$$\overline{\lim}_{t \rightarrow +\infty} \frac{\|\sum_{j=1}^2 c_j^T(t) A_{j2}(t) + \dot{c}_2^T(t)\|}{\sqrt{s_2(t)} g(t)} = \bar{\beta}_2 = 0, \quad (71)$$

we know that  $\bar{\alpha}_i = 0, \bar{\beta}_i = 0 (i = 1, 2)$  in assumption A3'. Since

$$\overline{\lim}_{t \rightarrow +\infty} \frac{2 \|P_1 A_{12}(t)\|}{\sqrt{s_1(t)} s_2(t)} = \bar{\gamma}_{12} = 0,$$

$$\overline{\lim}_{t \rightarrow +\infty} \frac{2 \|P_2 A_{21}(t)\|}{\sqrt{s_2(t)} s_1(t)} = \bar{\gamma}_{21} = \frac{1}{\sqrt{5}}, \quad (72)$$

we know that assumption A4' is satisfied and

$$\bar{G} = \begin{pmatrix} -1 & \bar{\gamma}_{12} & \bar{\alpha}_1 \\ \bar{\gamma}_{21} & -1 & \bar{\alpha}_2 \\ \bar{\beta}_1 & \bar{\beta}_2 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ \frac{1}{\sqrt{5}} & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (73)$$

It is easy to verify that  $\bar{G}$  is stable. In summary, according to Corollary 4 (or Corollary 8), the system is absolutely stable.

Let  $f(\sigma) = \sigma^3$  and  $x_1 = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, x_2 = y_3$ ; then system (1) can be written as

$$\dot{y}_1 = -2ty_1 + y_2 + y_3 - e^{t^2}(e^{t^2}y_1 + t^2y_2 + ty_3)^3,$$

$$\dot{y}_2 = ty_1 - 3ty_2 + t^2y_3 + 2t(e^{t^2}y_1 + t^2y_2 + ty_3)^3, \quad (74)$$

$$\dot{y}_3 = t^3y_1 - 2ty_2 - 5t^5y_3 - 6t^3(e^{t^2}y_1 + t^2y_2 + ty_3)^3.$$

Simulation results are shown in Figure 1. Although the coefficients of the large-scale Lurie direct control system with time-varying coefficients are norm-unbounded, we can clearly see from Figure 1 that the convergence rate of the system is very fast. This illustrates the availability of our results.

*Example 2.* Consider system (39) with

$$A_{11}(t) = \begin{pmatrix} -2t & 1 \\ t & -3t \end{pmatrix}, \quad A_{12}(t) = \begin{pmatrix} 1 \\ t^2 \end{pmatrix},$$

$$A_{21}(t) = (t^3 \quad -2t), \quad A_{22}(t) = -5t^5,$$

$$b_{11}(t) = \begin{pmatrix} -e^{t^2} \\ 2t \end{pmatrix}, \quad b_{12}(t) = \begin{pmatrix} -e^{t^2} \\ 0 \end{pmatrix},$$

$$b_{21}(t) = -6t^3, \quad b_{22}(t) = t^2, \quad (75)$$

$$c_{11}(t) = \begin{pmatrix} e^{t^2} \\ t^2 \end{pmatrix}, \quad c_{12}(t) = t,$$

$$c_{21}(t) = \begin{pmatrix} e^{t^2} \\ 0 \end{pmatrix}, \quad c_{22}(t) = 2t^2.$$

Similarly, we choose

$$T = 1, \quad P_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad P_2 = 1. \quad (76)$$

Then,

$$s_1(t) = 2t, \quad s_2(t) = 10t^5,$$

$$g_1(t) = e^{2t^2} - 2t^3, \quad g_2(t) = e^{2t^2} - 2t^4. \quad (77)$$

This implies that assumptions A5 and A6 are satisfied. By calculating the upper limits, we have

$$\begin{aligned}
 \overline{\lim}_{t \rightarrow +\infty} \frac{2 \|P_1 b_{11}(t)\|}{\sqrt{s_1(t) g_1(t)}} &= \bar{\alpha}_{11} = 0, \\
 \overline{\lim}_{t \rightarrow +\infty} \frac{2 \|P_1 b_{12}(t)\|}{\sqrt{s_1(t) g_2(t)}} &= \bar{\alpha}_{12} = 0, \\
 \overline{\lim}_{t \rightarrow +\infty} \frac{2 \|P_2 b_{21}(t)\|}{\sqrt{s_2(t) g_1(t)}} &= \bar{\alpha}_{21} = 0, \\
 \overline{\lim}_{t \rightarrow +\infty} \frac{2 \|P_2 b_{22}(t)\|}{\sqrt{s_2(t) g_2(t)}} &= \bar{\alpha}_{22} = 0, \\
 \overline{\lim}_{t \rightarrow +\infty} \frac{\|\sum_{j=1}^2 c_{1j}^T(t) A_{j1}(t) + \dot{c}_{11}^T(t)\|}{\sqrt{s_1(t) g_1(t)}} &= \bar{\beta}_{11} = 0, \\
 \overline{\lim}_{t \rightarrow +\infty} \frac{\|\sum_{j=1}^2 c_{1j}^T(t) A_{j2}(t) + \dot{c}_{12}^T(t)\|}{\sqrt{s_2(t) g_1(t)}} &= \bar{\beta}_{12} = 0, \\
 \overline{\lim}_{t \rightarrow +\infty} \frac{\|\sum_{j=1}^2 c_{2j}^T(t) A_{j1}(t) + \dot{c}_{21}^T(t)\|}{\sqrt{s_1(t) g_2(t)}} &= \bar{\beta}_{21} = 0, \\
 \overline{\lim}_{t \rightarrow +\infty} \frac{\|\sum_{j=1}^2 c_{2j}^T(t) A_{j2}(t) + \dot{c}_{22}^T(t)\|}{\sqrt{s_2(t) g_1(t)}} &= \bar{\beta}_{22} = 0.
 \end{aligned}
 \tag{78}$$

So, assumption A7' is satisfied. From

$$\begin{aligned}
 \overline{\lim}_{t \rightarrow +\infty} \frac{2 \|P_1 A_{12}(t)\|}{\sqrt{s_1(t) s_2(t)}} &= \bar{\gamma}_{12} = 0, \\
 \overline{\lim}_{t \rightarrow +\infty} \frac{2 \|P_2 A_{21}(t)\|}{\sqrt{s_2(t) s_1(t)}} &= \bar{\gamma}_{21} = \frac{\sqrt{2}}{\sqrt{5}}, \\
 \overline{\lim}_{t \rightarrow +\infty} \frac{|\sum_{j=1}^2 c_{1j}^T(t) b_{j2}(t)|}{\sqrt{g_1(t) g_2(t)}} &= \bar{\mu}_{12} = 0, \\
 \overline{\lim}_{t \rightarrow +\infty} \frac{|\sum_{j=1}^2 c_{2j}^T(t) b_{j1}(t)|}{\sqrt{g_2(t) g_1(t)}} &= \bar{\mu}_{12} = 0,
 \end{aligned}
 \tag{79}$$

we know that assumptions A8' and A9' are satisfied and

$$\begin{aligned}
 \bar{Q} &= \begin{pmatrix} -1 & \bar{\gamma}_{12} & \bar{\alpha}_{11} & \bar{\alpha}_{12} \\ \bar{\gamma}_{21} & -1 & \bar{\alpha}_{21} & \bar{\alpha}_{22} \\ \bar{\beta}_{11} & \bar{\beta}_{12} & -1 & \bar{\mu}_{12} \\ \bar{\beta}_{21} & \bar{\beta}_{22} & \bar{\mu}_{21} & -1 \end{pmatrix} \\
 &= \begin{pmatrix} -1 & 0 & 0 & 0 \\ \frac{\sqrt{2}}{\sqrt{5}} & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.
 \end{aligned}
 \tag{80}$$

It is easy to see that  $\bar{Q}$  is stable. Hence, according to Corollary 11, the system is absolutely stable.

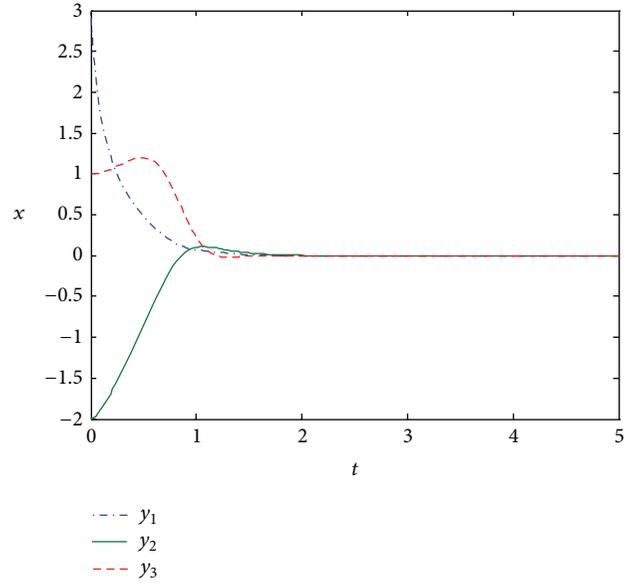


FIGURE 1: The state response of system (74).

Let  $f_1(\sigma_1) = \sigma_1^3$ ,  $f_2(\sigma_2) = \sigma_2^5$ , and  $x_1 = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ ,  $x_2 = y_3$ ; then system (39) can be written as

$$\begin{aligned}
 \dot{y}_1 &= -2ty_1 + y_2 + y_3 - e^{t^2} (e^{t^2} y_1 + t^2 y_2 + ty_3)^3 \\
 &\quad - e^{t^2} (e^{t^2} y_1 + 2t^2 y_3)^5, \\
 \dot{y}_2 &= ty_1 - 3ty_2 + t^2 y_3 + 2t (e^{t^2} y_1 + t^2 y_2 + ty_3)^3, \\
 \dot{y}_3 &= t^3 y_1 - 2ty_2 - 5t^5 y_3 - 6t^3 (e^{t^2} y_1 + t^2 y_2 + ty_3)^3 \\
 &\quad + t^2 (e^{t^2} y_1 + 2t^2 y_3)^5,
 \end{aligned}
 \tag{81}$$

as shown in Figure 2. For the large-scale Lurie direct control system with time-varying coefficients and multiple nonlinearities, although the coefficients are norm-unbounded, we can see from Figure 2 that the large-scale system is absolutely stable.

## 5. Conclusions

The absolute stability of large-scale Lurie direct control systems with time-varying coefficients and systems with multiple nonlinearities is studied in this paper. By restricting the relative magnitude of the time-varying coefficients and employing the decomposition theory of large-scale systems, some absolute stability criteria were obtained. The criteria, introduced in this paper, can be used not only in large-scale Lurie direct control systems with norm-unbounded coefficients but also in systems with norm-bounded coefficients. Two numerical examples are introduced to illustrate the availability of our results.

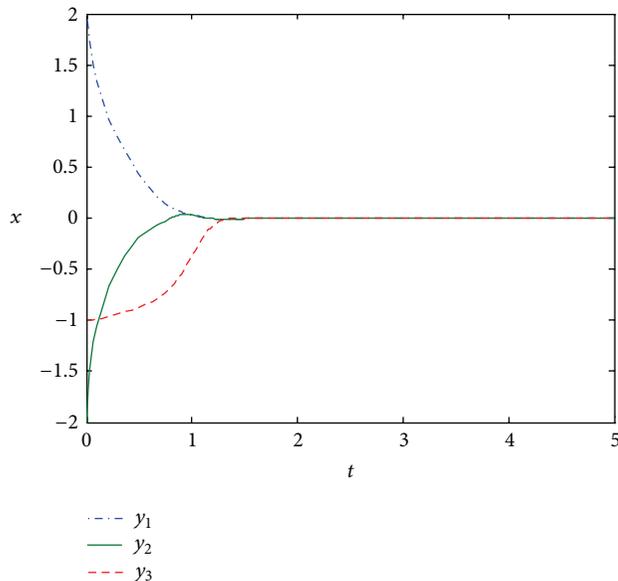


FIGURE 2: The state response of system (81).

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 61174209) and the Oriented Award Foundation for Science and Technological Innovation, Inner Mongolia Autonomous Region, China (2012).

## References

- [1] A. I. Lurie, *Some Non-Linear Problems in the Theory of Automatic Control*, Her Majesty's Stationery Office, 1957.
- [2] M. A. Aizerman and F. R. Gantmacher, *Absolute Stability of Regulator Systems*, Holden-Day, San Francisco, Calif, USA, 1964.
- [3] S. Lefschetz, *Stability of Nonlinear Control Systems*, Academic Press, New York, NY, USA, 1965.
- [4] S. Zhao, "On the absolute stability of direct control system," *Acta Mathematica Sinica*, vol. 22, pp. 404–419, 1979.
- [5] L. T. Grujić and D. Petkovski, "On robustness of Lurie systems with multiple non-linearities," *Automatica*, vol. 23, no. 3, pp. 327–334, 1987.
- [6] T. Wada, M. Ikeda, Y. Ohta, and D. D. Šiljak, "Parametric absolute stability of multivariable Lurie systems," *Automatica*, vol. 36, no. 9, pp. 1365–1372, 2000.
- [7] M. Arcak and A. Teel, "Input-to-state stability for a class of Lurie systems," *Automatica*, vol. 38, no. 11, pp. 1945–1949, 2002.
- [8] X. X. Liao and P. Yu, "Sufficient and necessary conditions for absolute stability of time-delayed Lurie control systems," *Journal of Mathematical Analysis and Applications*, vol. 323, no. 2, pp. 876–890, 2006.
- [9] N. Chen, W.-H. Gui, and B.-Y. Liu, "Parametric absolute stability of interconnected Lurie control systems," *Acta Automatica Sinica*, vol. 33, no. 12, pp. 1283–1289, 2007.
- [10] A. I. Zeević, E. Cheng, and D. D. Iljak, "Control design for large-scale Lur'e systems with arbitrary information structure constraints," *Applied Mathematics and Computation*, vol. 217, no. 3, pp. 1277–1286, 2010.
- [11] F. Hao, "Absolute stability of uncertain discrete Lur'e systems and maximum admissible perturbed bounds," *Journal of the Franklin Institute*, vol. 347, no. 8, pp. 1511–1525, 2010.
- [12] C. Yin, S.-M. Zhong, and W.-F. Chen, "Robust H<sub>∞</sub> control for uncertain Lur'e systems with sector and slope restricted nonlinearities by PD state feedback," *Nonlinear Analysis: Real World Applications*, vol. 12, no. 1, pp. 501–512, 2011.
- [13] K. Ramakrishnan and G. Ray, "Improved delay-range-dependent robust stability criteria for a class of Lur'e systems with sector-bounded nonlinearity," *Journal of the Franklin Institute*, vol. 348, no. 8, pp. 1769–1786, 2011.
- [14] S. J. Choi, S. M. Lee, S. C. Won, and J. H. Park, "Improved delay-dependent stability criteria for uncertain Lurie systems with sector and slope restricted nonlinearities and time-varying delays," *Applied Mathematics and Computation*, vol. 216, no. 3, pp. 1024–1027, 2010.
- [15] T. Li, A. Song, S. Fei, and T. Wang, "Global synchronization in arrays of coupled Lurie systems with both time-delay and hybrid coupling," *Communications in Nonlinear Science and Numerical Simulation*, vol. 16, no. 1, pp. 10–20, 2011.
- [16] H. Xie, *Theories and Applications of Absolute Stability*, Science Press, Beijing, China, 1986.
- [17] X. Liao, Z. Chen, F. Xu, and P. Yu, "Robust absolute stability of Lurie interval control systems," *International Journal of Robust and Nonlinear Control*, vol. 17, no. 18, pp. 1669–1689, 2007.
- [18] F. Liao, "The stability of a type of large-scale system with variable coefficients," *Journal of Mathematical Research and Exposition*, vol. 8, no. 3, pp. 411–415, 1988.
- [19] Y. Cui, F. Liao, and H. Wei, "Robust exponential control problem for linear large-scale systems with uncertainties and unbounded coefficients," *Control Engineering of China*, vol. 17, no. 5, pp. 27–31, 2010.
- [20] F. Liao, H. Yao, and H. Liu, "Absolute stability of large-scale Lurie indirect systems with unbounded coefficients," *Journal of Systems Science & Information*, vol. 3, no. 2, pp. 303–311, 2005.
- [21] F. Liao, L. Zhao, Y. Tang, and H. Liu, "Absolute stability of general Lurie indirect control large-scale systems," *Journal of Basic Science and Engineering*, vol. 14, no. 4, pp. 579–589, 2006.
- [22] F.-C. Liao, A.-G. Li, and F.-B. Sun, "Absolute stability of Lurie systems and Lurie large-scale systems with multiple operators and unbounded coefficients," *Journal of University of Science and Technology Beijing*, vol. 31, no. 11, pp. 1472–1479, 2009.
- [23] D. Wang and F. Liao, "Absolute stability of Lurie direct control systems with time-varying coefficients and multiple nonlinearities," *Applied Mathematics and Computation*, vol. 219, no. 9, pp. 4465–4473, 2013.
- [24] F. N. Bailey, "The application of Lyapunov's second method to interconnected systems," *Journal of the Society for Industrial and Applied Mathematics A*, vol. 3, no. 3, pp. 443–462, 1966.
- [25] A. N. Michel and R. K. Miller, *Qualitative Analysis of Large Scale Dynamical Systems*, Academic Press, New York, NY, USA, 1977.
- [26] X. Liao and P. Yu, *Absolute Stability of Nonlinear Control Systems*, Springer, 2008.

- [27] J. Lunze, *Feedback Control of Large-Scale System*, Prentice Hall, New York, NY, USA, 1992.
- [28] F. Liao, L. Zheng, T. Katayama, and K. Takaba, "Decentralized robust controller design for large-scale systems with uncertainties," in *Proceedings of the 4th Asian Control Conference*, pp. 174–179, Singapore, 2002.
- [29] Z. Ji-ye and S. Zhong-Zhou, "Necessary and sufficient criteria for absolute stability of the direct control system," *Applied Mathematics and Mechanics*, vol. 15, no. 3, pp. 259–266, 1994.
- [30] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York, NY, USA, 1979.
- [31] T. S. Shores, *Applied Linear Algebra and Matrix Analysis*, Springer, 2007.

## Research Article

# An Optimal Control Strategy for DC Bus Voltage Regulation in Photovoltaic System with Battery Energy Storage

Muhamad Zalani Daud,<sup>1,2</sup> Azah Mohamed,<sup>2</sup> and M. A. Hannan<sup>2</sup>

<sup>1</sup> School of Ocean Engineering, Universiti Malaysia Terengganu, 21030 Kuala Terengganu, Malaysia

<sup>2</sup> Department of Electrical, Electronic & Systems Engineering, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

Correspondence should be addressed to Muhamad Zalani Daud; zalani@umt.edu.my

Received 17 February 2014; Accepted 17 March 2014; Published 24 April 2014

Academic Editors: V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Muhamad Zalani Daud et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents an evaluation of an optimal DC bus voltage regulation strategy for grid-connected photovoltaic (PV) system with battery energy storage (BES). The BES is connected to the PV system DC bus using a DC/DC buck-boost converter. The converter facilitates the BES power charge/discharge to compensate for the DC bus voltage deviation during severe disturbance conditions. In this way, the regulation of DC bus voltage of the PV/BES system can be enhanced as compared to the conventional regulation that is solely based on the voltage-sourced converter (VSC). For the grid side VSC (G-VSC), two control methods, namely, the voltage-mode and current-mode controls, are applied. For control parameter optimization, the simplex optimization technique is applied for the G-VSC voltage- and current-mode controls, including the BES DC/DC buck-boost converter controllers. A new set of optimized parameters are obtained for each of the power converters for comparison purposes. The PSCAD/EMTDC-based simulation case studies are presented to evaluate the performance of the proposed optimized control scheme in comparison to the conventional methods.

## 1. Introduction

Despite the many advantages offered by photovoltaic- (PV-) based renewable energy (RE) generation, it suffers from unpredictable environmental conditions and abrupt changes in system loads. In addition, when the PV-based RE generation is grid connected, the possibility of utility grid fault at the point of common connection (PCC) might result in a system breakdown or the interruption of power supplied to critical loads. One of the typical challenges in integrating such a variable generation to utility grid is in controlling the DC bus voltage stability within the power conversion system [1, 2]. The disturbances, such as varying environmental conditions, system loads, and fault occurrences, cause the DC bus voltage to fluctuate, overshoot or undershoot, and sag or dip [3]. Poor regulation of the DC bus voltage may result in system instability and inferior efficiency of PV systems. The problems above can be attributed to the poor dynamics of PV control systems in comparison to the transient time from the disturbances [4].

The instability of DC bus voltage may propagate over the PV system network, where, in some cases, the requirement for fast dynamic compensation devices, such as diesel generators or the battery energy storage (BES) for power fluctuation management and fault ride by mitigation, is indispensable. Lead-acid-type battery is currently the most preferred storage method for most RE based integration issues because their mature technology provides a reasonable trade-off between cost and performance [5, 6]. The importance of DC bus voltage regulation based on BES is that it provides a constant DC bus voltage seen by the grid side voltage-sourced converter (G-VSC), resulting in efficient power conversion while protecting the DC bus capacitor and the G-VSC valves against overvoltage stress.

Technically, the variation in DC bus voltage can be mitigated by adjusting the duty cycle of the power converters responsible for voltage regulation, which is normally performed by the proportional integral (PI) compensator [7–10]. The PI controller is commonly characterized by its distinct parameters, namely, the proportional gain,  $k_p$ ,

and the integral gain,  $k_i$ , respectively. From the literature, several conventional tuning strategies that have been in use include the hand tuning or trial and error, Smith, Ziegler-Nichols, and pole placement methods [7]. However, these tuning strategies require the determination of system transfer function, which makes the method more complex [8]. A number of optimization methods have been developed for fine-tuning converter control parameters [9, 10], but only few applications used the simplex algorithm. From previous studies, applications of the simplex algorithm are in controlling FACTS devices [11, 12] and vehicular control systems that employ bidirectional converters [13, 14]. This paper addresses the issue of improving DC bus voltage regulation by using G-VSC and BES with a bidirectional DC/DC buck-boost converter for the efficient performance of grid-connected PV systems. The BES connected to the DC bus of the PV system can enhance the system's dynamic performance by controlling the charging/discharging of BES with a bidirectional converter when the system is subjected to varying disturbances. The optimal controls of the G-VSC and the bidirectional DC/DC converter are investigated in the proposed control methodologies. First, the control methods of the G-VSC, namely, the voltage-mode and current-mode controls, are assessed. Second, the control parameters of the G-VSC and the bidirectional DC/DC are optimized using the simplex optimization algorithm to obtain the optimum set of parameters for the control systems. Section 2 briefly describes the system configuration and operation principle of the considered grid-connected PV with BES. Section 3 provides the modeling details of the PV and BES. Section 4 describes the proposed control strategies for the G-VSC and the bidirectional DC/DC converter. Section 5 discusses the DC bus regulation strategy and the control parameter optimization in detail. Sections 6 and 7 present the simulated results and some concluding remarks, respectively.

## 2. System Configuration and Operation

The configuration of the considered single-stage grid-connected hybrid PV/BES system is as shown in Figure 1. The system consists of a PV array (PV gen.) interfaced to the DC bus through a buck DC/DC converter with a maximum power point tracking (MPPT) controller. The G-VSC facilitates the MPPT operation through regulation of the DC bus voltage as well as transfer of power from the DC bus to the utility grid. In addition, the G-VSC provides synchronization of the PV system with the grid during startup or reconnection after system islanding.

As can be seen from Figure 1, to enhance the DC bus voltage regulation, BES is used where it is interfaced via a bidirectional buck-boost converter (BES conv.) which controls the charge/discharge processes during severe operating conditions such as abrupt change in solar irradiation level and fault occurrences. From the G-VSC AC output terminals, the hybrid subsystem is connected to the utility grid at the PCC through a low-pass filter and an interconnection transformer that is represented by an inductor. These components are responsible for filtering harmonics and isolating the entire

system from the utility grid. The transformer steps up the voltage level of the PV system from 0.23 kV to 11 kV line-to-line RMS voltage. The PV/BES system injects total power,  $P_G$ , to the utility grid in which in this case the utility grid network is based on a standard medium voltage distribution system [15].

## 3. System Modeling

This section provides the detailed models of the system components which are simulated using the PSCAD/EMTDC transient simulator [16].

*3.1. Modeling of PV Array.* The inclusion of the PV model in PSCAD/EMTDC is based on the Norton equivalent circuit given in Figure 2 [17]. Assuming that the PV modules are of the same type and are subjected to similar environmental conditions, a PV array model (Figure 2(b)) can be represented by a combination of series-connected,  $N_s$ , and parallel-connected,  $N_p$ , modules. The current and resistance of such PV array model are given by

$$\begin{aligned} I_{eq,array} &= N_p I_{eq} = N_p I_i \frac{R_p}{R_s + R_p}, \\ R_{eq,array} &= \frac{N_s}{N_p} R_{eq} = \frac{N_s}{N_p} (R_s + R_p), \end{aligned} \quad (1)$$

where  $R_s$  and  $R_p$  are the series and parallel resistance of an ideal single diode PV cell model, representing the structural resistance and leakage current of the  $p$ - $n$  junction, respectively [18].  $I_i$  is the terminal current of an ideal PV module derived from the ideal PV cell model, which has the following form [17]:

$$I_i = I_g \left[ I_0 \left\{ \exp \left( \frac{\beta}{\alpha} \cdot \frac{R_p}{R_s + R_p} (V + R_s I) \right) - 1 \right\} \right], \quad (2)$$

where  $I_g$  is the cell's photocurrent,  $I_0$  is the dark current, and  $\alpha$  is the diode ideality factor.  $\beta$  is the inverse thermal voltage defined as  $\beta(T) = q/kT$ , where  $q$  is the electron charge,  $k$  is Boltzmann's constant, and  $T$  is the  $p$ - $n$  junction temperature. Except for  $\alpha$  (assumed in this study as 1.5), all of the circuit parameters given in (2) are functions of the PV device type obtained from the manufacturer datasheet [17, 19].

*3.2. Modeling of Lead-Acid Battery.* Many options are available for the selection of a suitable battery model for a range of applications. These options include the simple voltage source model, the Thevenin model [20], generic models [21, 22], and dynamic and more realistic models [23, 24] that consider the nonlinear characteristics of battery parameters. However, to avoid excessive complexity with the consideration of the dynamic behavior of the battery cell, this study considered a generic model described in [21]. The model for the lead-acid battery is constructed based on two important parameters, namely, the terminal voltage,  $V_{bat}$ , and state of charge, SOC,

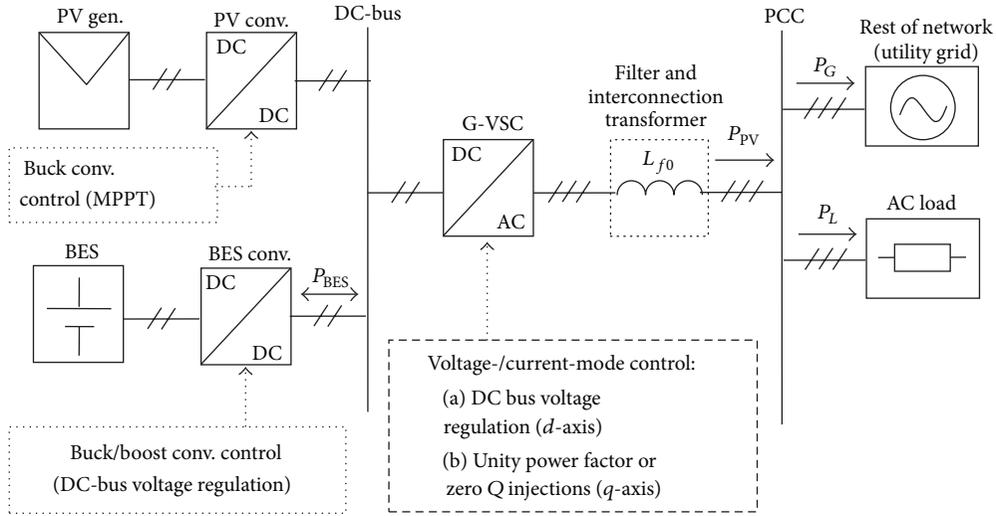


FIGURE 1: System configuration/control of the grid-connected PV with BES.

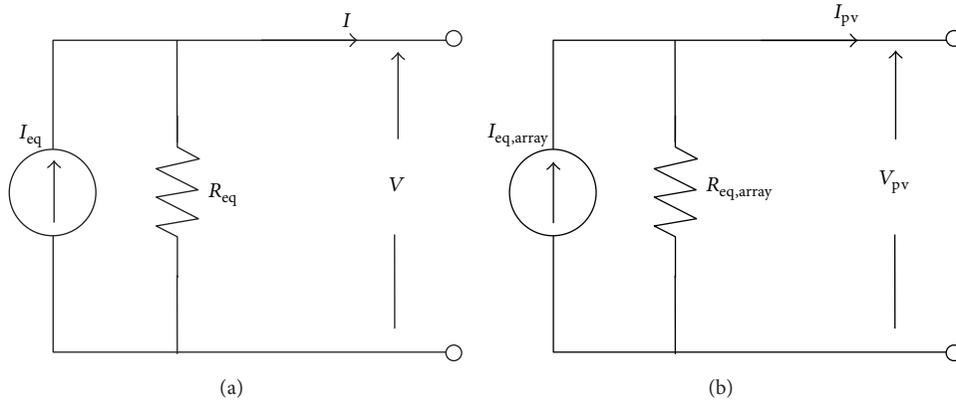


FIGURE 2: Norton equivalent circuit of the (a) PV module and (b) PV array.

which represent the behavior of the battery.  $V_{bat}$  and SOC can be calculated as functions of battery current,  $I_{bat}$ , as follows:

$$V_{bat} = E_{bat} - R_{int}I_{bat},$$

$$SOC = 100 \left( 1 - \frac{\int I_{bat} dt}{Q_{bat}} \right), \quad (3)$$

where  $E_{bat}$  is the open circuit voltage of the battery that is modeled using a controlled voltage source calculated as follows:

$$E_{bat} = E_0 - K \frac{1 - SOC}{SOC} Q + A \exp \{-B(1 - SOC) Q_{bat}\}. \quad (4)$$

In (3)–(4),  $R_{int}$  and  $Q_{bat}$  are the battery internal resistance and capacity, respectively.  $E_0$  represents the battery open circuit voltage that lies between the fully charged voltage and the exponential voltage of the battery discharge curve,  $K$  is the polarization voltage,  $A$  is the exponential voltage, and  $B$  is the exponential capacity. From (4), the model accounts for both

the normal voltage part and the exponential part represented by the second and third terms, respectively. Following the parameter extraction procedure provided in [21], all model parameters can be extracted from the manufacturer discharge curves, which are available in the battery datasheet [25].

#### 4. Control of Power Converters

As shown in Figure 1, the power converters are comprised of the DC/DC buck converter, G-VSC, and BES DC/DC buck-boost converter, respectively. Since G-VSC and BES DC/DC buck-boost controls directly influence the voltage regulation of PV system DC bus, the control schemes are described in detail in the following subsections.

**4.1. G-VSC Control.** In general, the aim of the G-VSC control scheme is to control the active and reactive power transfer between the PV system and the utility grid. The active current component ( $d$ -axis) of the G-VSC is used to regulate the DC bus voltage of the PV system, whereas the reactive

current component ( $q$ -axis) is used to ensure that the system injects reactive power at approximately unity power factor. The general operation of the PV system has to follow the regulations stated in the IEEE 1547 and UL1741, which states that the PV inverter is only allowed to inject power at unity power factor [26]. The G-VSC control developed is based on two different methods, namely, the voltage-mode [1] and current-mode [17] controllers, respectively.

**4.1.1. Voltage-Mode Control.** Figure 3 shows the detailed G-VSC circuit with the control block diagram based on the voltage-mode control scheme. The controller has merits, such as reduced control complexity and low number of control loops [1]. Since it is an open-loop control scheme, the measures taken to protect the G-VSC valves against instantaneous over current or network faults are by limiting the output references from PI controllers using the saturation blocks (limiter) shown in Figure 3(b). The principle for the active,  $P$ , and reactive,  $Q$ , power controls via G-VSC is expressed as follows:

$$P = \frac{V_{pcc} V_{dc}}{X_s} m \sin(\delta),$$

$$Q = \frac{V_{pcc}}{X_s} [V_{pcc} - m V_{dc} \cos(\delta)],$$
(5)

where  $V_{pcc}$  is the output fundamental component voltage at the PCC,  $V_{dc}$  is the DC bus voltage,  $X_s = \omega L_{f0}$  is the sum of inductances for filter inductance and coupling transformer inductance,  $\delta$  is the angle, and  $m$  is the modulation index. Notably, based on the PWM principle, the relationship between the output fundamental component voltage at the G-VSC terminal,  $V_{pwm}$ , and the DC bus voltage can be expressed as  $V_{pwm} = m V_{dc}$ .  $P$  and  $Q$  are proportional to  $\delta$  and  $m$ , respectively, when  $L_{f0}$  is optimally designed and  $\delta$  is less than  $10^\circ$  [27]. This control strategy results in a simple voltage-mode control system design where active power is controlled by the value of  $\delta$  and reactive power is regulated using  $m$  through PI1 and PI2, respectively.

As shown in Figure 3(b), PI1 generates a reference angle,  $\delta$ , for the PWM by processing the error of the DC bus voltage to ensure that the flow of power is constantly directed from the PV to the grid while regulating the DC bus voltage. Consequently, PI2 generates the output reference,  $m$ , that processes the error of reactive power from its reference value and keeps the net power injected to the grid at the unity power factor ( $Q_{ref} = 0$ ). Importantly, the value of  $m$  is controlled to make it remain less than unity (i.e., the linear modulation region), with the optimal value preferably close to unity, to ensure a reasonable efficient power conversion by the G-VSC. The output terminal voltage of G-VSC is regulated at  $\omega$  which is the reference fundamental frequency at the PCC.

**4.1.2. Current-Mode Control.** The current-mode control is a closed-loop current control method, in which the phase angle and magnitude of the VSC terminal voltage are controlled in a  $d$ - $q$  rotating reference frame. This control approach imposes direct regulation on the G-VSC current, which avoids the

vulnerability of G-VSC valves when they are subjected to large transient currents, such as during network faults and abrupt load changes. Referring to the G-VSC circuit shown in Figure 3(a), the control block diagram for the current-mode control scheme is as shown in Figure 4.

As can be seen in Figure 4, a phase-locked loop (PLL) is utilized to extract the angle,  $\theta$ , of the grid voltage, that is to be used by the  $abc$ -to- $dq0$  or  $dq0$ -to- $abc$  transformation blocks. In the PLL, the PCC voltage vector ( $v_{u-abc}$ ) is projected on the  $d$  and  $q$  axes of the  $d$ - $q$  frame and the  $d$ - $q$  frame is rotated in such a way that  $v_{uq}$  is forced to zero [17]. Settling  $v_{uq}$  at zero makes the  $d$ -axis of the  $d$ - $q$  frame aligned with the grid voltage vector, and the  $d$ - $q$  frame rotational speed becomes equal to the grid angular frequency [1]. Regarding this, the synchronization scheme based on PLL makes the value of  $P$  and  $Q$  proportional to and can be controlled by  $i_d$  and  $i_q$ , respectively, as follows:

$$P = \frac{3}{2} v_{ud} i_d,$$

$$Q = -\frac{3}{2} v_{ud} i_q.$$
(6)

In this case,  $P$  is represented as a DC bus voltage ( $V_{dc}$ ), in which the processing of the  $V_{dc}$  and  $Q$  errors is based on the PI compensators (PI3 and PI4) to generate the reference currents,  $i_{d,ref}$  and  $i_{q,ref}$ , respectively, which can be used by the closed-loop current controller.  $i_{d,ref}$  and  $i_{q,ref}$  can be expressed as follows:

$$i_{d,ref} = \left( k_{pd3} + k_{id3} \frac{1}{s} \right) (V_{dc} - V_{dc,ref}),$$

$$i_{q,ref} = \left( k_{pq4} + k_{iq4} \frac{1}{s} \right) (Q - Q_{ref}),$$
(7)

where  $k_p$  and  $k_i$  are the proportional and integral gains, respectively, used in the PI3 and PI4 corresponding to  $d$ - and  $q$ -axes components, respectively.

The current-mode control scheme within the inner current loop control of Figure 4 independently controls  $i_d$  and  $i_q$  to make them rapidly track their respective current commands ( $i_{d,ref}$  and  $i_{q,ref}$ ), which can be represented in  $d$ - $q$ -frame equivalents as follows:

$$L_{f0} \frac{di_d}{dt} = \omega L_{f0} i_q + 0.5 V_{dc} m_d - v_{ud},$$

$$L_{f0} \frac{di_q}{dt} = -\omega L_{f0} i_d + 0.5 V_{dc} m_q - v_{uq},$$
(8)

where  $L_{f0}$  is the equivalent filter inductance that represents the total inductance of the grid side filter and the transformer leakage reactance,  $V_{dc}$  is the DC bus voltage, and  $\omega$  is the grid frequency. It should be noted that the terms  $0.5 V_{dc} m_d$  and  $0.5 V_{dc} m_q$  in (8) represent the G-VSC AC-side terminal voltages,  $v_{td}$  and  $v_{tq}$ , respectively, in which  $m_d$  and  $m_q$  are the corresponding  $d$ -axis and  $q$ -axis control inputs (PWM modulation signals), while  $v_{ud}$  and  $v_{uq}$  are the exogenous inputs, respectively [17].

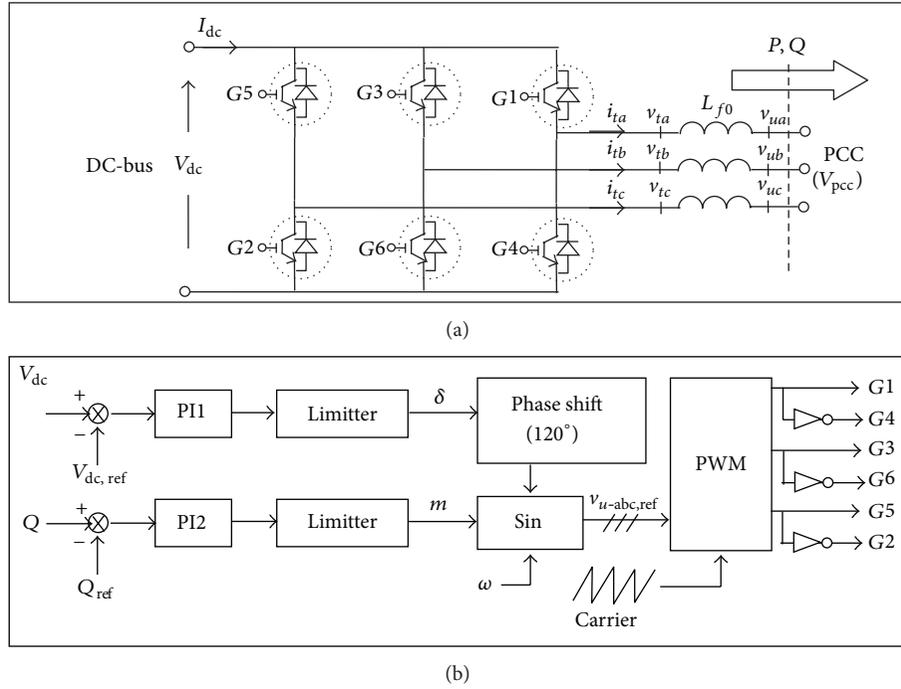


FIGURE 3: G-VSC control scheme, (a) G-VSC circuit, and (b) voltage-mode controller.

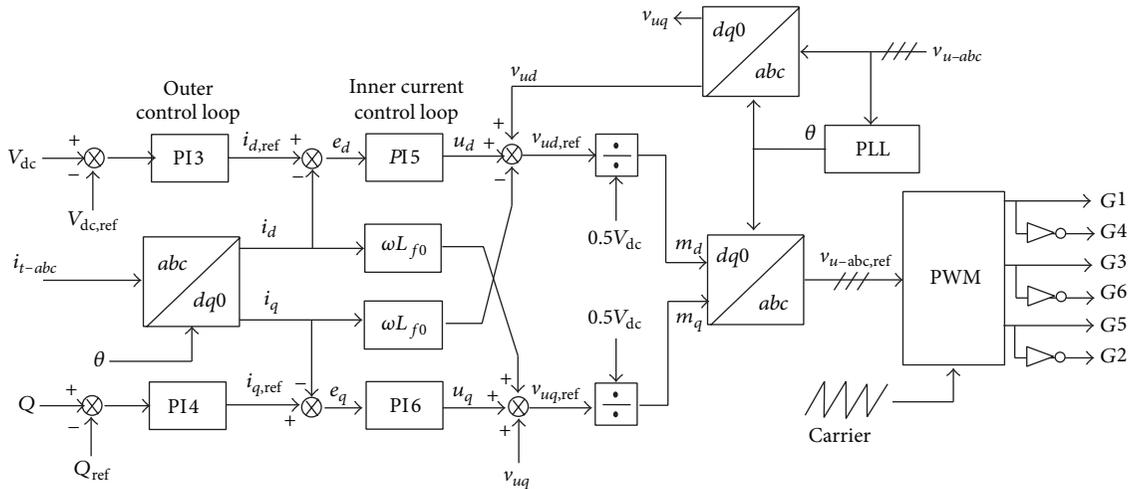


FIGURE 4: G-VSC control based on the current-mode scheme.

Rearranging (8) and considering  $u_d = L_{f0} di_d/dt$  and  $u_q = L_{f0} di_q/dt$ ,  $0.5V_{dc} m_d = v_{ud,ref}$ , and  $0.5V_{dc} m_q = v_{uq,ref}$  yield the following:

$$\begin{aligned} v_{ud,ref} &= v_{ud} + u_d - \omega L_{f0} i_q, \\ v_{uq,ref} &= v_{uq} + u_q + \omega L_{f0} i_d, \end{aligned} \quad (9)$$

where  $v_{ud,ref}$  and  $v_{uq,ref}$  are the output voltage references from the current controller,  $u_d$  is the output of the compensator PI5 ( $k_{pd5}$  and  $k_{id5}$ ) that processes the error signal of  $e_d = i_{d,ref} - i_d$ , and  $u_q$  is the output of compensator PI6 ( $k_{pd6}$  and  $k_{iq6}$ ) that processes the error signal of  $e_q = i_{q,ref} - i_q$ . The generated reference modulation index from the current controllers,  $m_d$

and  $m_q$ , is transformed back to the phase equivalents for use of the PWM generator using the  $dq0$ -to- $abc$  transformation block.

4.2. BES DC/DC Buck-Boost Converter Control. Control of the charging/discharging of BES is achieved by using a buck-boost converter circuit with two PI controllers, as shown in Figure 5, where PI7 processes the DC bus voltage discrepancies during disturbances to make the bus voltage follow the voltage set point set as  $V_{dc,ref} = 500$  V. The internal current control loop is also adopted for the battery current controller compensated by PI8. The output signal from PI8 is

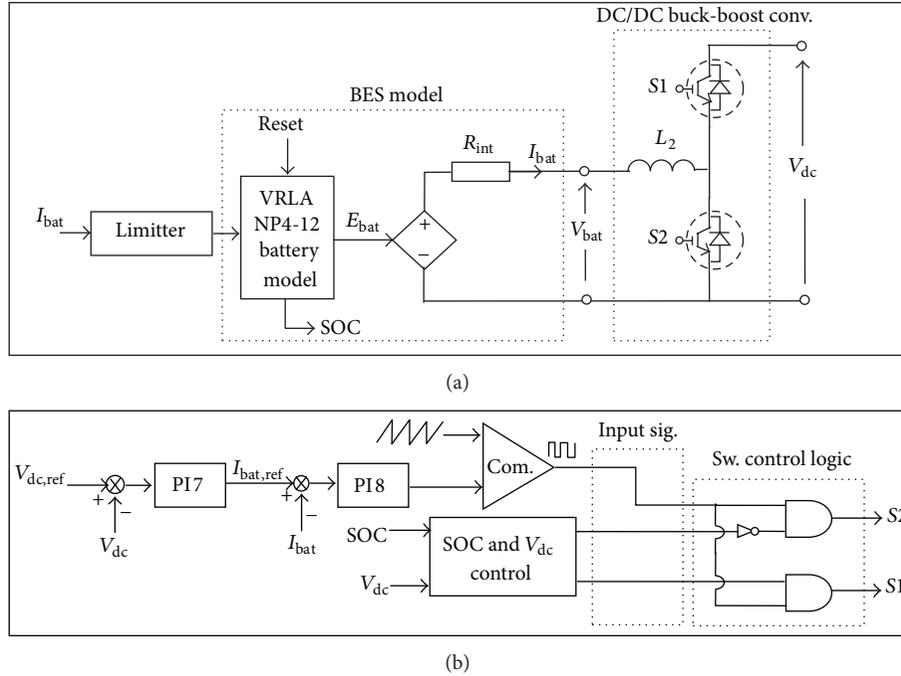


FIGURE 5: BES buck/boost converter control: (a) converter circuit and (b) control scheme.

passed to the PWM generation circuit where the logic circuit is used for the decisions, including the charge, discharge, or halt modes of operation. The switch S1 is triggered and S2 is zero during the boost (discharge) mode and vice versa during the buck (charge) mode to absorb power from the DC bus. Both S1 and S2 become zero (halt mode) when no regulation signal is transferred.

Figure 6 shows the algorithm developed for SOC and bus voltage control to ensure a secure and optimal operation of BES. The upper and lower limits of SOC denoted as  $SOC_H$  and  $SOC_L$  are defined to avoid overcharge or deep discharge of BES. The upper and lower SOC levels avoid operation at exponential capacity region which may affect the magnitude of current drawn at the battery terminal. Preferably, for lead-acid battery, the recommended useable capacity should not exceed 70% of the total capacity especially when considering continuous charge/discharge operation [25]. Consequently, the upper and lower boundaries of the DC bus voltage control are also considered to avoid unnecessary charge/discharge. In this case, the battery is set to charge/discharge only when the deviation of DC bus voltage exceeds a certain range denoted as  $V_{dc,H}$  and  $V_{dc,L}$ , respectively. The algorithm utilizes the input references of the SOC and the DC bus voltage where the output logic from the algorithm provides input for the control logic circuit for determining the operating modes (see Figure 5).

## 5. Optimal DC Bus Voltage Regulation Strategy

In conventional grid connected PV system, without BES, the regulation of DC bus voltage is achieved solely through the G-VSC. However, with BES connected to the DC bus, enhanced

control strategy can be achieved through regulation by both G-VSC and BES buck-boost converters. Regarding this, it is important to develop an optimal control scheme to boost the regulation performances of these power converters so as to minimize the stress received along the DC bus during severe disturbance condition. The effect of bus voltage fluctuation is related to power imbalance caused by fluctuations that originate from various sources of disturbances such as sudden change of solar irradiation, utility grid fault, and load step change in the vicinity. Such a power imbalance causes extra energy,  $\Delta E$ , which is related to the DC bus capacitor,  $C_{dc}$ , as follows:

$$\Delta E = \int_{T_s} \Delta P dt = \frac{1}{2} C_{dc} (V_{dc1}^2 - V_{dc0}^2), \quad (10)$$

where  $V_{dc0}$  and  $V_{dc1}$  represent the DC voltage at the start and end of the period  $T_s$ , respectively, and  $C_{dc}$  is the total DC capacitance. In (10),  $C_{dc}$  largely influences the variation in DC bus voltage, and thus the response of the system compensators for DC bus voltage regulation must be sufficiently fast.

From Figure 5(b), the battery current reference,  $I_{bat,ref}$ , for charge/discharge operation of buck-boost converter is provided by PI7 in which the DC bus voltage is regulated using the compensator as follows:

$$I_{bat,ref} = \left( k_{pb7} + k_{ib7} \frac{1}{s} \right) (V_{dc,ref} - V_{dc}), \quad (11)$$

where  $k_{pb7}$  and  $k_{ib7}$  are the control parameters.

To improve voltage regulation so as to increase the dynamic performance of the system, the control parameters for the BES DC/DC converter and the G-VSC need to be

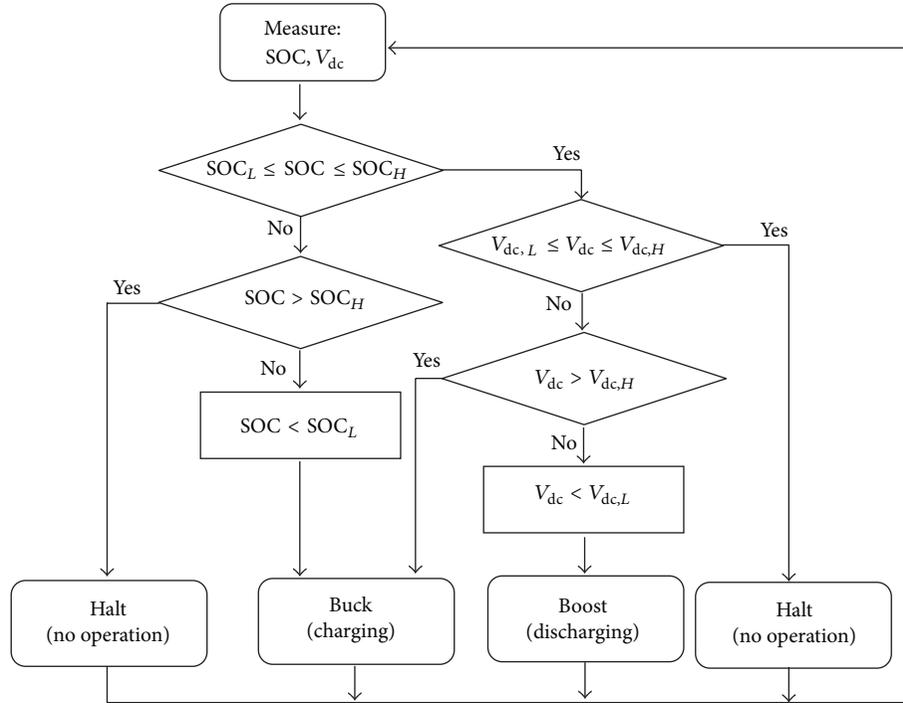


FIGURE 6: Flowchart for SOC and BES power charge/discharge controls.

optimized. The conventional hand-tuning method of the PI controllers used for the control system is improved by using the simplex optimization toolbox available in the PSCAD/EMTDC [28].

The simplex method is an optimization technique originally devised by Nelder and Mead [29] and is based on geometric considerations. The geometric figure whose vertices are defined by a set of  $n + 1$  in an  $n$ -dimensional space is called simplex. For example, for two variables, a simplex is a triangle. The simplex algorithm compares function values at the three vertices of the triangle where the worst vertex (e.g., the largest value for minimization problem) is discarded and replaced by a new one. By iterative process, a new triangle is formed and the search continues until the function values at the vertices become smaller and smaller. The size of the triangle is thus reduced until the coordinates of the optimum point are found [30]. The advantage of this method is its simplicity which needs only objective function evaluations and does not require first and higher order derivatives such as the gradient-based methods [31]. Moreover, it is computationally compact and effective with good convergence properties for most problems with parameters to be optimized less than ten [32]. The method is highly applicable for nonlinear multi-input multioutput systems without obtaining transfer function or mathematical formulation and hence suitable for finding local minimum of an objective function defined by several variables.

In the present study, the aim is to minimize the weighted integral squared error (WISE) products of the DC bus voltage controlled by the G-VSC and the BES buck-boost converters.

The considered performance index can be expressed in the following general form:

$$WISE = W_1 ISE_1 + W_2 ISE_2, \quad (12)$$

where  $W_1$  and  $W_2$  are the weighting factors used to signify the relative importance of the design specifications and  $ISE_1$  and  $ISE_2$  are the partial indices used to penalize the deviation of the DC bus voltage [31]. Based on the performance index in (12), the objective function to search for optimal parameters for the PI controllers (i.e., PI1, PI3, and PI7) responsible for DC bus voltage regulation is formulated. For example, the objective function for the case of BES-based regulation together with current-mode control scheme of G-VSC is described as follows:

$$OF(\mathbf{x}_{dc}) = \int_{T_0}^{T_1} W_1 (V_{dc} - V_{dc,ref})^2 dt + \int_{T_1}^{T_F} W_2 (V_{dc} - V_{dc,ref})^2 dt, \quad (13)$$

where  $\mathbf{x}_{dc}$  represents the control parameters for the PI3 and PI7 controller gains,  $T_0$  is the time at which the reference is changed within the entire length of the simulation time,  $T_F$ , and  $T_1$  is suitably selected intermediate point that permits the assignment of different weighting factors,  $W_1$  and  $W_2$ , to the initial and later portions of the response. The values of  $W_1$  and  $W_2$  are selected depending on the proportional and integral gains selected during initial trial and error tuning of the PI controllers, preferably  $W_1 > W_2$  to ensure stability of the system.

TABLE 1: PV module parameters.

Model parameter	Value	Symbols [17]
Cell type	6" multicrystalline silicon	—
Number of cells and connections	60 in series	—
Short circuit current at <sup>a</sup> STC	8.3 A	$I_{sc,r}$
Open circuit voltage at STC	36.7 V	$V_{oc,r}$
Maximum power current at STC	7.7 A	$I_{mpp,r}$
Maximum power voltage at STC	29.2 V	$V_{mpp,r}$
Current temperature coefficient	0.004 A/K	$k_I$
Voltage temperature coefficient	-0.1 V/K	$k_V$
Diode ideality factor	1.5	$\alpha$
Electron charge	$1.602 \times 10^{-19}$ C	$q$
Boltzmann's constant	$1.380 \times 10^{-23}$ J/K	$k$

<sup>a</sup>STC: standard test condition.

## 6. Results And Discussion

The effectiveness of the proposed control strategies in comparison to the conventional method is investigated through simulation in PSCAD/EMTDC. Validation of the system component models is first given, and then performance comparison of the simulation results with and without use of BES as well as with optimized control scheme is discussed.

**6.1. Model Validation.** Validation of the models is made by comparing the steady state PV module  $I$ - $V$  curves and the BES discharge voltage curves, respectively, with the corresponding manufacturer data.

**6.1.1. PV Model.** The parameter extraction method for the PV model is based on [17], with key equations given in (1)–(2). Table 1 provides the parameter values of the PV module used in the simulation of PV model in PSCAD/EMTDC.

The results from the simulated module's characteristic curves are compared with the Hyundai SG-series multicrystalline type solar module manufacturer datasheet [19]. Figure 7 shows the results of the comparison of the module's  $I$ - $V$  characteristics simulated at standard test conditions (STCs). As can be seen in Figure 7, the simulated results (Figure 7(a)) are compared well to the values in the manufacturer datasheet (Figure 7(b)). The module's nominal power output is measured approximately by 226 W, as shown in Figure 7(a). It is evident that the behavior of the simulated PV module current increases with the increase in solar irradiation at constant temperature (25°C). Consequently, the voltage decreases with the increase in the cell's temperature, when the irradiation is fixed at 100 W/m<sup>2</sup>. Having validated the PV simulation model, the PV modules are arranged in series and parallel connections to form a PV array (generator). The PV generator size considered in the simulation is 215 kW<sub>p</sub> which is a typical size for grid-connected PV system found in residential areas.

**6.1.2. Battery Model.** An NPC series valve regulated lead acid, namely, the Yuasa VRLA NP4-12cell (12 V, 4 Ah), is considered suitable for the cyclic application of the battery

TABLE 2: Battery module parameters (12 V, 4 Ah).

Model parameter	Value	Symbols [21]
Cell type	Lead-acid (VRLA)	—
Number of cells per module	6 in series	—
Rated capacity	4 Ah	$Q_{rated}$
Battery reserve	0.99 Ah	$Bat.resv$
Nominal capacity	0.85 Ah	$Q_{nom}$
Exponential capacity	0.67	$Q_{exp}$
Maximum voltage	12.15 V	$E_{full}$
Exponential voltage	12.05 V	$E_{exp}$
Nominal voltage	12 V	$E_{nom}$
Charge current	4 A	$I_{chg}$
Internal resistance	0.04	$R_{int}$
Efficiency	80%	$Eta$
<sup>a</sup> Series battery	1	$n_s$
<sup>a</sup> Parallel battery	1	$n_p$

<sup>a</sup>Increased combinations of  $n_s$  and  $n_p$  scales up the battery modules to form a BES.

model [25]. Parameter values of the battery module are detailed in Table 2.

Figure 8 shows a comparison of the 12 V battery module simulation results with the battery data from the manufacturer in terms of the terminal voltage behaviors versus discharge time. The discharge currents ranging from 0.8 to 4 A have been simulated as shown in Figure 8(a). From the figure, all the discharge current curves of the simulated battery module are generally in good agreement with the standard manufacturer's discharge test data [25]. From Figure 8(b), by superimposing the voltage curve of  $0.2 \times C_{bat}$  ampere (0.8 A), the curve matches very well with the actual test curve during almost 80% of the discharge.

The BES simulation was developed by increasing  $n_s$  and  $n_p$  values, that is, the series and parallel arrangement of the battery module's matrix. In the simulation, a total of 48 kWh BES was used which can be arranged from approximately 20 series connected battery modules with 50 parallel strings. The BES output terminal voltage is around 240 V and it is

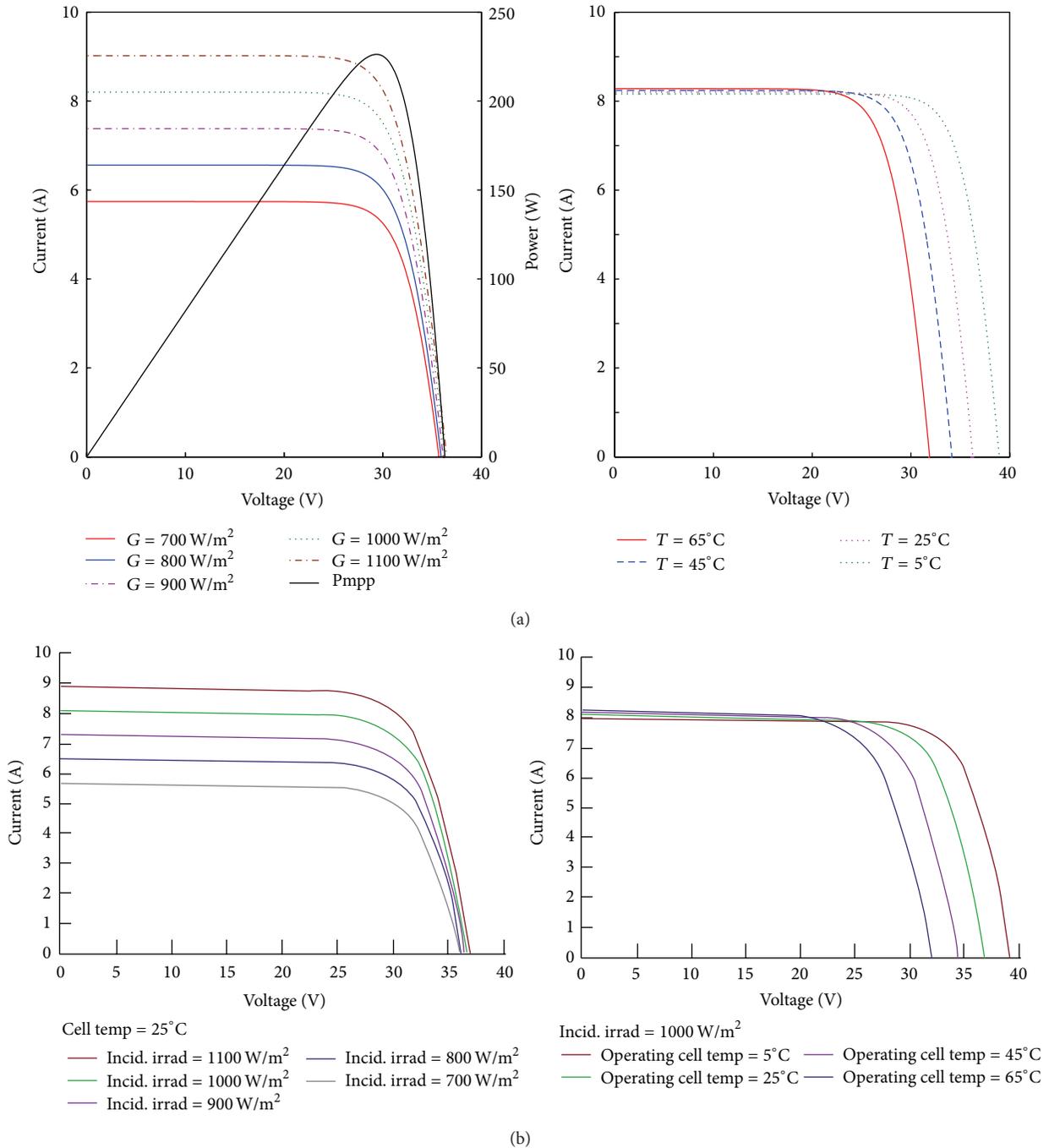


FIGURE 7: Comparison of the PV module *I-V* curves (a) simulated Hyundai M224SG and (b) manufacturer data.

interfaced with the DC/DC buck-boost converter that can step up the voltage to 500 V during boost mode. The battery bank is considered suitable for intermittent discharge to compensate for DC bus voltage discrepancies [28].

**6.2. Simulation Results.** The DC bus voltage regulation performance was evaluated by considering with and without the use of BES. The simulation parameters are as shown in Table 3.

**6.2.1. Evaluation of the Voltage-Mode Control Scheme of G-VSC.** Figure 9 shows the dynamic performance of the PV/BES system being subjected to critical a condition of  $\pm 50\%$  step change of irradiation and a voltage sag at the PCC. As shown in Figure 9(a), an abrupt change in irradiation is applied at 2 s and lasted for 0.5 s which causes the array power to drop suddenly about 50%. Similarly, in Figure 9(b), when a voltage sag is applied at 3 s for a duration of 0.5 s, the PCC RMS voltage is found to be 5.5 kV, hence causing a 50% drop from the nominal 11 kV value.

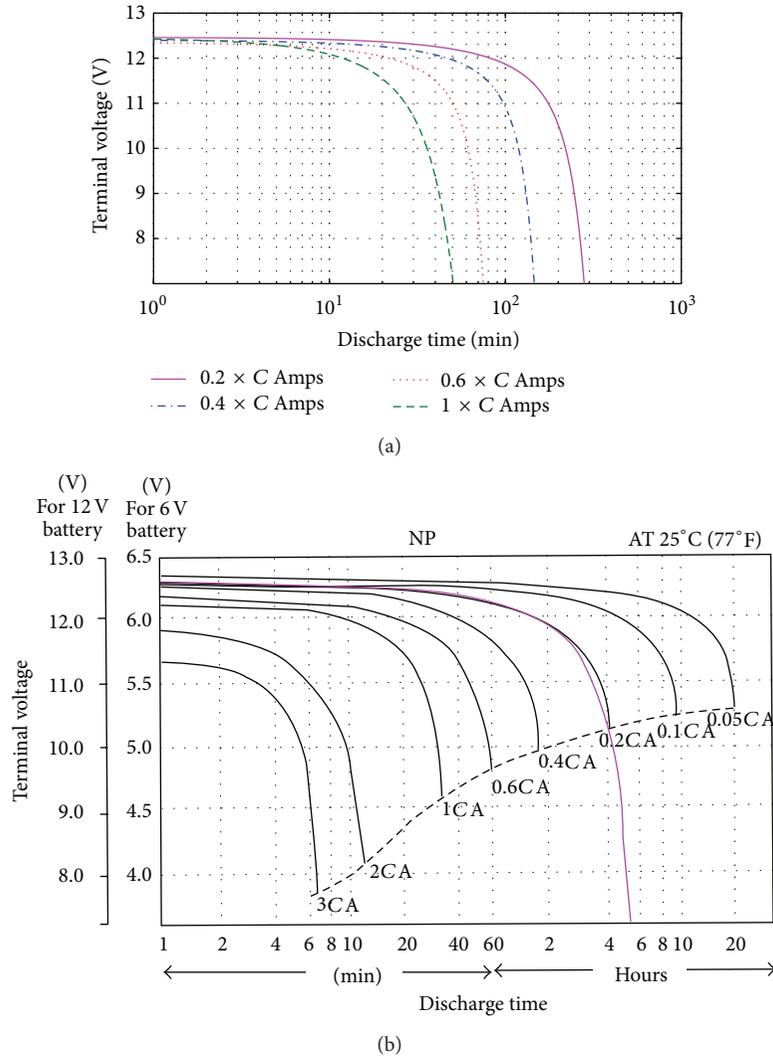


FIGURE 8: Comparison of the battery module discharge curves at different current magnitudes: (a) simulated Yuasa NP4-12 discharge curve and (b) manufacturer data of the Yuasa NP series battery (note that battery module’s capacity  $\ast C = C_{bat} = 4 \text{ Ah}$ ). Note. C: given capacity as stated on each battery in Ah.

Overall, as can be observed from the DC bus voltage profile of Figure 9(c), voltage fluctuation decreases when BES is connected as evident from the variation of transient peak surges and the corresponding transient settling times. Without BES, the transient spikes are measured around 10%–10.8% and 20%–23.2% (sag and spike), during irradiation disturbance and utility grid fault, respectively. However, when BES is connected, the spikes and sag are reduced to 6.6%–7.4% and 13.2%–18% (spikes with sag around 6.4%), respectively. These results imply that the BES compensation for transient overshoot and undershoot at the DC bus voltage results in minimal power fluctuation for the total power delivered to the utility grid (see Figure 9(d)).

Severe disturbance such as utility grid fault prevents the voltage-mode control method for the G-VSC to restore the DC bus voltage to its reference 500 V value, causing subsequent voltage sag for approximately 20% without BES. With BES connected, the voltage regulation is improved

to a certain degree of approximately 6.4% sag because of the compensation through the BES converter. As shown in Figure 9(d), for the reactive power control, BES cannot overcome poor regulation of the voltage-mode control scheme, thus causing fluctuation of the reactive power during both disturbance conditions. This is due to the fact that BES compensation in this case only applies for active power regulation.

During the simulation period of Figure 9, the BES output power and SOC profiles are recorded as shown in Figures 10(a) and 10(b), respectively. The switching states of the DC/DC buck-boost converter during regulation service are also shown in Figure 10(c).

As shown in Figures 10(a) and 10(b), BES discharged and charged its power to compensate for the DC bus voltage fluctuation, thus causing decrease/increase of SOC at 2 s and 2.5 s, respectively. However, more power is discharged during utility grid fault causing SOC to drop about 9%

TABLE 3: Simulation parameters.

Model parameter	Value	Symbols/comments
PV system	200 kWp	$P_{PV}$
BES system	48 kWh	$P_{BES}$
BES terminal voltage	240 V	$V_{bes}$
BES total capacity	200 Ah	$C_{bes}$
System loads	100 kW	$P_L$
PCC system voltage	11 kV	$V_{pccLL, RMS}$
G-VSC AC side voltage	0.23 kV	$V_{iLL, RMS}$
DC bus voltage	500 V	$V_{dc}$
Capacitor	10,000 uF, 50,000 uF	$C_{PV}, C_{dc}$
Inductor	10 mH, 0.12 mH, 0.4 mH	$L_1, L_2, L_{f0}$
Transformer (TR1)	0.1 MVA, 0.1 p.u.	Rating and leakage inductance
PI1, PI2 controller gains (G-VSC voltage mode)	1, 0.005, 0.2, 1	$k_{pd1}, k_{id1}, k_{pq2}, k_{iq2}$
PI3, PI4, PI5, PI6 controller gains (G-VSC current mode)	2, 0.001, 0.6, 0.05, 0.1, 20, 5, 0.05	$k_{pd3}, k_{id3}, k_{pq4}, k_{iq4}, k_{pd5}, k_{id5}, k_{pq6}, k_{iq6}$
PI7, PI8 controller gains (BES DC/DC converter)	1.2, 0.2, 1, 0.5	$k_{pb7}, k_{ib7}, k_{p8}, k_{i8}$
SOC limits	Max 0.95 p.u. Min 0.5 p.u.	<sup>a</sup> SOC <sub>H</sub> , SOC <sub>L</sub>
$V_{dc}$ limits	Max 0.525 Min 0.475	<sup>b</sup> $V_{dc,H}, V_{dc,L}$
Switching frequency	2500 Hz	DC/DC (MPPT/buck-boost) and G-VSC
Simulation time step and running duration	0.05 ms, 4 s	$\Delta t, T_F$

<sup>a</sup>Maximum 45% useable capacity, <sup>b</sup>5%  $V_{dc}$  droop characteristic.

(see Figure 10(b)) in order to restore the resultant 20% drop of voltage of the DC bus in between 3 s and 3.5 s. It is evident that, although BES can slightly improve DC bus voltage regulation in the cases discussed, the voltage-mode control scheme of G-VSC shows poor performance and needs extra energy from the BES in order to enhance the DC bus voltage regulation.

**6.2.2. Evaluation of the Current-Mode Control Scheme of G-VSC.** The performance of PV/BES system using the current-mode control scheme of G-VSC is evaluated based on the similar disturbance conditions as previously discussed. Figure 11 shows the overall system performance for the cases with and without BES.

From Figure 11, it is evident that the closed-loop control response of the current-mode scheme is relatively faster than that of the open-loop voltage-mode control scheme. The DC bus voltage profile as shown in Figure 11(a) is improved even without BES connected to the DC bus with 6.4%–6.8% maximum spikes during irradiation step change and 8.2%–10.2% range of spikes during utility grid fault. No voltage sag is observed during the latter disturbance condition. When BES is connected, the voltage profiles in terms of transient peaks and settling times (Figure 11(a)) are improved with only approximately 4.2%–4.8% and 5.8%–8.2% range of spikes during irradiation disturbance and utility grid fault,

respectively. The results indicate the capability of BES in performing DC bus voltage regulation.

Overall comparison of the voltage-mode and current-mode control performances reveals that regulation by using the current-mode control scheme is superior, particularly when the system is subjected to utility grid faults. A more stable and fast response of system controller can be seen for the current-mode control method between 3 s and 3.5 s simulations as shown in Figures 11(a) and 11(b). It is important to note in Figure 11(b) that no change is observed for reactive power regulation with or without BES because BES is only set to compensate the bus voltage discrepancies using its active power, while G-VSC reactive power control works at unity power factor. However, it is obvious that reactive power regulation through current-mode control scheme of G-VSC is relatively better compared to the voltage-mode control scheme in controlling constant zero reactive power injection to the utility grid.

Figure 12 shows the BES power and SOC profiles with the corresponding switching states of the DC/DC buck-boost converter. It is evident that improvement in system stability ensures optimal charging/discharging through the BES converter, in which minimum power from batteries is consumed/released during compensation, as shown in Figure 12(a). Small changes in SOC (Figure 12(b)) at the instant of disturbances applied indicate that the BES can be resized to a more optimal size which can be much smaller.

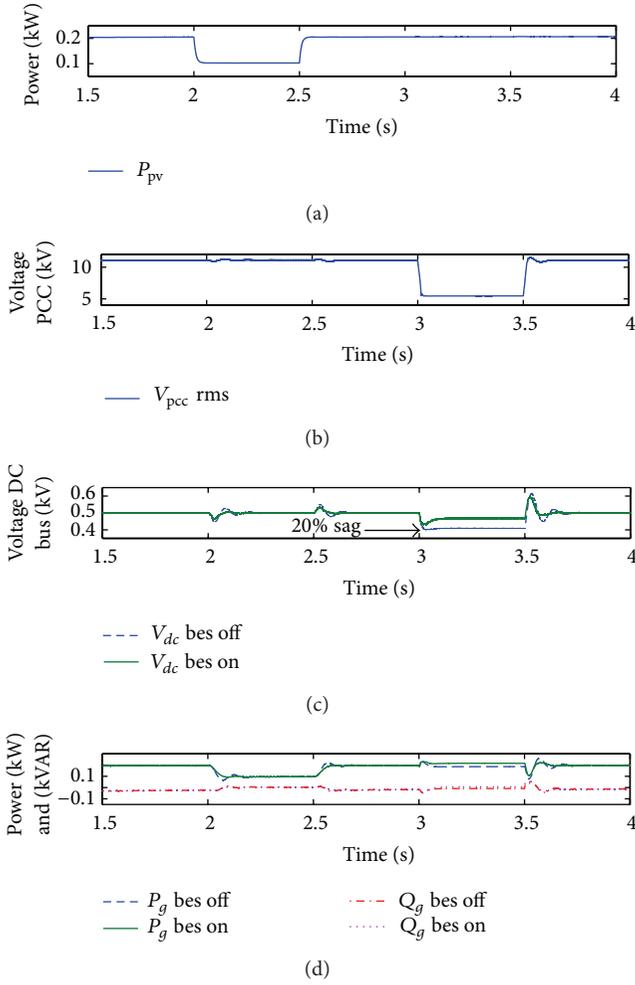


FIGURE 9: Results of DC bus voltage regulation and active and reactive power profiles for the case with and without BES and using voltage-mode controlled G-VSC.

Overall simulation results show that a combination of BES with current-mode control of G-VSC can provide superior DC bus voltage control performance of the grid-connected PV system.

**6.2.3. Results of DC Bus Voltage Regulation Using Simplex Optimized Control Schemes.** To compare the effectiveness of the proposed optimized control scheme with the traditional hand-tuning method of the power converters, several simulation cases were considered for the voltage- and current-mode controlled G-VSC with BES connected to the DC bus. The results of the parametric optimization problem using the voltage-mode and current-mode control schemes are as shown in Figure 13. From the figure, the objective function that is the WISE performance index of the DC bus voltage deviations has been reduced and the parameters become optimal after less than 100 numbers of iterations.

From the simulation results of Figure 13, a summary of the optimized control parameters used for the power converters described in Figure 1 is shown in Table 4.

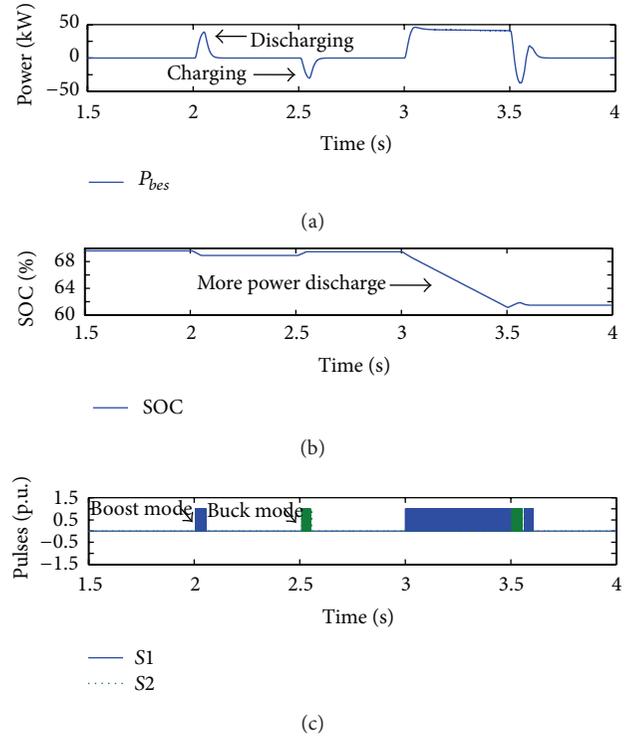


FIGURE 10: BES profiles during voltage regulation with voltage-mode control scheme of G-VSC.

Using the optimized controller gains shown in Table 4, simulations were carried out to compare the DC bus voltage regulation performance of the cases with BES using voltage-mode and current-mode controlled G-VSC. Figure 14 shows the simulation results for the purpose of the comparison. In general, it is evident that, with optimized control schemes, the DC bus voltage can be further regulated at a certain level providing a more stable DC bus voltage over the period of varying disturbances conditions. As shown in Figure 14(a), when the system is subjected to changing atmospheric conditions, the optimized controller of the voltage-mode scheme can fairly improve voltage regulation. The undesired spikes are reduced from the range 6.6%–7.4% to 2.46%–2.7% at 2 s and 2.5 s, respectively. However, during voltage sag at the PCC, the DC bus voltage cannot be restored to its reference value. Again, this can be attributed to poor regulation of the voltage-mode control scheme which was based on the open-loop voltage control strategy [1]. Another reason is that BES operation is subjected to a maximum BES current discharge (e.g., at maximum of  $1 \times C_{bes}$  rate) in order to ensure safe operation so as to obey the current rating of the BES DC/DC converter.

However, with current-mode control scheme as shown in Figure 14(b), the optimized control parameters can significantly increase the regulation performance. The measured voltage spikes over the period of simulation are very marginal. From the figure, the voltage spikes have been reduced from the range 4.2%–4.8% to 1.82%–1.96% during irradiation disturbance and from 5.8%–8.2% to 2.12%–2.96%

TABLE 4: Initial and optimised PI controller gains for the power converters based on simplex optimum run.

Power converter	Compensator	Hand tuning		<sup>a</sup> Optimized	
		$k_p$	$k_i$	$k_p$	$k_i$
G-VSC (voltage mode)	PI1	1	0.005	<b>1.015</b>	<b>0.0008713</b>
	PI2	0.2	1	0.2	1
	PI3	2	0.001	<b>1.957</b>	<b>0.0002511</b>
G-VSC (current mode)	PI4	0.6	0.005	0.6	0.005
	PI5	0.1	20	0.1	20
	PI6	5	0.05	5	0.05
DC/DC buck-boost (BES)	PI7 (voltage-mode)	1.2	0.2	<b>1.147</b>	<b>0.2438</b>
	PI7 (current-mode)	1.2	0.2	<b>1.212</b>	<b>0.2026</b>
	PI8	1	0.5	1	0.5

<sup>a</sup>Minimize WISE performance index with  $W_1 = 0.07, W_2 = 0.008$ , total number of runs = 200.

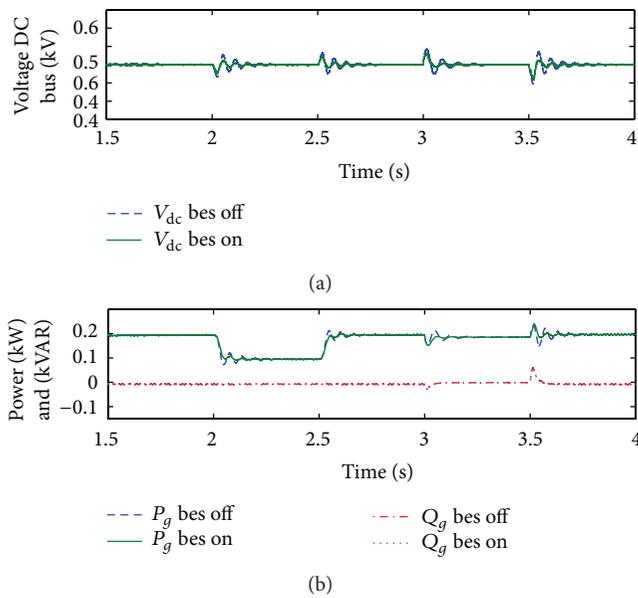


FIGURE 11: Results of DC bus voltage regulation and active and reactive power profiles for the case with and without BES and using current-mode controlled G-VSC.

when voltage sag occurs at the PCC. This can significantly improve the performance of the G-VSC operation since the input voltage DC bus is now more stable and constant, thus contributing to minimizing loss of power within the power converter and a robust control system design.

### 7. Conclusion

This paper presented an assessment of the optimal control for DC bus voltage regulation by using a voltage-sourced converter (VSC) and a battery energy storage (BES) DC/DC buck-boost converter. The voltage-mode control method has a low number of control loops compared to the current-mode control scheme, making it simple in practice. However, the poor regulation performance of the voltage-mode control method results in greater energy requirement of

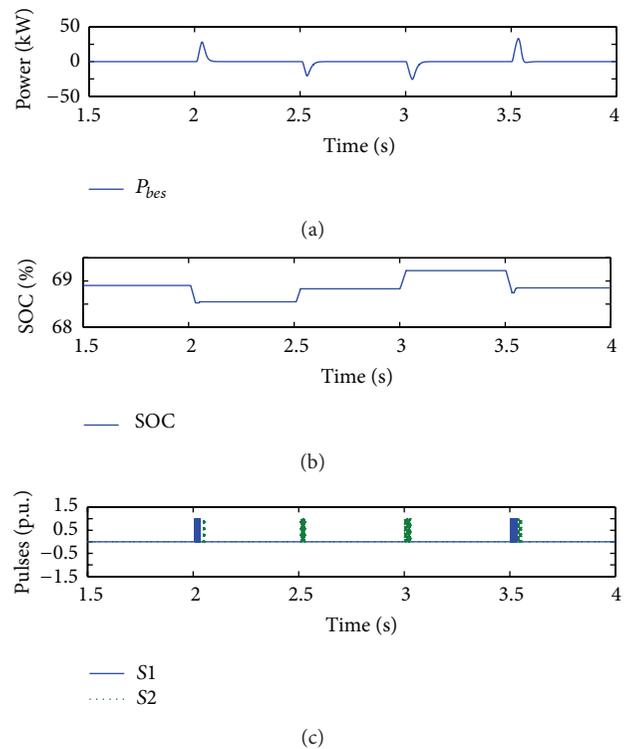


FIGURE 12: BES profiles during voltage regulation service of the system with current-mode control scheme of G-VSC.

batteries for compensation, which consequently increases storage cost. Using the BES compensation and voltage-mode controlled G-VSC, the DC bus voltage fluctuated at approximately 6.6% to 7.4% and 13.2% to 18% during the 50% step changes in the irradiation and utility grid faults (voltage sags), respectively. On the other hand, the overall performance was improved using the current mode control scheme, with fluctuations minimized to the range of 4.2% to 4.8% and 5.8% to 8.2% in the irradiation and utility grid faults (voltage sags), respectively. Moreover, the latter case causes no voltage sag to occur at the DC bus during the utility

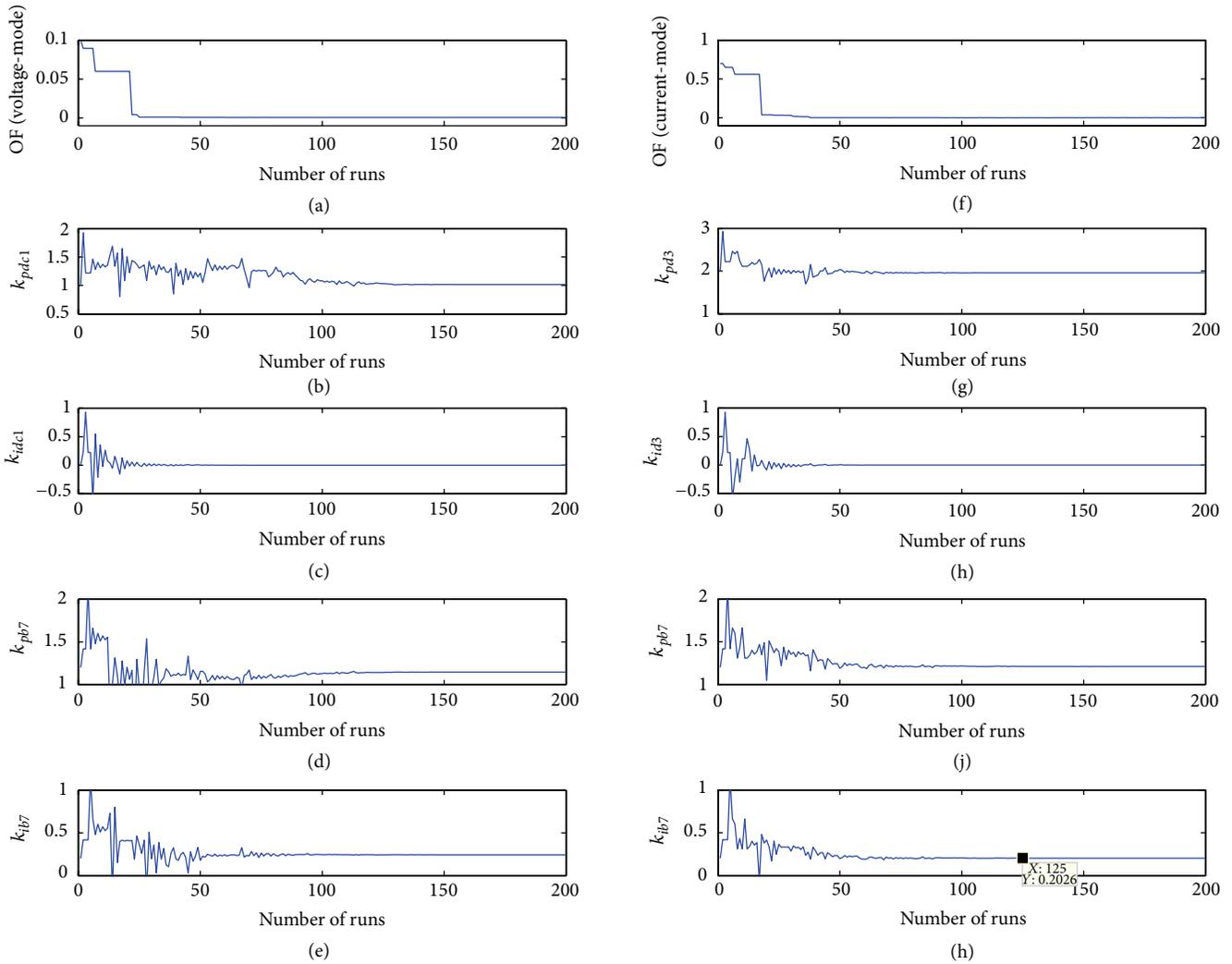


FIGURE 13: Optimization results of the BES control parameters and the parameters of voltage-mode and current-mode control schemes that are responsible for DC bus voltage regulation.

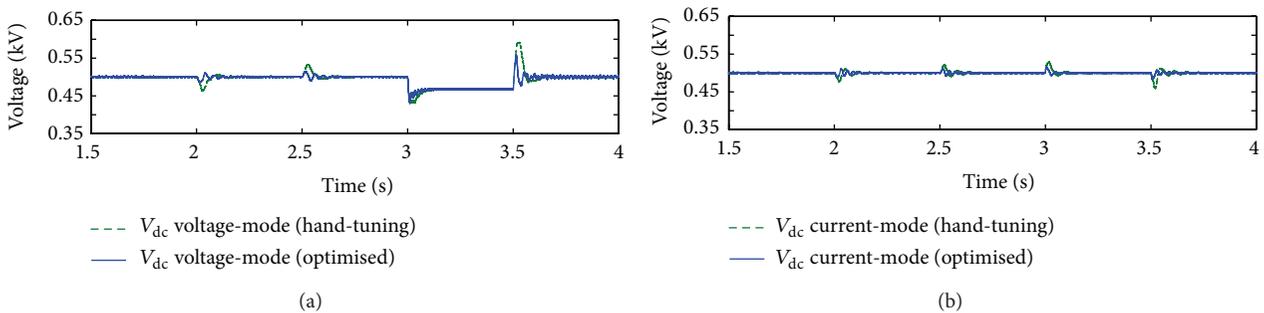


FIGURE 14: Comparison of DC bus voltage profiles for the cases without and with optimized control parameters: (a) voltage mode and (b) current mode.

grid fault. These results can be attributed to the capability DC-coupled BES to restore DC bus voltage to 1 p.u. quickly. Furthermore, using the optimized current-mode control scheme, the DC bus voltage overshoot and undershoot only fluctuated in approximately 1.82% to 2.96% range during

all disturbance cases. Overall, the results show that the proposed optimization method for PI control parameters can contribute to the optimal control of the power converters of the PV/BES system and improve the overall system efficiency.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work is supported by Universiti Kebangsaan Malaysia under the research Grants ERGS/1/2012/TK02/UKM01/3 and DIP-2012-30. Scholarships provided by Universiti Malaysia Terengganu and the Ministry of Education, Malaysia, are also gratefully acknowledged.

## References

- [1] A. Yazdani and R. Iravani, *Voltage-Sourced Converters in Power Systems*, John Wiley & Sons, Hoboken, NJ, USA, 2010.
- [2] N. A. Ahmed, A. K. Al-Othman, and M. R. Alrashidi, "Development of an efficient utility interactive combined wind/photovoltaic/fuel cell power system with MPPT and DC bus voltage regulation," *Electric Power Systems Research*, vol. 81, no. 5, pp. 1096–1106, 2011.
- [3] A. Hajizadeh and M. A. Golkar, "Control of hybrid fuel cell/energy storage distributed generation system against voltage sag," *International Journal of Electrical Power and Energy Systems*, vol. 32, no. 5, pp. 488–497, 2010.
- [4] M. A. Tankari, M. B. Camara, B. Dakyo, and C. Nichita, "Ultra-capacitors and batteries integration for power fluctuations mitigation in wind/PV/diesel hybrid system," *International Journal of Renewable Energy Research*, vol. 1, pp. 86–95, 2011.
- [5] M. Zalani Daud, A. Mohamed, and M. A. Hannan, "A review of the integration of energy storage systems for utility grid support," *Przegląd Elektrotechniczny*, vol. 10, pp. 185–191, 2012.
- [6] X. Tan, Q. Li, and H. Wang, "Advances and trends of energy storage technology in Microgrid," *International Journal of Electrical Power & Energy Systems*, vol. 44, pp. 179–191, 2013.
- [7] K. J. Astrom and T. Hagglund, *PID Controllers: Theory, Design and Tuning*, The Instrumentation, Systems, and Automation Society, Philadelphia, Pa, USA, 1995.
- [8] L. Wang and N. Ertugrul, "Selection of PI compensator parameters for VSC-HVDC system using decoupled control strategy," in *Proceedings of the 20th Australasian Universities Power Engineering Conference*, pp. 1–7, Christchurch, New Zealand, December 2010.
- [9] D. C. Das, A. K. Roy, and N. Sinha, "GA-based frequency controller for solar thermal-diesel-wind hybrid energy generation/energy storage system," *Electrical Power and Energy Systems*, vol. 43, pp. 262–279, 2012.
- [10] F. D. Bianchi, A. Egea-Alvarez, A. Junyent-Ferré, and O. Gomis-Bellmunt, "Optimal control of voltage source converters under power system faults," *Control Engineering Practice*, vol. 20, no. 5, pp. 539–546, 2012.
- [11] A. Luo, Z. Shuai, W. Zhu, and Z. J. Shen, "Combined system for harmonic suppression and reactive power compensation," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 2, pp. 418–428, 2009.
- [12] A. M. Vural and K. C. Bayindir, "Optimization of parameter set for STATCOM control system," in *Proceedings of the IEEE PES Transmission and Distribution Conference and Exposition: Smart Solutions for a Changing World*, pp. 1–6, Orleans, La, USA, April 2010.
- [13] D. R. Northcott, S. Filizadeh, and A. R. Chevretils, "Design of a bidirectional buck-boost dc/dc converter for a series hybrid electric vehicle using PSCAD/EMTDC," in *Proceedings of the 5th IEEE Vehicle Power and Propulsion Conference (VPPC '09)*, pp. 1561–1566, Dearborn, Mich, USA, September 2009.
- [14] S. Filizadeh, A. R. Chevretils, and D. R. Northcott, "Analysis and design of vehicular power systems using PSCAD/EMTDC," in *Proceedings of the IEEE Vehicle Power and Propulsion Conference (VPPC '07)*, pp. 463–468, Arlington, Tex, USA, September 2007.
- [15] K. Strunz, "Benchmark systems for network integration of renewable and distributed energy resources," Cigre Task Force C6.04.02.
- [16] C. Muller and R. Jayasinghe, *PSCAD/EMTDC User's Guide*, Manitoba HVDC Research Centre, 2010.
- [17] A. Yazdani, A. R. di Fazio, H. Ghoddami et al., "Modeling guidelines and a benchmark for power system simulation studies of three-phase single-stage photovoltaic systems," *IEEE Transactions on Power Delivery*, vol. 26, no. 2, pp. 1247–1264, 2011.
- [18] R. C. Neville, *Energy Conversion: The Solar Cell*, Elsevier, New York, NY, USA, 1995.
- [19] Hyundai Solar, "Hyundai SG-series solar module," 2011.
- [20] C.-J. Zhan, X. G. Wu, S. Kromlidis et al., "Two electrical models of the lead-acid battery used in a dynamic voltage restorer," *IEEE Proceedings Generation, Transmission and Distribution*, vol. 150, no. 2, pp. 175–182, 2003.
- [21] O. Tremblay, L.-A. Dessaint, and A.-I. Dekkiche, "A generic battery model for the dynamic simulation of hybrid electric vehicles," in *Proceedings of the IEEE Vehicle Power and Propulsion Conference (VPPC '07)*, pp. 284–289, Arlington, Tex, USA, September 2007.
- [22] J. B. Copetti, F. Lorenzo, and F. Chenlo, "A general battery model for PV system simulation," *Progress on Photovoltaics: Research and Application*, vol. 1, pp. 283–292, 1993.
- [23] S. Barsali and M. Ceraolo, "Dynamical models of lead-acid batteries: implementation issues," *IEEE Transactions on Energy Conversion*, vol. 17, no. 1, pp. 16–23, 2002.
- [24] M. Ceraolo, "New dynamical models of lead-acid batteries," *IEEE Transactions on Power Systems*, vol. 15, no. 4, pp. 1184–1190, 2000.
- [25] Yuasa Battery, "Yuasa NP valve regulated lead-acid battery manual," 1999.
- [26] Y. Liu, J. Bebic, B. Kroposki, J. de Bedout, and W. Ren, "Distribution system voltage performance analysis for high-penetration PV," in *Proceedings of the IEEE Energy 2030 Conference*, pp. 1–8, Atlanta, Ga, USA, November 2008.
- [27] S. Peng, G. Shi, Y. Cao, and X. Cai, "Voltage and frequency control of an islanded power system based on wind-battery generation," *International Review of Electrical Engineering*, vol. 6, no. 4, pp. 2002–2012, 2011.
- [28] M. Zalani Daud, A. Mohamed, M. Z. C. Z. Wanik, and M. A. Hannan, "Performance evaluation of grid-connected photovoltaic system with battery energy storage," in *Proceedings of the IEEE Conference on Power and Energy (PECON '12)*, pp. 396–401, 2012.
- [29] J. A. Nelder and R. A. Mead, "A simplex method for function minimization," *Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965.

- [30] H. M. John and D. F. Kurtis, *Numerical Methods Using Matlab*, Prentice Hall, New York, NY, USA, 2004.
- [31] A. M. Gole, S. Filizadeh, R. W. Menzies, and P. L. Wilson, "Electromagnetic transients simulation as an objective function evaluator for optimization of power system performance," in *Proceedings of the International Conference on Power System Transients (IPST '03)*, pp. 1-6, 2003.
- [32] A. M. Gole, S. Filizadeh, and P. L. Wilson, "Inclusion of robustness into design using optimization-enabled transient simulation," *IEEE Transactions on Power Delivery*, vol. 20, no. 3, pp. 1991-1997, 2005.

## Research Article

# Risk Intelligence: Making Profit from Uncertainty in Data Processing System

Si Zheng, Xiangke Liao, and Xiaodong Liu

*School of Computer, National University of Defense Technology, China*

Correspondence should be addressed to Si Zheng; [si.zheng1009@gmail.com](mailto:si.zheng1009@gmail.com)

Received 27 February 2014; Accepted 19 March 2014; Published 24 April 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Si Zheng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In extreme scale data processing systems, fault tolerance is an essential and indispensable part. Proactive fault tolerance scheme (such as the speculative execution in MapReduce framework) is introduced to dramatically improve the response time of job executions when the failure becomes a norm rather than an exception. Efficient proactive fault tolerance schemes require precise knowledge on the task executions, which has been an open challenge for decades. To well address the issue, in this paper we design and implement RiskI, a profile-based prediction algorithm in conjunction with a riskaware task assignment algorithm, to accelerate task executions, taking the uncertainty nature of tasks into account. Our design demonstrates that the nature uncertainty brings not only great challenges, but also new opportunities. With a careful design, we can benefit from such uncertainties. We implement the idea in Hadoop 0.21.0 systems and the experimental results show that, compared with the traditional LATE algorithm, the response time can be improved by 46% with the same system throughput.

## 1. Introduction

Nowadays we are witnessing extreme scale data processing systems that employ thousands of servers to coordinately accomplish jobs [1]. In systems of such a large scale, hardware failure becomes a norm rather than an exception and the fault tolerance becomes an indispensable part of the system scheduling [2]. For example, in an early report in March 2006, Google reported that at least one server would fail every day, and each job would experience five failures. Besides, a single failure on job could cause completion time to increase by up to 50% [3, 4]. When the system scales, the problem becomes even more severe.

In traditional approaches, the fault tolerance is provided in a reactive manner. Failed jobs are reexecuted until they succeed. As the systems scale up, this simple reactive scheduling scheme becomes inefficient because the failures will become increasingly more frequent. In the extreme case, job may be executed forever due to continuous failures.

To well address this issue, a more promising approach is the proactive fault tolerance scheme. For example, in the MapReduce computational model [5, 6], the speculative execution is employed. Speculative execution will monitor

the execution status of the subjobs (called tasks) and predict the execution failures. Once a task is obviously slower than the others (called straggler [6] or outlier [7]), a backup task will be executed to ensure the successful execution of the task. Similar proactive mechanisms are employed by other large scale data processing systems such as Mesos [8], Dryad [9], and Spark [10].

Speculative execution can dramatically reduce the execution time of tasks and bring great advantages on system response time. It is, however, on the cost of system throughput. To allow speculative executions, certain amount of the system resources will be allocated to execute backup tasks rather than new tasks, demoting the system throughput that depends on the allocated resources. In other words, there is a fundamental tradeoff between the overall system throughput and the response time of individual tasks. A well designed fault tolerance scheme should strike the best tradeoff between the two objectives.

Efficient fault tolerance schemes raise great challenges to the system designers. First, the optimal scheduling requires the precise knowledge on the execution time for each task [5, 6], which is very difficult, if not impossible in practice [11]. Robust scheduling algorithms can tolerate the inaccurate

```

Input: Task  $x$ , Node  $a$ ;
Output: Prediction time  $[t_{\min}, t_{\max}]$ ; weighting scheme  $W_k$ ;
(1)  $d \leftarrow 0$ , Task  $z$ ;
(2) for each Task  $y$  in database profile do
(3)   compute  $d(x, y)$ ;
(4)   if  $d(x, y) > d$  then
(5)      $d \leftarrow d(x, y)$ ;
(6)      $z \leftarrow y$ ;
(7)   end if
(8) end for
(9)  $T(z) \leftarrow [t_{\min}, t_{\max}]$ 
(10) Running Task  $x$ ;
(11)  $T(x) \leftarrow t$ ;
(12) compute  $R_{ij}$ ;
(13) while not the minimum  $\sum_{i=1}^n \sum_{j=i+1}^n (L_{ij}^2 + M_{ij}^2)$  do
(14)   adjust weighting scheme  $W_k = (w_i, w_j, w_k, w_l, \dots)$ ;
(15) end while

```

ALGORITHM 1: Algorithm for task execution time prediction.

execution time information, but on the cost of the scheduling optimality. The widely adopted virtualization technology (e.g., [12]) makes the problem even more challenging.

Second, our later empirical experiments (in Section 3) will demonstrate that the task execution time in such a large scale is not a precise time but naturally contains uncertainties. It can only be confined to a range of time even when the same task is executed in the same environment.

Third, assuming the execution time is available for every task, a slower task is not necessary the straggler because of the different launch time of tasks [5]. The very nature of the MapReduce environment makes trouble in finding correct stragglers. The successful identification of stragglers requires much more design intelligence [6].

To well address these challenges and provide an efficient fault tolerance scheme, in this paper we propose a novel approach that exploits the very nature of the MapReduce-like large scale data processing systems. Our approach, called RiskI, is inspired by two observations. First, in such extreme scale data processing systems, jobs are periodically executed and thus we can collect the history information of a task to predict its future execution time. Second, the task execution time is naturally uncertain, and thus we can explore the risk management theory to assign tasks. For example, a faster node in average is not necessary to be a better node unless it can reduce the execution time for the whole job. To summarize, the main contributions of this paper are as follows.

First, we collect real traces from a MapReduce production system and find the uncertain nature of task execution time in such systems. We conduct comprehensive experiments to investigate the impact factors of the task uncertainties.

Second, we define a similarity measure to quantify the difference between executed tasks in system, based on which we design a task execution time prediction algorithm (Algorithm 1). Note that the output of the prediction is a range of time with probability distribution rather than a precise time.

Third, being aware of the task execution uncertainty, we design a risk management based task assignment algorithm and implement it on top of a Hadoop 0.21.0 production system. We conduct comprehensive experiments to evaluate the performance. Compared with the state-of-art scheduling algorithm LATE [6], the job response time can be improved by up to 46% with the same system throughput and compared with the native MapReduce scheduling with no fault tolerance capability, the performance degradation of our scheduling is neglected small.

The remainder of the paper is organized as follows. In Section 2, we give a review of the related works on performance prediction and task scheduling in distributed systems. In Section 3, we first introduce the background information of modern extreme scale data processing system and motivation of this work. In Section 4, we present our empirical study results revealing the very nature of the task execution uncertainty. We present our design on task execution time prediction and the risk management based scheduling in Sections 5 and 6, respectively. We will present the performance evaluation in Section 7 and draw a conclusion with the future work directions in the last section.

## 2. Related Work

Fault tolerance and task assignments have been widely studied, especially in traditional distributed system [13]. Research topics mainly include performance modeling and prediction [11], task assignment [14], heuristic algorithm [15], resource allocation [8], and fairness [16]. In our work, we mainly focus on two aspects, namely, the performance prediction and task assignment.

Performance modeling is a conventional way of predicting the execution time of tasks in advance of executions [17]. The design philosophy is to have a comprehensive understanding on machine capabilities and application features so that the execution behavior of the machines can be well characterized. Snively et al. [11] proposed a framework for

performance modeling and prediction on large HPC system. Their work focused on characterizing machine profiles and application signatures. Prediction is made by a proper convolution method. Modeling-based prediction works well under the condition of comprehensive understanding the machines and applications, complicated convolution method. These factors above may have the effect on prediction accuracy. Some works employed queuing network models [18–21] to represent and analyze resource sharing systems. The model is a collection of interacting services centers representing system resources and a set of customers representing the users sharing the resources. Another model is a simulation model [22] and it is actually the most flexible and general analysis technique. The main drawback is its development and execution cost. Petri nets can also be used to answer performance-related questions since they can verify the correctness of synchronization between various activities of concurrent systems [20, 23, 24].

Our work differs significantly from this literature as we make our prediction based on historical execution record. Because of considering a heterogeneous environment in MapReduce, getting such information is much costly and difficult. First, collecting and analyzing sufficient characterizations of the ability of the machine are an inefficient work. It needs to run low-level benchmarks for gathering performance attributes. It may be applicable for HPC system where the environment is homogeneous, but it is not practical in practical heterogeneous environments. Second, virtualization in computer makes some resource sharing invisible. Third, capturing and analyzing application features will introduce extra overhead. So, it is rarely possible to employ modeling-based method to predict execution time of task in MapReduce anymore.

Task assignment problem is also a very active research area, especially in traditional distributed system [25–28]. Much of works leverage precise time prediction for scheduling decision making [29, 30]. Other works concentrate on the target of load balance [31, 32]. However, in heterogeneous environment, the blind pursuit of load balance does not bring better performance any more.

Real-time scheduling [29] is relative work. Most works employ the worst time for scheduling to meet the requirement of real time and avoid the loss of uncertainty. Our work is also related to coscheduling [30] in multicore system, particularly with processor heterogeneity. Jiang et al. proposed a scheduling police based on prior knowledge that accurate threads execution times were known to the system. With this advanced knowledge, they can make better scheduling decisions. In contrast, our scheduling has no such precise and global information. Instead of making scheduling decisions based on precise time estimation, we take into account the fact that our prediction is a range of time.

Task assignment problem in MapReduce is still in the beginning since MapReduce adopted a simple scheduling police which is easy to implement. The master will assign a task to a tasktracker without considering its ability and possible execution time. It is only to lunch the task as early as possible. Run-time stragglers have been identified by past work. For the sake of reducing the response time of

job, LATE [6] schedules a duplicate task for the straggler, which is named as speculative execution. However, since no distinction is made between tasktrackers, LATE cannot guarantee the speculative execution to be successful. So the effect of speculative execution is by chance totally. Delay scheduling [33] is another closed work we know of to our own. They begin to distinguish the tasktrackers with data locality. It schedules task as far as possible to the tasktracker who meets the need of data locally. Mantri [7] uses the progress rate instead of the task duration to remain agnostic to skews in work assignment among tasks.

Our work is different from above as we recognized the distinction of performance between working nodes. First, we investigated how to estimate the response time for a task on different working nodes. Instead of modeling-based method, we make our prediction through learning from experience. Second, since our prediction is a range of time which combines with confidence interval, we employed the method from risk of decision making to help us make scheduling decisions.

### 3. Motivation

In this section, we provide some background information of the MapReduce-like computing environments as the background information. We first describe the working mechanism of the MapReduce and elaborate how the speculative execution mechanism works. With simple examples, we point out the limitations in existing systems.

*3.1. MapReduce Background.* MapReduce is a typical representative of the modern parallel computing environment that provides extreme scale support. Hadoop is an open source implementation of MapReduce. Besides MapReduce and Hadoop, there are many other similar MapReduce-like computing frameworks such as the Mesos [8], Dryad [9], and Spark [10]. All these frameworks share a common idea.

Roughly speaking, a computational system contains a master node and a number of working nodes. The master node splits the job to a number of nearly independent tasks and assigns them to working nodes. As these tasks are nearly independent, the parallelism can be fully exploited. As illustrated in Figure 1, a computation job mainly has three stages: the Map tasks for processing the raw data, the shuffle stage is to transfer the intermediate value, and the Reduce tasks that come out with the final results. Each working node has a number of slots for running the tasks. When a task finishes, the slot will be freed for the other task usage.

In extreme scale data processing systems, the traditional reactive fault tolerance scheme becomes ineffective because of the large number of failures and the long response. And thus, MapReduce framework adopts a proactive fault tolerance scheme. More specifically, when a task in execution is much slower than other tasks in the same job, speculative execution will be invoked to cope with this slowest task (called straggler [6]). In fact, because there is little dependence between tasks, the response time of a job is only determined by its straggler. To reduce the job response time and provide efficient proactive fault tolerance, we only need

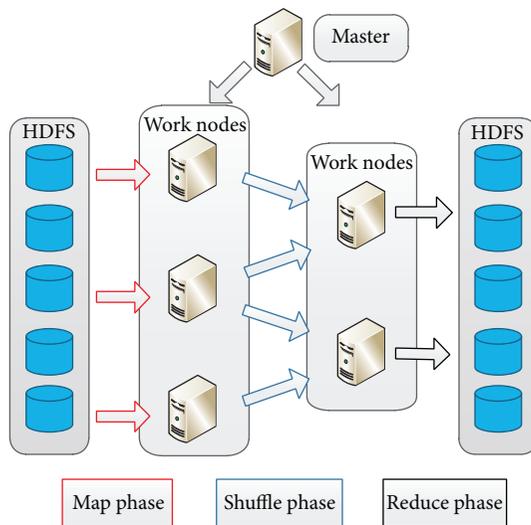


FIGURE 1: MapReduce computation framework that the job is split to two kinds of tasks, Map and Reduce. Different tasks are nearly independent so that different tasks can be executed in a parallel manner.

to accelerate its straggler while not delaying the other tasks too much so that new stragglers appear.

Speculative execution is based on a simple principle. When the straggler is identified, a backup task will be executed for the straggler, hoping the backup can be finished earlier than the straggler. As long as one wins, that is, either the original straggler or the backup one finishes, the other will be killed. It is essential to trade the system resources (and thus the overall throughput) to fuel the executions of individual jobs.

**3.2. Challenges.** In practice, efficient speculative execution faces many challenges. A crucial fact is that the backup task is not necessary to be the winner (finishes earlier). Speculative execution may fail, meaning that though the backup is invoked, the original straggler finishes earlier than the backup. When this happens, we in effect waste the resource allocated to backup task but make no benefit on response time. In other words, not all the speculative executions are profitable. We thus argue that the efficiency of the speculative execution highly depends on the winning rate of the backup tasks, defined as the percentage of the backup task that is finished earlier than the straggler. The main objective of this paper then becomes to increase the winning rate of backup tasks.

This is a very challenging issue in practice. First, stragglers can only be identified when they have been executed for a while, and thus the backup task starts always at later time. The backup must be sufficiently faster than the straggler to be the winner. This requires the accurate knowledge on the execution time of both stragglers and backup tasks. In literature, however, accurate prediction for task executions has been an open challenge for decades in traditional parallel systems. It is partially because, in traditional approaches [11], it is assumed that we have the comprehensive understanding

of the machine capability and the job workload. This requires a great deal of measurement efforts, which are prohibitively high in a highly dynamic and heterogeneous environment as we are in. The application of virtualization makes the problem even challenging as resources become virtualized in an invisible manner. The real capability of each machine (virtual machine or virtual storage) is hard to measure.

Second, stragglers are difficult to identify. We are facing the dilemma that, on one hand, we desire an early identification of the stragglers so that the backup tasks have a larger improvement space and their winning becomes easier. On the other hand, later identifications of stragglers are more likely to be accurate as the stragglers may change during the execution.

Third, optimal speculative execution requires the precise execution time of tasks; for example, a task will be finished in 15 seconds. Unfortunately, uncertainty is a nature in MapReduce. The execution time of a task is within a range of time even under the same computing environments. Such an uncertainty is a double-blade sword. On the one hand, there will be no optimal assignment schemes in prior execution and thus to find the optimal one in advance is impossible. On the other hand, recall that the straggler determines the response time of jobs, and thus there will be multiple assignment schemes that perform exactly the same as the optimal assignment. In other words, even when our assignment scheme is not the optimal one, the performance may not be degraded. This property greatly eases the design challenges. In later sections, we will show how to benefit from this by applying the theory of decision making under risk and uncertainty.

## 4. Observation

In this section, we study the fundamental characteristics of the task execution in MapReduce environments. We first use experimental results to reveal the uncertainty nature of task executions, and then demonstrate the repeatable execution property of tasks. In the last, we will show that by a simple similarity-based scheme the degree of execution time uncertainty can be dramatically reduced.

**4.1. Uncertainty of Task Executions.** In this part, we use experimental results to reveal the uncertainty nature of the MapReduce. We configure a standard Hadoop system version 0.21.0 (Hadoop is an open sourced implementation of MapReduce) with four machines and run a WordCount application [34]. In our experiments, 128 Map tasks and one reduce task are executed and the results are presented in Figure 2.

Figure 2 shows that the execution time ranges from 39 s to 159 s, with the average of 80 s. The standard deviation (Std) is 27.42 s. Note that the MapReduce system will equally split the Map tasks, and thus the workload of these 128 map tasks is nearly identical. We thus argue that uncertainty is a nature in MapReduce. On the other hand, even if identical tasks were executed in the identical machines or the same task repeatedly running under the same environment, the execution time is not identical but stable. Figure 4(a)

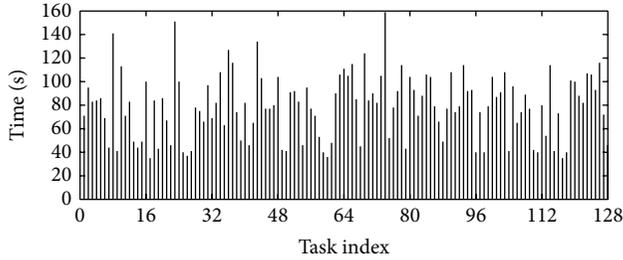


FIGURE 2: Execution time of tasks. There are 128 map tasks in total and each execution time of task was plotted.

shows the CDF of response time of the task which was repeatedly running for 20 times in the same condition (we choose two tasks from two different jobs). We can figure that the execution time is also uncertain. By this nature, the traditional model-based task execution prediction algorithm is costly and will fail as the implicit assumption of this algorithm is that a task will behave the same under the same environment. However, we also find that the execution time is stable, which means most time is within a small range of time. So we argue that a reasonable prediction for the task should be a range of time with probability distribution.

**4.2. Repeated Execution of MapReduce Tasks.** In MapReduce environments, most of the jobs are batch for extreme scale data processing. The spirit is that moving the program is less costly than moving the data [5]. Consequently, the same jobs are repeatedly invoked to process different data [35]. For example, Facebook, Yahoo!, and eBay process terabytes of data and event logs per day on their MapReduce clusters for spam detection, business intelligence, and various optimizations. Moreover, such jobs rarely update. An interesting observation here is that the executions of a job in the past can be of great value for future. We can build profiles for each task and refer such profile to predict the job future executions.

Besides the similarity between tasks of different jobs, in some instance, similarity also exists in tasks of the same job. Because the number of tasks may be more than the number of slots, tasks may start and finish in “waves” [35]. Figure 6 shows the progress of the Map and Reduce tasks of a WordCount job with 8 GB data. The  $x$ -axis is the time and the  $y$ -axis is for the 34 map slots and 10 reduce slots. The block size of the file is 64 MB and there are  $8 \text{ GB}/64 \text{ MB} = 128$  input splits. As each split is processed by a different map task, the job consists of 64 map tasks. As seen in the figure, since the number of map tasks is greater than the number of provided map slots, the map stage proceeds in multiple rounds of slot assignment and the reduce stage proceeds in 3 waves. From the figure we can find that the tasks assigned to the same working node have the close execution time. The similarity exists between different waves on each node in map stage.

In Figure 2, we reorganize the execution time of tasks in Figure 3 while, this time, tasks executed in the same machine will be grouped together. Table 1 gives the maximal and minimal average and Std of these tasks in each group. Compared with that in the original configuration, the Std reduces

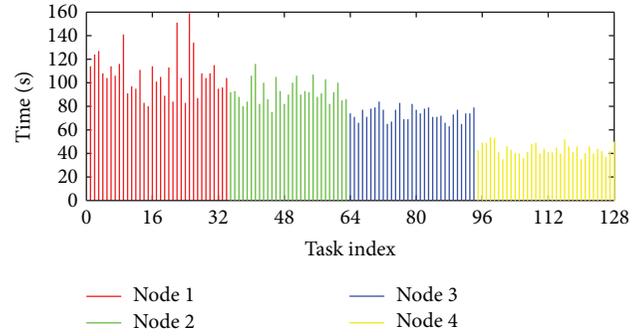


FIGURE 3: Reorganized execution time of tasks. Tasks in the same node were plotted together.

TABLE 1: Execution time statistics for 128 map tasks from job word-count.

	Avg.	Std	Max.	Min.
Total	80	27.42	159	32
Node 1	108.61	18.46	159	80
Node 2	92.66	9.68	116	75
Node 3	73.61	5.61	84	63
Node 4	43.5	4.94	54	32

dramatically, for example, by 33% as from 27.42 to 18.46. This is because the four machines have the different hardware and tasks in different machines will experience the different resource sharing conditions. Nevertheless, tasks in the same machine will experience a more “similar” environment. This shows that a simple approach, with the similarity of tasks taken into account, can dramatically reduce the uncertainty. Next, we will show how to exploit such similarity.

It should be pointed out that the execution environment we mentioned above includes two parts. Except from the hardware configurations from different machines, tasks which are running parallel on the same execution unit (as known as tasktracker in Hadoop) can also affect execution environment. Based on different configuration, one or more tasks will run on the same tasktracker, which will cause resource contention and impact on execution time.

**4.3. Design Overview.** Keeping the uncertainty and tasks similarity nature in mind, in this paper we propose a novel proactive fault tolerance scheme in MapReduce environment. As illustrated in Figure 5, the design, named as RiskI, is constituted of two components: (i) a profile-based execution time prediction scheme and (ii) a risk-aware backup task assignment algorithm. The basic idea is that we build a profile for each executed task in history and refer the profile of a similar task (in terms of its own property and the environment) to predict a new task in execution. As the uncertainty is a nature that cannot be completely avoided, we design a risk-aware assignment algorithm that can benefit from such an uncertainty by exploring the theory of decision making under uncertainty. Our goal is to maximize the assignment profiting and minimize the risk meanwhile.

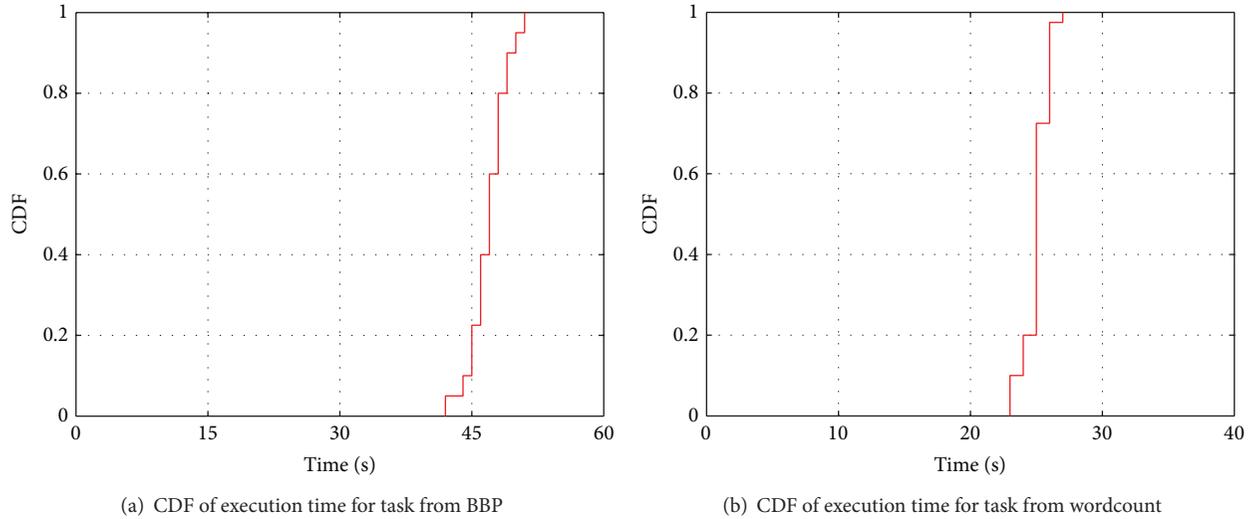


FIGURE 4: Execution time.

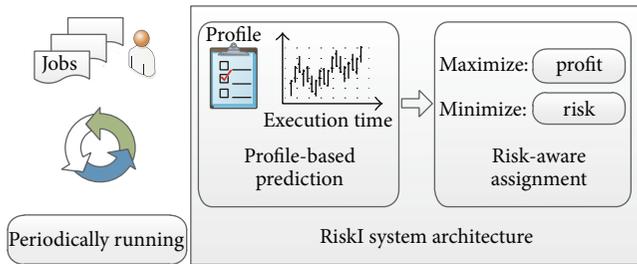


FIGURE 5: RiskI has two main parts, a profile-based execution algorithm for task execution time prediction and a risk-aware backup task assignment algorithm used to make scheduling based on range of time.

### 5. Execution Time Prediction

In this section, we will introduce how to leverage historical information to predict the execution time for tasks. We first present the architecture of the prediction algorithm and then introduce the components, respectively. In the last, we use a pseudocode to present details of the design.

**5.1. Prediction Algorithm Architecture.** The prediction of the job execution is based on a simple idea. For each executed task, we will maintain a profile to record its execution information. When predicting a new task, we refer to these historical profiles and look for the most similar task. Its execution time will then be our prediction result. Notice that due to the uncertainty nature of the job execution, the output of our prediction is also a range of time, indicating the maximum and minimum of the execution time, rather than a single value.

The prediction algorithm mainly consists of four steps, as illustrated in Figure 7. We first make a similarity search and look for the most similar task in history. Tasks will then be assigned and executed according to our predictions. As

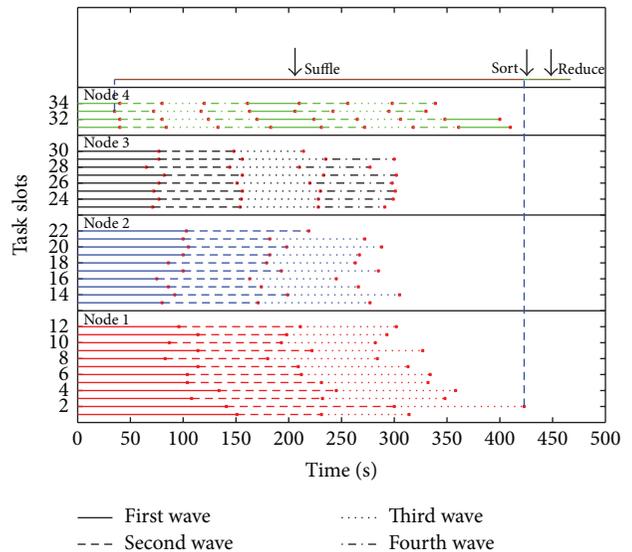


FIGURE 6: “Waves” in execution. Since the number of tasks is more than task slots, tasks were executed in “waves.” In our example, there are 9 waves in node 4 and 3 waves in node 1.

long as the execution is finished, we will adjust some critical control parameters and maintain the profiles for the next tasks. Noticing that there is no absolute the same between tasks, the key in the prediction algorithm design is the similarity of definition for any given two tasks that can yield the best prediction results. Next, we will give details on these steps.

**5.2. Similarity Scheme.** The similarity of two tasks can be affected by many factors such as the number of read bytes and the number of write bytes. In our work we apply all these factors, while our method has the adaptive capability that, for factors of little impact, they will be ignored automatically.

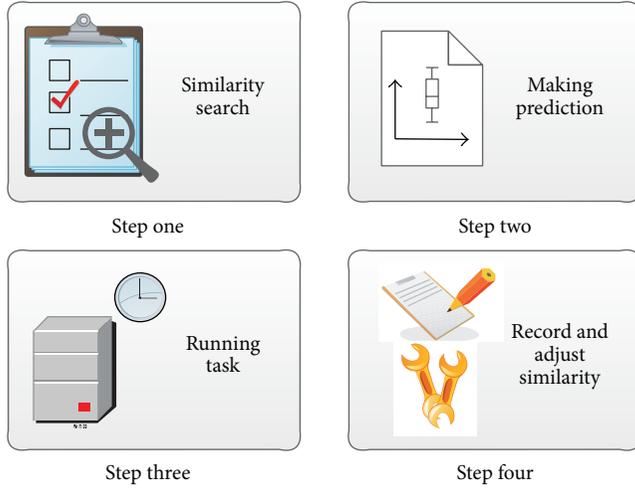


FIGURE 7: Our prediction has four steps. Upon a new task is in prediction, we firstly look for the most similar task in history and then make a prediction based on similar task. Then, we wait for the task running and record the running information and adjust weight setting scheme.

This is done by applying an appropriate similarity scheme, that is, how to obtain the similarity based on these factors.

In literature, there are several approaches to measure the similarity that measures the similarity between tasks in different aspects. For example, the weighted Euclidean distance (WED) [36] measures the actual distance between two vectors, the cosine similarity measures the direction information, and the Jaccard similarity measures the duplication ratio. Considering the requirement of our work, we adopt WED and leave more design options for future work.

*Definition 1.* Given two tasks  $X$  and  $Y$  with their affecting factor vectors  $x_j$  and  $y_j$  and a weight scheme  $w_j$ , the similarity between the two tasks can be calculated as follows:

$$d(X, Y) = \sqrt{\sum_{i=1}^n w_i (x_i - y_i)^2}, \quad (1)$$

where  $n$  is the number of affecting factors and  $w_j$  is a weight scheme that reflects the importance of the different affecting factors on the similarity.

**5.3. Weighting Scheme.** Weighting scheme determines the impact of the affecting factors on the similarity measure. In practice, different machines may have different weighting schemes because of the different hardware. For example, for a CPU-rich machine, the IO may become the bottleneck of the task execution and therefore IO-related factors are more important, while IO-rich machines may desire CPU-related factors. To well address the issue, we apply a quadratic programming method to dynamically set the weighting.

Quadratic programming is a mature method for weight scheme setting in machine learning [15]. Our aim is to minimize the differences between the similarities calculated from our WED and the real differences obtained

TABLE 2: A rough weight scheme for job wordcount.

Factors	Weight
Job name	
Work node	
Input size	34.5%
Environment	
Task 1 (randomwrite)	33.6%
Task 2 (wordcount map)	15.8%
Task 3 (BBP)	14.1%
Others	2%

by comparing the real execution time. Supposing there are  $n$  task records in the node and each has  $q$  affecting factors, the constrains in the problem can be presented in formulas (2) and (3), in which  $S_{ijk}$  is the similarity on the  $k$ th attribute between task  $i$  and task  $j$ ,  $\sum_{k=1}^q S_{ijk} W_k$  is the similarity between task  $i$  and task  $j$  calculated using formula (1), and  $R_{ij}$  is the real similarity between task  $i$  and task  $j$  calculated by real execution time.  $L_{ij}$  is the value by which the calculated similarity is less than the real similarity, and  $M_{ij}$  is the value by which the calculated similarity is greater than the real similarity:

$$\text{Minimize: } \sum_{i=1}^n \sum_{j=i+1}^n (L_{ij}^2 + M_{ij}^2) \quad (2)$$

$$\text{Subject to: } \sum_{k=1}^q S_{ijk} W_k + L_{ij} - M_{ij} = R_{ij} \quad (3)$$

$$(i, j = 1 \dots n, i < j)$$

$$R_{ij} = 1 - \frac{\min\{|t - t_{\max}|, |t_{\min} - t|\}}{\max\{t, t_{\text{avg}}\}}, \quad (4)$$

where  $t$  is the real execution time and the prediction time is  $[t_{\min}, t_{\max}]$ .  $t_{\text{avg}}$  presents the average time of prediction.

As there is no comprehensive analysis of the complexity of quadratic programming problems, our analysis only can indicate how large this quadratic programming problem can be, but not the actual complexity of the problem. From the above formalization, the quadratic programming problem has  $n * (n - 1) + q$  variables and  $n * (n - 1)$  constraints.

The general factors used in calculating similarity include job name, node name, input size, workload, output size, and the parallel task that is running on the same node. We figure a rough weight scheme for job wordcount in one node. We showed it in Table 2. From the table we can see that the factor input size and IO-heavy task like randomwrite influence the execution time most.

## 6. Risk-Aware Assignment

In this section, we introduce our risk-aware assignment algorithm that attempts to benefit from the uncertainty of task executions. We first characterize the unique features and design overview of task assignment with uncertainty

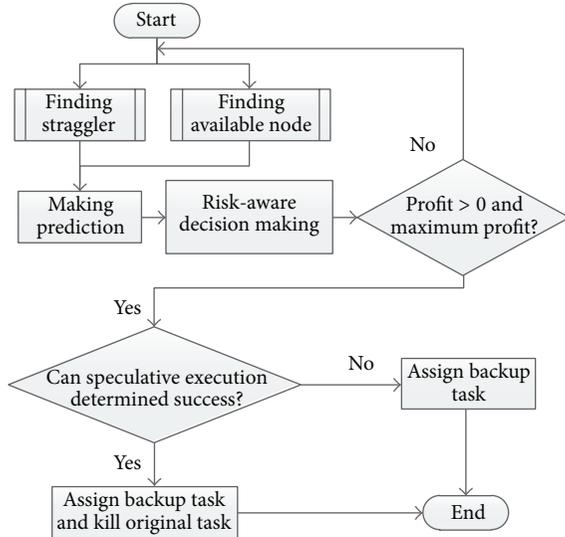


FIGURE 8: Flowchart of risk-aware assignment.

and then present the theory of decision making under uncertainty, which is the theoretical foundation for our risk-aware assignment. At last, we introduce our assignment algorithm based on the theory.

**6.1. Design Overview.** The method is presented as follows: in the first step, by comparing the average execution time of all the running tasks and the passed running time, we find out the straggler; in step two, we check out all the available work nodes and make the execution time prediction; in step three, we make the execution time prediction for the straggler running on the available node, in step four, with the help of the theory of decision making under risk, the most expected profiting assignment was chosen for speculative execution. If the new assignment can guarantee the speculative success, the original task can be killed. Figure 8 is a flowchart of our risk-aware assignment design. The main challenge is how to leverage theory of risk decision to help us assign task.

**6.2. Theory of Risk Decision.** Since our prediction is a range of time, how to compare two ranges of time is another issue to help decision making. A direct method is comparing the average time of two ranges, and the distance of two average values is the time profit or time loss. The average time works well in the condition where time profit is uniformed per unit of time. However, this assumption is broken in MapReduce. As we mentioned in Section 3, because of the existence of straggler, the time profit is only related to the last finish time, but care nothing about the time before last finish time. In other words, may be some tasks would finish faster, it is still helpless for the execution time of the job, because the job would be finished after all the tasks were executed. Under this condition, time profit is different per unit time and the decision making under risk was introduced.

In many fields, risky choice and the selection criterion are what people seek to optimize. There are two parts

included in risk of decision making, risk assessment and risk management. Risk assessment is in charge of evaluating risk or benefit and risk management takes responsibility for taking appropriate actions to control risk and maximize profit. Expected value criterion is also called Bayesian principle. It incorporates the probabilities of the states of nature, computes the expected value under each action, and then picks the action with the largest expected value. We can calculate the potential value of each option with the following equation:

$$E(X) = \sum_x P(x)U(x), \quad (5)$$

where  $E(X)$  stands for the expected value. The probability  $P(x)$  of an event  $x$  is an indication of how likely that event is to happen. In the equation,  $U(x)$  denotes the profit and loss value of  $x$ . The summation range includes every number of  $x$  that is a possible value of the random variable  $X$ .

**6.3. Profit Function and Unique Feature.** First, the function  $P(x)$  is easy to get from historical statistic. As mentioned before,  $P(x)$  is the probability of the occurrence of state  $x$ . In our case,  $P(t_i, t_j)$  represents the probability that the real execution time falls in the area between  $[t_i, t_j]$ . This probability can learn from statistic information in history. For example, the output of prediction is a set  $T = \{t_{\min}, t_i, \dots, t_j, t_{\max}\}$  in which the  $t_i$  was sorted by its value. Since the prediction is a range of time with a probability distribution, we can exploit the probability distribution to compute the  $P(t_i, t_j)$  with the following equation:

$$P(t_i, t_j) = \frac{|[t_i, t_j] \cap T|}{|T|}, \quad (6)$$

where  $|T|$  is the number of elements in the set  $T$  and  $|[t_i, t_j] \cap T|$  is the number of elements falling into range  $[t_i, t_j]$ .

Next, we will take an example to explain our definition for profit function  $U(x)$ . As shown in Figure 9(a), the execution time of task is a nature uncertainty and the assignment scheme should be aware of the potential risks. In Figure 9(b), because of the existence of straggler, the time profit is only related to the last finish time which was denoted as  $t_s$ . Under this condition, time profit is different per unit time and is divided into two parts which are part one  $[t_i, t_s]$  and part two  $[t_s, t_j]$  and we plotted them in Figure 9(c). The profit function  $U(x)$  can be calculated as

$$U(t_i, t_j) = \begin{cases} \text{avg}(t_i, t_s) - t_s & [t_i, t_s] \\ \text{avg}(t_i, t_j) - t_s & [t_s, t_j] \end{cases}, \quad (7)$$

where  $\text{avg}(t_i, t_s)$  is the average time of  $[t_i, t_s]$  and  $\text{avg}(t_i, t_j)$  is the average time of  $[t_i, t_j]$ .

Let  $T$  denote an assignment option. Let  $T_1$  and  $T_2$  denote two ranges of time divided by  $t_s$ . The expected profit can be calculated as

$$E(T) = P(T_1)U(T_1) + P(T_2)U(T_2). \quad (8)$$

```

Input: available Nodes  $N = (n_1, n_2, n_3, \dots)$ ;
(1) Finding the straggler task  $s$  by prediction;
(2) for each available node  $n_i$  in  $N$  do
(3)   Make prediction  $T = [t_{\min}, \dots, t_i, \dots, t_{\max}]$ ;
(4)   Compute profit  $P_i \leftarrow E(T)$ ;
(5)   Record the maximal  $P_i$  and node  $n$ ;
(6) end for
(7) if remaining time  $t_{\text{left}} >$  the worst reexecution time  $t_{\text{max}}$  then
(8)   kill task  $\text{task} \cdot \text{kill}(s)$ ;
(9) end if
(10) assign backup task  $\text{Taskassign}(s, n)$ ;

```

ALGORITHM 2: Algorithm for risk-aware assignment.

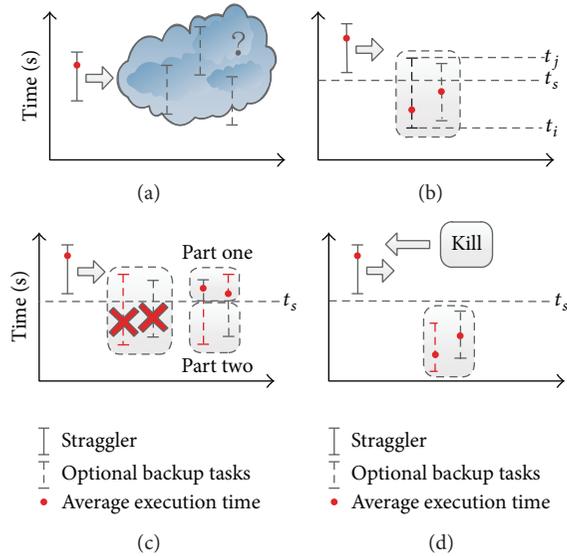


FIGURE 9: An example for different time profit per unit. The time profit is divided into two parts. So we considered separately these two parts.

If  $E(T) > 0$ ,  $E(T)$  would be a better option since it is likely more close to reduce the completion time of a job. If there are more options for backup task assignment, the max  $E(T)$  would be the best choice.

A special case is that  $t_{\text{max}}$  is less than the best time of straggler, as shown in Figure 9(d). In this case, the straggler can be killed immediately to release resources.

The detailed algorithm is shown in Algorithm 2. We first compare the average execution time of all the running tasks and the passed running time and then we find out the straggler. And then we check out all the available work nodes and make the execution time prediction. In step three, we make the execution time prediction for the straggler running on the available node. In step four, with the help of theory of decision making under risk, the most expected profiting assignment was chosen for speculative execution. If the new assignment can guarantee the speculative success, the original task can be killed.

TABLE 3: Our testbed used in evaluation.

VMs	Map slots	Reduce slots	VM configuration
Node 1	0	1	1 core, 2 G RAM
Node 2	4	0	4 core, 2 G RAM
Node 3	4	4	4 core, 2 G RAM
Node 4	4	4	4 core, 2 G RAM
Node 5	1	1	4 core, 2 G RAM
Total	13	10	17 core, 10 G RAM

## 7. Evaluation

We have conducted a series of experiments to evaluate the performance of RiskI and LATE in a variety combination of different job size. Since the environmental variability may result in high variance in the experiment result, we also performed evaluations on different heterogeneous environment configuration. We ran our first experiment on a small private cluster which has 5 working nodes and a master. Our private cluster occupies two whole Dell PowerEdge R710 machines. Each machine has 16 Xeon E5620 2.4 GHz CPUs, 12 GB DDR3 RAM, and  $4 \times 2.5T$  disks. We use Xen virtualization software to manage virtual machines. Table 3 lists the hardware configurations of our testbed. We installed our modified version of Hadoop 0.21.0 to cluster and the block size was configured as 64 MB. Different numbers of task slots were set to each node according to diversity of VM configurations.

**7.1. Performance on Different Job Size.** To identify the predominance of RiskI, we first ran some simple workloads. In this configuration, all jobs were set with one size in a round. We performed evaluation on three different types of job size, which are small, medium, and large, respectively. The small job only contains 2 maps, the medium job has 10 maps, and the large job has 20 maps. We chose Bailey-Borwein-Plouffe (BBP) job for evaluation and there are 20 jobs in total with arrival interval 60 s.

Figure 10(a) shows a CDF of job execution time for large jobs. We see that about 70% of the large jobs are significantly improved under our scheduling. Figure 10(b) illustrates the corresponding running details, including the total number of speculative executions, those succeeded and being killed.

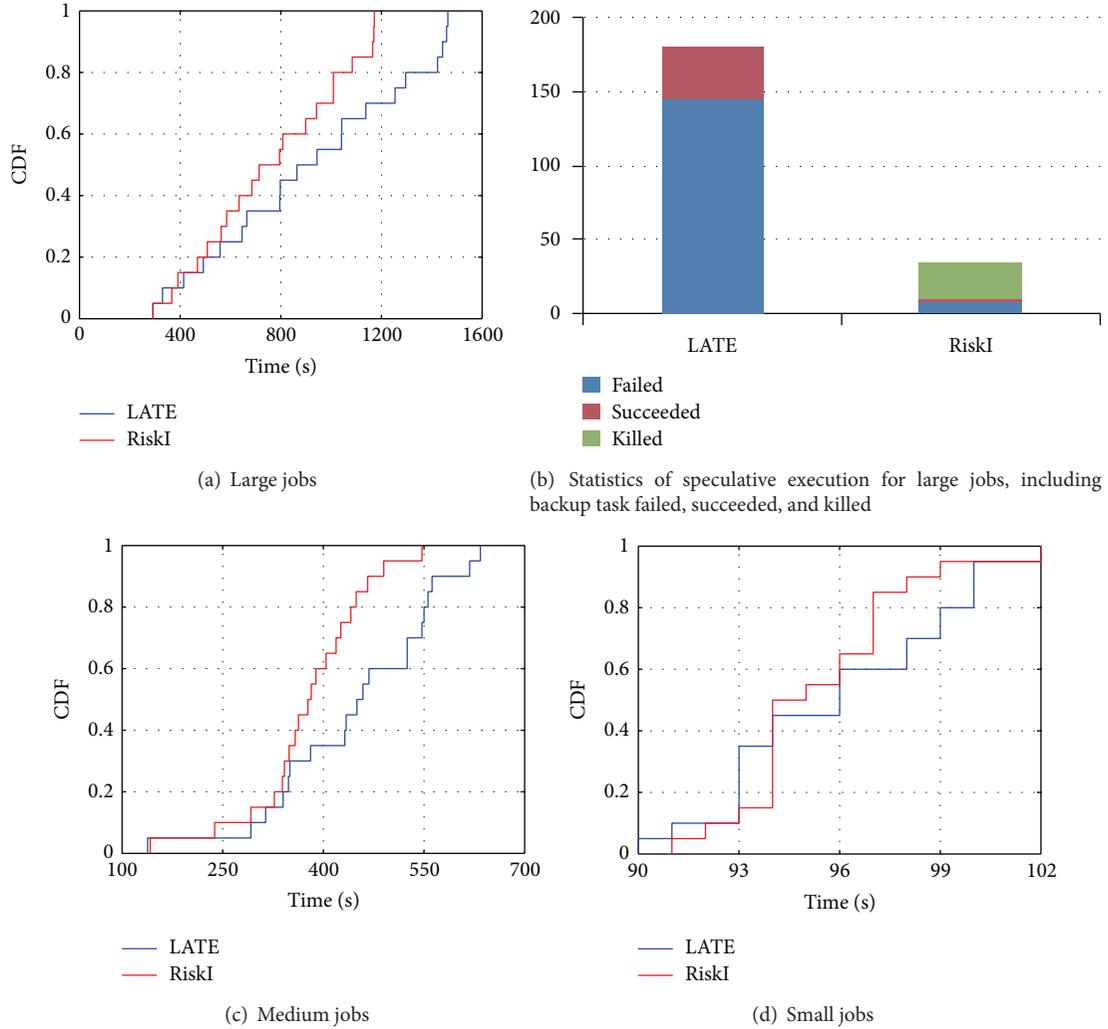


FIGURE 10: CDFs of execution times of jobs in various job sizes. RiskI greatly improves performance for all size of jobs, especially for large and medium jobs.

From Figure 10(b), we can find that RiskI has two main benefits on scheduling. First, with the help of estimation, we are able to cut off some unnecessary speculative executions, thus release more resources. Since most speculative executions were failed in LATE, the total speculative executions were reduced from 181 to 34 in RiskI. Second, the successful ratio of speculative increased. In our scheduler only 6 speculative executions failed out of 34 and the success ratio is 81%.

Figures 10(c) and 10(d) plotted the results on medium job size and small job size. The performance on large job size is distinguished from those on smaller jobs because the small jobs cause less resource contention for low system workload.

**7.2. Performance on Mix Job Size.** In this section, we consider a workload with more complex job combination. We use Gridmix2 [37], a default benchmarks for Hadoop cluster, which can generate an arrival sequence of jobs with the input of a mix of synthetic jobs considering all typical possibility.

In order to simulate the real trace for synthetic workload, we generated job sets that follow the trace from Facebook that

TABLE 4: Distribution of job sizes.

Job size	# Map	# Reduce	# Jobs	% Jobs
Small	2	2	19	38%
Medium	10	5	19	38%
Large	20	10	12	24%
Total	468	253	50	

were used in the evaluation of delay scheduling [33]. Delay scheduling divided the jobs into 9 bins on different job size and assigned different proportion, respectively. For the sake of simplicity, we generalize these job sizes into 3 groups, small, medium, and large, respectively. Jobs in bin 1 take up about 38% in that trace and we assign the same proportion to the small jobs. We grouped bins 2–4 as medium jobs which also hold 38% in total. The rest of bins are grouped into large jobs with about 24% amount. Table 4 summarizes the job size configuration.

1	2	3	4	5	6	7	8	9	10	11	12	13				
M	M	M	M	M	M	M	M	M	M	M	M	M				
14	15	16	17	18	19	20	21	22	23							
S	S	S	S	S	S	S	S	S	S							
24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
M	S	S	L	S	L	S	L	S	L	S	L	S	L	S	L	S
41	42	43	44	45		46	47	48	49	50						
L	L	L	L	L		M	M	M	M	M						

FIGURE 11: The Gridmix job has 50 jobs in total. The first 13 jobs are medium size. The jobs from 14 to 23 are small. The small job and large job are alternate arrived from job 26 to job 40. The large jobs appear consecutive form job 41 to job 45 and the last 5 jobs are medium.

TABLE 5: The configuration details for BBP, including start digit, end digit, and the workload per map task.

Job size	Start Digit	ndigit	Workload/map
Small	1	25000	156283174
Medium	1	55000	151332558
Large	1	70000	122287914

Except for the job size proportion, arrival interval is also very important in generating a job arrival trace, since it may lead to different schedule behaviors and subsequently affect system throughput. In our experiment, we set the arrival interval to 60 s and 120 s. Figure 11 gives the generated job sequences.

Finally, we ran the workload described above under two schedulers: LATE (Hadoop default scheduler FIFO) and RiskI. We submitted each job by only one user.

7.2.1. *Result for BBP.* Our detailed configuration of BBP jobs was shown in Table 5. Figure 12 plotted the execution time of each job with interval 60 s.

For ease of analysis, we divided the job sequence into 5 parts. The first part is from job 1 to job 13, where the jobs are medium. The second part is from job 14 to job 23. The jobs in the second part are all small. Job 24 to job 50 is the third part mainly consists of small and large jobs. The last 10 jobs are equally divided into two parts and the jobs are large and medium.

In the first part M, the response time of jobs is accumulated from the beginning, because the workload is becoming heavier gradually. The resource contention led to more stragglers and more speculative executions. More resources were used for backup task. The following jobs had to wait until resources were released. However, because of lack of delicate design of scheduling, most backup tasks failed. The system sacrificed the throughput but did not get benefit. From Figure 12(b) we can see that there are 48 backup tasks and only 19 succeeded. Compared to LATE, RiskI cut the number of backup task from 48 to 13. The static result was

plotted in Figure 12(c). Many invalid speculative executions are avoided and saved the throughput. Since some backup tasks are certainly finished earlier, we killed the first attempt of task immediately in order to save resources. As a result, only 2 backup tasks failed. The same conditions occurred in the left 4 parts.

However, the situation is a little different when the job arrival interval is 120 s. We also plotted the execution time of each job with interval 120 s in Figure 13. There are little different exits in the first 4 parts. The reason is that the system workload is light because each job has 120 s before the next job arrives. The system has much free resources for backup task running. We can see from Figure 13(b) that no backup task succeeded in the first two parts. Even if the backup task failed, it will not cause any loss. But in RiskI, no backup task was launched for speculative execution which was plotted in Figure 13(c). Because none of them would succeed, most differences lied in the part L where we reduced the speculative execution count from 39 to 12 and only 1 speculative execution failed.

As a result, our scheduler performs much better in interval 60 than interval 120.

7.2.2. *Result for Streamsort.* We evaluated the job of stream-sort by running configured gridmix2 also. But we have to modify the Hadoop 0.21.0 for different arrival interval. The performance of sort is also improved better when workload is heavy since RiskI not only shortened the response time of job, but also killed some tasks to release resource earlier. When workload is light, our scheduler only takes advantage of reducing job response time.

Another conclusion is that the average performance of BBP is better than sort. The main reason is that, compared to BBP, prediction of task for streamsort has longer range which means more uncertainty.

7.3. *Impact of Heterogeneous Degree.* We change the testbed and Hadoop configurations to evaluate RiskI in different execution environments. We run our experiments in two other different environments. The original environment above is denoted as testbed 1. And the new environments

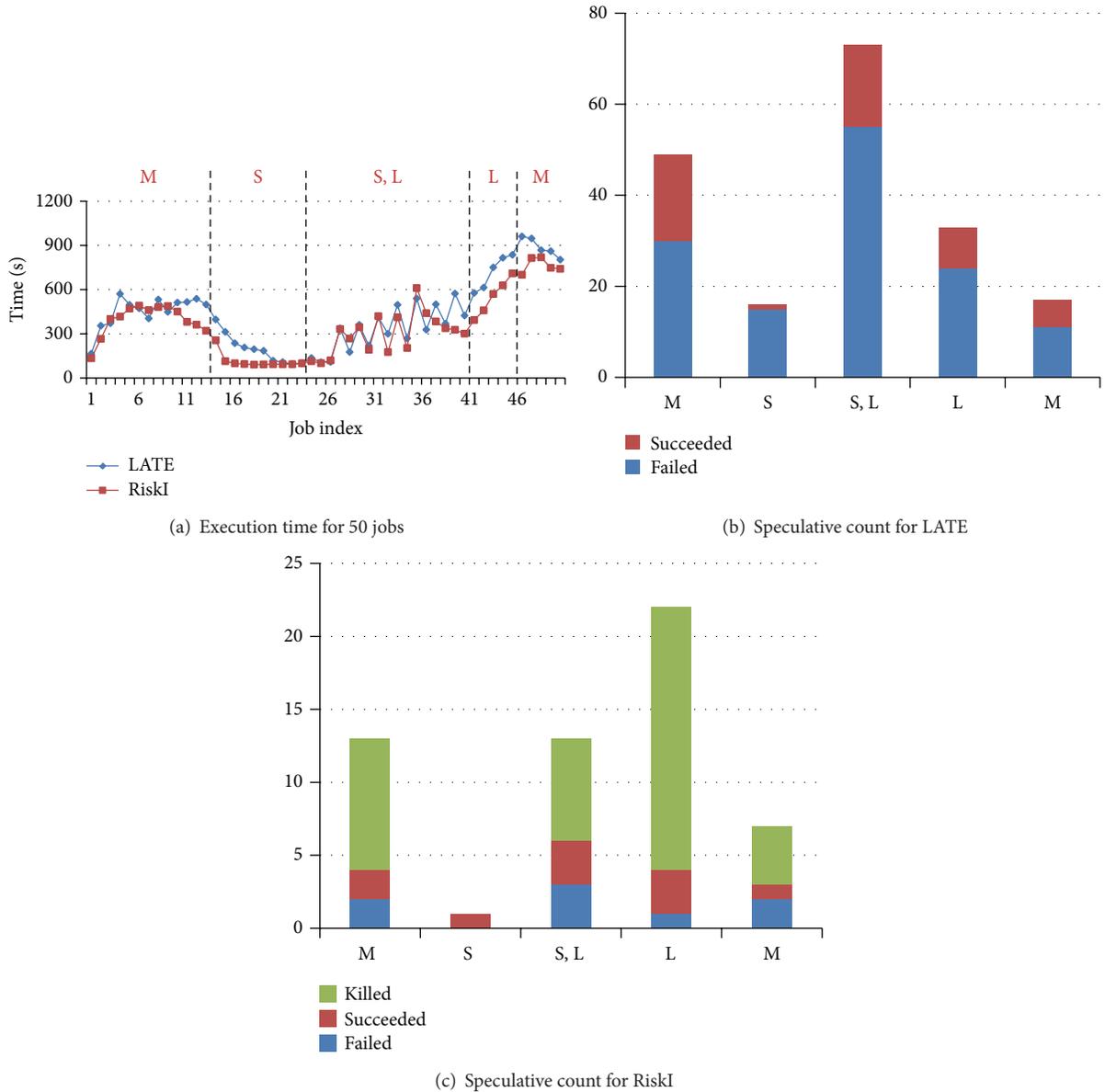


FIGURE 12: Details of execution information with interval 60. (a) is static of execution time for each job. (b and c) are statistics of speculative execution count in different parts.

were denoted as testbeds 2 and 3. The testbed 3 is 4 times more heterogeneous than testbed 2 and 8 times than testbed 1. We measure heterogeneity by its number of virtual machines on the same physical machine. The result was shown in Figure 14. When the environment is more heterogeneous, RiskI perform degrades much less than LATE.

**7.4. Stable of Prediction.** Since we make our prediction by searching similar executions of task in history, we have to demonstrate that the similar executions have stable response time. It is obvious that stable response time will make our prediction significant. Otherwise, our prediction is helpless

for scheduling. Therefore, we did the following experiments to proof our prediction.

First, we selected a four-slot work node which implies at most four tasks that can run on the node in parallel. And then we investigated the task running time under all the possible execution conditions. These conditions consist of various combinations of tasks which are selected from wordcount, randomwrite, and Bailey-Borwein-Plouffe. So there are 9 conditions in total. Figure 15(a) shows our results on running time of task from BBP job that the task ran 20 times in every environment. We can see that the execution time is relatively stable in a range time. For further analysis, Figure 4(a) shows the CDF of response time distribution between the longest time and the shortest time in one

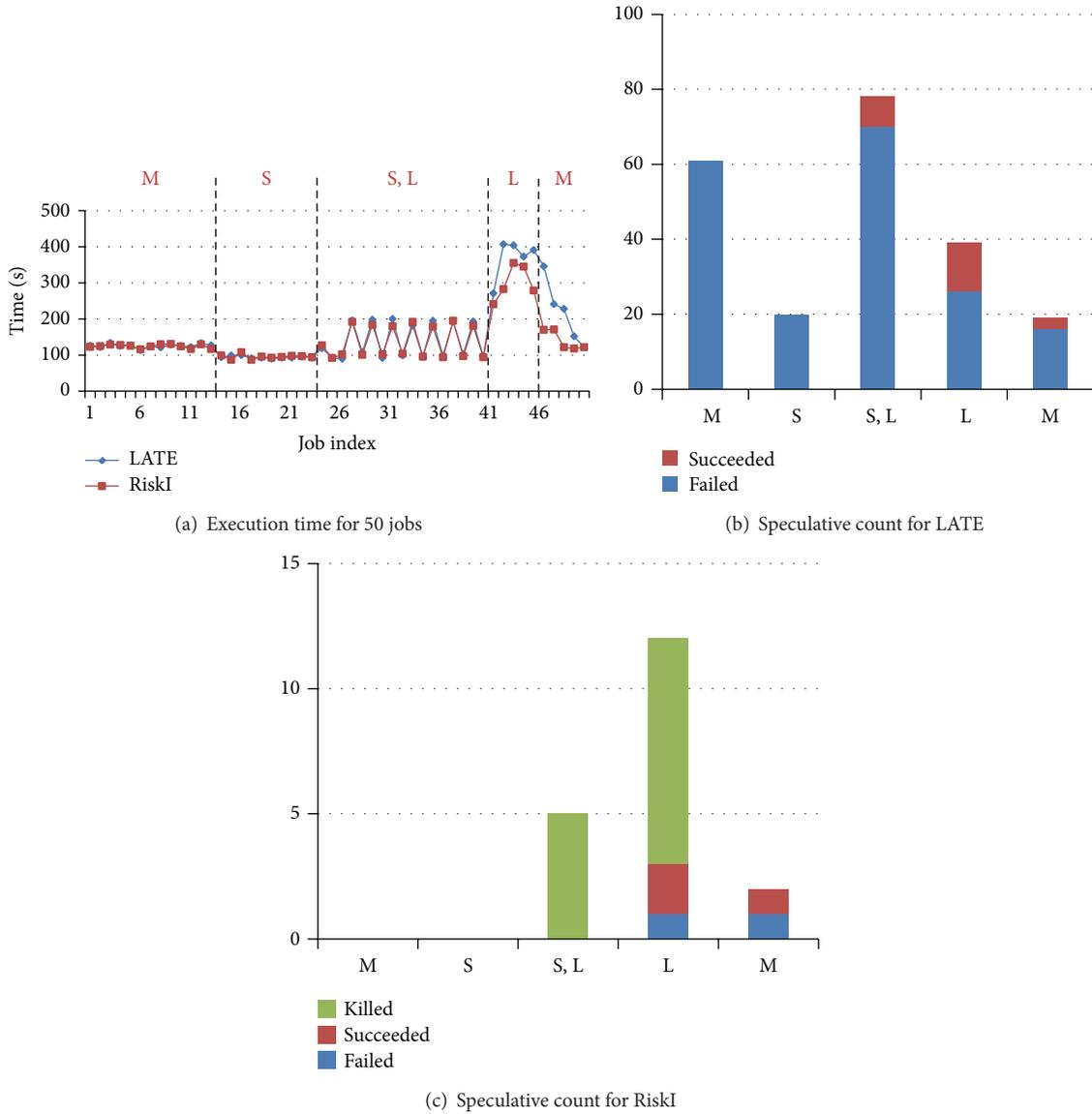


FIGURE 13: Details of execution information with interval 120. (a) is static of execution time for each job. (b and c) are statistics of speculative execution count in different parts.

condition. It demonstrated that most times are centralized distribution. The result of task from wordcount is shown in Figures 15(b) and 4(b). The conclusion can be also certified.

We showed the prediction evolution with the execution times increased in Figure 16. We selected two conditions where the prediction is most stable and most unstable. The evolutionary process of prediction is plotted.

### 8. Conclusion

We propose a framework named RiskI for task assignment in a fault tolerance data processing system. We suggest that, by finding the most similar task in history, profile-based

method can predict the execution time with a range of time. Then it makes risk-aware assignment for task to seek more profit while reducing the risk introduced by uncertainty. Extensive evaluations have shown that RiskI can provide good performance in all conditions. And RiskI performs much better when the system puts up more heterogeneity. We believe that the performance of RiskI will improve further if we understand the behavior of system and find better similar algorithm.

### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

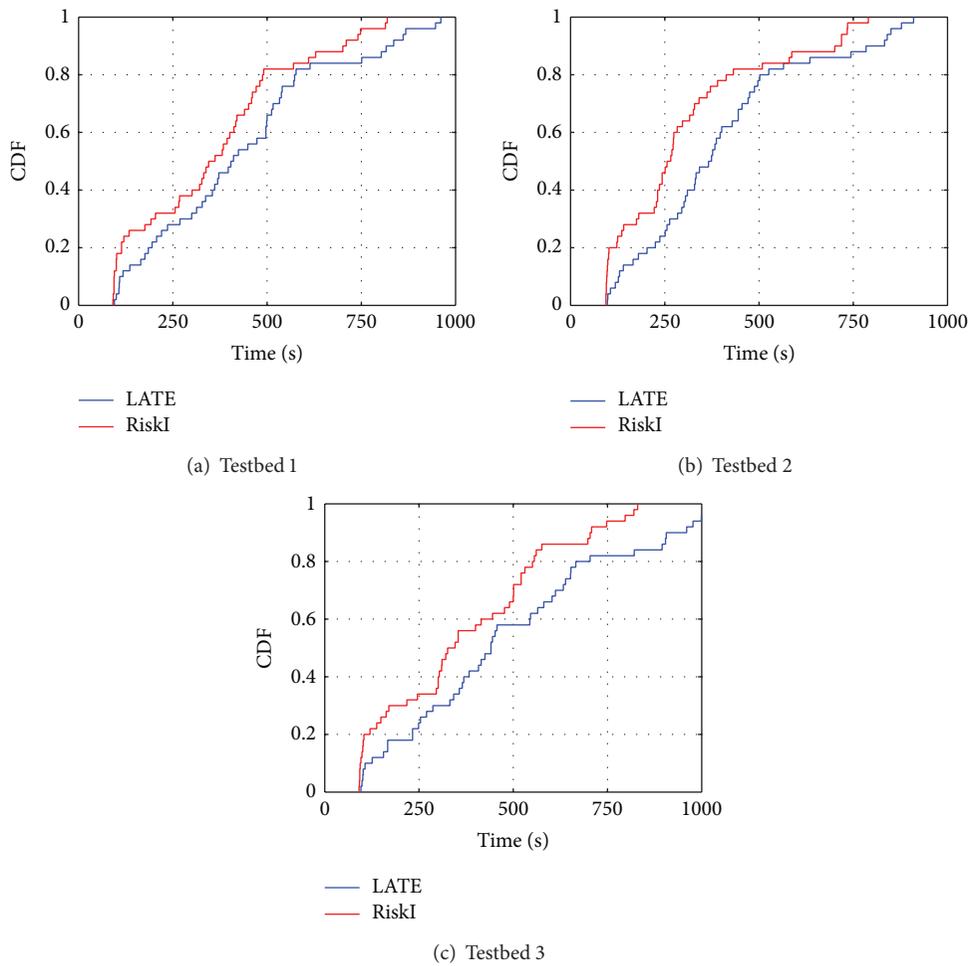


FIGURE 14: CDFs of execution time in different testbeds.

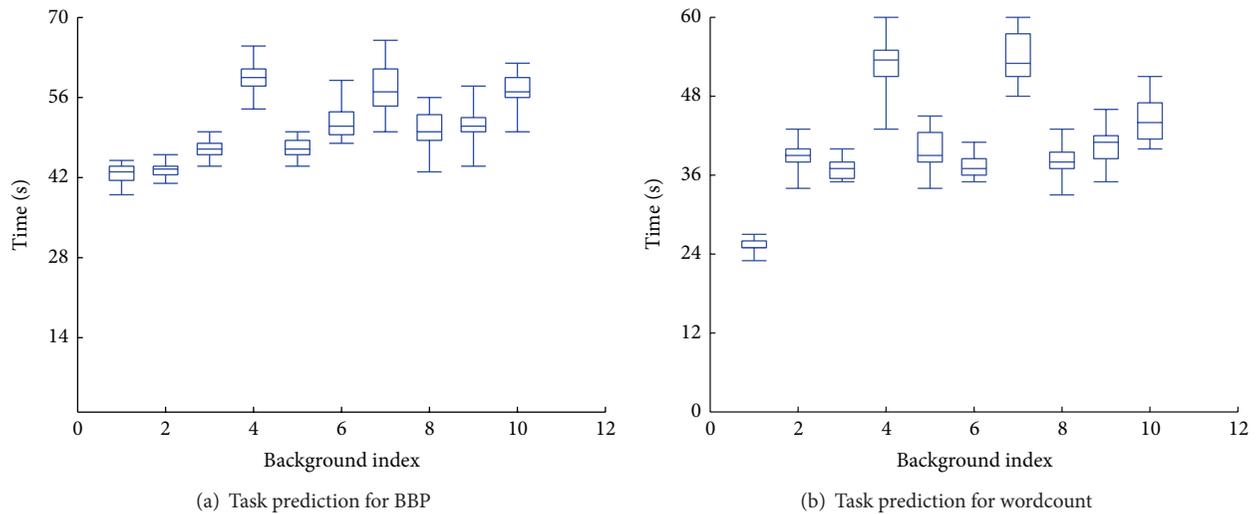


FIGURE 15: Execution time prediction.

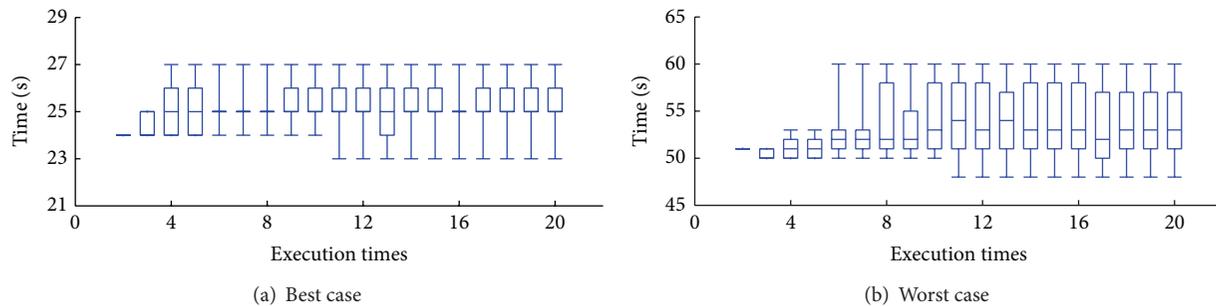


FIGURE 16: Prediction evolution.

## Acknowledgments

This paper is supported by NSFC no. 61272483 and no. 61202430 and National Lab Fund no. KJ-12-06.

## References

- [1] K. V. Vishwanath and N. Nagappan, "Characterizing cloud computing hardware reliability," in *Proceedings of the 1st ACM Symposium on Cloud Computing (SoCC '10)*, pp. 193–203, Citeseer, June 2010.
- [2] P. Gill, N. Jain, and N. Nagappan, "Understanding network failures in data centers: measurement, analysis, and implications," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 4, pp. 350–361, 2011.
- [3] Q. Zheng, "Improving MapReduce fault tolerance in the cloud," in *Proceedings of the IEEE International Symposium on Parallel and Distributed Processing, Workshops and Phd Forum (IPDPSW '10)*, pp. 1–6, IEEE, April 2010.
- [4] Q. Zhang, L. Cheng, and R. Boutaba, "Cloud computing: state-of-the-art and research challenges," *Journal of Internet Services and Applications*, vol. 1, no. 1, pp. 7–18, 2010.
- [5] J. Dean and S. Ghemawat, "MapReduce: simplified data processing on large clusters," *Communications of the ACM*, vol. 51, no. 1, pp. 107–113, 2008.
- [6] M. Zaharia, A. Konwinski, A. Joseph, R. Katz, and I. Stoica, "Improving mapreduce performance in heterogeneous environments," in *Proceedings of the 8th USENIX Conference on Operating Systems Design and Implementation*, pp. 29–42, 2008.
- [7] G. Ananthanarayanan, S. Kandula, A. Greenberg et al., "Reining in the outliers in map-reduce clusters using mantri," in *Proceedings of the 9th USENIX Conference on Operating Systems Design and Implementation*, pp. 1–16, USENIX Association, 2010.
- [8] B. Hindman, A. Konwinski, M. Zaharia et al., "Mesos: a platform for fine-grained resource sharing in the data center," in *Proceedings of the 8th USENIX Conference on Networked Systems Design and Implementation*, pp. 22–22, USENIX Association, 2011.
- [9] M. Isard, M. Budiu, Y. Yu, A. Birrell, and D. Fetterly, "Dryad: distributed data-parallel programs from sequential building blocks," *ACM SIGOPS Operating Systems Review*, vol. 41, no. 3, pp. 59–72, 2007.
- [10] M. Zaharia, M. Chowdhury, M. Franklin, S. Shenker, and I. Stoica, "Spark: cluster computing with working sets," in *Proceedings of the 2nd USENIX Conference on Hot Topics in Cloud Computing*, pp. 10–10, USENIX Association, 2010.
- [11] A. Snaveley, N. Wolter, and L. Carrington, "Modeling application performance by convolving machine signatures with application profiles," in *Proceedings of the IEEE International Workshop on Workload Characterization (WWC-4 '01)*, pp. 149–156, IEEE, 2001.
- [12] R. Uhlig, G. Neiger, D. Rodgers et al., "Intel virtualization technology," *Computer*, vol. 38, no. 5, pp. 48–56, 2005.
- [13] P. Jalote, *Fault Tolerance in Distributed Systems*, Prentice-Hall, 1994.
- [14] A. Salman, I. Ahmad, and S. Al-Madani, "Particle swarm optimization for task assignment problem," *Microprocessors and Microsystems*, vol. 26, no. 8, pp. 363–371, 2002.
- [15] M. Frank and P. Wolfe, "An algorithm for quadratic programming," *Naval Research Logistics Quarterly*, vol. 3, no. 1-2, pp. 95–110, 2006.
- [16] R. Jain, D. Chiu, and W. Hawe, *A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Computer System*, Eastern Research Laboratory, Digital Equipment Corporation, 1984.
- [17] S. Balsamo, A. Di Marco, P. Inverardi, and M. Simeoni, "Model-based performance prediction in software development: a survey," *IEEE Transactions on Software Engineering*, vol. 30, no. 5, pp. 295–310, 2004.
- [18] L. Kleinrock, *Queueing Systems. Volume 1: Theory*, Wiley-Interscience, 1975.
- [19] E. Lazowska, J. Zahorjan, G. Graham, and K. Sevcik, *Quantitative System Performance: Computer System Analysis Using Queueing Network Models*, Prentice-Hall, 1984.
- [20] K. Kant and M. Srinivasan, *Introduction to Computer System Performance Evaluation*, McGraw-Hill College, 1992.
- [21] K. Trivedi, *Probability and Statistics with Reliability, Queuing, and Computer Science Applications*, Wiley, New York, NY, USA, 2002.
- [22] J. Banks and J. Carson, *Event System Simulation*, Prentice-Hall, 1984.
- [23] M. Marsan, G. Balbo, and G. Conte, *Performance Models of Multiprocessor Systems*, MIT Press, 1986.
- [24] F. Baccelli, G. Balbo, R. Boucherie, J. Campos, and G. Chiola, "Annotated bibliography on stochastic petri nets," in *Performance Evaluation of Parallel and Distributed Systems Solution Methods*, vol. 105, CWI Tract, 1994.
- [25] M. Harchol-Balter, "Task assignment with unknown duration," *Journal of the ACM*, vol. 49, no. 2, pp. 260–288, 2002.
- [26] D. W. Pentico, "Assignment problems: a golden anniversary survey," *European Journal of Operational Research*, vol. 176, no. 2, pp. 774–793, 2007.

- [27] C. Shen and W. Tsai, "A graph matching approach to optimal task assignment in distributed computing systems using a minimax criterion," *IEEE Transactions on Computers*, vol. 100, no. 3, pp. 197–203, 1985.
- [28] M. Harchol-Balter, M. E. Crovella, and C. D. Murta, "On choosing a task assignment policy for a distributed server system," *Journal of Parallel and Distributed Computing*, vol. 59, no. 2, pp. 204–228, 1999.
- [29] R. I. Davis and A. Burns, "A survey of hard real-time scheduling for multiprocessor systems," *ACM Computing Surveys*, vol. 43, no. 4, article 35, 2011.
- [30] Y. Jiang, X. Shen, J. Chen, and R. Tripathi, "Analysis and approximation of optimal co-scheduling on chip multiprocessors," in *Proceedings of the 17th International Conference on Parallel Architectures and Compilation Techniques (PACT '08)*, pp. 220–229, ACM, October 2008.
- [31] R. Bunt, D. Eager, G. Oster, and C. Williamson, "Achieving load balance and effective caching in clustered web servers," in *Proceedings of the 4th International Web Caching Workshop*, April 1999.
- [32] P. B. Godfrey and I. Stoica, "Heterogeneity and load balance in distributed hash tables," in *Proceedings of the IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '05)*, vol. 1, pp. 596–606, IEEE, March 2005.
- [33] M. Zaharia, D. Borthakur, J. Sen Sarma, K. Elmeleegy, S. Shenker, and I. Stoica, "Delay scheduling: a simple technique for achieving locality and fairness in cluster scheduling," in *Proceedings of the 5th European Conference on Computer Systems (EuroSys '10)*, pp. 265–278, ACM, April 2010.
- [34] <http://wiki.apache.org/hadoop/WordCount>.
- [35] A. Verma, L. Cherkasova, and R. H. Campbell, "ARIA: automatic resource inference and allocation for mapreduce environments," in *Proceedings of the 8th ACM International Conference on Autonomic Computing*, pp. 235–244, ACM, June 2011.
- [36] J. de Leeuw and S. Pruzansky, "A new computational method to fit the weighted euclidean distance model," *Psychometrika*, vol. 43, no. 4, pp. 479–490, 1978.
- [37] <http://hadoopapache.org/docs/mapreduce/current/gridmix.html>.

## Research Article

# Effect Analysis of Design Variables on the Disc in a Double-Eccentric Butterfly Valve

**Sangmo Kang, Da-Eun Kim, Kuk-Kyeom Kim, and Jun-Oh Kim**

*Department of Mechanical Engineering, Dong-A University, Busan 604-714, Republic of Korea*

Correspondence should be addressed to Sangmo Kang; kangsm@dau.ac.kr

Received 12 January 2014; Accepted 20 February 2014; Published 17 April 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Sangmo Kang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We have performed a shape optimization of the disc in an industrial double-eccentric butterfly valve using the effect analysis of design variables to enhance the valve performance. For the optimization, we select three performance quantities such as pressure drop, maximum stress, and mass (weight) as the responses and three dimensions regarding the disc shape as the design variables. Subsequently, we compose a layout of orthogonal array (L16) by performing numerical simulations on the flow and structure using a commercial package, ANSYS v13.0, and then make an effect analysis of the design variables on the responses using the design of experiments. Finally, we formulate a multiobjective function consisting of the three responses and then propose an optimal combination of the design variables to maximize the valve performance. Simulation results show that the disc thickness makes the most significant effect on the performance and the optimal design provides better performance than the initial design.

## 1. Introduction

The design process of a valve system demands a variety of requirements such as low pressure drop, no leakage, high structural stability, and low manufacturing cost. Such design requirements indicate conditions or performances needed by the whole or a part of the system to meet certain design criteria. A butterfly valve is generally used for isolating or regulating flow and, particularly, allows for quick shutoff of flow. In the present study, we want to discuss a double-eccentric butterfly valve that is being widely used at the industrial sites due to the small space taken up by it and its small weight and low maintenance cost. Note that the standard pressure grades (ISO 10631, ASME B16.34, AWWA C504, KS B 2333, and so on), inspection methods, and pressure tests (ISO 5208, KS B 2304, and so on) are currently used in accordance with the design requirements for the use of butterfly valves in the industrial field.

Ding et al. [1] optimized a sealing structure in a triple-eccentric butterfly valve, while Ejab and Samir [2] optimized a butterfly valve housing to reduce its weight. Since then, a variety of optimal design techniques have been applied to enhance the performance and stability of industrial valves [3–6]. In this context, the present study aims to select design

variables needed to optimize an industrial double-eccentric butterfly valve, analyze the effect of the design variables on the valve performance, and then propose an optimal combination of the design variables to provide the best performance.

## 2. Flow and Structural Analyses

*2.1. Analysis Model: Butterfly Valve.* Figure 1 shows the external appearance of the industrial butterfly valve considered in the present study, which is drawn using a commercial 3D modeling package, Solidworks. The valve consists mainly of disc, shaft, body, and bush. The opening and closing operation of the valve is run by rotating the disc, which is positioned in the center of the pipe. The valve is called the double-eccentric butterfly valve because of the following two offsets on the valve; the shaft of the disc is offset from the center line of the disc seat and body seal (offset one) and from the center line of the bore (offset two). Therefore, it has much less wear than the concentric butterfly valve and is excellent in the flow control [3]. In the present study, we attempt to perform a shape optimization of the disc installed in a double-eccentric butterfly valve using the effect analysis of design variables to enhance the valve performance.

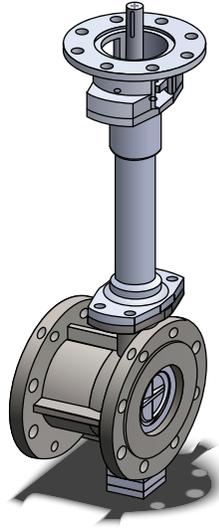


FIGURE 1: External appearance of the double-eccentric butterfly valve considered in the present study.

TABLE 1: Operating conditions for the flow analysis and materials for the structural analysis.

Operating conditions	
Working fluid	water
Working Temperature	20° C
Design pressure	1 MPa
Materials	
Body, disc	CF8M
Shaft, bush	SUS316

For the optimization of the disc, we select three performance quantities such as pressure drop, maximum stress, and mass (weight) as the responses and three dimensions regarding the disc shape as the design variables. Subsequently, we compose a layout of orthogonal array (L16) by performing numerical simulations on the flow and structure of the valve using a commercial package, ANSYS v13.0.

**2.2. Flow Analysis.** Figure 2(a) shows the computational domain constructed for the flow analysis and the grid system created using a commercial package, ANSYS v13.0. The domain is composed of the valve and the inlet and outlet pipes connected to it on both sides. The pipes have lengths of  $10D$  and  $20D$  ( $D$  is the pipe diameter), respectively, so that the flow becomes fully developed or stabilized. To calculate the pressure drop across the valve, the pressures are measured at  $2D$  and  $6D$  positions, respectively, before and behind the valve. For the turbulence model, the shear stress transport (SST) model is used. The SST model is a combination of  $k-\omega$  and  $k-\epsilon$  models produced to make the best use of the merit of each of both models; the former is more accurate near the wall, whereas the latter is away from the wall [7]. The working conditions applied to the flow analysis are shown in Table 1.

At first, we perform a numerical simulation on the flow through the initial design (model) of the valve and then

TABLE 2: Design variables and their levels.

Design variable	Unit	Level 1	Level 2	Level 3	Level 4
Disc thickness ( $x_1$ )	mm	7.65	8.22	8.79	9.35
Distance ( $x_2$ )	mm	5.4	5.8	6.2	6.6
Diameter ( $x_3$ )	mm	16.2	17.4	18.6	19.8

present one of its results in Figure 3(a). Figure 3(a) shows the pressure distribution generated around the valve for the initial model. It is observed that the pressure sharply drops across the valve and then gradually recovers downstream behind it.

**2.3. Structural Analysis.** Figure 2(b) shows the model and grid system used for the structural analysis. The maximum allowable working pressure or design pressure, 1 MPa, is applied to the back of the valve disc as a boundary condition (see Table 1) and the constraint conditions are given to the top and bottom of the valve shaft and the contact surface (or the metal seat on the valve disc) between the disc and body. To reduce the analysis time, we do not attempt to perform the structural analysis on the whole valve but instead on each of the components comprising the valve. In other words, we establish the boundary conditions and constraints by following the studies of Shin [8] executed on the butterfly valve. The materials necessary for the structural analysis are also presented in Table 1.

Subsequently, we perform a numerical simulation on the structure of the initial model and then present the stress distribution on the valve disc, as shown in Figure 3(b). The maximum stress occurs on the contact surface between the disc and body to which a constraint condition is given.

### 3. Shape Optimization of the Disc

**3.1. Design of Experiments.** As mentioned above, we chose three performance quantities such as pressure drop, maximum stress, and mass (weight) as the responses for the optimization of the butterfly valve. Next we also select three dimensions regarding the disc shape as the design variables, which are expected to significantly influence the valve performance or the responses. Figure 4 shows the schematic details on the three design variables, which include the thickness of the valve disc ( $x_1$ ), the distance between the bush and valve disc boss ( $x_2$ ), and the diameter of the valve shaft ( $x_3$ ). To make the design of experiments, four uniformly spaced levels are assigned to each design variable and presented in Table 2. Here, the bottom and top limits (levels 1 and 4) are determined, respectively, through the decrease and increase by 10% from the initial value and then levels 2 and 3 are appropriately distributed with a uniform space.

Subsequently we make the design of experiments based on the L16 orthogonal array to perform the effect analysis of the design variables on the valve performance and then propose their optimal combination. For the effect analysis, we define the multiobjective function consisting of the three

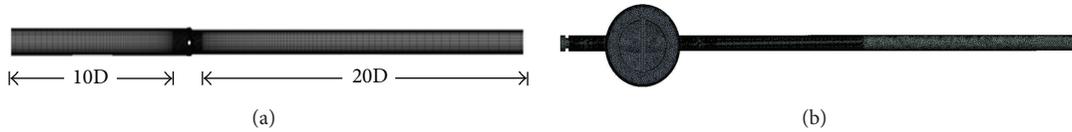


FIGURE 2: Models and grid systems generated for the structural and flow analyses: (a) flow and (b) structure.

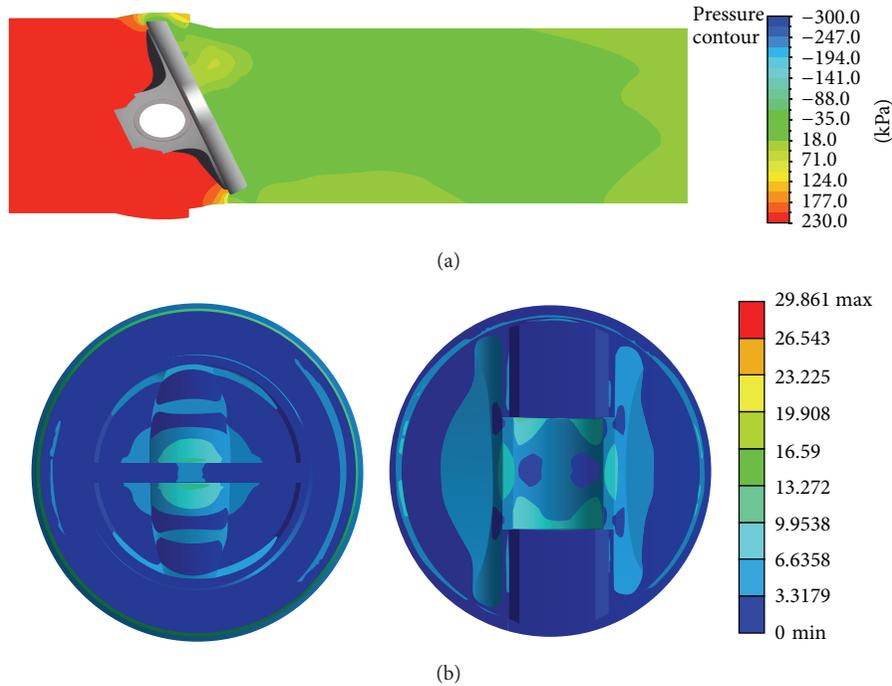


FIGURE 3: Results of the numerical simulations on the flow and structure of the initial model: (a) pressure distribution around the valve and (b) stress distribution on the valve disc.

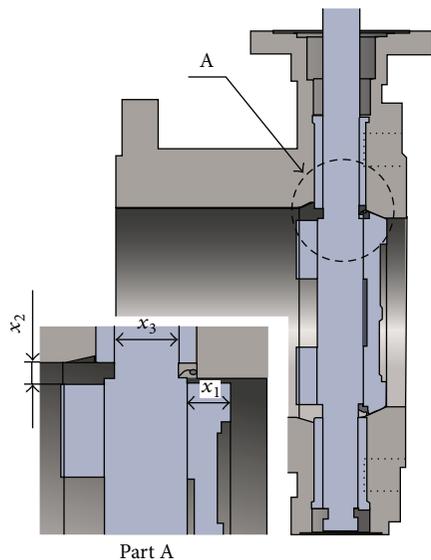


FIGURE 4: Design variables selected for the shape optimization of the valve disc.

responses, pressure drop, maximum stress, and mass, as follows:

$$\Phi = -10 \log \left( \frac{1}{n} \sum_{i=1}^n y_i^2 \right), \quad (1)$$

where  $y_i$  denotes the response and  $n$  is the response number (here  $n = 3$ ). In other words, we attempt to determine an optimal combination of the design variables in a manner to maximize the above defined objective function (1). Table 3 shows the layout of the L16 orthogonal arrays composed for the design of experiments and the responses obtained by the full numerical simulations (flow and structural analyses).

**3.2. Effect Analysis and Shape Optimization.** Finally we make the effect analysis of the design variables on the responses and then attempt to propose their optimal combination to enhance the valve performance. Here, the effect analysis is made by evaluating how much each design variable affects the responses according to the level through the analysis of the multiobjective function. The analysis results according to

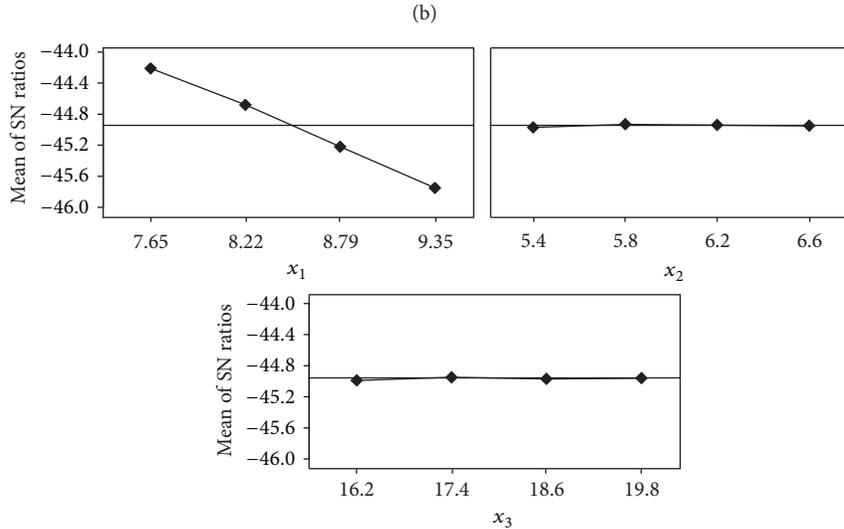
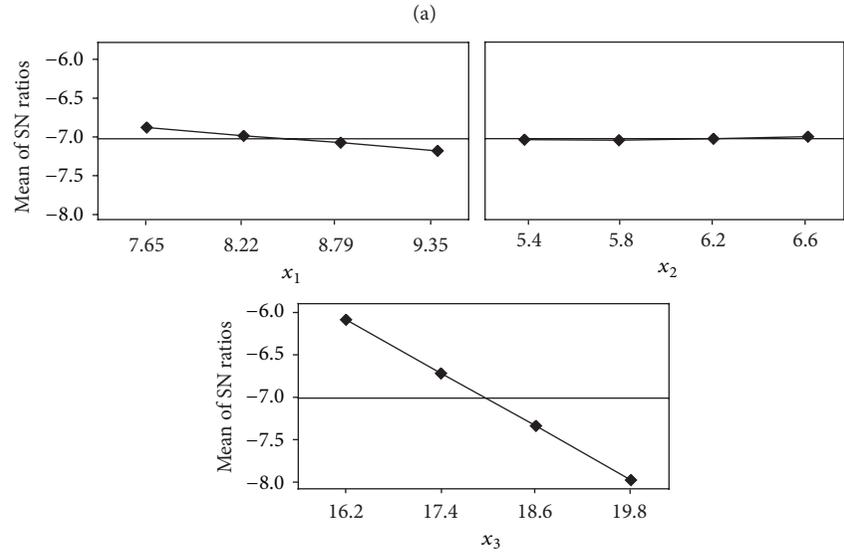
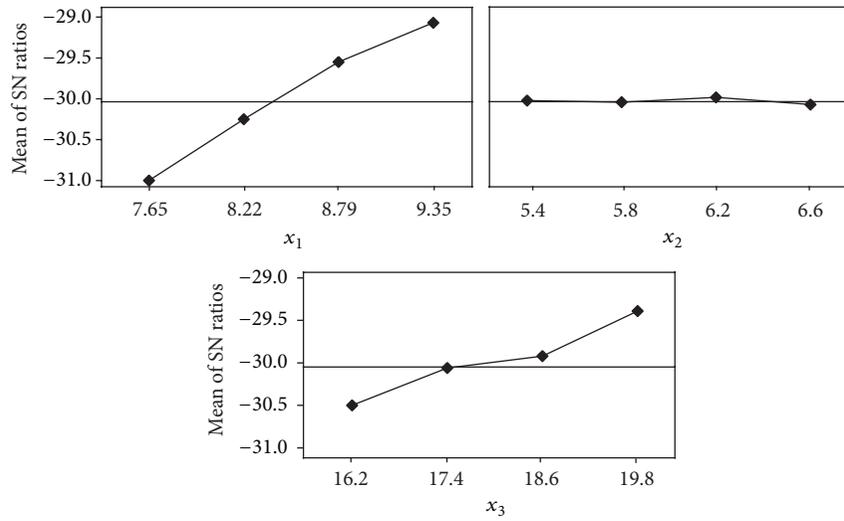


FIGURE 5: Plots to represent the main effects of the design variables on (a) pressure drop, (b) maximum stress, and (c) mass.

TABLE 3: Layout of the L16 orthogonal array based on the design of experiments and the responses obtained by the full numerical simulations.

Exp	$x_1$	$x_2$	$x_3$	Stress (MPa)	Mass (kg)	$\Delta p$ (kPa)
1	7.65	5.4	16.2	36.47	1.98	163.8
2	8.22	5.8	17.4	32.56	2.16	171.0
3	8.79	6.2	18.6	29.22	2.34	182.4
4	9.35	6.6	19.8	26.59	2.54	194.0
5	8.22	6.2	19.8	29.91	2.49	171.6
6	7.65	6.6	18.6	35.82	2.28	162.8
7	9.35	5.4	17.4	28.79	2.21	194.1
8	8.79	5.8	16.2	32.52	2.03	182.7
9	8.79	6.6	17.4	30.63	2.17	182.2
10	9.35	6.2	16.2	30.84	2.05	193.9
11	7.65	5.8	19.8	33.93	2.47	161.8
12	8.22	5.4	18.6	33.44	2.32	171.2
13	9.35	5.8	18.6	27.59	2.37	193.6
14	8.79	5.4	19.8	27.96	2.52	182.5
15	8.22	6.6	16.2	34.43	2.00	171.7
16	7.65	6.2	17.4	35.83	2.13	161.4

the level are plotted in Figure 5 and tabulated in Tables 4, 5, and 6. Note that such level based effect analysis provides very effective information to determine the design direction [9]. Following the effect analysis results (see Figure 5 and Tables 4–6), we can find that the performance (or responses) can become improved or worsened depending on the level of each design variable. Therefore, we can produce an optimal combination of the design variables by quantifying the effect by the level of each design variable on the performance. In this paper, we calculate the level based effects of the design variables and, as a result, find that the disc thickness ( $x_1$ ) has the greatest effect on the pressure drop at 95.7% and on the maximum stress at 61.7%, while the shaft diameter ( $x_3$ ) has an effect on the mass at 84.4%. On the contrary, the distance between the bush and boss ( $x_2$ ) has a relatively trivial effect on all the responses.

From the effect analysis results, we can draw an optimal combination of the design variables, that is, the disc thickness at 7.65 mm, the distance between the boss and bush at 6.2 mm, and the shaft diameter at 17.4 mm, which is corresponding to the 16th experiment in the orthogonal array table (Table 3). Table 7 shows a comparison between the initial model and the optimal model; the optimal model is improved by 2.74% in the mass and by 6.16% in the pressure drop compared with the initial model. However, there is no improvement found in the maximum stress, but it is still within the material yield strength, 205 MPa.

#### 4. Conclusions

In the present study, we have performed a shape optimization of the disc in an industrial double-eccentric butterfly valve using the effect analysis of design variables to enhance the valve performance. For the optimization, we select three performance quantities such as pressure drop, maximum stress,

TABLE 4: Analysis results on the effect of the design variables on the pressure drop.

	Level 1	Level 2	Level 3	Level 4	Deviation	%	Rank
$x_1$	-44.21	-44.68	-45.22	-45.75	1.54	95.7	1
$x_2$	-44.99	-44.96	-44.95	-44.97	0.03	1.8	3
$x_3$	-44.99	-44.95	-44.97	-44.96	0.04	2.5	2
Total					1.61	100	

TABLE 5: Analysis results on the effect of the design variables on the maximum stress.

	Level 1	Level 2	Level 3	Level 4	Deviation	%	Rank
$x_1$	-31.01	-30.25	-29.55	-29.07	1.93	61.7	1
$x_2$	-29.96	-29.98	-29.93	-30.01	0.09	2.8	3
$x_3$	-30.50	-30.06	-29.92	-29.39	1.11	35.5	2
Total					3.13	100	

TABLE 6: Analysis results on the effect of the design variables on the mass.

	Level 1	Level 2	Level 3	Level 4	Deviation	%	Rank
$x_1$	-6.896	-6.997	-7.101	-7.203	0.307	13.7	2
$x_2$	-7.067	-7.060	-7.046	-7.024	0.043	1.9	3
$x_3$	-6.105	-6.736	-7.366	-7.991	1.886	84.4	1
Total					2.236	100	

TABLE 7: Comparison of the performance between the initial and optimal models.

	Stress (MPa)	Mass (kg)	$\Delta p$ (kPa)
Initial model	29.86	2.19	172.0
Optimal model	35.83	2.13	161.4

and mass (weight) as the responses and three dimensions regarding the disc shape as the design variables. Here, the effect analysis is made by evaluating how much each design variable affects the responses according to the level through the design of experiments made with the L16 orthogonal array and the analysis of the multiobjective function.

Results indicate that the disc thickness has the greatest effect on the flow and structural performance of the butterfly valve considered in the present study. Through the shape optimization of the disc using the effect analysis, the performance of the optimal design becomes improved by 6.16% in the pressure drop and by 2.74% in the mass compared with the initial design. On the other hand, there is no improvement found in the maximum stress. Although only one design variable plays a dominant role in the valve performance in this study, systematic design with more diverse variables regarding the disc shape in addition to the disc thickness is expected to bring better findings in later studies.

#### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work has been supported in part by the Valve Center from the Regional Innovation Center (RIC) Program and by the Human Resources Development Program (no. 20134010200550) of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant. Note that both programs are commonly funded by the Ministry of Trade, Industry and Energy (MOTIE), Republic of Korea.

## References

- [1] Q. Ding, G. Sun, H. Zhao, and S. Zhang, "Optimization design of metallic seal pair structure for triple offset butterfly valve," in *Proceedings of the International Technology and Innovation Conference (ITIC '09)*, October 2009.
- [2] A. Ejab and L. Samir, "Weight optimization of the butterfly valve housing," in *Proceedings of the 10th International Research/Expert Conference*, 2006.
- [3] S. M. Yang, S. H. Baek, and S. Kang, "Shape design for disc of a double-eccentric butterfly valve using the topology optimization technique," *Journal of Computational Fluids Engineering*, vol. 17, no. 1, pp. 61–69, 2012.
- [4] C. C. Tsai, C. Y. Chang, and C. H. Tseng, "Optimal design of metal seated ball valve mechanism," *Structural and Multidisciplinary Optimization*, vol. 26, no. 3–4, pp. 249–255, 2004.
- [5] X. G. Song, L. Wang, S. H. Baek, and Y. C. Park, "Multidisciplinary optimization of a butterfly valve," *ISA Transactions*, vol. 48, no. 3, pp. 370–377, 2009.
- [6] P. Duda and R. Dwornicka, "Optimization of heating and cooling operations of steam gate valve," *Structural and Multidisciplinary Optimization*, vol. 40, no. 1–6, pp. 529–535, 2010.
- [7] B. S. Lee, "Analysis of unsteady turbulent flow fields around 2 dimensional square cylinder using  $k-\omega$  SST turbulence model," in *Proceedings of the Korean Society for Aeronautical and Space Sciences*, pp. 104–107, 1998.
- [8] M. S. Shin, "A study on structural analysis of butterfly valve components by pressure testing of the industrial standard," *Journal of Fluid Machinery*, vol. 14, no. 3, pp. 5–9, 2011.
- [9] S. H. Baek, "Optimization of process parameters for mill scale recycling using Taguchi method," *Journal of Korean Society for Precision Engineering*, vol. 25, no. 2, pp. 88–95, 2008.

## Research Article

# Optimized Scheduling Technique of Null Subcarriers for Peak Power Control in 3GPP LTE Downlink

Soobum Cho<sup>1</sup> and Sang Kyu Park<sup>2</sup>

<sup>1</sup> Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA

<sup>2</sup> Department of Electronics and Computer Engineering, Hanyang University, Seoul 133-791, Republic of Korea

Correspondence should be addressed to Sang Kyu Park; [skpark@hanyang.ac.kr](mailto:skpark@hanyang.ac.kr)

Received 28 February 2014; Accepted 18 March 2014; Published 17 April 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 S. Cho and S. K. Park. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Orthogonal frequency division multiple access (OFDMA) is a key multiple access technique for the long term evolution (LTE) downlink. However, high peak-to-average power ratio (PAPR) can cause the degradation of power efficiency. The well-known PAPR reduction technique, dummy sequence insertion (DSI), can be a realistic solution because of its structural simplicity. However, the large usage of subcarriers for the dummy sequences may decrease the transmitted data rate in the DSI scheme. In this paper, a novel DSI scheme is applied to the LTE system. Firstly, we obtain the null subcarriers in single-input single-output (SISO) and multiple-input multiple-output (MIMO) systems, respectively; then, optimized dummy sequences are inserted into the obtained null subcarrier. Simulation results show that Walsh-Hadamard transform (WHT) sequence is the best for the dummy sequence and the ratio of 16 to 20 for the WHT and randomly generated sequences has the maximum PAPR reduction performance. The number of near optimal iteration is derived to prevent exhausted iterations. It is also shown that there is no bit error rate (BER) degradation with the proposed technique in LTE downlink system.

## 1. Introduction

The fields of mobile communication techniques have been rapidly developed in recent decades. One of the development results is the 3rd generation partnership project (3GPP) long term evolution (LTE), which has been deployed all over the world. Downlink transmission of the LTE is based on the use of multiple access technology: orthogonal frequency division multiple access (OFDMA), which is a modification of orthogonal frequency division multiplexing (OFDM) for the multiple access [1, 2]. Recent advances of digital signal processing (DSP) technique have accelerated the popularity of the OFDM. The technique has a lot of tolerances to frequency selective fading and multipath interference; therefore it has been adapted to numerous international standards for wired and wireless communication systems such as very-high-bit-rate digital subscriber line (VDSL) [3], power line communication (PLC) [4], wireless local area network (WLAN) [5], and ultrawideband (UWB) [6]. It is

also attracting a lot of interest in visible light communication (VLC) and optical wireless communications [7].

However, together with its advantages, still some challenging issues remain for the OFDM access technology design. One of the major drawbacks is high peak-to-average power ratio (PAPR) of transmitted signals. Therefore, the detection efficiency of the OFDM receiver is very sensitive to the nonlinear devices such as digital-to-analog converter (DAC) and high power amplifier (HPA). That may severely diminish the system performance because of the detection efficiency degradation and induced spectral regrowth. Most of the transmitters of wireless communication systems employ the HPA to obtain sufficient transmit power. The HPA usually operates near the saturation region to achieve the maximum output power efficiency; thus the memoryless nonlinear distortions occur in the communication channels due to the high PAPR of the input signals. If the HPA does not operate within linear region with power back-off (PBO), it is difficult to keep the out-of-band power below the specified

limits. This situation leads to very inefficient amplification and expensive transmitters [8]. Therefore, it is important to do research on the characteristics of the PAPR including its reduction in order to use the features of the OFDM.

To deal with the PAPR problem, various approaches have been proposed such as deliberate clipping [9], partial transmit sequence (PTS) [10], selected mapping (SLM) [11], interleaving [12], coding [13], tone reservation (TR) [14], active constellation extension (ACE) [15], and dummy sequence insertion (DSI) [16]. These techniques reduce the PAPR by the trade-off among signal power, data rate, system complexity, and bit error rate (BER) performance. The DSI scheme can be simply implemented by scheduling some dummy sequences which are used only for PAPR reduction. However, the large usage of additional subcarriers for the dummy sequences can directly cause the reduction of transmission efficiency. The number of dummy sequences needed for the desired PAPR reduction level depends on the feature of the communication systems. In [17], the number of unused subcarriers was calculated in the LTE single-input single-output (SISO) system and it was used for the dummy sequences with cyclic shifted sequences scheme. However, the PAPR reduction performance was still high, since dummy sequences were not optimally scheduled. Furthermore, BER performances were not compared with the conventional LTE system even though the BER is a very important aspect.

In this paper, the null subcarriers for the dummy sequences are derived in LTE SISO,  $2 \times 2$ , and  $4 \times 4$  multiple-input multiple-output (MIMO) systems, respectively, and transmission efficiencies are calculated. The optimal design of the dummy sequences is derived and it is scheduled to control the peak power. Finally, the simulation result shows the comparison of the BER performances which demonstrates that the proposed method can reduce the PAPR considerably without BER performance degradation.

This paper is organized as follows. Section 2 describes the PAPR and complementary cumulative distribution function (CCDF) definitions. In Section 3, the null subcarriers of the LTE SISO,  $2 \times 2$ , and  $4 \times 4$  MIMO systems are derived. Section 4 shows the simulation results and analyzes the performances. Finally, Section 5 offers our conclusions and future works.

## 2. PAPR and CCDF Definitions

Several drawbacks arise in OFDM, the most severe of which is the highly nonconstant envelope of the transmitted signals, that is, the PAPR, making the OFDM very sensitive to non-linear components in the transmission path. The use of HPA can be one of the solutions. However, owing to cost, design, and, most importantly, power efficiency considerations, the HPA cannot resolve the dynamics of the transmitted signal. A clipping method inevitably cuts off the signal at some point, which causes additional in-band distortion and adjacent channel interference. The power efficiency penalty is certainly the major obstacle in implementing OFDM systems in low-cost applications. Moreover, in power limited regimes determined by regulatory bodies, the average power is reduced

in comparison to single-carrier systems. The main goal of peak power control is to diminish the influence of high peaks in transmit signals on the performance of the transmission system. The PAPR of the transmit signal can be defined as

$$\text{PAPR} = \frac{\max_{0 \leq k \leq N-1} |x_k|^2}{E[|x_k|^2]}, \quad (1)$$

where  $E[\cdot]$  denotes mathematical expectation.

The CCDF which denotes the probability that the PAPR of a data block exceeds a given threshold is one of the most frequently used performance measures for PAPR reduction techniques. If the number of subcarriers is large enough, magnitudes of real and imaginary parts of output signal have Gaussian distribution with mean of zero and variance of  $1/2$  by central limit theorem. Thus, the amplitude of the OFDM signal follows Rayleigh distribution while the power distribution of OFDM signal is central chi-square distribution with two degrees of freedom and a mean of zero. The CCDF of the PAPR of a data block with Nyquist rate sampling is derived as

$$\Pr(\text{PAPR} > \text{PAPR}_0) = 1 - (1 - \exp(-\text{PAPR}_0))^N, \quad (2)$$

where  $\text{PAPR}_0$  is the threshold PAPR. This expression assumes that the  $N$  time domain signal samples are mutually independent and uncorrelated. However, when oversampling is applied, the assumption is no longer valid. The CCDF of the PAPR for  $N$  subcarriers with oversampling is given by

$$\Pr(\text{PAPR} > \text{PAPR}_0) = 1 - (1 - \exp(-\text{PAPR}_0))^{\alpha N}, \quad (3)$$

where  $\alpha$  is a certain number expressing the effect of oversampling.

## 3. Calculation of the Null Subcarriers of the LTE Downlink System

**3.1. Frame Structure.** There are two radio frame structures for LTE, that is, frame structure type 1 (FS1) for full and half duplex frequency division duplex (FDD) and frame structure type 2 (FS2) for time division duplex (TDD). This paper focuses on FDD. In FDD, because uplink and downlink transmissions are separated in the frequency domain, the frame structure is the same in the uplink and downlink in terms of frame, subframe, and slot duration. FS1 is shown in Figure 1.

The size of various fields in the time domain is expressed as a number of time units,  $T_s$ . This structure consists of ten 1 ms subframes, each composed of two 0.5 ms slots ( $T_{\text{slot}} = 15360T_s = 0.5$  ms) for a total duration of 10 ms ( $T_f = 307200T_s = 10$  ms).

**3.2. Downlink Physical Resource Elements.** One symbol on one subcarrier is defined as the resource element, which is the smallest time-frequency unit used for downlink transmission. A group of twelve contiguous subcarriers in frequency and one slot in time is called a resource block (RB) [19], which is shown in Figure 2.

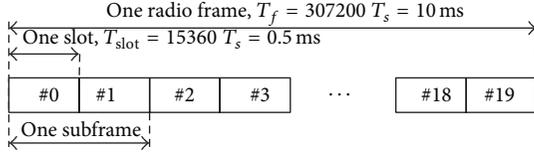


FIGURE 1: Frame structure type 1 [18].

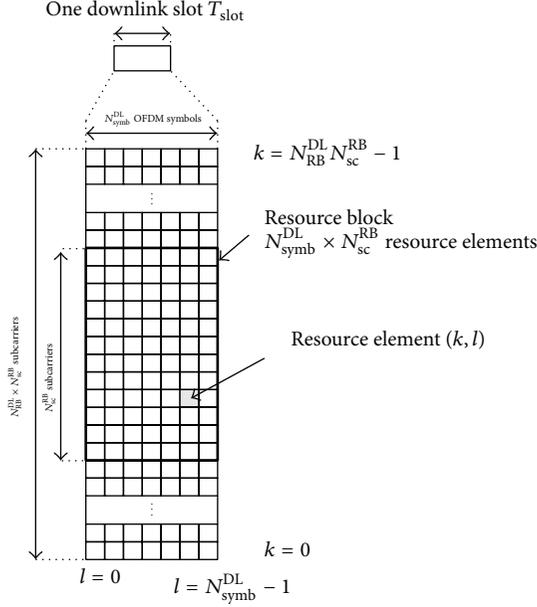


FIGURE 2: Downlink resource grid [18].

TABLE 1: Physical resource block parameters [18].

Configuration	$\Delta f$	$N_{sc}^{RB}$	$N_{symb}^{DL}$
Normal cyclic prefix	$\Delta f = 15 \text{ kHz}$	12	7
Extended cyclic prefix	$\Delta f = 15 \text{ kHz}$	12	6
Extended cyclic prefix	$\Delta f = 7.5 \text{ kHz}$	24	3

A physical RB consists of  $N_{symb}^{DL} \times N_{sc}^{RB}$  resource elements, where  $N_{symb}^{DL}$  is the number of symbols per slot and  $N_{sc}^{RB}$  is the number of subcarriers per RB.

One downlink slot using the normal cyclic prefix (CP) length contains seven symbols. Variations on this configuration for FSI are summarized in Table 1. The CP is chosen to be slightly longer than the longest expected delay spread in the radio channel.

**3.3. Null Subcarriers in PDSCH of SISO, 2 × 2, and 4 × 4 MIMO Systems.** Firstly, the number of null subcarriers per frame,  $N_{spaces,1 \times 1}^{FRM}$ , is calculated in the physical downlink shared channel (PDSCH) of the LTE SISO system [17]. The number of data per RB,  $N_{data,1 \times 1}^{RB}$ , can be obtained as

$$N_{data,1 \times 1}^{RB} = N_{symb}^{DL} \times N_{sc}^{RB} - N_{pilot}^{RB} - N_{u-pilot,1 \times 1}^{RB} \quad (4)$$

$$= 7 \times 12 - 4 - 0 = 80,$$

where  $N_{symb}^{DL}$  is the number of symbols per slot, which is 7 in normal CP,  $N_{sc}^{RB}$  is the number of subcarriers per RB, which is 12,  $N_{pilot}^{RB}$  is the number of reference signals (RSs) per RB, which is 4, and  $N_{u-pilot,1 \times 1}^{RB}$  is the number of RSs used in other antenna ports per RB, which is 0 in SISO system. The number of data per subframe,  $N_{data,1 \times 1}^{S-FRM}$ , can be calculated as

$$N_{data,1 \times 1}^{S-FRM} = N_{data,1 \times 1}^{RB} \times 2 - N_{PDCCH,1 \times 1}^{RB} \quad (5)$$

$$= 80 \times 2 - 34 = 126,$$

where  $N_{PDCCH,1 \times 1}^{RB}$  is the number of symbols for the physical downlink control channel (PDCCH) per RB, which is 34 in SISO system, and two means there are two slots in one subframe. The number of data per frame,  $N_{data,1 \times 1}^{FRM}$ , can be calculated as

$$N_{data,1 \times 1}^{FRM} = N_{data,1 \times 1}^{S-FRM} \times N_{RB}^{LEN} \times N_{mod}^{ODR} \times TB_{CR} \quad (6)$$

$$= 126 \times 5 \times 2 \times \frac{1}{3} = 420,$$

where  $N_{RB}^{LEN} (\leq N_{RB}^{DL})$  is the number of RBs, assumed to be five in this paper.  $N_{RB}^{DL}$  is the maximum number of RBs for fixed transmission bandwidth, which is 25 at the 5 MHz bandwidth.  $N_{mod}^{ODR}$  is the modulation order, which is 2, because this paper assumes quadrature phase shift keying (QPSK) modulation.  $TB_{CR}$  is the turbo coding rate, which is defined as 1/3 in PDSCH. Therefore, the number of data created in the data source step,  $N_{data,1 \times 1}^{DS}$ , can be calculated as

$$N_{data,1 \times 1}^{DS} = (N_{data,1 \times 1}^{FRM} - N_{CW,1 \times 1} \times N_{chksun}^{CRC} - N_{PBCH,1 \times 1}^{FRM} - N_{S-SS}^{FRM} - N_{P-SS}^{FRM}) \times (N_{CW,1 \times 1})^{-1} \quad (7)$$

$$= \frac{420 - 1 \times 24 - 23 - 6 - 6}{1} = 361,$$

where  $N_{CW,1 \times 1}$  is the number of code words in SISO system, which is 1,  $N_{chksun}^{CRC}$  is the number of cyclic redundancy check (CRC) checksums, which is 24 bits,  $N_{PBCH,1 \times 1}^{FRM}$  is the number of symbols for the physical broadcast channel (PBCH) per frame in SISO,  $N_{S-SS}^{FRM}$  is the number of secondary synchronization signals per frame, and  $N_{P-SS}^{FRM}$  is the number of primary synchronization signals per frame. The number of data after the CRC encoding,  $N_{data,1 \times 1}^{CRC}$ , can be calculated as

$$N_{data,1 \times 1}^{CRC} = N_{data,1 \times 1}^{DS} + N_{chksun}^{CRC} = 361 + 24 = 385, \quad (8)$$

where  $N_{chksun}^{CRC}$  is the number of CRC checksums. Then, the number of data after turbo encoding,  $N_{data,1 \times 1}^{TURBO}$ , can be calculated as

$$N_{data,1 \times 1}^{TURBO} = \text{TableSearch} (N_{data,1 \times 1}^{CRC}) \times 3 + 12 \quad (9)$$

$$= 492 \times 3 + 12 = 1188,$$

where  $\text{TableSearch}(\cdot)$  is the turbo encoding table. The number of data after QPSK modulation and layer mapping,  $N_{\text{data},1 \times 1}^{\text{MOD}}$ , can be calculated as

$$N_{\text{data},1 \times 1}^{\text{MOD}} = \frac{N_{\text{data},1 \times 1}^{\text{TURBO}}}{N_{\text{mod}}^{\text{ODR}}} = \frac{1188}{2} = 594, \quad (10)$$

where  $N_{\text{mod}}^{\text{ODR}}$  is the modulation order. Finally, the number of null subcarriers after the resource element mapping step,  $N_{\text{spaces},1 \times 1}^{\text{FRM}}$ , can be calculated as

$$N_{\text{spaces},1 \times 1}^{\text{FRM}} = N_{\text{data}}^{\text{N-FRM}} - N_{\text{data},1 \times 1}^{\text{MOD}} = 630 - 594 = 36, \quad (11)$$

where  $N_{\text{data}}^{\text{N-FRM}}$  is the number of data in the normal frame of SISO system. Therefore, there are 36 null subcarriers following the resource element mapping step in PDSCH of the LTE SISO system and it can be used for inserting dummy sequences.

Secondly, the number of null subcarriers per frame is obtained in the PDSCH of the LTE  $2 \times 2$  MIMO system. The number of data per RB in  $2 \times 2$  MIMO,  $N_{\text{data},2 \times 2}^{\text{RB}}$ , can be calculated as

$$\begin{aligned} N_{\text{data},2 \times 2}^{\text{RB}} &= N_{\text{symp}}^{\text{DL}} \times N_{\text{sc}}^{\text{RB}} - N_{\text{pilot}}^{\text{RB}} - N_{\text{u-pilot},2 \times 2}^{\text{RB}} \\ &= 7 \times 12 - 4 - 4 = 76, \end{aligned} \quad (12)$$

where  $N_{\text{u-pilot},2 \times 2}^{\text{RB}}$  is the number of RSs used in other antenna ports per RB, which is 4 in  $2 \times 2$  MIMO system. The number of data per subframe,  $N_{\text{data},2 \times 2}^{\text{S-FRM}}$ , is obtained as

$$N_{\text{data},2 \times 2}^{\text{S-FRM}} = N_{\text{data},2 \times 2}^{\text{RB}} \times 2 - N_{\text{PDCCH},2 \times 2}^{\text{RB}} = 76 \times 2 - 32 = 120, \quad (13)$$

where  $N_{\text{PDCCH},2 \times 2}^{\text{RB}}$  is the number of symbols for the PDCCH per RB in  $2 \times 2$  MIMO system. The number of data per frame in  $2 \times 2$  MIMO,  $N_{\text{data},2 \times 2}^{\text{FRM}}$ , is derived as

$$\begin{aligned} N_{\text{data},2 \times 2}^{\text{FRM}} &= N_{\text{TX}} \times N_{\text{data},2 \times 2}^{\text{S-FRM}} \times N_{\text{RB}}^{\text{LEN}} \times N_{\text{mod}}^{\text{ODR}} \times \text{TB}_{\text{CR}} \\ &= 2 \times 120 \times 5 \times 2 \times \frac{1}{3} = 800, \end{aligned} \quad (14)$$

where  $N_{\text{TX}}$  is the number of antenna which is 2 in  $2 \times 2$  MIMO system. Therefore, the number of data created in the data source step,  $N_{\text{data},2 \times 2}^{\text{DS}}$ , can be obtained as

$$\begin{aligned} N_{\text{data},2 \times 2}^{\text{DS}} &= \left( N_{\text{data},2 \times 2}^{\text{FRM}} - N_{\text{CW},2 \times 2} \times N_{\text{chksum}}^{\text{CRC}} \right. \\ &\quad \left. - N_{\text{PBCH},2 \times 2}^{\text{FRM}} - N_{\text{S-SS}}^{\text{FRM}} - N_{\text{P-SS}}^{\text{FRM}} \right) \times (N_{\text{CW},2 \times 2})^{-1} \\ &= \frac{800 - 1 \times 24 - 22 - 6 - 6}{1} = 742, \end{aligned} \quad (15)$$

where  $N_{\text{CW},2 \times 2}$  is the number of code words in  $2 \times 2$  MIMO system, which is 1, and  $N_{\text{PBCH},2 \times 2}^{\text{FRM}}$  is the number of symbols for PBCH per frame in  $2 \times 2$  MIMO, which is 22. The number

of data after the CRC encoding in  $2 \times 2$  MIMO,  $N_{\text{data},2 \times 2}^{\text{CRC}}$ , can be calculated as

$$N_{\text{data},2 \times 2}^{\text{CRC}} = N_{\text{data},2 \times 2}^{\text{DS}} + N_{\text{chksum}}^{\text{CRC}} = 742 + 24 = 766. \quad (16)$$

Then, the number of data after turbo encoding,  $N_{\text{data},2 \times 2}^{\text{TURBO}}$ , can be derived as

$$\begin{aligned} N_{\text{data},2 \times 2}^{\text{TURBO}} &= \text{TableSearch} \left( N_{\text{data},2 \times 2}^{\text{CRC}} \right) \times 3 + 12 \\ &= 768 \times 3 + 12 = 2316. \end{aligned} \quad (17)$$

The number of data after modulation and layer mapping,  $N_{\text{data},2 \times 2}^{\text{MOD}}$ , is calculated as

$$N_{\text{data},2 \times 2}^{\text{MOD}} = \frac{N_{\text{data},2 \times 2}^{\text{TURBO}} / N_{\text{mod}}^{\text{ODR}}}{N_{\text{TX}}} = \frac{2316/2}{2} = 579. \quad (18)$$

Finally, the number of null subcarriers after the resource element mapping step,  $N_{\text{spaces},2 \times 2}^{\text{FRM}}$ , can be obtained as

$$N_{\text{spaces},2 \times 2}^{\text{FRM}} = N_{\text{data},2 \times 2}^{\text{N-FRM}} - N_{\text{data},2 \times 2}^{\text{MOD}} = 600 - 579 = 21, \quad (19)$$

where  $N_{\text{data},2 \times 2}^{\text{N-FRM}}$  is the number of data in the normal frame of  $2 \times 2$  MIMO system. Therefore, there are 21 null subcarriers.

Thirdly,  $N_{\text{spaces},4 \times 4}^{\text{FRM}}$ , is derived in the PDSCH of the LTE  $4 \times 4$  MIMO system. The number of data per RB is derived as

$$\begin{aligned} N_{\text{data},4 \times 4}^{\text{RB}} &= N_{\text{symp}}^{\text{DL}} \times N_{\text{sc}}^{\text{RB}} - N_{\text{pilot}}^{\text{RB}} - N_{\text{u-pilot},4 \times 4}^{\text{RB}} \\ &= 7 \times 12 - 4 - 8 = 72, \end{aligned} \quad (20)$$

where  $N_{\text{u-pilot},4 \times 4}^{\text{RB}}$  is the number of RSs used in other antenna ports per RB in  $4 \times 4$  MIMO system. The number of data per subframe,  $N_{\text{data},4 \times 4}^{\text{S-FRM}}$ , can be obtained as

$$N_{\text{data},4 \times 4}^{\text{S-FRM}} = N_{\text{data},4 \times 4}^{\text{RB}} \times 2 - N_{\text{PDCCH},4 \times 4}^{\text{RB}} = 72 \times 2 - 28 = 116, \quad (21)$$

where  $N_{\text{PDCCH},4 \times 4}^{\text{RB}}$  is 28 in  $4 \times 4$  MIMO system. The number of data per frame,  $N_{\text{data},4 \times 4}^{\text{FRM}}$ , in  $4 \times 4$  MIMO can be obtained as

$$\begin{aligned} N_{\text{data},4 \times 4}^{\text{FRM}} &= N_{\text{TX}} \times N_{\text{data},4 \times 4}^{\text{S-FRM}} \times N_{\text{RB}}^{\text{LEN}} \times N_{\text{mod}}^{\text{ODR}} \times \text{TB}_{\text{CR}} \\ &= 4 \times 116 \times 5 \times 2 \times \frac{1}{3} = 1546, \end{aligned} \quad (22)$$

where  $N_{\text{TX}}$  is 4. Therefore, the number of data created in the data source step,  $N_{\text{data},4 \times 4}^{\text{DS}}$ , in  $4 \times 4$  MIMO, is derived as

$$\begin{aligned} N_{\text{data},4 \times 4}^{\text{DS}} &= \left( N_{\text{data},4 \times 4}^{\text{FRM}} - N_{\text{CW},4 \times 4} \times N_{\text{chksum}}^{\text{CRC}} \right. \\ &\quad \left. - N_{\text{PBCH},4 \times 4}^{\text{FRM}} - N_{\text{S-SS}}^{\text{FRM}} - N_{\text{P-SS}}^{\text{FRM}} \right) \times (N_{\text{CW},4 \times 4})^{-1} \\ &= \frac{1546 - 2 \times 24 - 20 - 6 - 6}{1} = 733, \end{aligned} \quad (23)$$

TABLE 2: Parameters of the computer simulations.

Parameter	Value
Carrier frequency $f_0$	2140 MHz
Channel bandwidth	2.5 MHz
FFT size	512
Duplex mode	FDD
Cyclic shift	Normal
Modulation type	QPSK
Doppler frequency	119 MHz (velocity = 60 Km/h)
CRC	24 bit
Forward error correction (FEC)	1/3 turbo coding
Number of dummy bit	36

where  $N_{CW,4 \times 4}$  is 2 and  $N_{PBCH,4 \times 4}^{FRM}$  is 20 in  $4 \times 4$  MIMO system. The number of data after the CRC encoding,  $N_{data,4 \times 4}^{CRC}$ , is obtained as

$$N_{data,4 \times 4}^{CRC} = N_{data,4 \times 4}^{DS} + N_{chksun}^{CRC} = 733 + 24 = 757. \quad (24)$$

Then, the number of data after turbo encoding,  $N_{data,4 \times 4}^{TURBO}$ , in  $4 \times 4$  MIMO system can be obtained as

$$\begin{aligned} N_{data,4 \times 4}^{TURBO} &= \text{TableSearch} \left( N_{data,4 \times 4}^{CRC} \right) \times 3 + 12 \\ &= 768 \times 3 + 12 = 2316. \end{aligned} \quad (25)$$

The number of data after QPSK modulation and layer mapping,  $N_{data,2 \times 2}^{MOD}$ , can be derived as

$$N_{data,4 \times 4}^{MOD} = \frac{N_{data,4 \times 4}^{TURBO} / N_{mod}^{ODR}}{N_{TX}} = \frac{2316/2}{4} = 297. \quad (26)$$

Therefore,  $N_{spaces,4 \times 4}^{FRM}$  in  $4 \times 4$  MIMO can be derived as

$$N_{spaces,4 \times 4}^{FRM} = N_{data,4 \times 4}^{N-FRM} - N_{data,4 \times 4}^{MOD} = 300 - 297 = 3, \quad (27)$$

where  $N_{data,4 \times 4}^{N-FRM}$  is the number of data in the normal frame of  $4 \times 4$  MIMO system. Finally, there are 3 null subcarriers following the resource element mapping step in PDSCH of the LTE  $4 \times 4$  MIMO system.

We have derived the null subcarriers for SISO,  $2 \times 2$ , and  $4 \times 4$  MIMO systems, respectively. Since this paper focuses on SISO LTE system, we assume that the maximum dummy subcarrier is 36, which does not decrease the transmission efficiency. When MIMO LTE system is applied, we may sacrifice some decrease in transmission efficiency, which can be defined as

$$\text{transmission efficiency} = \frac{N - (36 - N_{spaces}^{FRM})}{N} \times 100 [\%], \quad (28)$$

where  $N$  is the number of subcarriers and  $N_{spaces}^{FRM}$  is the number of null subcarriers. Therefore, the transmission efficiencies of the  $2 \times 2$  and  $4 \times 4$  MIMO LTE systems are 97% and 94%, respectively, when  $N$  is 512.

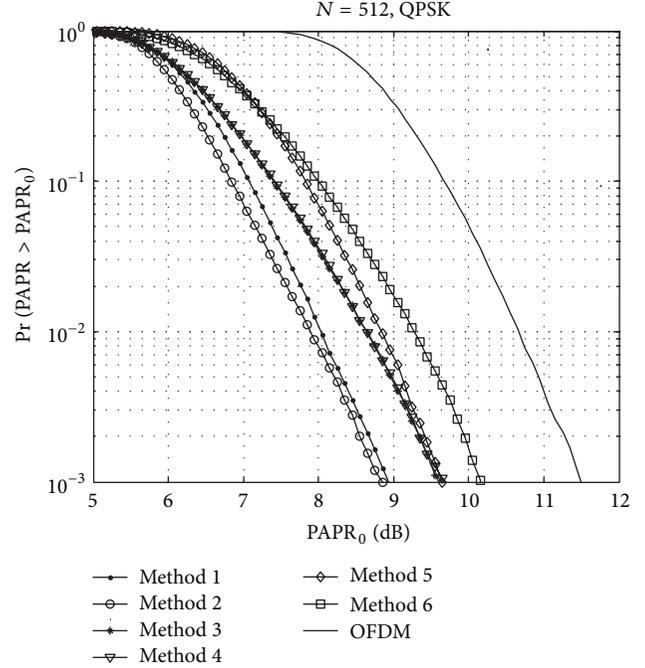


FIGURE 3: CCDF comparison of DSI methods.

#### 4. Simulation Results and Discussion

In the proposed scheme, there is a trade-off between the type and pattern of dummy sequence and the iteration time for the cyclic shift. Therefore, consideration of these elements is an important aspect of PAPR reduction performance and suitable system complexity. In this section, we find the near optimum values for the DSI method, the ratio of dummy sequence for the null subcarriers, and the number of iterations by various simulation results. The simulations are performed under the 3GPP LTE physical layer standard [18, 19]. Table 2 lists the parameters of our simulations.

For the suitable DSI, we compare the PAPR reduction performances of the well-known DSI methods. The DSI methods are briefly introduced as follows.

- (i) Method 1: complementary sequences and correlation sequences corresponding to the first bits of each partitioned subblock are inserted as dummy sequences before the inverse fast Fourier transform (IFFT) stage [16].
- (ii) Method 2: WHT is inserted as a dummy sequence before the IFFT stage [20].
- (iii) Method 3: time-frequency domain swapping algorithm and flipping technique are used to optimize the phase of dummy sequences [21].
- (iv) Method 4: every initial dummy sequence is “0” and employs bit flipping method to generate dummy sequences for next branch [22].
- (v) Method 5: the total sequences consist of  $L$  length data sequences and  $M$  length dummy sequences. A dummy bit is inserted at the end of the sequences for

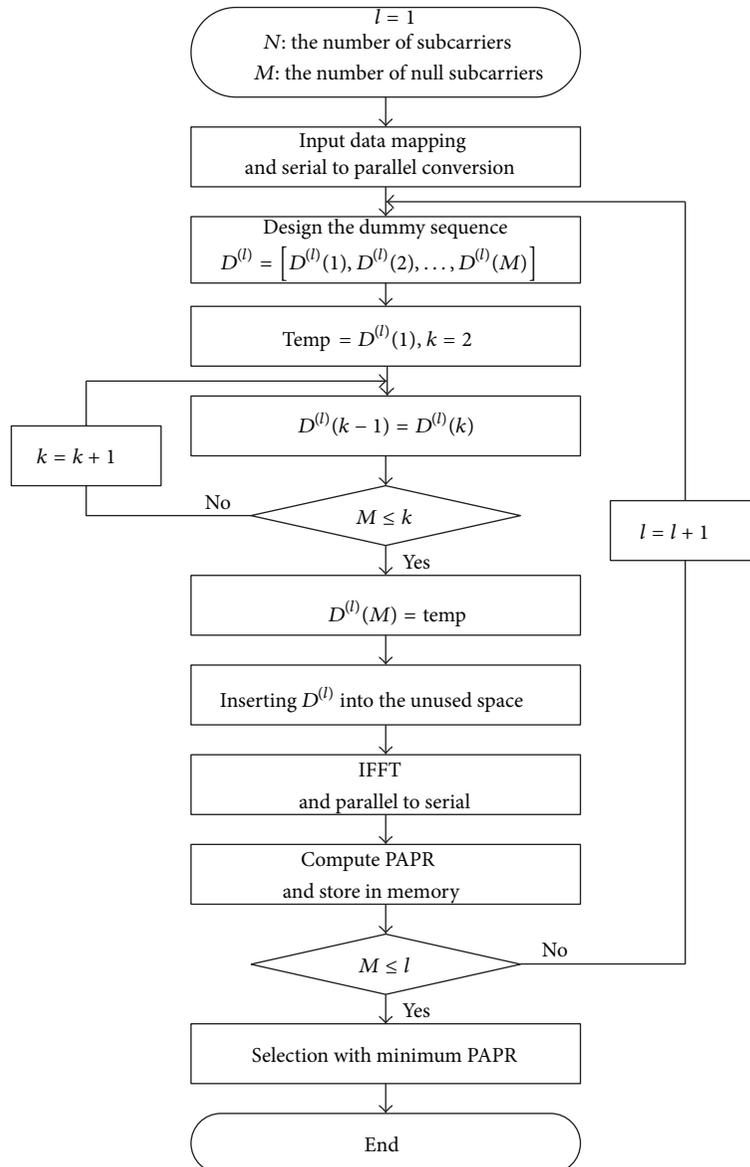


FIGURE 4: Flow chart of scheduling the null subcarriers for peak power reduction.

binary addition between adjacent bits  $(L + M)/2 + 1$  and  $L + M$ . At the same time, the  $N$ -point IFFT block is divided into two sub-IFFT blocks to reduce the IFFT complexity [23].

- (vi) Method 6: a partial DSI method is a combination of the DSI and the PTS. The original data sequences are partitioned and zero padded, and a “0” or “1” dummy sequence is inserted into each subblock. The time domain waveforms are summed after IFFT, and the sequence with the lowest PAPR is selected and transmitted [24].

Figure 3 shows the CCDF comparison among the six kinds of the DSI methods. Considering PAPR reduction performance, we can conclude that Method 2, that is, WHT sequences, is the best choice for the dummy sequence.

As we analyzed in Section 3, there are 36 null subcarriers following the resource element mapping step in the PDSCH of the LTE SISO system. The essence of the proposed scheme is making full use of the null subcarriers, which have to be designed for optimal PAPR reduction performance. It is derived that WHT is the most suitable for the dummy sequences. Since the length of WHT is  $2^n = 2, 4, 8, 16, 32 \leq 36$ ,  $(n = 1, 2, 3, 4, 5)$ , the 36 null subcarriers are partitioned into two parts. The first part is for inserting WHT and the second part is for inserting randomly generated sequences with  $\{-1, 1\}$  elements as a dummy sequence. Then, 36 null subcarriers are cyclic shifted with  $l$  time iterations to find the minimum PAPR. Figure 4 shows the flow chart of the proposed method.

Since PAPR performance is affected by the patterns of the WHT, the PAPR performances of the proposed method are

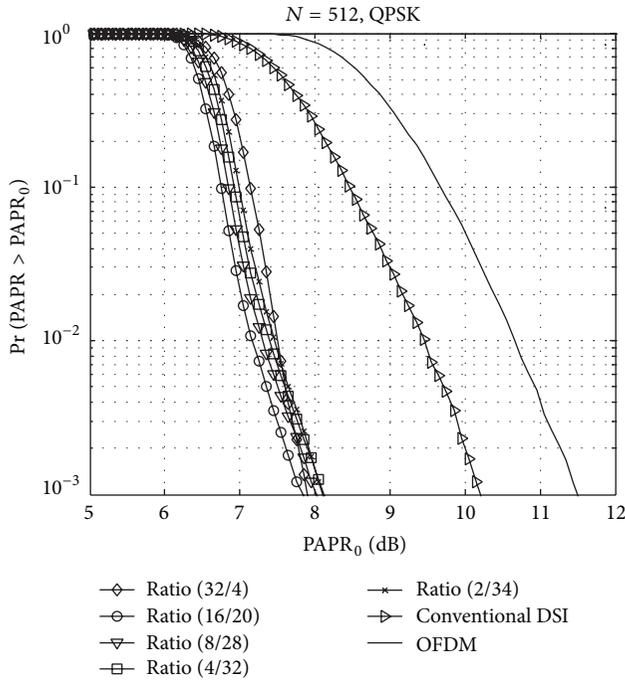


FIGURE 5: CCDF comparison as a function of WHT and randomly generated sequence ratio.

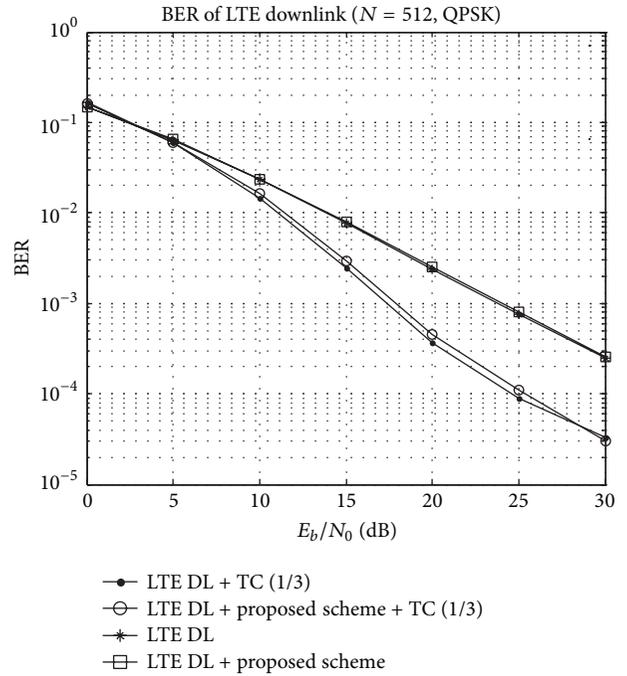


FIGURE 7: BER performance of the proposed scheme and LTE system.

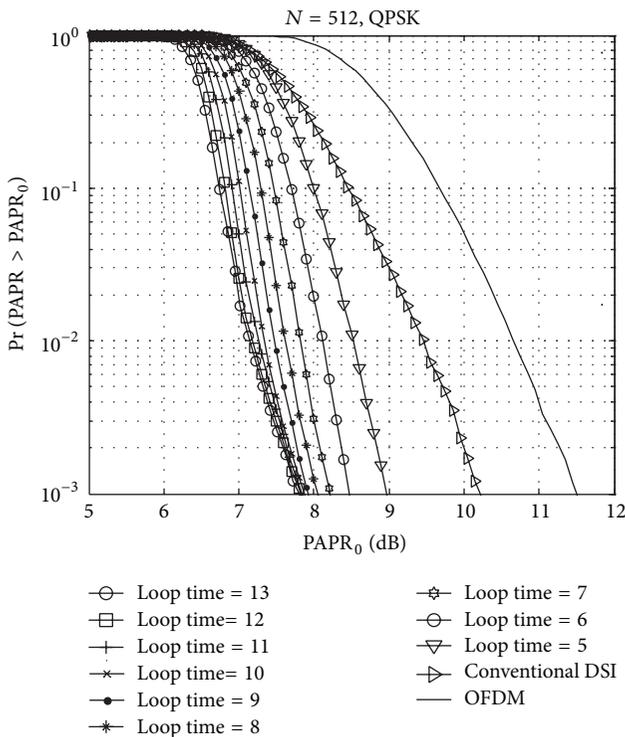


FIGURE 6: CCDF comparison over the number of iterations.

compared by the ratio of the WHT and randomly generated sequence. The five patterns of the WHT and randomly generated sequence can be in the ratio of 32/4, 16/20, 8/28, 4/32, and 2/34.

Figure 5 shows the CCDF of the PAPR of the five kinds of null subcarrier design with full iteration ( $l = M$ ). Since the ratio of 16/20 has the best PAPR reduction performance, we can conclude that the ratio of 16/20 is the optimal choice for the proposed method.

In order to approach a more efficient PAPR reduction within the limited null subcarriers, cyclic shifting is used with WHT sequences. Multiple iteration operations for the cyclic shifting, however, cause the high computational complexity of the LTE system. Therefore, we need to consider the cyclic shift loop times to approach the minimum computational complexity. The PAPR performances of various cyclic shift loop times are compared in Figure 6. Obviously, the bigger cyclic shift loop time has more PAPR reduction performance. Nevertheless, PAPR performances are saturated to about 7.9 dB. Therefore, we can conclude that 9 may be the nearly optimal number of iterations.

In addition to the PAPR comparison, we examine the BER performance of the proposed method in the LTE downlink system. Simulation is performed under Rayleigh fading channel, and the turbo coding is used with a coding rate of  $R = 1/3$ . As shown in Figure 7, the conventional LTE downlink system and the proposed method have nearly the same BER performance. Therefore, the proposed scheme can reduce the PAPR for the LTE downlink system considerably without the degradation of the BER performance.

## 5. Conclusions

This paper proposed a novel DSI scheme for the LTE downlink system. For the application of the DSI to LTE system, the null subcarriers were obtained in LTE SISO,  $2 \times 2$ , and  $4 \times 4$  MIMO systems, respectively, and each transmission efficiency was calculated. The dummy sequence was designed by scheduling the ratio between WHT and random sequences. The number of near optimal iteration and BER performances were derived, which showed that exhausted iterations could be prevented and proposed DSI can reduce PAPR without BER degradation.

The future works will derive the number of subcarriers in LTE-Advanced and  $8 \times 8$  MIMO systems. To overcome the PAPR reduction performance with the limited null subcarrier, the other dummy sequences will be applied. In addition, new algorithm will be researched to reduce the iteration time or eliminate it completely for the more realistic system.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

- [1] F. Khan, *LTE for 4G Mobile Broadband*, Cambridge University Press, New York, NY, USA, 2009.
- [2] R. Prasad, *OFDM for Wireless Communication Systems*, Artech House, Boston, Mass, USA, 2004.
- [3] M. D. Nava and G. S. Ökvist, "The zipper prototype: a complete and flexible VDSL multicarrier solution," *IEEE Communications Magazine*, vol. 40, no. 12, pp. 92–105, 2002.
- [4] H. Kubota, K. Suzuki, I. Kawakami, M. Sakugawa, and H. Kondo, "High frequency band dispersed-tone power line communication modem for networked appliances," *IEEE Transactions on Consumer Electronics*, vol. 52, no. 1, pp. 44–50, 2006.
- [5] L. Li, L. Ma, Y. Xu, and Y. Fu, "Motion adaptive vertical handoff in cellular/WLAN heterogeneous wireless network," *The Scientific World Journal*, vol. 2014, Article ID 341038, 7 pages, 2014.
- [6] H. Liu and Z. Xu, "Design of UWB monopole antenna with dual notched bands using one modified electromagnetic-bandgap structure," *The Scientific World Journal*, vol. 2013, Article ID 917965, 9 pages, 2013.
- [7] S. Rajagopal, R. D. Roberts, and S.-K. Lim, "IEEE 802.15.7 visible light communication: modulation schemes and dimming support," *IEEE Communications Magazine*, vol. 50, no. 3, pp. 72–82, 2012.
- [8] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wireless Communications*, vol. 12, no. 2, pp. 56–65, 2005.
- [9] X. Li and L. J. Cimini Jr., "Effects of clipping and filtering on the performance of OFDM," *IEEE Communications Letters*, vol. 2, no. 5, pp. 131–133, 1998.
- [10] S. H. Müller and J. B. Huber, "OFDM with reduced peak-to-average power ratio by optimum combination of partial transmit sequences," *Electronics Letters*, vol. 33, no. 5, pp. 368–369, 1997.
- [11] R. W. Bäuml, R. F. H. Fischer, and J. B. Huber, "Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping," *Electronics Letters*, vol. 32, no. 22, pp. 2056–2057, 1996.
- [12] A. D. S. Jayalath and C. Tellambura, "Reducing the peak-to-average power ratio of orthogonal frequency division multiplexing signal through bit or symbol interleaving," *Electronics Letters*, vol. 36, no. 13, pp. 1161–1163, 2000.
- [13] A. E. Jones, T. A. Wilkinson, and S. K. Barton, "Block coding scheme for reduction of peak to mean envelope power ratio of multicarrier transmission schemes," *Electronics Letters*, vol. 30, no. 25, pp. 2098–2099, 1994.
- [14] J. Tellado, *Peak to average power reduction for multicarrier modulation [Ph.D. dissertation]*, Stanford University, 2000.
- [15] B. S. Krongold and D. L. Jones, "PAR reduction in OFDM via active constellation extension," *IEEE Transactions on Broadcasting*, vol. 49, no. 3, pp. 258–268, 2003.
- [16] H.-G. Ryu, J.-E. Lee, and J.-S. Park, "Dummy Sequence Insertion (DSI) for PAPR reduction in the OFDM communication system," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 1, pp. 89–94, 2004.
- [17] S. Cho, S. K. Park, and D. J. Kwon, "Utilization of null subcarriers for PAPR reduction in 3GPP LTE downlink," in *Proceedings of the 3rd IEEE International Conference on Network Infrastructure and Digital Content (IC-NIDC '12)*, pp. 54–56, September 2012.
- [18] 3GPP TS 36.211, "Physical channels and modulation (Release 9)," December 2009.
- [19] 3GPP TS 36.201, "LTE physical layer—general description (Release 9)," December 2009.
- [20] S.-W. Kim, J.-K. Chung, and H.-G. Ryu, "PAPR reduction of the OFDM signal by the SLM-based WHT and DSI method," in *Proceedings of the 10th IEEE Region Conference (TENCON '06)*, pp. 1–4, November 2006.
- [21] P. Boonsrimuang, K. Mori, T. Paungma, and H. Kobayashi, "PAPR reduction method for OFDM signal by using dummy sub-carriers," in *Proceedings of the 1st International Symposium on Wireless Pervasive Computing (ISWPC '06)*, pp. 1–5, January 2006.
- [22] S. W. Kim, H. S. Byeon, J. K. Kim, and H.-G. Ryu, "An SLM-based real-time PAPR reduction method using dummy sequence insertion in the OFDM communication," in *Proceedings of the 5th International Conference on Information, Communications and Signal Processing (ICICS '05)*, pp. 258–262, December 2005.
- [23] J.-K. Lee, J.-S. Park, and J.-U. Kim, "Modified dummy sequence insertion method for PAPR reduction of OFDM signal," in *Proceedings of the 66th IEEE Vehicular Technology Conference (VTC '07)*, pp. 1265–1268, October 2007.
- [24] C.-M. Li, J.-C. Wu, C.-C. Tseng, I.-T. Tang, and Y.-C. Chang, "Performance comparisons of PAPR reduction methods for the OFDM system," in *Proceedings of the IEEE International Symposium on Industrial Electronics (ISIE '09)*, pp. 1413–1416, July 2009.

## Research Article

# Balance Maintenance in High-Speed Motion of Humanoid Robot Arm-Based on the 6D Constraints of Momentum Change Rate

Da-song Zhang, Rong Xiong, Jun Wu, and Jian Chu

National Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, Zhejiang 310027, China

Correspondence should be addressed to Rong Xiong; rxiong@iipc.zju.edu.cn

Received 18 January 2014; Accepted 9 February 2014; Published 17 April 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Da-song Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Based on the 6D constraints of momentum change rate (CMCR), this paper puts forward a real-time and full balance maintenance method for the humanoid robot during high-speed movement of its 7-DOF arm. First, the total momentum formula for the robot's two arms is given and the momentum change rate is defined by the time derivative of the total momentum. The author also illustrates the idea of full balance maintenance and analyzes the physical meaning of 6D CMCR and its fundamental relation to full balance maintenance. Moreover, discretization and optimization solution of CMCR has been provided with the motion constraint of the auxiliary arm's joint, and the solving algorithm is optimized. The simulation results have shown the validity and generality of the proposed method on the full balance maintenance in the 6 DOFs of the robot body under 6D CMCR. This method ensures 6D dynamics balance performance and increases abundant ZMP stability margin. The resulting motion of the auxiliary arm has large abundance in joint space, and the angular velocity and the angular acceleration of these joints lie within the predefined limits. The proposed algorithm also has good real-time performance.

## 1. Introduction

Humanoid robot is characterized as a high-order, nonlinear, and multiple DOF unstable system. Research topics such as motion planning and balance maintenance for humanoids remain a great challenge. After years of development, there have already been several humanoid robot systems with outstanding performance, such as WABIAN of Waseda University [1], ASIMO of Honda Inc. [2], QRIO of Sony Inc. [3], HRP of AIST [4], and HUBO of South Korea [5]. With the advancing progress in both theoretic study and engineering practice, humanoid robot is evolving towards a higher level of intelligence and mobility. Nowadays humanoids can display more complex motions, such as pushing a table [6, 7], fetching and manipulating small objects [8], assembling under industrial environment [9], lifting a heavy object with human help [10], and dancing featuring with whole-body movement [11]. These various complicated dynamic operations bring the robot body time-varying impact in both force and moment, which makes the balance maintenance a critical and important research issue for the humanoid robots.

Balance maintenance is the key to successful humanoid robot, which is fundamentally unstable. Instinct makes us believe that balance comes along with low velocity; in other words, low momentum can guarantee balance. However, we also notice that high-speed trains, airplanes, and rockets can operate steadily while their momentum is large. Why do we have the illusion that balance comes along with lower momentum? As a matter of fact, objects with lower momentum usually need a lower momentum change rate when they switch from a moving state into stationary within a unit period of time or vice versa. Put in other words, the momentum change rate is smaller. In fact, the momentum change rate of high-speed trains, airplanes, and rockets in their moving process is either quite low or is rigorously controlled within certain safe limits, even though their momentum is extremely large. For instance, the high-speed trains will cover a long distance to accelerate or decelerate when it is starting or braking, which actually is to decrease the acceleration or the deceleration; in other words, the momentum change rate of the train in the whole process of its drive is controlled within a small range. Moreover, planes and rockets are often equipped with accelerometers, which are used to

monitor their acceleration and deceleration and thus help control them within a certain safe range. Serious accidents caused by full braking are in fact resulted by the overly great deceleration; in other words, the momentum change rate is too big. Therefore, the key to perform balanced and steady motion lies in acceleration control or the momentum change rate control more strictly. For humanoid robots, the essence of balance maintenance also depends on the regulation of the momentum change rate. Equally speaking, the key is the control of the ground reaction force and moment to the robot.

Based on the constraints of momentum change rate, this paper sets out to study the real-time balance maintenance utilizing the auxiliary arm of the robot during high-speed arm motions. Different from the industrial robot, which is located fixed on the ground, the humanoid robot is “pressed” to the ground by gravity. Thus, the 6D time-varying inertia impact generated by the high-speed motion of the robot arm might cause the robot tilt, slide, or even jump, which leads to the shifting of the end point of the operating arm and failure of the manipulator task. The force and moment which the robot exerts to the ground should be controlled within a safe range by properly designed motion of the auxiliary arm and the resultant force and moment impact due to the fact that high-speed motion of the operating arm should be eliminated or reduced, so as to guarantee the accurate and stable operation of the humanoid robot.

A balance maintenance method based on resolved momentum control has been adopted in prior studies [12, 13]; however, the high-speed motion of the robot arm only makes the method applicable to the balance maintenance on no more than two dimensions and inapplicable on three or more dimensions. According to Newton’s second law and momentum conservation principle, the authors put forward an optimal balance maintenance method that generates proper motions of the auxiliary arm’s joint based on the constraints of momentum change rate (CMCR). From the perspective of CMCR, the authors first give the formula to calculate the total momentum of the motion of the robot’s two arms. Then the fundamental relation between 6D momentum change rate and full balance is specifically examined. In addition, the discretization process of full balance maintenance under the constraint of 6D momentum change rate is provided along with the convex quadratic programming equation constrained by inequalities. The angular velocity and angular acceleration constraints on the joints of the auxiliary arm are considered at the same time. The angular velocity vector of optimal joint motion of the auxiliary arm under CMCR and the constraints of auxiliary arm’s motion are also introduced. Lastly, the authors explain the methods taken to increase the algorithm speed. The simulation experiments show that the algorithm is effective and has good real-time performance.

## 2. Related Works

The balance maintenance method in the early years was centered on static balance maintenance, which keeps the projection of center of gravity within the support area. In the year of 1972, the Yugoslav scholar Dr. Vukobratovic put

forward the ZMP (Zero-Moment Point) theory [14], which has become the basic theoretic foundation of the dynamic balance maintenance of humanoid robots. At present, ZMP control method has been widely applied in the dynamic walking balance maintenance of humanoid robot, for example, the most famous ASIMO. ASIMO, equipped with the world’s most advanced technologies, features a systematic ZMP control scheme, including ground reactive force control, model ZMP control, and foot landing position control [2].

In recent years, researchers from different fields have undertaken deep studies on motion planning and balance maintenance from various perspectives as humanoid robot become more and more popular. In particular, balance maintenance of humanoid robot based on momentum control has caught great attention. As for linear momentum control, balance maintenance generally requires that the projection of the center of gravity falls within the support area by controlling the linear momentum [15, 16]. From the perspective of angular momentum control, it has been learned that the central nervous system of human beings would intentionally adjust the angular momentum with respect to the center of gravity while walking, and thus the method of minimizing the angular momentum is proposed [17, 18]. Besides, the stability of the robot is also evaluated using the criterion that the rate of angular momentum change should be zero [19].

To achieve better balance maintenance performance, Professor Shuji Kajita at AIST (National Institute of Advanced Industrial Science and Technology) of Japan put forward in 2003 the resolved momentum control scheme considering both the linear momentum and angular momentum based on his long-time research results and technological accumulations [13]. He applied this control strategy to achieve whole-body autonomous balance maintenance of the humanoid robot HRP-2 under remote operation. Several years later, he succeeded in autonomous balance of the robot when it picked objects up from the ground [12].

However, it has been found in practice that (1) when the resolved momentum control is applied to high-speed 6D full balance maintenance, the angular velocity of the auxiliary arm may exceed its limit; and the overrun is so great that it further leads to the overrun of angular acceleration of the joints immediately; (2) when the resolved momentum control is applied to high-speed and low-dimensional balance maintenance problem, the effect of the balance maintenance is excellent in the specified dimensions, but has no positive result on the other dimensions. For instance, when ZMP is regarded as the performance index of the balance maintenance problem, in other words the forward and backward tilt and left and right tilt dimensions are considered, the best ZMP balance can be achieved using the resolved momentum control and the tilting of the robot in these dimensions is well controlled. Nevertheless, the balance of other dimensions not considered in the scheme is totally lost, which leads the robot to slip and jump. The close relation between humanoid robot’s balance and its momentum change rate is clearly illustrated by Newton’s second law. The authors propose an optimal, real-time, all-dimensional balance maintenance method, based on CMCR, to meet the needs of the 7-DOF arm of the robot to perform high-speed dynamic operation safely.

### 3. Balance Maintenance of the Auxiliary Arm

3.1. *Total Momentum Calculation.* This paper assumes that the humanoid robot moves only its two arms to perform certain tasks while the other parts of the body are fixed with no movement. Therefore, the total momentum of the robot equals the total momentum of its two arms. As shown in Figure 1, the origin of the reference coordinate is selected as the central point  $w$  of the two feet. By referring to the resolved momentum control, the total momentum, including linear momentum  $P$  and angular momentum  $L$ , can be defined as [13]

$$\begin{bmatrix} P \\ L \end{bmatrix} = \begin{bmatrix} M_{\text{jobarm}} \\ H_{\text{jobarm}} \end{bmatrix} \dot{\theta}_{\text{jobarm}} + \begin{bmatrix} M_{\text{aidarm}} \\ H_{\text{aidarm}} \end{bmatrix} \dot{\theta}_{\text{aidarm}}. \quad (1)$$

Specifically,  $P = [P_x \ P_y \ P_z]^T$  denotes the three-dimensional linear momentum of the robot;  $L = [L_x \ L_y \ L_z]^T$  denotes the three-dimensional angular momentum.  $M_{\text{jobarm}}$  and  $H_{\text{jobarm}}$  denote the inertia matrixes of the operating arm and  $\dot{\theta}_{\text{jobarm}}$  denotes the velocity vector of the joints in the operating arm;  $M_{\text{aidarm}}$  and  $H_{\text{aidarm}}$  denote the inertia matrixes of the auxiliary arm while  $\dot{\theta}_{\text{aidarm}}$  denotes the velocity vector of the joints in the auxiliary arm.

3.2. *Full Balance and Six-Dimensional CMCR.* The humanoid robot has six degrees of freedom; and, thus, the robot can achieve full balance only if it is well balanced and stable at every DOF. In other words, the robot is not expected to tilt or slide backward and forward or left and right, to jump vertically, or to rotate around the vertical axis.

The 6D CMCR is exactly corresponded to the 6 DOF balance maintenance of the humanoid robot. In the following part, the authors will illustrate the relation between the 6D CMCR and the 6 DOF balance maintenance, as well as how the full balance maintenance is achieved. It is straightforward that according to Newton's second law the momentum change rate equals the external force exerted by system. Since Newton's second law can also be applied to the generalized 6D form of force and moment, as discussed in this case, the momentum change rate of the humanoid robot has 6 dimensions of which the number equals the ground reactive force and moment acts on the robot feet. Specially, the robot will lose its balance and fails at its task, when the planned momentum change rate exceeds the maximum force and moment that can be exerted by the ground. The total momentum change rate of the robot can be formulated by time derivative of total momentum by using

$$\begin{aligned} \frac{d \begin{bmatrix} P \\ L \end{bmatrix}}{dt} &= \begin{bmatrix} \frac{dP_x}{dt} & \frac{dP_y}{dt} & \frac{dP_z}{dt} & \frac{dL_x}{dt} & \frac{dL_y}{dt} & \frac{dL_z}{dt} \end{bmatrix}^T \\ &= [\bar{F}_x \ \bar{F}_y \ \bar{F}_z \ \bar{M}_x \ \bar{M}_y \ \bar{M}_z]^T. \end{aligned} \quad (2)$$

Specifically,  $t$  denotes time;  $dP_x/dt$ ,  $dP_y/dt$ ,  $dP_z/dt$ ,  $dL_x/dt$ ,  $dL_y/dt$ , and  $dL_z/dt$  denote the momentum change rate of the six DOFs when the robot is moving.  $\bar{F}_x, \bar{F}_y, \bar{F}_z, \bar{M}_x, \bar{M}_y$ , and  $\bar{M}_z$  represent the forces and moments the ground

should provide to the robot feet. Obviously, the maximum force and moment the ground can provide to the robot by its feet is limited within a certain range, which is the CMCR range of the six DOFs. Note that the momentum change rate and the force and moment have the same dimension and are fundamentally identical. The constraint range in this paper is described by force and moment style while the momentum change rate style is also used. Therefore, the authors set out to regulate the force and moment within its constraints by calculating the momentum and controlling the momentum change rate, with the ultimate goal to achieve motion balance of the robot.

The authors will discuss and analyze the importance of momentum change rate at every dimension on balance, so as to find out the fundamental links between balance and momentum change rate.

(1)  $dP_z/dt$  denotes the vertical component of the momentum change rate, which is equal to  $\bar{F}_z$  that the force the ground acts on the robot, representing the robot's vertical overweight and weight loss. The pressing force of the robot on the ground and the supporting force of the ground on the robot are a pair of interaction force. Specifically, the pressing force is  $G + dP_z/dt$  with  $G$  being the weight of the robot. If  $dP_z/dt > 0$ , the pressing force exceeds the robot's weight, and the robot presses the ground with its overweight. In this situation, the robot can provide greater static friction and can resist a greater tilting moment. Generally speaking, overweight is conducive to the robot to be steady; that is to say, the greater the  $dP_z/dt$  the better. However, it has also been found that the greater overweight can exert negative effect on the robot, for it could enlarge the force on the robot structure and the moment on the joints. Thus,  $dP_z/dt < 0.5G$  is defined.

If  $dP_z/dt < 0$ , the pressing force is less than the weight of the robot; the robot loses its weight and attempts to escape from the ground. In this situation, the maximum static friction the ground can provide and the tilting moment it can resist will both decrease. Particularly, when  $dP_z/dt = -G$ , the pressing force of the robot on the ground is zero. In this situation, the robot will leave the ground for weightlessness and thus is out of control. Since weightlessness is harmful to the balance of the robot, it needs to be under more rigorous control.  $dP_z/dt > -0.3G$  is suggested.

Therefore, the vertical component of momentum change rate is constrained within the range as

$$-0.3G < \frac{dP_z}{dt} < 0.5G. \quad (3)$$

(2)  $dL_x/dt$  and  $dL_y/dt$  denote the tilting components of the momentum change rate in the directions of left and right and forward and backward, which are equal to moments  $\bar{M}_x$  and  $\bar{M}_y$  that the ground acts to the robot, representing to what degree the robot will tilt in the directions of left and right and forward and backward. When the tilting component of the momentum change rate is too large, the robot tends to tilt or even overturn. The degree of tilting is related to the tilting moment and the support margin in the specific direction. In practice, due to the flexibility of the robot structure and its joints, the tilting component of the momentum change rate

will cause the robot swinging forward and backward or left and right. And the robot will actually tilt when it swings too violently.

The influence of tilting component of the momentum change rate on the robot's stability can be demonstrated by the disturbance on the ZMP point of the robot and defined as follows:

$$\begin{aligned}\Delta ZMP_x &= \frac{-dL_y/dt}{G + dP_z/dt}, \\ \Delta ZMP_y &= \frac{dL_x/dt}{G + dP_z/dt}.\end{aligned}\quad (4)$$

Specifically,  $\Delta ZMP_x$  and  $\Delta ZMP_y$  are the shift of the ZMP point on the ground according to the tilting component of the momentum change rate. When the shift is so great that the ZMP point locates out of the support area on the ground, the robot is unstable, according to ZMP stability principle. In practice, the closer the ZMP point is to the edge of the support area, the less stable the robot will be. It is obvious that the dramatic changes of the tilting moment component will cause the ZMP point to deviate greatly from its original position. And at the same time, the robot will begin to shake, which will lead to actual tilting if the shake is too strong.

Assume that initially the ground projection of the robot's center of gravity is at the center of the support area and the length of the feet's support area in left-right plane and the forward-backward plane are defined as  $A$  and  $B$ ; then

$$\begin{aligned}-\frac{B}{2} < \Delta ZMP_x < \frac{B}{2}, \\ -\frac{A}{2} < \Delta ZMP_y < \frac{A}{2}.\end{aligned}\quad (5)$$

The constraint range of the tilting component of the momentum change rate is shown in (6) by taking (4) into (5):

$$\begin{aligned}-\frac{A}{2} \left( G + \frac{dP_z}{dt} \right) < \frac{dL_x}{dt} < \frac{A}{2} \left( G + \frac{dP_z}{dt} \right), \\ -\frac{B}{2} \left( G + \frac{dP_z}{dt} \right) < \frac{dL_y}{dt} < \frac{B}{2} \left( G + \frac{dP_z}{dt} \right).\end{aligned}\quad (6)$$

(3)  $dP_x/dt$  and  $dP_y/dt$  refer to the transitional sliding components of the momentum change rate to the forward and backward and the left and right, which are equal to  $\tilde{F}_x$  and  $\tilde{F}_y$  that the ground acts to the robot, representing to what degree the robot will transitionally slide on the ground. When the horizontal component of the momentum change rate is too great, the maximum static friction the ground can provide is not sufficient to support the horizontal momentum change. In this situation, the robot will slide transitionally.

The rigorous conditions on which the robot will not slide transitionally are defined as follows:

$$\left( \frac{dP_x}{dt} \right)^2 + \left( \frac{dP_y}{dt} \right)^2 < \mu^2 \left( G + \frac{dP_z}{dt} \right)^2. \quad (7)$$

Specifically,  $\mu$  is the friction coefficient between the robot's feet and the ground. When the transitional sliding

components of the momentum change rate do not satisfy this condition, the robot body will slide transitionally.

According to the equation, (7) is a nonlinear inequality constraint and should be linearized for further discussion. According to common practice, the inequality circle constraint is replaced by the internal access square constraint, as shown in Figure 2.

Therefore, (7) can be replaced by the following linear inequality. It is obvious that the linearization is conservative, and the constraint range of transitional sliding of the momentum change rate is

$$\begin{aligned}-\frac{\sqrt{2}}{2} \mu \left( G + \frac{dP_z}{dt} \right) < \frac{dP_x}{dt} < \frac{\sqrt{2}}{2} \mu \left( G + \frac{dP_z}{dt} \right), \\ -\frac{\sqrt{2}}{2} \mu \left( G + \frac{dP_z}{dt} \right) < \frac{dP_y}{dt} < \frac{\sqrt{2}}{2} \mu \left( G + \frac{dP_z}{dt} \right).\end{aligned}\quad (8)$$

(4)  $dL_z/dt$  is the rotating component of momentum change rate around the vertical axis, which is equal to moment  $\tilde{M}_z$  that the ground acts to the robot, representing to what degree the robot will rotate around the vertical axis. When the rotating component of momentum change rate is too great, the maximum static friction the ground can provide is not enough to support the robot's momentum change in the rotation direction. In this situation, the robot will rotate around the vertical axis.

The accurate calculation of the maximum friction moment the ground can provide is fairly complex. As an engineering practice, the distributed pressing force the robot's two feet exert to the ground is assumed to be equal to the concentrated force exerted through the two feet's center point acts to the ground. Thus, the maximum friction force allowed is defined as

$$\frac{\mu C}{2} \left( G + \frac{dP_z}{dt} \right). \quad (9)$$

Specifically,  $C$  is the distance between the centers of the two feet of the robot. Thus, the constraint range of the rotating component of the momentum change rate is

$$-\frac{\mu C}{2} \left( G + \frac{dP_z}{dt} \right) < \frac{dL_z}{dt} < \frac{\mu C}{2} \left( G + \frac{dP_z}{dt} \right). \quad (10)$$

Based on the analyses above, the inequalities of the 6D momentum change rate to realize full balance maintenance of the humanoid robot are defined as

$$\begin{aligned}-\frac{\sqrt{2}}{2} \mu \left( G + \frac{dP_z}{dt} \right) < \frac{dP_x}{dt} < \frac{\sqrt{2}}{2} \mu \left( G + \frac{dP_z}{dt} \right) \\ -\frac{\sqrt{2}}{2} \mu \left( G + \frac{dP_z}{dt} \right) < \frac{dP_y}{dt} < \frac{\sqrt{2}}{2} \mu \left( G + \frac{dP_z}{dt} \right) \\ -0.3G < \frac{dP_z}{dt} < 0.5G\end{aligned}$$

$$\begin{aligned}
 -\frac{A}{2} \left( G + \frac{dP_z}{dt} \right) &< \frac{dL_x}{dt} < \frac{A}{2} \left( G + \frac{dP_z}{dt} \right) \\
 -\frac{B}{2} \left( G + \frac{dP_z}{dt} \right) &< \frac{dL_y}{dt} < \frac{B}{2} \left( G + \frac{dP_z}{dt} \right) \\
 -\frac{\mu C}{2} \left( G + \frac{dP_z}{dt} \right) &< \frac{dL_z}{dt} < \frac{\mu C}{2} \left( G + \frac{dP_z}{dt} \right).
 \end{aligned}
 \tag{11}$$

It can be concluded from the inequalities of full balance conditions above that (1) either the tilting stableness represented by traditional ZMP or the jumping and sliding stableness is related to the momentum change rate of the corresponding dimension and (2) balance in any dimension is confined to a zone instead of a point. In other words, instability will occur when the momentum change rate is out of the zone. With smaller zone locating at the center of the support area, the bigger balance margin will bring stronger stability in practice. When the zone is narrowed down to a point, it falls into the traditional stability maintenance concept. For instance, in the direction of tiling, when the zone is narrowed down to a point, it becomes the classical ZMP control. Most interestingly, though jumping in one direction, tiling in two directions, and sliding in three directions look like different phenomena, they can be defined in the same mathematical and physical constraint inequality.

By observing inequalities (11), we can find the overweight and weight loss of the robot coupled with the other five dimensions. To get rid of the coupling and to take into account the robot modeling errors and external disturbance, we will further tighten the constraint based on specific situations to get bigger balance margin. At first, weight loss coefficient  $\alpha$  is added, and  $-\alpha G \leq dP_z/dt \leq 0.4G$  and  $0 \leq \alpha < 0.3$  are defined.  $dP_z/dt = -\alpha G$  is then taken into the five other constraints to get the constraint inequalities of the practical 6D constraints of momentum change rate:

$$\begin{aligned}
 -\frac{\sqrt{2}}{2} \mu (1 - \alpha) G &\leq \frac{dP_x}{dt} \leq \frac{\sqrt{2}}{2} \mu (1 - \alpha) G \\
 -\frac{\sqrt{2}}{2} \mu (1 - \alpha) G &\leq \frac{dP_y}{dt} \leq \frac{\sqrt{2}}{2} \mu (1 - \alpha) G \\
 -\alpha G &\leq \frac{dP_z}{dt} \leq 0.4G \\
 -\frac{A}{2} (1 - \alpha) G &\leq \frac{dL_x}{dt} \leq \frac{A}{2} (1 - \alpha) G \\
 -\frac{B}{2} (1 - \alpha) G &\leq \frac{dL_y}{dt} \leq \frac{B}{2} (1 - \alpha) G \\
 -\frac{\mu C}{2} (1 - \alpha) G &\leq \frac{dL_z}{dt} \leq \frac{\mu C}{2} (1 - \alpha) G.
 \end{aligned}
 \tag{12}$$

Obviously, if inequalities in (12) are satisfied, inequalities in (11) must hold. And inequalities in (12) provide larger stability margins to get better balance performance.

The inequalities in (12) are transformed as follows for the convenience of further mathematical processing:

$$-\begin{bmatrix} F^- \\ M^- \end{bmatrix} \leq \frac{d \begin{bmatrix} P \\ L \end{bmatrix}}{dt} \leq \begin{bmatrix} F^+ \\ M^+ \end{bmatrix},
 \tag{13}$$

where

$$\begin{aligned}
 F^- &= \begin{bmatrix} \frac{\sqrt{2}}{2} \mu (1 - \alpha) G & \frac{\sqrt{2}}{2} \mu (1 - \alpha) G & \alpha G \end{bmatrix}^T, \\
 F^+ &= \begin{bmatrix} \frac{\sqrt{2}}{2} \mu (1 - \alpha) G & \frac{\sqrt{2}}{2} \mu (1 - \alpha) G & 0.4G \end{bmatrix}^T,
 \end{aligned}
 \tag{14}$$

where  $F^-$  and  $F^+$  are 3D forces. Consider

$$M^- = M^+ = \begin{bmatrix} \frac{A}{2} (1 - \alpha) G & \frac{B}{2} (1 - \alpha) G & \frac{\mu C}{2} (1 - \alpha) G \end{bmatrix}^T,
 \tag{15}$$

where  $M^-$  and  $M^+$  are 3D moments.

$F^+$  and  $M^+$  denote the index of dynamic balance constraints of the robot's linear and angular momentum increase rates, which can also be called positive index of dynamic balance constraints.  $F^-$  and  $M^-$  are the index of dynamic balance constraints of the robot's linear and angular momentum decrease rates, which can be referred to as negative index of dynamic balance constraints.

### 3.3. Discretization and Optimization Solution of the CMCR.

The formula of the CMCR needs to be discretized for convenient processing on the computer.

Firstly, the derivative of momentum in inequalities (13) is discretized with a sample time of  $T$ . Consider

$$-\begin{bmatrix} F^- \\ M^- \end{bmatrix} \leq \frac{\begin{bmatrix} P \\ L \end{bmatrix}_i - \begin{bmatrix} P \\ L \end{bmatrix}_{i-1}}{T} \leq \begin{bmatrix} F^+ \\ M^+ \end{bmatrix}.
 \tag{16}$$

Then, the robot's total momentum at the  $i$ th cycle is taken into the two arms' momentum formula. Consider

$$\begin{aligned}
 -T * \begin{bmatrix} F^- \\ M^- \end{bmatrix} &\leq \left( \begin{bmatrix} M_{\text{jobarm}} \\ H_{\text{jobarm}} \end{bmatrix} \dot{\theta}_{\text{jobarm}} + \begin{bmatrix} M_{\text{aidarm}} \\ H_{\text{aidarm}} \end{bmatrix} \dot{\theta}_{\text{aidarm}} \right)_i \\
 -\begin{bmatrix} P \\ L \end{bmatrix}_{i-1} &\leq T * \begin{bmatrix} F^+ \\ M^+ \end{bmatrix}.
 \end{aligned}
 \tag{17}$$

Lastly, the auxiliary arm's momentum formula at the  $i$ th cycle is kept in the middle of the inequalities. Since the inertia matrix of the auxiliary arm is not square matrix, let alone nonsingular matrix, and thus it does not need to be simplified. Therefore

$$\begin{aligned}
 \begin{bmatrix} P \\ L \end{bmatrix}_{i-1} - \left( \begin{bmatrix} M_{\text{jobarm}} \\ H_{\text{jobarm}} \end{bmatrix} \dot{\theta}_{\text{jobarm}} \right)_i - T * \begin{bmatrix} F^- \\ M^- \end{bmatrix} \\
 \leq \left( \begin{bmatrix} M_{\text{aidarm}} \\ H_{\text{aidarm}} \end{bmatrix} \dot{\theta}_{\text{aidarm}} \right)_i \\
 \leq \begin{bmatrix} P \\ L \end{bmatrix}_{i-1} - \left( \begin{bmatrix} M_{\text{jobarm}} \\ H_{\text{jobarm}} \end{bmatrix} \dot{\theta}_{\text{jobarm}} \right)_i + T * \begin{bmatrix} F^+ \\ M^+ \end{bmatrix}.
 \end{aligned}
 \tag{18}$$

The angular velocity vector of the joints in the auxiliary arm that fits with the constraint inequality (18) can realize all-dimensional CMCR at the six degrees of freedom; in other words, the robot can achieve overall balance and stability. However, multiple solutions of the angular velocity vector of the joints in the auxiliary arm that fits with the constraint inequalities exist. Thus, it is necessary to set the angular velocity of the joints in the auxiliary arm as the optimal goal by means of convex quadratic programming, in order to get the only optimal vector. The mathematic description of the inequalities constrained by the convex quadratic programming is defined as

$$\begin{aligned} \min_x \quad & \frac{1}{2}x^T Qx, \\ \text{subject to} \quad & b_1 \leq A_0 x \leq b_2. \end{aligned} \quad (19)$$

Specifically,  $Q$  is the optimal weight matrix of the angular velocity of the joints in the auxiliary arm; usually unit diagonal matrix is competent. Consider

$$\begin{aligned} x &= (\dot{\theta}_{\text{aidarm}})_i, \\ A_0 &= \begin{pmatrix} M_{\text{aidarm}} \\ H_{\text{aidarm}} \end{pmatrix}_i, \\ b_1 &= \begin{bmatrix} P \\ L \end{bmatrix}_{i-1} - \begin{pmatrix} M_{\text{jobarm}} \\ H_{\text{jobarm}} \end{pmatrix}_i \dot{\theta}_{\text{jobarm}} - T * \begin{bmatrix} F^- \\ M^- \end{bmatrix}, \\ b_2 &= \begin{bmatrix} P \\ L \end{bmatrix}_{i-1} - \begin{pmatrix} M_{\text{jobarm}} \\ H_{\text{jobarm}} \end{pmatrix}_i \dot{\theta}_{\text{jobarm}} + T * \begin{bmatrix} F^+ \\ M^+ \end{bmatrix}. \end{aligned} \quad (20)$$

**3.4. Motion Constraint of the Auxiliary Arm Joints.** In practice, it is necessary to take into account the output angular velocity and acceleration of the motor and reducer of the joints in the auxiliary arm. If not, the motor will overspeed and be overheated for the large current, and the reducer will overload. Since the constraint conditions of the defined convex quadratic programming are open, more reasonable constraints can be added. In the following part, the motion constraint of the joints in the auxiliary arm will be formulated.

Assume that the angular velocity constraint of the joints in the auxiliary arm is  $\dot{\theta}_{\text{max}}$  and the velocity constraint on the angular acceleration is  $a$ , and then the velocity constraint on the joints is defined as

$$-\dot{\theta}_{\text{max}} \leq x \leq \dot{\theta}_{\text{max}}. \quad (21)$$

The constraint on the angular acceleration of the joints will be

$$(\dot{\theta}_{\text{aidarm}})_{i-1} - a * T \leq x \leq (\dot{\theta}_{\text{aidarm}})_{i-1} + a * T. \quad (22)$$

Above all, the motion constraint of the joints in the auxiliary arm is

$$c_1 \leq x \leq c_2. \quad (23)$$

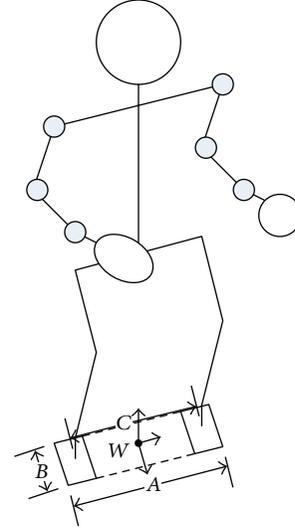


FIGURE 1: Humanoid robot model.

Specifically,  $c_1 = \max(-\dot{\theta}_{\text{max}}, (\dot{\theta}_{\text{aidarm}})_{i-1} - a * T)$  means that the maximum in the lower limits of the angular velocity constraints of every joint while  $c_2 = \min(\dot{\theta}_{\text{max}}, (\dot{\theta}_{\text{aidarm}})_{i-1} + a * T)$  refers to the minimum in the upper limits of the angular velocity constraints of every joint.

**3.5. Algorithm.** The convex quadratic programming falls into the category of classical mathematic programming with mature algorithms. Since the problem scale is small, the following steps have been taken to improve the speed of algorithm.

- (1) The kinematic chain of calculation is shortened. The desired motion and balance maintenance of the robot is and can only be performed by the two arms with other joints in the body remaining fixed. Thus, only the motion of the two arms needs to be calculated.
- (2) Array structure, instead of tree structure, is used in the algorithm data structure. The traverse calculation can be completed by one simple linear loop.
- (3) Recursive algorithm is replaced by nonrecursive algorithm, in order to save the calculation overhead and increase the speed.

Based on the steps above, the effectiveness and real-time performance of the algorithm have been tested on a computer platform with the configuration of Intel Core 2 Duo 2.93 GHz CPU, 2 GB RAM, Windows Xp. The computational time for each sample cycle on Matlab R2008b is 5.6 ms on average and 8.6 ms at most. Thus, the computation speed of the algorithm basically meets the requirement of real-time applications.

## 4. Simulation Results

We have simulated the high-speed motion of the humanoid robot's right arm for further experiments and analyses. The robot model we used in the experiment is shown in Figure 1. The robot weighs 55 kg, with a height of 165 cm and 30 degrees

TABLE 1: Detailed parameters of the arms.

Arms	Length (m)	Mass (kg)	Equivalent cross-section size (m * m)
Upper arm	0.25	3.5	0.1 * 0.1
Forearm	0.25	2.5	0.08 * 0.08
Hand	0.20	0.5	0.05 * 0.05

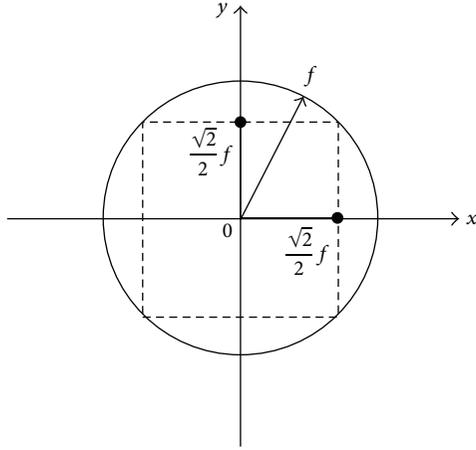


FIGURE 2: Linearization of transitional friction constraint.

of freedom. Specifically, the robot has 7 degrees of freedom in its two arms, respectively, 3 in its shoulders, 1 in its elbow, and 3 in its wrist. Figure 3 shows the structure of the two arms, with detailed parameters in Table 1 and the motion parameters of each joint of the arms in Table 2.

The constraint parameters are set as follows. Based on previous experience, the weight loss coefficient is  $\alpha = 0.15$  and then the dynamic balance constraint index has been calculated as follows:

$$\begin{aligned}
 F^- &= [97 \ 97 \ 80]^T \text{ N} \\
 M^- &= [103 \ 54 \ 20]^T \text{ Nm} \\
 F^+ &= [97 \ 97 \ 215]^T \text{ N} \\
 M^+ &= [103 \ 54 \ 20]^T \text{ Nm}.
 \end{aligned}
 \tag{24}$$

According to test results, though the dynamic balance constraint index can guarantee basic balance and stableness, the constraint is still too loose. To achieve better balance and stability, the constraint is further tightened as

$$\begin{aligned}
 \begin{bmatrix} F^- \\ M^- \end{bmatrix} &= [97 \ 97 \ 80 \ 80 \ 40 \ 20]^T, \\
 \begin{bmatrix} F^+ \\ M^+ \end{bmatrix} &= [97 \ 97 \ 188 \ 80 \ 40 \ 20]^T.
 \end{aligned}
 \tag{25}$$

For uniformity and writing convenience, the force and moment are unified into 6D column vector with their units removed.

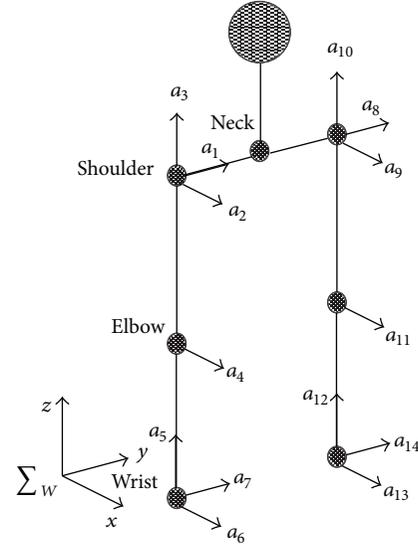


FIGURE 3: Model structures of the robot arms.

The angular velocity constraint and angular acceleration constraint of the joints in the auxiliary arm are the maximum angular velocity and maximum angular acceleration in Table 2.

In the simulation experiment, the task of the humanoid robot's operating arm (the right arm) performs the motion as follows: at first, the operating arm returns to the initial posture shown in Figure 4(a); then, the operating arm moves to the preparatory posture in Figure 4(b) with the time duration of 0.3 s; later, the operating arm waits for the task for 0.3 s in this case. Generally, the task means that the operating arm moves from the preparatory posture to the designated target posture within specified time duration at the designated target speed. In this paper, the set time is 0.3 s and the designated target posture is as shown in Figure 4(c). Moreover, the designated target speed includes linear velocity and angular velocity; specifically, the linear velocity is  $[1.76 \ 0 \ 0.38]^T$  m/s, the angular velocity  $[0 \ 0 \ 0]^T$  rad/s, the synthesized linear velocity 1.8 m/s, and synthesized angular velocity 0 rad/s. After the designated task gets finished, the operating arm returns to its preparatory posture from the target posture, waiting for the next task.

Relatively great continuous inertia impact is exerted to the robot body when the robot's arm accelerated from static to the synthesized linear velocity up to 1.8 m/s in a short time of 0.3 s. However, the greatest impact does not occur when the maximum synthesized linear velocity is reached; instead, it happens after the arm reaches the maximum synthesized linear velocity and when it is ready to withdraw its arm. At this moment, the main joints of the operating arm switch the maximum angular acceleration to the opposite maximum angular acceleration in an instant. Large jerk value is required bringing instantaneous impact to the humanoid robot body, and therefore the balance maintenance of the robot will play a vital role. As a result, it can be concluded that balance maintenance of the auxiliary arm is needed only when the operating arm performs and withdraws. From initial posture

TABLE 2: Motion parameters of the arm joints.

	Motion range of the joint angle (deg)		Maximum angular velocity (deg/s)	Maximum angular acceleration (deg/s <sup>2</sup> )
	Minimum	Maximum		
Pitch of shoulder	-120	40	430	5730
Roll of shoulder	-130	10	430	5730
Yaw of shoulder	-170	90	573	5730
Roll of elbow	-20	110	573	5730
Yaw of wrist	-130	130	573	2865
Roll of wrist	-90	90	487	2865
Pitch of wrist	-60	60	487	2865

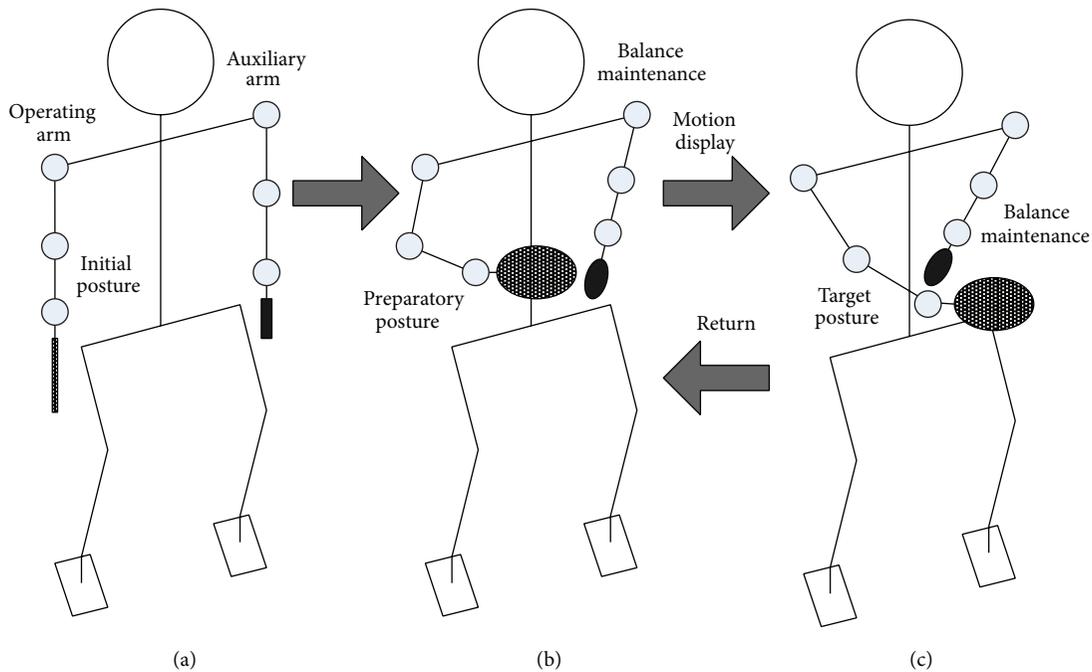


FIGURE 4: Operation and balance maintenance processing of the humanoid robot.

to preparatory posture, the time interval is long enough and the motion is mild, so this time interval is not the focus of this paper. Therefore, at this interval, the auxiliary arm is not used for balance maintenance but for moving to an advantageous initial posture.

From Figures 6 and 7, the angular accelerations and angular curves of the 7 joints of the robot's right arm show the motion of the operating arm in the whole process of the task. To put it more specifically, in Figure 6, the operating arm reaches the preparatory posture, the planned target posture, and the withdrawal preparatory posture at times 0 s, 0.3 s, and 0.8 s, respectively. At the same time, balance maintenance of the robot's auxiliary arm is executed simultaneously to guarantee the balance of the whole body.

Figure 5 demonstrates that the 6D constraints of momentum change rate of the robot is up to the expectation, for the fact that the 6D momentum change rate is constrained within the planned range of balance. Here we present several detailed problems: (1)  $F_x$  and  $F_z$  are not greatly influenced by the existence of balance maintenance, and these two dimensions

are still within the stability range; (2)  $M_y$  and  $M_z$ , which are the two dimensions to which the constraint of momentum change rate exerts active and effective effect, remain within the stability range after the balance maintenance has been executed. (3) Unfortunately,  $F_y$  and  $M_x$  increase so greatly after balance maintenance is executed that their value almost reaches the range of stability constraints, but they are still controlled within the constraint range. This shows the concept proposed in this paper that balance maintenance is a matter of range and only needs to be constrained only when the robot attempts to exceed the range of stability. In addition, it is clear that the humanoid robot is a severely nonlinear and strong coupling system. Under the worst circumstances, changes in some parts will lead other parts to change dramatically or even to worsen. It is necessary to take into consideration this feature when designing the balance maintenance strategy.

Figure 6 shows the curve of the angular velocity of the operating arm's joints and the auxiliary arm's joints generated by the full balance maintenance. It can be seen that the auxiliary arm moves only near at 0.3 s and 0.6 s, which are

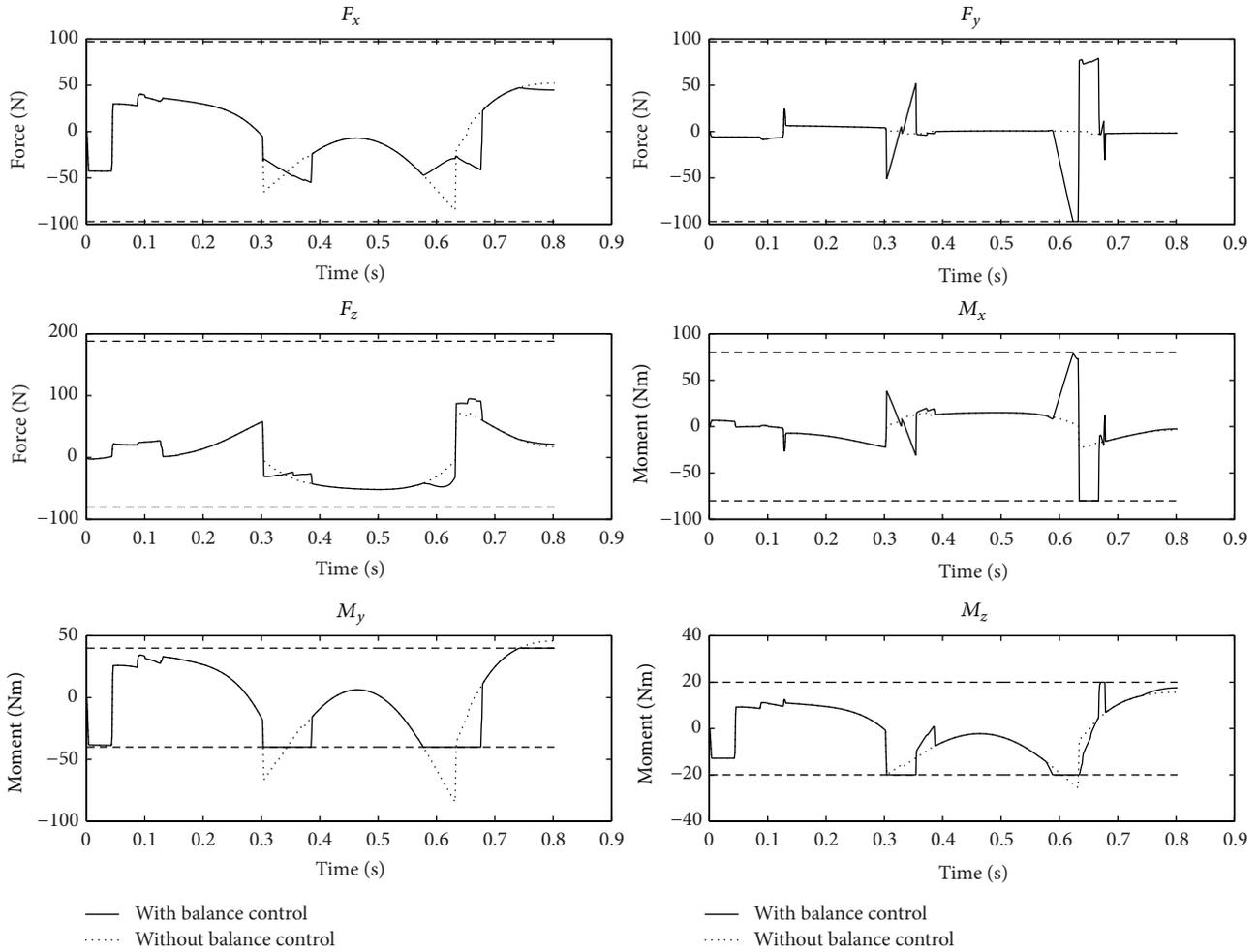


FIGURE 5: Momentum change rate curve of the robot.

the exact moments when the angular velocity of the operating arm's joints reaches the maximum and the angular acceleration switches to the opposite direction. This fully shows that the impact generated by the operating arm is not only related to the angular velocity of the joints but, more importantly, also greatly linked to the angular acceleration. Thus, it is more necessary to focus on the angular acceleration, instead of the angular velocity of the joints in the auxiliary arm, so as to resist the impact of the operating arm. As a result, the angular velocity curve of the auxiliary arm's joints is steeper.

Figure 7 shows the angular curve of the operating arm's joint as well as the angular curve of the auxiliary arm's joint generated by the balance maintenance. The angular curve of the joint is the integration of its angular velocity curve, so it is relatively smooth. Note that the auxiliary arm helps balance maintenance within a short period of time, and thus the angular change of the joints in the auxiliary arm is minor with great angular margin.

Figure 8 is the comparison of the ZMP curves under the robot's feet when the robot is performing the same task as

mentioned and after full balance maintenance is exerted. The solid line shows the ZMP curve without balance maintenance of the auxiliary arm. From this curve, ZMP under the robot's feet obviously exceeds the support area. According to ZMP theory [14], the operation of the robot is unstable; in other words, the robot will fall and fail at the motion when the ZMP exceeds. The dotted line shows the ZMP curve under the robot's feet when the operation and auxiliary arms work together. With balance maintenance, ZMP is fully controlled within the support area, with a certain distance kept from the edge of the support area. According to ZMP theory [14], the robot's operation is stable, with a certain amount of stability margin.

The ZMP curve after balance maintenance has been executed can be observed from the other aspect. In Figure 5,  $F_z$ ,  $M_x$ , and  $M_y$  are successfully constrained within the range of stability; in other words, they respectively meet the third, fourth, and fifth conditions in constraint inequalities (12), which can guarantee the ZMP stability of the robot. Therefore, the traditional ZMP stability is just a subset of the full balance maintenance method put forward in this

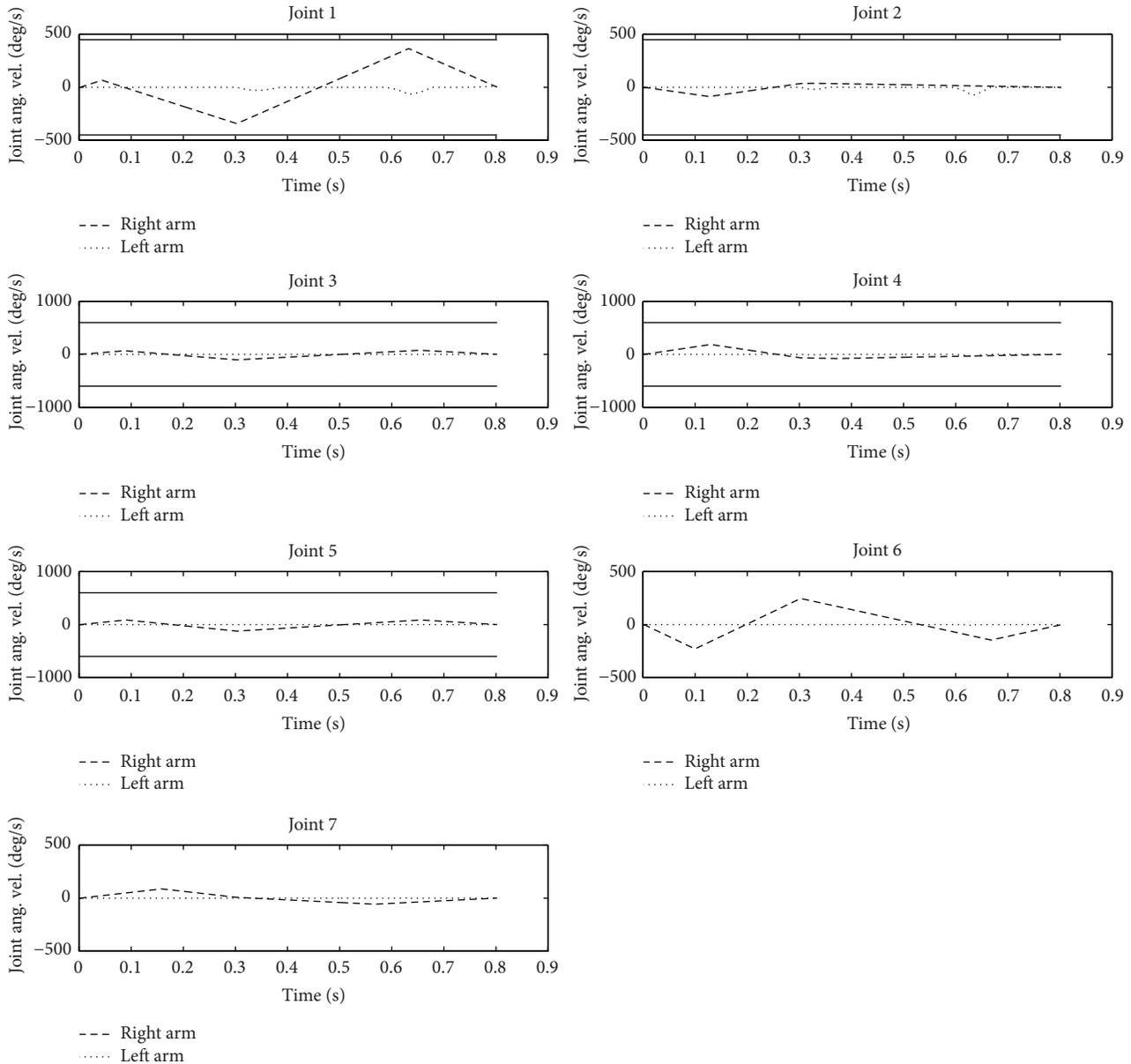


FIGURE 6: Angular velocity curve of the robot's arms.

paper. It is evident that the full balance maintenance proposed in this paper includes ZPM stability, as well as the balance on the six degrees of freedom which can make sure of the comprehensive stability of the robot.

We have selected ten groups of typical motion programming data of the operating arm. The expected effect has been achieved in these tests by means of the full balance maintenance method, employed with the constraints of momentum change rate and joint motion of the auxiliary arm. The test involves 12 dynamic balance constraint indexes, angular acceleration constraints, and angular velocity constraints of the seven joints in the robot's auxiliary arm, as well as the optimal diagonal weight matrix parameters. As a result, we can conclude the following.

Firstly, balance maintenance during high-speed motion of the robot under the constraint of the momentum change rate is feasible; and the framework of full balance maintenance is open and extensive. For instance, angular velocity and angular acceleration constraints on the joints of the auxiliary arm have been added in this paper. In other words, other reasonable requirements can be added to the solution, as long as they can be written in the form of constraint inequalities. For example, collision and interference constraints of the auxiliary arm can be added if necessary.

Secondly, the multiple parameters in the convex quadratic programming constrained by the inequalities on the one hand stand in the way of full balance maintenance; but, on the other hand, they bring great flexibility to the realization

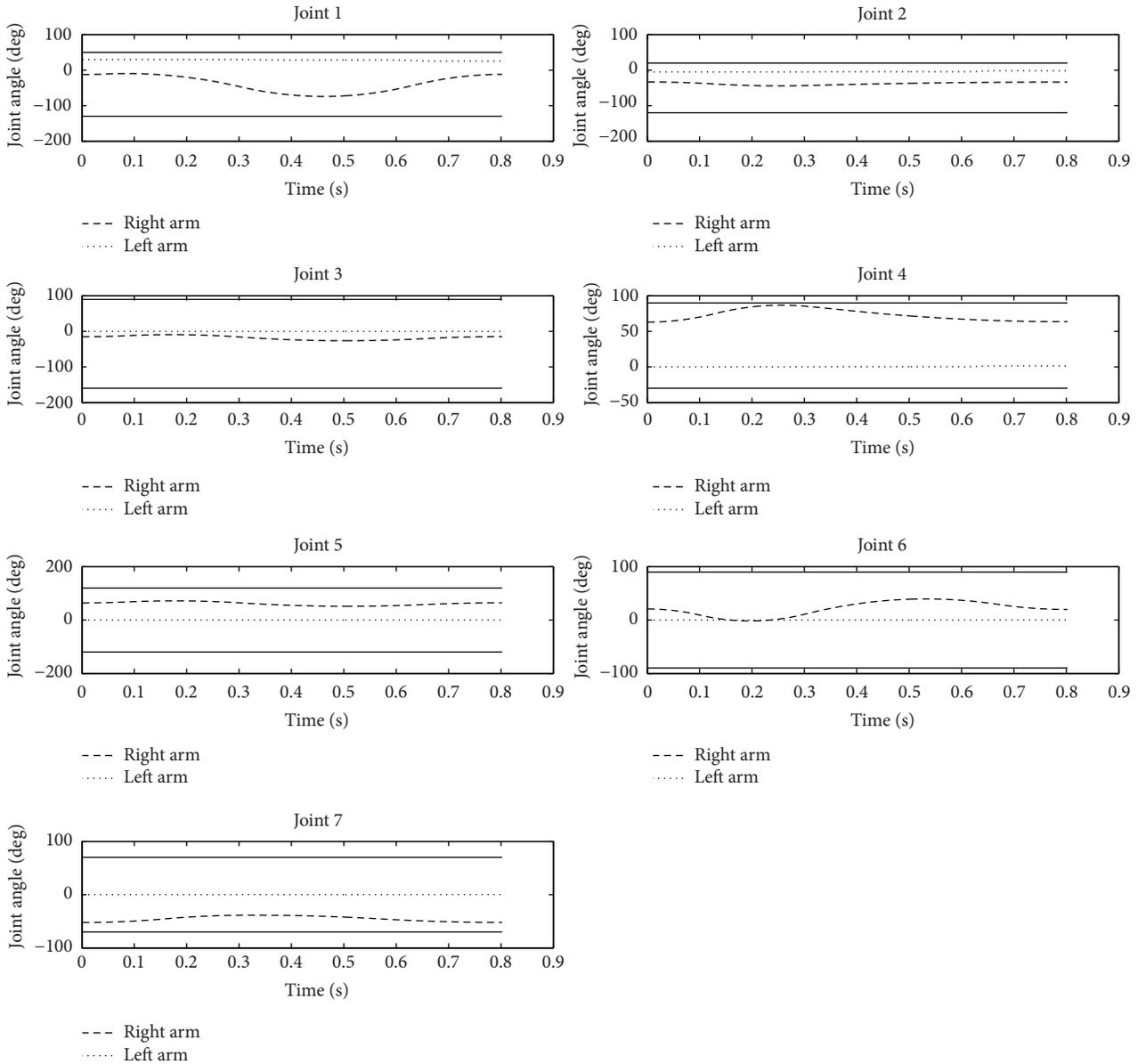


FIGURE 7: Joint angular curve of the robot’s arms.

of the convex quadratic programming. For instance, these parameters can be organically linked by means of neural network or even learned by iterative evolution algorithm, since these parameters can be adjusted online at real time.

Lastly, full balance maintenance is a concept and constraint of momentum change rate is an idea. The convex quadratic programming under the constraint of the inequalities is one of the mathematic methods under this concept and this idea. It is possible that a better and more advanced mathematical tool can be sought to achieve better balance between balance maintenance and the motion constraint of the auxiliary arm’s joints.

All in all, the concept of full balance and the idea of constraint of momentum change rate is a new solution to achieve dynamic balance at the six degrees of freedom

when the humanoid robot is performing high-speed motion. The solution has been specifically illustrated and briefly experimented in the paper and shows a good generality. In addition, this solution is highly open, extensive, and flexible in the parameter adjustments and realization methods. Thus, the authors have created a new framework for the balance and stability of the robot, with great value for future studies.

To test the balance performance when the robot operating arm’s end is at a higher speed, we gradually increase the linear speed at the end of the operating arm within the range of the robot’s actual capabilities. It has been found that, when the linear speed is as high as 3.2m/s, the robot can still remain balanced. Moreover, the bottleneck that occurs first is the overrunning of the angular velocity of the joints to their limits in the operating arm, rather than those in the

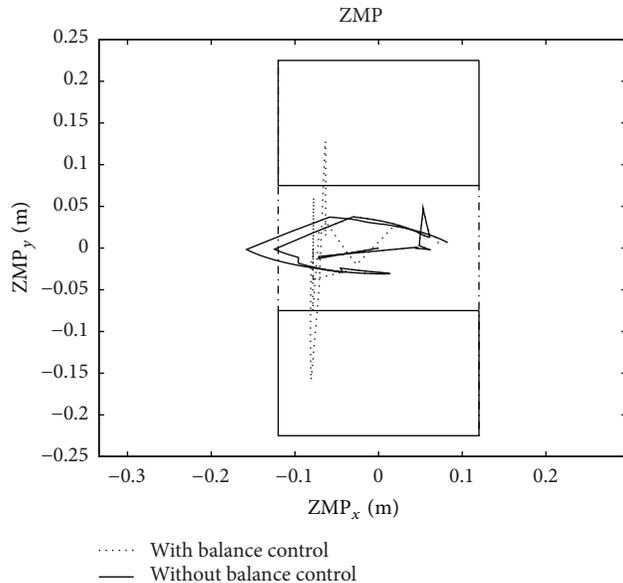


FIGURE 8: ZMP curve of the robot.

auxiliary arm. As a matter of fact, the auxiliary arm modifies its angular acceleration to adjust robot momentum change rate for balance only when the momentum change rate tries to exceed the dynamic balance constraint index, and this process will not last long. Thus, the angular velocity of the auxiliary arm basically will not overrun. However, the angular acceleration of the joints in the auxiliary arm should not be too large or last with a high speed for too long. The robot's arm end can perform with balance at a higher speed, if the actual performance capacities of the robot arm are further improved.

## 5. Conclusion

This paper puts forward the concept of full balance maintenance, as well as the idea of constraint of momentum change rate for the first time, and the concept and idea have been employed to achieve balance maintenance of the high-speed motion of the humanoid robot. Firstly, the total momentum formula of the robot's two arms is provided. Then momentum change rate is defined by the time derivative of the total momentum. In addition, the authors illustrate the concept of full balance, so as to find out the physical meaning of the constraint of the momentum change rate at the six dimensions, as well as the relation between full balance and 6D CMCR. Lastly, for the convenience of computer processing, CMCR is discretized and the convex quadratic programming is employed to solve motion constraint of the joints in the auxiliary arm. Besides, the actual algorithm is optimized to get real-time performance. The simulation results show the validity to achieve balance maintenance at the six degrees of freedom of the robot body by means of 6D CMCR. The resulted balance motion of the auxiliary arm has large angular margin, maintains angular velocity and angular acceleration within the range, and archives

a comprehensive dynamic balance performance and good ZMP stability margin. The actual tests have a good real-time performance. For future research, the real-time full balance maintenance utilizing the motion of the humanoid's waist and legs' joints will be studied.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work is supported by the National 863 Plan (Grant no. 2008AA042602), the National Nature Science Foundation of China (Grant no. 61075078), and the project of Key Science and Technology Innovation Team of Zhejiang Province (2011R09014).

## References

- [1] Y. Ogura, H. Aikawa, K. Shimomura et al., "Development of a new humanoid robot WABIAN-2," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '06)*, pp. 76–81, Orlando, Fla, USA, May 2006.
- [2] K. Hirai, M. Hirose, Y. Haikawa, and T. Takenaka, "Development of Honda humanoid robot," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 1321–1326, Leuven, Belgium, May 1998.
- [3] T. Ishida, "Development of a small biped entertainment robot QRIO," in *Proceedings of the International Symposium on Micro-NanoMechatronics and Human Science (MHS '04)*, pp. 23–28, Nagoya, Japan, November 2004.
- [4] K. Kaneko, K. Harada, F. Kanehiro, G. Miyamori, and K. Akachi, "Humanoid robot HRP-3," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '08)*, pp. 2471–2478, Nice, France, September 2008.
- [5] J.-H. Oh, D. Hanson, W.-S. Kim, I. Y. Han, J.-Y. Kim, and I.-W. Park, "Design of android type humanoid robot Albert HUBO," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '06)*, pp. 1428–1433, Beijing, China, October 2006.
- [6] T. Takubo, K. Inoue, and T. Arai, "Pushing an object considering the hand reflect forces by humanoid robot in dynamic walking," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 1706–1711, Piscataway, NJ, USA, April 2005.
- [7] Q. Huang, X. C. Chen, J. Y. Gao et al., "Analysis of pushing manipulation by humanoid robot during dynamic walking," *Information-An International Interdisciplinary Journal*, vol. 12, no. 1, pp. 217–234, 2009.
- [8] N. Vahrenkamp, S. Wieland, P. Azad, D. Gonzalez, T. Asfour, and R. Dillmann, "Visual servoing for humanoid grasping and manipulation tasks," in *Proceedings of the 8th IEEE-RAS International Conference on Humanoid Robots (Humanoids '08)*, pp. 406–412, Daejeon, South Korea, December 2008.
- [9] C. Ott, O. Eiberger, W. Friedl et al., "A humanoid two-arm system for dexterous manipulation," in *Proceedings of the 6th IEEE-RAS International Conference on Humanoid Robots (HUMANOIDS '06)*, pp. 276–283, Genoa, Italy, December 2006.

- [10] K. Yokoyama, H. Handa, T. Isozumi et al., "Cooperative works by a human and a humanoid robot," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 2985–2991, Taipei, Taiwan, September 2003.
- [11] S. Nakaoka, A. Nakazawa, F. Kanehiro et al., "Learning from observation paradigm: leg task models for enabling a biped humanoid robot to imitate human dances," *International Journal of Robotics Research*, vol. 26, no. 8, pp. 829–844, 2007.
- [12] E. S. Neo, K. Yokoi, S. Kajita, and K. Tanie, "Whole-body motion generation integrating operator's intention and robot's autonomy in controlling humanoid robots," *IEEE Transactions on Robotics*, vol. 23, no. 4, pp. 763–775, 2007.
- [13] S. Kajita, F. Kanehiro, K. Kaneko et al., "Resolved momentum control: Humanoid motion planning based on the linear and angular momentum," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1644–1650, Las Vegas, Nev, USA, October 2003.
- [14] M. Vukobratovic and J. Stepanenko, "On the stability of anthropomorphic systems," *Mathematical Biosciences*, vol. 15, no. 1-2, pp. 1–37, 1973.
- [15] S. Kagami, F. Kanehiro, Y. Tamiya et al., "AutoBalancer: an online dynamic balance compensation scheme for humanoid robots," in *Proceedings of the International Workshop on the Algorithmic Foundations of Robotics*, pp. 329–339, Springer, Hanover, NH, USA, 2000.
- [16] T. Sugihara, Y. Nakamura, and H. Inoue, "Realtime humanoid motion generation through ZMP manipulation based on inverted pendulum control," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 1404–1409, Washington, DC, USA, May 2002.
- [17] Z. Li, Q. Huang, K. Li, and X. Duan, "Stability criterion and pattern planning for humanoid running," in *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 1059–1064, Beijing, China, May 2004.
- [18] M. B. Popovic, A. Goswami, and H. Herr, "Ground reference points in legged locomotion: definitions, biological trajectories and control implications," *International Journal of Robotics Research*, vol. 24, no. 12, pp. 1013–1032, 2005.
- [19] A. Goswami and V. Kalleem, "Rate of change of angular momentum and balance maintenance of biped robots," in *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 3785–3790, May 2004.

## Research Article

# An Optimization Method for Condition Based Maintenance of Aircraft Fleet Considering Prognostics Uncertainty

Qiang Feng,<sup>1</sup> Yiran Chen,<sup>1</sup> Bo Sun,<sup>1</sup> and Songjie Li<sup>2</sup>

<sup>1</sup> School of Reliability and Systems Engineering, Beihang University, Beijing, China

<sup>2</sup> Sichuan jiuzhou Aerocont Technologies Co. Ltd., Mianyang, China

Correspondence should be addressed to Bo Sun; sunbo@buaa.edu.cn

Received 28 February 2014; Accepted 19 March 2014; Published 17 April 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Qiang Feng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An optimization method for condition based maintenance (CBM) of aircraft fleet considering prognostics uncertainty is proposed. The CBM and dispatch process of aircraft fleet is analyzed first, and the alternative strategy sets for single aircraft are given. Then, the optimization problem of fleet CBM with lower maintenance cost and dispatch risk is translated to the combinatorial optimization problem of single aircraft strategy. Remain useful life (RUL) distribution of the key line replaceable Module (LRM) has been transformed into the failure probability of the aircraft and the fleet health status matrix is established. And the calculation method of the costs and risks for mission based on health status matrix and maintenance matrix is given. Further, an optimization method for fleet dispatch and CBM under acceptable risk is proposed based on an improved genetic algorithm. Finally, a fleet of 10 aircrafts is studied to verify the proposed method. The results shows that it could realize optimization and control of the aircraft fleet oriented to mission success.

## 1. Introduction

Prognostic and health management (PHM) technology has a rapid development and been widely used in aeronautical equipment in recent years. The failure position and remain useful life (RUL) of equipment could be predicted by PHM. Further, it can be used in aircraft condition based maintenance (CBM) [1]. However, due to the uncertainty of prognostics, there are certain risks in the maintenance decisions based on the prediction of RUL [2, 3].

The aircraft usually performs mission in fleet manner and shares limited support resource. So, there will be a tradeoff range for fleet CBM. This means each aircraft can choose strategy among dispatching strategy, standby strategy, and maintenance strategy or their combination when the RUL has been obtained, and the synthetic strategy for fleet (combination of each aircraft's strategy) should meet the mission requirement.

There are three forms of RUL in PHM, and each form includes some uncertainty. First is the point value of the time of potential failure. Second is the interval value of the

time of potential failure [4–7]. Third is the RUL distribution of the device [8–12]. The third form has the maximum information and the highest availability but is the most difficult in acquisition and application.

Two methods can be used in reducing the impact of prognostics uncertainty on CBM decision. One is to reduce the uncertainty of failure prediction directly so that the decision risk will decrease [13–15]. The other one is to take the prediction uncertainty into account and make the optimum decision under acceptable risk [16–21]. Because the uncertainty of failure prediction could not be completely eliminated, the latter is more useful in engineering applications.

Most research about RUL for CBM is about the life cycle maintenance optimization decisions on single aircraft and researchers would rather consider the maintenance decisions than think about the mission requirements and the dispatched strategy. The research on fleet CBM oriented to mission successes is few. Agent technology and the heuristic algorithm are used to fleet CBM in article [22, 23], but sample point values of RUL were used only.

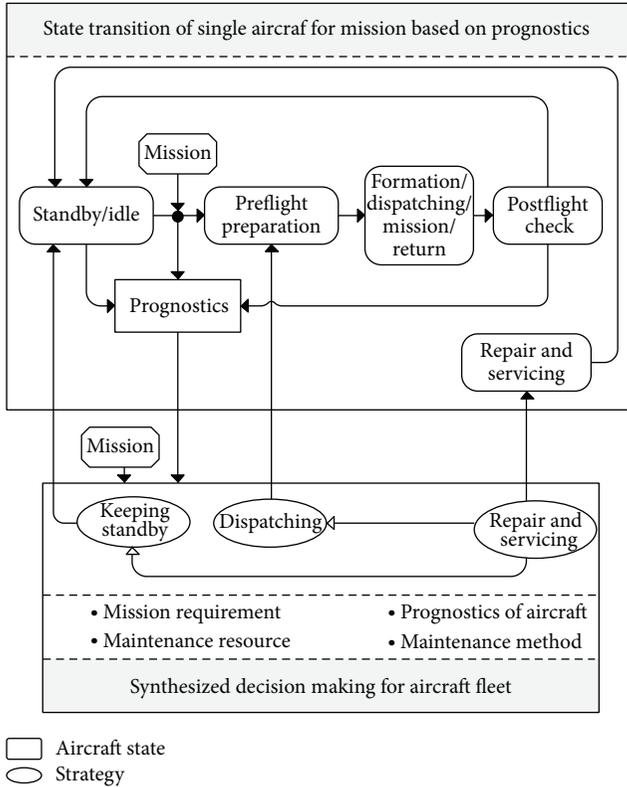


FIGURE 1: CBM for aircraft fleet.

An optimal aircraft fleet CBM method for aviation unit maintenance is proposed in the paper considering dispatch, mission, and resource constraints. Moreover, the RUL distribution of the key LRM has been transformed into the failure probability of the aircraft, and the calculation method of the costs and risks for mission is given. Then, an optimization decision making method for fleet dispatch and CBM under acceptable risk is proposed based on an improved genetic algorithm.

## 2. Analysis of Aircraft Fleet CBM

**2.1. Basic Process Analysis.** Consider an aircraft fleet containing  $m$  aircrafts and  $k$  integrated support stations (ISS) facing continuous combat missions ( $k < m$ ), in which a single mission requires  $l$  aircrafts ( $l \leq m$ ). Each aircraft contains  $p$  LRMs of which RUL can be estimated. The mission preparation period starts at time  $t_0$ , while the mission period is from time  $t_1$  to  $t_2$ . The basic process of the fleet CBM decisions, which is mission success oriented, is given in Figure 1. There are two kinds of single strategies (keeping standby and dispatching) and two kinds of mixed strategies (dispatching after maintenance and standby after maintenance) before making synthesized decision. The fleet CBM decisions consist of these single strategies that should meet the requirements of missions, cost, and risk.

**2.2. Assumptions.** The basic assumptions of the problem are listed below in order to define the problem.

- (1) The aircraft fails when any key LRM fails.
- (2) The RUL distribution of LRM which is given at the time  $t_0$  is  $F(t)$  of which probability density function is  $f(t)$ .
- (3) Assume the maintenance method of the LRM is renew, which means the LRM will be as good as new after maintenance, considering the field maintenance of aviation unit maintenance.
- (4) Only one aircraft can be repaired in each ISS simultaneously. But the total number of the aircraft maintenance may be more than one from the time  $t_0$  to  $t_1$ .
- (5) Different LRM in the same aircraft can be replaced at the same time for renew is served as a maintenance method.
- (6) The maintenance cost of the different LRM varied while the same LRM cost is the same. The maintenance cost of the LRM on  $j$  class is  $C_j$ .
- (7) Each aircraft malfunction will cause the mission to fail when the fleet is on mission. The consequences of the economic loss will not be taken into consideration.
- (8) Spare parts are plentiful.

## 3. Modeling Method to Aircraft Fleet CBM Considering Prognostics Uncertainty

**3.1. Modeling Framework.** The main work of the optimization decision making method for fleet CBM considering prognostics uncertainty includes the following steps: (1) the definition of the fleet initial health status based on the RUL prognostic, (2) maintenance program generating, (3) maintenance time and cost estimation, and (4) mission risk assessment. Based on the objects above, the optimal CBM and maintenance program through the rational optimization algorithm is obtained in the paper. The modeling framework is given in Figure 2.

**3.2. The Definition of the Fleet Initial Health Status Based on the RUL Prognostic.** Assume the distribution of the  $j$ th key LRU $_{ij}$  on the  $i$ th aircraft is  $F_{ij}(t)$ . In the mission period ( $t_1$ - $t_2$ ), the probability of failure can be got by

$$P_{ij}(a) = \int_{t_1}^{t_2} f_{ij}(t) dt \quad i = 1, 2, \dots, m \quad j = 1, 2, \dots, n \quad (1)$$

where  $f_{ij}(t)$  is the probability density function of the  $F_{ij}(t)$ .

Considering an aircraft fleet containing  $m$  aircrafts and each include  $n$  LRMs, the initial health status matrix of all LRM is  $P(a)$  that is given by

$$P(a) = \begin{bmatrix} p_{11}(a) & \cdots & p_{1n}(a) \\ \cdots & \cdots & \cdots \\ p_{m1}(a) & \cdots & p_{mn}(a) \end{bmatrix}_{m \times n} \quad (2)$$

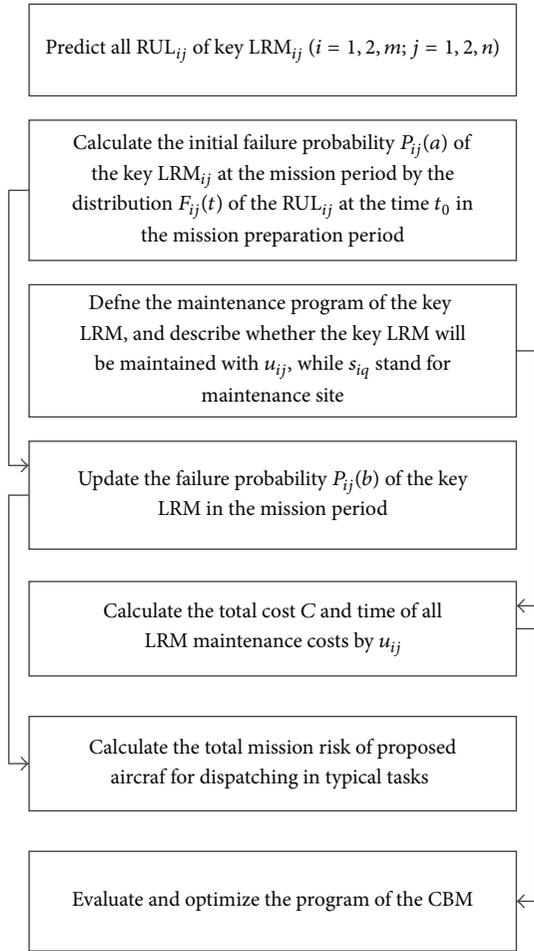


FIGURE 2: Modeling framework for CBM of aircraft fleet.

3.3. *Maintenance Program Generating.* Maintenance program considers whether a certain LRM should be maintained and the selection of the ISSs.

The maintenance matrix  $U$  of fleet can be described as

$$U = \begin{bmatrix} u_{11} & \cdots & u_{1n} \\ \cdots & \cdots & \cdots \\ u_{m1} & \cdots & u_{mn} \end{bmatrix}_{m \times n}, \quad (3)$$

where  $u_{ij} = 1$  means that the  $j$ th LRM of the  $i$ th aircraft needs to be repaired; otherwise,  $u_{ij} = 0$ .

According to the assumption (4), the LRM can be maintained at the same place; however, many the LRMs fails. Therefore, the ISS matrix  $S$  of the fleet is shown as

$$S = \begin{bmatrix} s_{11} & \cdots & s_{1k} \\ \cdots & \cdots & \cdots \\ s_{m1} & \cdots & s_{mk} \end{bmatrix}_{m \times k}, \quad (4)$$

where  $s_{iq} = 1$  means that the  $i$ th aircraft should be maintained at the  $q$ th ISS; otherwise,  $u_{ij} = 0$ .

3.4. *Maintenance Time and Cost Estimation.* Assume repairing the  $j$ th LRM spends time  $T_j$  and needs cost  $C_j$ . According

to the assumption (5), the total maintenance time  $Tm_i$  of the  $i$ th aircraft is given as

$$Tm_i = \max(u_{ij} \times T_j), \quad j = 1, 2, \dots, n. \quad (5)$$

There may be more than one aircraft that should be repaired at ISS  $q$ , so the total maintenance time of all aircrafts can be calculated as

$$\sum_{i=1}^m s_{iq} \times Tm_i \quad i = 1, 2, \dots, m. \quad (6)$$

The total maintenance cost of all aircrafts can be got as

$$C = \sum_{i=1}^m \sum_{j=1}^n [u_{ij} \times C_j] \quad i = 1, 2, \dots, m \quad j = 1, 2, \dots, n. \quad (7)$$

3.5. *Mission Risk Assessment*

*Step 1* (modify the health matrix of the fleet). Whether the aircraft is “dispatching” or “dispatching after maintenance” should be taken into consideration when calculating the mission risk of the fleet. The status of the aircraft should be updated if the single strategy of the fleet is “dispatching after maintenance.” Then, the modified health status matrix  $P(b)$  of the fleet can be built according to the Assumption (3). Consider  $p_{ij}(b) = 0$  after the LRM<sub>ij</sub> on the  $i$ th aircraft was renewed; otherwise,  $p_{ij}(b) = p_{ij}(a)$  without renew. The elements in the matrix can be obtained by

$$p_{ij}(b) = p_{ij}(a) \times (1 - u_{ij}). \quad (8)$$

*Step 2* (estimate the failure probability of the single aircraft and rank). The failure probability of the single aircraft could be estimated after modifying the health matrix of the fleet. According to the first assumption, “the aircraft fails when any key LRM fails”; the failure probability  $P_i(a)$  of the  $i$ th aircraft can be given as

$$P_i(a) = 1 - \prod_{j=1}^n [1 - p_{ij}(b)]. \quad (9)$$

Formula (10) can be obtained according to (1), (8), and (9):

$$P_i(a) = 1 - \prod_{j=1}^n \left[ 1 - \int_{t_1}^{t_2} f_{ij}(t) (1 - u_{ij}) \right] \quad i = 1, 2, \dots, m \quad (10)$$

Then,  $\text{Pro}(i) = 1, 2, \dots, m$  can be obtained by sorting the failure probability of single aircraft in ascending order. The ordered failure probability of the aircraft is given as

$$P_{\text{pro}(i)}(b) = P_i(a). \quad (11)$$

Suppose  $P_2(a)$  is the smallest  $\text{Pro}(i)$ . Then, set  $\text{Pro}(2) = 1$ , and let  $P_1(b) = P_2(a)$  after reordering. Pick up aircrafts of which  $\text{Pro}(i) = 1, 2, \dots, l$  when the mission needs dispatch  $l$  aircrafts.

Step 3 (calculate the mission risk of the fleet). Assume the serious consequences of the mission that failed are similar without taking the economic losses into account. The failure probability of the fleet mission can be calculated by the following according to (7):

$$P_F = 1 - \prod_{i=1}^l [1 - P_i(b)], \tag{12}$$

where  $P_F$  is the mission risk.

### 4. Optimization Problem and Algorithms Design

4.1. *Problem Description.* The optimization problem in the paper is to find a fleet CBM strategy with acceptable risk and lowest cost considering prognostics uncertainty.

Thus, describe the objective of the optimization as  $\text{Min } C = \sum_{i=1}^m \sum_{j=1}^n [u_{ij} \times C_j]$ .

The constraints that should be considered about involve the maintenance ability constraint  $R_A$  of the site, the time constraint  $R_B$ , the security risk constraint  $R_C$ , the mission risk constraint  $R_D$ , and the variable constraint  $R_E$ .

For first constraint  $R_A$ , set  $s_{iq} = 0$  ( $q = 1, 2, \dots, k$ ) and  $u_{ij} = 0$  ( $j = 1, 2, \dots, n$ ) if none of LRMs need maintenance. Else if any  $\text{LRM}_{ij}$  requires maintenance, then  $u_{ij} = 1$  and the corresponding  $s_{iq} = 1$  while the other  $s_{iq} = 0$ . Thus, the  $R_A$  can be described as

$$R_A : \sum_{q=1}^k s_{iq} + \prod_{j=1}^n (1 - u_{ij}) = 1. \tag{13}$$

All maintenance of site  $q$  should be finished before the mission starts. Thus, the  $R_B$  can be given as

$$R_B : Tm_i \leq t_1 - t_0 \mid s_{iq} = 1 \quad i = 1, 2, \dots, m \quad q = 1, 2, \dots, k. \tag{14}$$

Assume that the total number of the aircraft which maintained at site  $q$  is  $\sum_{i=1}^m s_{iq} = x > 1$  (where the number of aircraft is  $1, 2, \dots, x$ ). If  $\sum_{i=1}^{x-1} s_{iq} \times Tm_i \leq t_1 - t_0 < \sum_{i=1}^x s_{iq} \times Tm_i$ , which means the maintenance for the  $x$ th aircraft could not be finished before the mission start time, this aircraft will not be taken into account when the maintenance decisions is “dispatching after maintenance.”

It will not be allowed to dispatch if the failure probability is too high for security risk existing in single aircraft. Consider a mission need  $l$  aircrafts, and the  $R_C$  can be described as (15). Moreover, the mission will fail if (15) could not be met. We have

$$R_C : P_l(b) < P_{sl}. \tag{15}$$

According to (12),  $R_D$  can be written as (16) considering the mission risk for fleet. We have

$$R_D : P_F = 1 - \prod_{i=1}^l [1 - P_i(b)] < P_m, \tag{16}$$

where  $P_m$  is the objective of the mission risk.

The variable constraint, which means that the variables should be in a certain range, is described as

$$R_E : C_j > 0, \quad u_{ij} \in \{0, 1\}, \quad s_{iq} \in \{0, 1\} \\ i = 1, 2, \dots, m \quad j = 1, 2, \dots, n \quad q = 1, 2, \dots, k. \tag{17}$$

The conceptual model for aircraft fleet condition based maintenance and dispatch is given as follows:

$$\text{Min } C = \sum_{i=1}^m \sum_{j=1}^n [u_{ij} \times C_j] \tag{18}$$

s.t.  $R_{A \sim E}$  is satisfied.

4.2. *Optimization Algorithms Design.* The optimization problem cannot meet the KKT (*Karush-Kuhn-Tucker*) conditions, and the dimension of decision making variables which can be written as  $m \times n + m \times k$  is relatively large. So an improved genetic algorithm was proposed in this paper for the problem instead of traditional mathematical methods.

The optimization model can be simplified as

$$\min C(U, S) \\ \text{s.t. } \begin{cases} g(U, S) \leq 0 \\ h(U, S) = 0 \\ u_{ij} \in \{0, 1\} \\ s_{ij} \in \{0, 1\} \end{cases}, \tag{19}$$

where  $U$  is the maintenance matrix while the  $S$  is the ISS matrix.

The problem has more variables and constraints, so the solution quality of problem and the convergence rate could not be satisfied. Therefore, the improvement strategy of the genetic algorithm is given in Figure 3.

4.2.1. *Define the Initial Population of the Maintenance Matrix U.* According to the multifactor and 2-level orthogonal experimental design in order to cover widely, define the initial population of the maintenance matrix  $U$ . The initial population should be filtered so as to make the convergence faster. Moreover, the number of the aircraft needs to be repaired in the population which should be less than  $l$  considering the dispatched requirements and the cost of the maintenance. The relationship among those factors is shown as

$$\sum_{i=1}^m \left[ \prod_{j=1}^n (1 - u_{ij}) \right] \geq m - l. \tag{20}$$

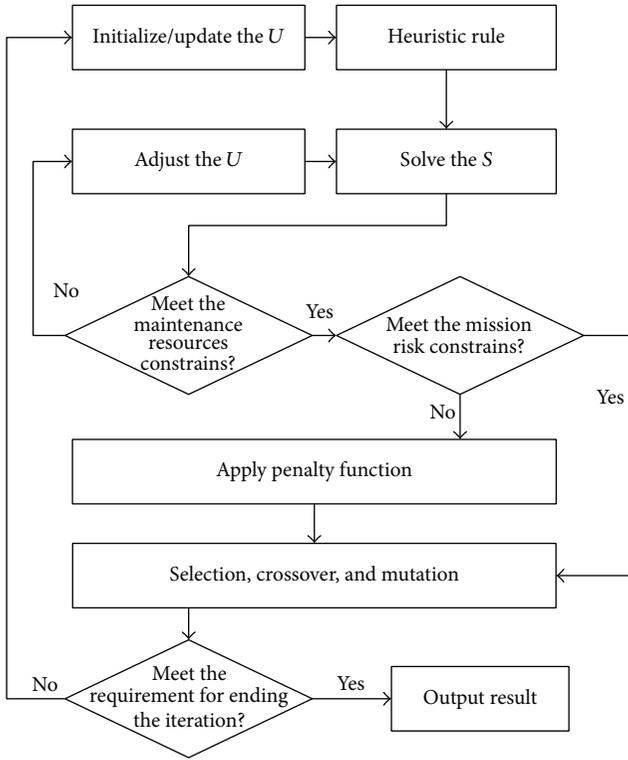


FIGURE 3: Improved strategies for genetic algorithm.

4.2.2. *Solve the ISS Matrix S.* It is necessary to find a set of feasible solutions which meet the constraint of the ability of ISS  $R_A$  and the maintenance time  $R_B$  on the basis of a certain  $U$ . The following heuristic rules can be used in order to reduce the amount of computation, increasing the efficiency of solving.

*Step 1.* According to formula (13), an initial value of the  $S$  can be given with the certain  $U$ . If  $\prod_{j=1}^n (1 - u_{ij}) = 1$ , the aircraft  $i$  need not be repaired and all  $s_{iq} (q = 1, 2, \dots, k) = 0$ ; otherwise, the aircraft  $i$  needs to be repaired. Then, the determining condition is described as  $\sum_{q=1}^k s_{iq} = 1$  and  $s_{iq} \in \{0, 1\}$ .

*Step 2.* Consider that there are  $y$  aircrafts need not be repaired. Remove  $y$  rows which stand for these aircrafts. Then, a new  $(m - y) \times k$  matrix  $S'$  which represents the new relationship between the ISS and the aircraft that needs to be repaired can be built as the reduced cycle matrix of  $S$ .

*Step 3.* The maintenance time  $Tm_i$  of the aircraft needs to be repaired in matrix  $S'$  which can be obtained by formula (5). Then, the average maintenance time AMT of the ISSs is given by  $AMT = \sum_{i=1}^{m-y} Tm_i / k$ . It can be determined not meet the time constrain if  $\max(Tm_i) > t_1 - t_0$  or  $AMT > t_1 - t_0$ , then turn to Step 6. Otherwise, turn to Step 4.

*Step 4.* Initialize the matrix  $S'$ , and set  $s_{iq} = 0 (i = 1, 2, \dots, m - y, q = 1, 2, \dots, k)$ .

*Step 5.* Set the value of the matrix  $S'$  from the first row to the  $k$ th row. The method of the  $q$ th is described as follows.

- (a) Calculate the value of  $|Tm_i - AMT| (i = 1, 2, \dots, m - y)$ . If the aircraft  $z$  makes the  $|Tm_z - AMT| = \min |Tm_i - AMT|$ , then  $s_{zq} = 1$ . Furthermore, the aircraft can be selected in random if there is more than one aircraft that meets this formula.
- (b) Remove the line in which the aircraft  $z$  is in to build a new reduced cycle matrix  $S'$ . Update the remaining maintenance time  $TG_q = t_1 - t_0 - Tm_z$  of the site  $q$ .
- (c) Compare the  $TG_q$  and the  $Tm_i$  for the  $S'$ . If the formula " $\min(Tm_i) | i = o \leq TG_q$ " can be met by a parameter  $o$ , form a new reduced cycle matrix and set  $s_{oq} = 1$ . Moreover, the remaining available reference time should be updated as  $TG_q = TG_q(b) - Tm_o$ . This work should be repeated until the  $\min(Tm_i) > TG_q$ ; then, turn to the  $(q + 1)$ th row.

*Step 6.* The matrix  $U$  should be adjusted if the constraints of resource maintenance cannot be met. Consider that the maintenance cost should be as low as possible and the requirement of the mission risk should be satisfied; the elements which  $u_{ij} = 1$  should find  $p_{ij}(a)$  corresponded and the  $\min\{p_{ij}(a)\}$ ; then, set  $u_{ij} = 0$ . Return to Step 1 and repeat after finishing the update for the  $U$  until meeting constrains  $R_A$  and  $R_B$ .

4.2.3. *Deal with the Constrains of the Mission Risk.* Some matrix  $U$  which is initial or got by adjusting, crossover, and mutation may not meet the requirement of the mission risk constrain. The penalty method can be used in the method followed to solve this problem.

The energy function for every  $U$  can be written as

$$E(U, S) = C(U, S) + F(U, S) \cdot M^T, \quad (21)$$

where  $F(U, S)$  is the vector of the penalty function and the  $F_i(U, S) = \max\{0, g_i(U, S)\}$ , while  $M$ , which is the penalty factor vector, is a large positive number.

*Step 1 (fitness function design).* The fitness function is given as follows in order to minimize the objective function:

$$f(U, S) = 1 - \frac{E(U, S) - E_{\min}}{E_{\max} - E_{\min}}, \quad (22)$$

where  $E_{\max}$  and  $E_{\min}$  are the maximum and the minimum values of the energy function in the population.

*Step 2 (selection, crossover, and mutation).* Proportional selection, single-point crossover, and the basic alleles can be used in solving this problem.

This problem can be dealt with by some method written in the article [24–26] in order to avoid the premature and the stalling that appear in the genetic algorithms.

TABLE 1: The RULs of the LRM.

Number	1	2	3	4	5	6	7	8	9	10
LRM <sub>A</sub>	(25, 7.7)	(19, 4.4)	(29, 8.3)	(29, 9.1)	(13, 3.8)	(20, 5.7)	(21, 6.3)	(20, 6.2)	(9, 2.5)	(21, 5.9)
LRM <sub>B</sub>	(28, 7.4)	(3, 0.7)	(29, 6.6)	(15, 3.3)	(28, 8.5)	(2, 0.5)	(23, 5.6)	(6, 1.3)	(2, 0.5)	(10, 2.8)
LRM <sub>C</sub>	(4, 1.0)	(9, 2.9)	(5, 1.2)	(25, 5.7)	(24, 7.1)	(26, 7.9)	(23, 6.1)	(22, 7.2)	(3, 0.8)	(29, 9.1)
LRM <sub>D</sub>	(28, 7.5)	(17, 5.6)	(30, 7.9)	(5, 1.2)	(29, 7.9)	(29, 7.1)	(12, 3.5)	(1, 0.3)	(25, 6.4)	(2, 0.6)

Simulate the annealing stretching for fitness before selecting the operator as follows:

$$f_i = \frac{e^{f_i/T}}{\sum_{j=1}^N e^{f_j/T}} \quad T = T_0 \times c^{g-1} \quad 0 < c < 1, \quad (23)$$

where  $N$  is the size of the population and  $g$  is the genetic algebra, while  $T_0$  is the initial temperature and  $f_i$  is the fitness of the  $i$ th individual.

$P_e$  and  $P_f$  can be defined as (24) in order to make the crossover and mutation probability changing dynamic with the fitness, which means that if the fitness of each individual is consistent,  $P_e$  and  $P_f$  will increase; otherwise, they will decrease:

$$P_e = \begin{cases} \frac{k_1 (f_{\max} - f')}{(f_{\max} - f_{\text{avg}})} & f' \geq f_{\text{avg}} \\ k_2 & f' < f_{\text{avg}} \end{cases} \quad (24)$$

$$P_f = \begin{cases} \frac{k_3 (f_{\max} - f')}{(f_{\max} - f_{\text{avg}})} & f' \geq f_{\text{avg}} \\ k_4 & f' < f_{\text{avg}} \end{cases}$$

where  $f_{\max}$  and  $f_{\text{avg}}$  are the maximum fitness and the average fitness in the populations and the  $f'$  is the maximum fitness of the parent. The  $k_1, k_2, k_3, k_4$  are all constant.

### 5. Case Study

Consider a fleet containing 10 aircrafts and each aircraft includes 4 LRM (A, B, C, D) of which life can be predicted. Assume the RUL following Gaussian distributions  $N(\mu, \sigma^2)$ , and the mean  $\mu$  and the variance  $\sigma^2$  are given in Table 1.

The mission requires dispatch 8 aircrafts one hour later and lasting two hours.  $P_{sl}$  should be below the  $10^{-8}$  while  $P_{sl}$  should be below  $10^{-6}$ .

Assume there are 3 ISSs, of which ability of the maintenance are the same, being in charge of all aircrafts' maintenance. The maintenance time and cost of each LRM is given in Table 2.

Consider that there are 100 individuals in populations, and one of these individuals is described as follows:

$$U = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}^T. \quad (25)$$

Set up the  $k_1 = k_2 = 0.97, k_3 = k_4 = 0.02$ . The result is described in Figure 4 after 250 iterations.

TABLE 2: The maintenance time of the LRM.

LRM	A	B	C	D
Maintenance time	20 min	25 min	11.6 min	16.6 min
Maintenance cost	2348.2	2843	1297.3	1009.2

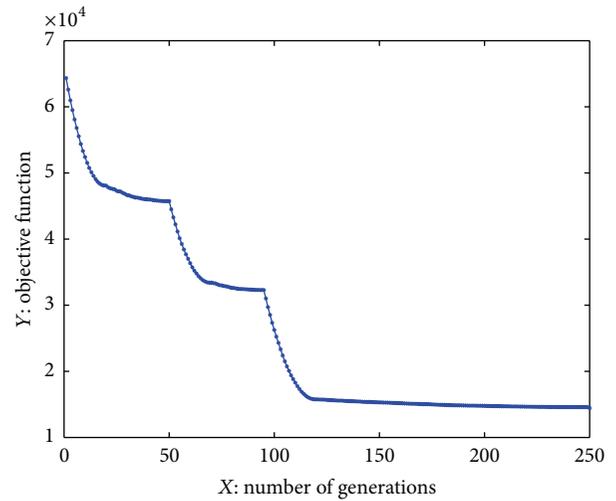


FIGURE 4: Result of calculation.

The total cost of the maintenance is 14439.3 and the mission risk is  $8.95 \times 10^{-07}$  which meet the requirement.

Then, optimal maintenance program can be written as Table 3:

$$U = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}^T, \quad (26)$$

$$S = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T.$$

Then, the optimal scheme of aircraft CBM and dispatching are described completely in Table 3, where the elements in the table such as LRM<sub>C</sub>, LRM<sub>B/D</sub> are the LRMs that need to be repaired. There are six aircrafts and eight LRMs that need to be repaired, and the numbers of the aircrafts that need dispatch are 1, 3, 4, 5, 6, 7, 8, and 10.

### 6. Conclusion

This paper researches optimization decision method for aircraft fleet CBM oriented to mission success considering

TABLE 3: The optimal scheme for aircraft fleet CBM and dispatching.

Aircraft number	1	2	3	4	5	6	7	8	9	10
ISS 1	LRM <sub>C</sub>	/	/	/	/	LRM <sub>B</sub>	/	/	/	/
ISS 2	/	/	LRM <sub>C</sub>	/	/	/	/	LRM <sub>B/D</sub>	/	/
ISS 3	/	/	/	LRM <sub>D</sub>	/	/	/	/	/	LRM <sub>B/D</sub>
Dispatch?	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes

prognostics uncertainty and the resource constrain. The CBM and dispatch process of fleet is analyzed; the modeling method and an improved genetic algorithm for the problem are given, and the method is verified by case about fleet with 10 aircrafts.

The main advantages of this method are shown as follows.

- (1) The alternative strategy sets for single aircraft are defined; then, the optimization problem of fleet CBM is translated to the combinatorial optimization problem of single aircraft strategy. The relationship between maintenance strategy and mission risk is established, and the problem becomes easier to solve.
- (2) This paper used the RUL distribution, which has the maximum information and the highest in prognostics. It has more accurate description of the uncertainty compared with others.
- (3) The optimization decision with risk for fleet CBM is realized. The fleet mission risk is quantitatively assessed, and the optimal CBM strategy for fleet could satisfy the requirement of lowest maintenance cost and acceptable risk.

This paper presents a theoretical approach for fleet CBM considering prognostics uncertainty. Some factors have been simplified, such as the cost of risk, the consequences of risk mission, the effect of the CBM process form ability of maintenance personnel, and the effect of random failures. The focus of further work is a more detailed and comprehensive model considering all above factors.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## References

- [1] B. Sun, S. Zeng, R. Kang, and M. G. Pecht, "Benefits and challenges of system prognostics," *IEEE Transactions on Reliability*, vol. 61, no. 1, pp. 323–335, 2012.
- [2] B. Sun, S. Liu, L. Tong, L. Shunli, and F. Qiang, "A cognitive framework for analysis and treatment of uncertainty in prognostics," *Chemical Engineering Transactions*, vol. 33, pp. 187–192, 2013.
- [3] I. Lopez and N. Sarigul-Klijn, "A review of uncertainty in flight vehicle structural damage monitoring, diagnosis and control: challenges and opportunities," *Progress in Aerospace Sciences*, vol. 46, no. 7, pp. 247–273, 2010.
- [4] J. Fang, M. Xiao, Y. Zhou, and Y. Wang, "Optimal dynamic damage assessment and life prediction for electronic products," *Chinese Journal of Scientific Instrument*, vol. 32, no. 4, pp. 807–812, 2011.
- [5] I. Barlas, G. Zhang et al., *Confidence Metrics and Uncertainty Management in Prognosis*, MARCON, Knoxville, Tenn, USA, 2003.
- [6] B. P. Leão and J. P. P. Gomes, "Improvements on the offline performance evaluation of fault prognostics methods," in *Proceedings of the IEEE Aerospace Conference*. IEEE Computer Society, pp. 1–6, 2011.
- [7] B. P. Leão, T. Yoneyama, G. C. Rocha, and K. T. Fitzgibbon, "Prognostics performance metrics and their relation to requirements, design, verification and cost-benefit," in *Proceedings of the International Conference on Prognostics and Health Management (PHM '08)*, October 2008.
- [8] A. Saxena, J. Celaya, B. Saha, S. Saha, and K. Goebel, "Evaluating prognostics performance for algorithms incorporating uncertainty estimates," in *Proceedings of the IEEE Aerospace Conference*, March 2010.
- [9] I. A. Raptis and G. Vachtsevanos, "An adaptive particle filtering-based framework for real-time fault diagnosis and failure prognosis of environmental control systems," in *Proceedings of the Prognostics and Health Management*, 2011.
- [10] L. Tang, J. Decastro, G. Kacprzynski, K. Goebel, and G. Vachtsevanos, "Filtering and prediction techniques for model-based prognosis and uncertainty management," in *Proceedings of the Prognostics and System Health Management Conference (PHM '10)*, January 2010.
- [11] B. Saha and K. Goebel, "Uncertainty management for diagnostics and prognostics of batteries using Bayesian techniques," in *Proceedings of the IEEE Aerospace Conference (AC '08)*, March 2008.
- [12] G. Xuefei, H. Jingjing, J. Ratneshwar et al., "Bayesian fatigue damage and reliability analysis using Laplace approximation and inverse reliability method," in *Proceedings of the Prognostics and Health Management Society Conference (PHM Society '11)*, 2011.
- [13] L. Tang, G. J. Kacprzynski, K. Goebel, and G. Vachtsevanos, "Methodologies for uncertainty management in prognostics," in *Proceedings of the IEEE Aerospace Conference*, March 2009.
- [14] A. Coppe, R. T. Haftka, and N.-H. Kim, "Least squares-filtered Bayesian updating for remaining useful life estimation," in *Proceedings of the 51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, April 2010.
- [15] M. Orchard, G. Kacprzynski, K. Goebel, B. Saha, and G. Vachtsevanos, "Advances in uncertainty representation and management for particle filtering applied to prognostics," in *Proceedings of the International Conference on Prognostics and Health Management (PHM '08)*, October 2008.
- [16] M. L. Neves, L. P. Santiago, and C. A. Maia, "A condition-based maintenance policy and input parameters estimation for deteriorating systems under periodic inspection," *Computers and Industrial Engineering*, vol. 61, no. 3, pp. 503–511, 2011.

- [17] P. A. Sandborn and C. Wilkinson, "A maintenance planning and business case development model for the application of prognostics and health management (PHM) to electronic systems," *Microelectronics Reliability*, vol. 47, no. 12, pp. 1889–1901, 2007.
- [18] Q. Feng, H. Peng, and D. W. Coit, "A degradation-based model for joint optimization of burn-in, quality inspection, and maintenance: a light display device application," *International Journal of Advanced Manufacturing Technology*, vol. 50, no. 5-8, pp. 801–808, 2010.
- [19] B. Wu, Z. Tian, and M. Chen, "Condition, based maintenance optimization using neural network, based health condition prediction," *Quality and Reliability Engineering International*, vol. 29, no. 8, pp. 1151–1163, 2013.
- [20] K. T. Huynh, A. Barros, and C. Berenguer, "Maintenance decision-making for systems operating under indirect condition monitoring: value of online information and impact of measurement uncertainty," *IEEE Transactions on Reliability*, vol. 61, no. 2, pp. 410–425, 2012.
- [21] R. Flage, D. W. Coit, J. T. Luxhøj, and T. Aven, "Safety constraints applied to an adaptive Bayesian condition-based maintenance optimization model," *Reliability Engineering and System Safety*, vol. 102, pp. 16–26, 2012.
- [22] Q. Feng, S. Li, and B. Sun, "A multi-agent based intelligent predicting method for fleet spare part requirement applying condition based maintenance," in *Proceedings of the 5th International Conference on Multimedia Information Networking and Security*, pp. 808–811, IEEE Computer Society, 2013.
- [23] Q. Feng, S. Li, and B. Sun, "An intelligent fleet condition-based maintenance decision making method based on multi-agent," *International Journal of Prognostics and Health Management*, vol. 3, no. 1, pp. 1–11, 2012.
- [24] M. Srinivas and L. M. Patnaik, "Adaptive probabilities of crossover and mutation in genetic algorithms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 24, no. 4, pp. 656–667, 1994.
- [25] P. Vasant, "A novel hybrid genetic algorithms and pattern search techniques for industrial production planning," *International Journal of Modeling, Simulation, and Scientific Computing*, vol. 3, no. 4, pp. 1–19, 2012.
- [26] P. Vasant, "Hybrid mesh adaptive direct search genetic algorithms and line search approaches for fuzzy optimization problems in production planning," *Intelligent Systems Reference Library*, vol. 38, pp. 779–799, 2013.

## Research Article

# Joint Power and Multiple Access Control for Wireless Mesh Network with Rose Projection Method

Meiqin Tang,<sup>1</sup> Lili Shang,<sup>2</sup> Yalin Xin,<sup>1</sup> Xiaohua Liu,<sup>1</sup> and Xinjiang Wei<sup>1</sup>

<sup>1</sup> School of Mathematics and Statistics Science, Ludong University, Yantai 264025, China

<sup>2</sup> School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 270072, China

Correspondence should be addressed to Meiqin Tang; meiqin@ualberta.ca

Received 27 February 2014; Accepted 19 March 2014; Published 16 April 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Meiqin Tang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper investigates the utility maximization problem for the downlink of the multi-interface multichannel wireless mesh network with orthogonal frequency division multiple access. A cross-layer joint power and multiple access control algorithm are proposed. Rosen projection matrix is combined with Solodov projection techniques to build a three-memory gradient Rosen projection method, which is applied to solve this optimization problem. The convergence analysis is given and simulations show that the proposed solution achieves significant throughput compared with existing approaches.

## 1. Introduction

Both the existing standards and future development trends of broadband wireless communication system make their way forwards to resource allocation policy based on orthogonal frequency division multiple access (OFDMA). In OFDMA technology, each user's effective time and frequency resources are orthogonal to each other in radio link, which can avoid the interference between users and the effect produced by multipath attenuation. For the fixed portable applications, wireless channel is slowly changing. The essential difference of OFDMA technology and other multiple access technologies is that OFDMA can realize multiuser diversity in frequency selective fading channel and at the same time take advantage of the time difference in various channels. As an important part of the new generation of internet technology, wireless mesh networks (WMN) have gradually got into the civil commercial development and application stage, providing last few miles connectivity. We consider the resource control algorithm for WMN-based OFDM. Assuming that all users share the same bandwidth, and all channel state information has been collected in the BS of multilink and all subcarrier wave attenuation channels, as a result, user scheduling and routing selective problem can be solved by way of a joint resolution.

The utility function and economic models have been introduced into networks since the publication of the seminal paper [1] by Kelly et al. in 1998. The utility in the models represents the degree of the user's satisfaction when it acquires certain amount of the resource and the price is the cost per unit resource which the user must pay for this resource. In [2], column generation with greedy pricing was used to maximize the minimum throughput among all flows, and larger cases are considered which compute nearly optimal solutions. A mathematical programming model and assignment algorithms are developed for minimizing the schedule length in adaptive power and adaptive rate link scheduling in spatial-TDMA wireless networks in [3]. The objective in [4] is to maximize the aggregate utility of traffic flows in a multihop wireless network, and the constraints are imposed both due to self-interference and minimum rate requirements. In [5], resource management and admission control schemes for renewable energy sources in wireless mesh network are proposed. The goal is to maximize the energy sustainability of the network.

In this paper we develop a utility maximization problem of joint power control and multiple access resource allocation in wireless mesh network multicarrier orthogonal frequency division. Dual decomposition is used to decouple problem into power control and the multihop orthogonal frequency

division multiple access scheduling in the transport layer and the MAC/PHY layer, which adjusts base cohort arrival rate. On the other hand, scheduling decides the subcarrier link rate and network modulation rate distribution. After transporting layer problem decoupled into speed control problem and the MAC/PHY layer into channel perception and queue scheduling problem, the maximum utility function is given combined with the control station to the other nodes in the output rate. Rosen projection matrix is combined with the Solodov projection techniques, namely, constructing a hyper plane using a point, which separates the current iteration point and the optimal solution of the problem sets. And then the current iteration points in this hyper plane projection, which structures a point list and gets set distance single drop of the optimal solution, thereby establish three-hybrid term memory gradient projection algorithm [6–10]. The convergence of the proposed algorithm is also given in this paper.

## 2. System Model

In this section, the network architecture and radio propagation model are first given, and then the elaborate network utility maximization model is given. The transmission is assumed in a time slot, the channel state information to maintain effective, but each data flow transmission is arbitrary and independent. Each relay node is equipped with a radio interface to accomplish the same time, send, and receive signals. Therefore, sending and receiving the subcarrier can be simultaneously conducted, and only one subcarrier can transmit data along the specified channel.

The network is considered as a flow set  $\varphi$  ( $\varphi = \{s : s = 1, 2, \dots, S\}$ ) from the source node to the base station  $w_s$  ( $s \in \varphi$ ). Each flow transmits from the source node to the destination node path  $L_s$  with rate  $v_s$ , where path  $L_s$  is the set from the source node to destination node pathway, expressed as  $j \in L_s$ . The link capacity vector is expressed as  $C = \{c_{ij}\}$  ( $c_{ij}$  is the channel capacity of the link) which is determined by the scheduling scheme. Scheduling scheme is connected by a set of orthogonal frequency division multiple cross-subcarrier  $\Omega = \{k : k = 1, 2, \dots, K\}$ . The node transmit power is expressed as  $P_i$ .

The noise on the link  $ij$  is  $r_{ij}^k$ , which can be expressed as

$$r_{ij}^k = \frac{|R_{ij}^k|}{H}, \quad (1)$$

where  $H$  represents the noise power density and  $R_{ij}^k$  is channel gain which is determined by path loss, disorder, and decay. We think that adaptive rate can allocate a limited set of modulation rate  $V = \{0, 1, 2, \dots, N\}$ . The corresponding subcarrier  $k$  in the link  $ij$  transmission rate can be expressed as

$$c_{ij}^k = \min \left\{ \left\lfloor \log_2 \left( 1 + r_{ij}^k \right) \right\rfloor, N \right\}. \quad (2)$$

The wireless mesh network from a gateway to the source node is composed of  $n$  nodes. If the distance of adjacent nodes is  $D$ , each node of the required data flow is  $\sigma_i$ . The actual

arrival of the maximum data flow is  $\zeta_i$ , and assume that the actual data traffic on the link is  $F_i$ . We set

$$\frac{\zeta_i}{\sigma_i} = \frac{\zeta_j}{\sigma_j} \quad (3)$$

which is to ensure the fairness of all node data flow. The network throughput is

$$\zeta = \sum_{i=1}^n \phi_i. \quad (4)$$

Nodes access network in a similar way with the MAC protocol in CSMA/CA. When the destination node receives the ACK packets from the source node, a data transmission is finished. Each node transmitting data power is  $P$ ,  $P \in [0, P_{\max}]$ , and  $P_{\max}$  is the maximum of power.  $M(d)$  is the path transmission gain,  $\tau$  is the path attenuation index,  $d$  is the data transmission distance,  $f_c$  is carrier, and  $c$  is light speed; channel attenuation can be expressed as

$$M(d) = \frac{1}{(4\pi f_c/c)^2 d^\tau}. \quad (5)$$

If each node in wireless mesh network can assembly  $K$  interface, and the network has  $N$  orthogonal channel available, and the data can only be transmitted in adjacent nodes, each node can detect whether the adjacent nodes send data, then

$$P_{\min} M(D) > CS_{\text{TH}}. \quad (6)$$

Each node transmitting data power is  $P$ , and the power constraint is

$$P \geq \frac{CS_{\text{TH}}}{M(D)}, \quad P \leq P_{\max}, \quad (7)$$

if  $R = \{r_{BS}^s \geq 0 : s \in \varphi\}$  is long-term average arrival rate to the base station set. Definite  $\Psi = \{C : C = [c_1, c_2, \dots, c_{ij}, \dots]\}$  as all feasible line capacity vector collection, and  $F = \{f_{ij}^{(s)} \geq 0 : s \in \varphi, ij \in L_s\}$  expresses connection set, where  $f_{ij}^{(s)}$  represents  $ij$  link to  $s$  power capacity.

The utility  $U_s$  in the models represents the degree of the user's satisfaction with service quality and cost in terms of resource consumption, which are continuously differentiable nonlinear functions. The optimization problem is formulated as follows:

$$\max_{R, f, C} \sum_{s:n=BS} (v_n^{(s)} \lambda P) \quad (8)$$

s.t.

$$P \leq P_{\max}, \quad P \geq \frac{CS_{\text{TH}}}{M(D)} \quad (9)$$

$$v_n^{(s)} + \sum_{i:n \in L_s} f_{in}^{(s)} \leq \sum_{i:nj \in L_s} f_{nj}^{(s)}, \quad \forall s, n \neq w_s \quad (10)$$

$$f_{ij} = \sum_{s:ij \in L_s} f_{nj}^{(s)} \leq C_{ij}, \quad \forall ij \quad (11)$$

$$f_{ij}^{(s)} = 0, \quad \text{if } ij \notin L_s \quad \forall s, ij \quad (12)$$

$$C \in \Psi, \quad (13)$$

where  $\lambda$  is the weighted factor, which can improve the speed and power of fairness and eliminate the dimension. Equation (10) can ensure that the system reaches the objective of the total efficiency; if  $n \neq BS$ , then

$$v_n^{(s)} = 0, \quad (14)$$

otherwise,

$$\sum_{i:in \in L_s} f_{in}^{(s)} = 0. \quad (15)$$

Equation (11) is the rate constraint. And (12) expresses that if the signal is without the link  $ij$ , the link rate is 0; and (13) expresses the link capacity constraint vector in the feasible region.

### 3. The Proposed Algorithm-Based Mixing Three-Term Memory Gradient Projection Algorithm

Assume

$$x = (v_n^{(s)}, \lambda P). \quad (16)$$

The original problem

$$\max_{R, f, C} \sum_{s:n=BS} (v_n^{(s)} \lambda P) \quad (17)$$

can be translated into

$$\min_{R, f, C} - \sum_{s:n=BS} U_s(x). \quad (18)$$

Assume the utility function of variables in the feasible domain of  $R$ ,

$$R = \left\{ x \in E^n \mid - \sum_{s:n=BS} U_{sj}(x) \leq 0, j \in L = \{1, 2, \dots, s+1\} \right\}, \quad (19)$$

where

$$- \sum_{s:n=BS} U_s(x) \quad (20)$$

is assumed as a nonlinear function:

$$g(x) = -\nabla \left( - \sum_{s:n=BS} U_{s0}(x) \right) \quad (21)$$

$$a_j(x) = \nabla \left( - \sum_{s:n=BS} U_{s0}(x) \right), \quad j \in L.$$

The index set  $j \in L, |J|$  indicates the  $J$  index number, and note that

$$A(x) = (a_j(x), j \in J) = A_J$$

$$J_\delta(x) = \left\{ j \in L \mid - \left( \sum_{s:n=BS} U_{sj}(x) \right) \leq \delta \right\}. \quad (22)$$

Starting from point  $x^k$ , take the function  $x^k$  at the fastest decline in direction as the search direction [10].

First, combine Rosen projection matrix with the Solodov projection techniques to construct a hyper plane, which can separate the current iteration and get the optimal solution of the proposed problem. Secondly, project the current iteration in the hyper plane, which constructs a sequence of points. Then the optimal solution can be obtained as a set distance of single point drop. Thereby establishing the solution of linear or nonlinear constrained optimal problem is converted to three-hybrid term memory gradient projection algorithm.

Assume  $M$  as the K-T point set of the problem of  $\min_{R, f, C} - \sum_{s:n=BS} U_s(x)$ , which satisfies the  $M$  condition mentioned:

$$\begin{aligned} & \nabla \sum_{s:n=BS} U_s(x) + \xi_1 \nabla (P - P_{\max}) + \xi_2 \nabla \left( \frac{CS_{TH}}{G(D)} - P \right) \\ & + \xi_3 \nabla \left( r_n^{(s)} + \sum_{i:in \in L_s} f_{in}^{(s)} \leq \sum_{j:nj \in L_s} f_{nj}^{(s)} \right) \\ & + \xi_4 \nabla \left( \sum_{s:ij \in L_s} f_{ij}^{(s)} - C_{ij} \right) = 0 \\ & f_{ij}^{(s)} = 0, \quad \text{if } ij \notin L_s \quad \forall s, ij \\ & C \in \psi, \end{aligned} \quad (23)$$

where  $\xi_i, i = 1, 2, 3, 4$  is real number.

**Lemma 1.** If  $x^k \in R$  is the non-K-T point of

$$\min_{R, f, C} - \sum_{s:n=BS} U_s(x), \quad (24)$$

where  $\beta_k^1$  meets (4), (5) and  $\beta_k^2$  meets (6), (7), and (8), then

$$\|S_k^2\| \leq \left( 1 + \frac{1}{\Delta_1} + \frac{1}{\Delta_2} \right) \|P_{J_k}(x^k) g(x^k)\|. \quad (25)$$

**Lemma 2.** If  $x^k \in R$  is the non-K-T point of

$$\min_{R, f, C} - \sum_{s:n=BS} U_s(x), \quad (26)$$

where  $\beta_k^1$  meets (3), (4), and (5) and  $\beta_k^2$  meets (8), (9), and (10), then

$$g(x^k)^T S_k^2 \geq \prod_{r=1}^2 \frac{1 + \Delta_r}{2 + \Delta_r} \|P_{J_k}(x^k) g(x^k)\|^2. \quad (27)$$

*Proof.* By the definition of  $S_k^1$  and (1), we can get

$$\begin{aligned} g^{kT} S_k^1 &= g^{kT} P_{J_k}(x^k) (g^k + \beta_k^1 d^{k-1}) \\ &\geq \|P_{J_k}(x^k) g^k\|^2 - |\beta_k^1 g^k P_{J_k}(x^k) d^{k-1}| \\ &\geq \|P_{J_k}(x^k) g^k\|^2 - \frac{1}{1 + \Delta_1} g^{kT} S_k^1. \end{aligned} \tag{28}$$

Based on

$$g^{kT} S_k^1 \geq \frac{1 + \Delta_1}{2 + \Delta_1} \cdot \|P_{J_k}(x^k) g^k\|^2, \tag{29}$$

then by the definition and (2), we have

$$g^{kT} S_k^2 \geq \frac{1 + \Delta_2}{2 + \Delta_2} \cdot g^{kT} S_k^1 \geq \prod_{r=1}^2 \frac{1 + \Delta_r}{2 + \Delta_r} \|P_{J_k}(x^k) g^k\|^2. \tag{30}$$

□

Assuming that

$$\varphi_i(\cdot) : E_+^1 E_+^1 \quad (i = 1, 2), \tag{31}$$

is two continuous functions:

$$\varphi_i(\lambda) = 0 \implies \lambda = 0, \quad i = 1, 2, \tag{32}$$

$$E_+^1 = \{\lambda \mid \lambda \geq 0, \lambda \in E^1\}, \tag{33}$$

we set  $\phi_1(x)$  as nonnegative continuous functions of  $R$  and  $\phi_2(x)$  as right continuous function of  $R$ . The detailed steps of the proposed algorithm are as follows.

*Step 1.* Consider  $\forall x^0 \in R, \delta_0 > 0, \beta > 1, \Delta_1 > 0, \Delta_2 > 0, c > 0, 0 < \tau < 1, d^{0-1} = d^{0-2} = 0$ . Assume

$$k := 0. \tag{34}$$

*Step 2.* Set

$$J_k = J_{\delta_k}(x^k). \tag{35}$$

If

$$|\det(A_{J_k}^T(x^k) A_{J_k}(x^k))| \geq \delta_k, \tag{36}$$

then go to Step 3; otherwise, set

$$\begin{aligned} x^{k+1} &= x^k, \\ \delta_k &= \frac{\delta_k}{\beta}, \quad k := k + 1. \end{aligned} \tag{37}$$

*Step 3.* Calculate  $B_{J_k}, P_{J_k}, U_{J_k}^k$ . If

$$P_{J_k} g^k = 0, \quad (\mu_{J_k}^k)^T \left( - \sum_{s:n=BS} U_{sJ_k}(x^k) \right) = 0, \quad \mu_{J_k}^k \geq 0, \tag{38}$$

then stop, where  $x^k$  is the K-T point of

$$\min_{R, f, C} \sum_{s:n=BS} U_s(x). \tag{39}$$

Otherwise go to Step 4.

*Step 4.*

$$\begin{aligned} d^k &= (1 + |\mu_{J_k}^{kT} W_{J_k}^T|) P_{J_k} \left( g^k + \sum_{r=1}^2 \beta_k^r d^{k-r} \right) + \beta_{J_k}^T \\ &\times \left[ (1 + |\mu_{J_k}^{kT} W_{J_k}^k|) V_{J_k}^k - \left( g^{kT} P_{J_k} \left( g^k + \sum_{r=1}^2 \beta_k^r d^{k-r} \right) \right. \right. \\ &\quad \left. \left. + \mu_{J_k}^{kT} V_{J_k}^k \right) W_{J_k}^T \right], \end{aligned} \tag{40}$$

$$V_{J_k}^k = (V_j^k, j \in J_k)^T$$

$$: V_j^k = \begin{cases} -\varphi_1(-\mu_j^k), & \mu_j^k \leq 0 \\ -\varphi_2(\mu_j^k) - \sum_{s:n=BS} U_{sj}(x^k), & -\mu_j^k > 0 \end{cases}$$

$$W_{J_k}^k = (W_j^k, j \in J_k)^T : W_j^k = \begin{cases} \phi_1(x^k), & j \in L_1 \cap J_k \\ \phi_2(x^k), & j \in L_2 \cap J_k, \end{cases} \tag{41}$$

where  $\beta_k^1$  meets (3), (4), and (5) and  $\beta_k^2$  meets (6), (7), and (8).

*Step 5.* Set  $\lambda \leftarrow 1$ ; if

$$W_j^k = \begin{cases} x^{k+1} = x^k + \lambda_k d^k \in R \\ \nabla \left( - \sum_{s:n=BS} U_{s0}(x^k + \lambda d^k)^T \right) d^k \\ \leq \gamma \nabla \left( - \sum_{s:n=BS} U_{s0}(x^k)^T \right)^T d^k \end{cases} \tag{42}$$

then if

$$\lambda_k = \lambda, \quad y^k = x^k + \lambda d^k \tag{43}$$

go to Step 6; otherwise, take

$$\lambda_{\text{new}} \in [\bar{\sigma}_1 \lambda, \bar{\sigma}_2 \lambda], \quad \lambda \leftarrow \lambda_{\text{new}} \tag{44}$$

and go to Step 5.

*Step 6.* Set

$$J_K = J_{\delta_k}(y^k). \tag{45}$$

If

$$\det(A_{J_K}^T(y^k) A_{J_K}(y^k)) \geq \delta_k, \tag{46}$$

then calculate  $B_{J_K}, P_{J_K}, \mu_{J_K}$ . If

$$P_{J_K} g^k = 0, \quad (\mu_{J_K}^k)^T \left( - \sum_{s:n=BS} U_{sJ_K}(y^k) \right) = 0, \quad \mu_{J_K}^k \geq 0 \tag{47}$$

then stop.  $y^k$  is the K-T point of

$$\min_{R, f, C} - \sum_{s:n=BS} U_s(x), \tag{48}$$

or else go to Step 7. If

$$\det(A_{J_k}^T(y^k)A_{J_k}(y^k)) \geq \delta_k \tag{49}$$

is not established, we set

$$y^{k+1} = y^k, \quad \delta_k := \frac{\delta_k}{\beta} \delta, \quad k := k + 1, \tag{50}$$

and then go to Step 6.

Step 7. If

$$\begin{aligned} v_k &= \nabla \left( - \sum_{s:n=BS} U_{s0}(y^k) \right), \\ \bar{x}_{k+1} &= x^k - \frac{(v_k, x^k - y^k)}{\|v_k\|^2}. \end{aligned} \tag{51}$$

Set

$$x^{k+1} = \bar{x}_{k+1}, \quad k = k + 1, \tag{52}$$

and then go to Step 2.

**Lemma 3.** If  $x^k \in R$  and meets the non K-T point of the

$$\min_{R,f,C} - \sum_{s:n=BS} U_s(x), \tag{53}$$

then  $d^k$  is the descent direction of

$$- \sum_{s:n=BS} U_{s0}(x) \tag{54}$$

at  $x^k$ .

### 4. Convergence Analysis of the Proposed Algorithm

Set  $R^*$  as the optimal solution set of

$$\min_{R,f,C} - \sum_{s:n=BS} U_s(x). \tag{55}$$

Suppose  $R^*$  is convex: the global convergence results are as follows.

**Theorem 4.** Located on

$$\left( - \sum_{s:n=BS} U_{s0}(\cdot) \right) \tag{56}$$

is continuously differentiable pseudorandom function, and  $\{x^k\}$  A is produced by the algorithm (PTMG) to produce infinite iterative sequence; then

$$\lim_{k \rightarrow \infty} x^k = x^*, \tag{57}$$

where  $x^* \in R^*$ .

*Proof.* As we know, sequence  $\{\|x^k - x^*\|\}$  is monotonically decreasing. Therefore,  $x^*$  is bounded; according to Lemma 1,

we can get that  $\{d^*\}$  is bounded. According to the definition of knowledge,  $\{y^k\}$  is bounded. So there exists  $M > 0$  satisfying  $\forall k \in N, \|v_k\| \leq M$ . Thus, by the Step 7, we can get

$$\begin{aligned} \frac{\gamma^2 \lambda_k^2 (\nabla f_0(x^k), d^k)^2}{M^2} &\leq \frac{\lambda_k^2 (\nabla f_0(x^k), d^k)^2}{M^2} \\ &\leq \frac{(v_k, x^k - y^k)^2}{\|v_k\|^2} \leq \|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2. \end{aligned} \tag{58}$$

That is,

$$\begin{aligned} \frac{\gamma^2}{M^2} \sum_{k=1}^{\infty} \lambda_k^2 (\nabla f_0(x^k), d^k)^2 \\ \leq \sum_{k=1}^{\infty} (\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2) < +\infty. \end{aligned} \tag{59}$$

Therefore,

$$\lim_{k \rightarrow \infty} \lambda_k^2 (\nabla f_0(x^k), d^k) = 0, \tag{60}$$

and we get

$$\lim_{k \rightarrow \infty} (\nabla f_0(x^k), d^k) = 0. \tag{61}$$

Actually, if there exists  $\varepsilon_0$ , an infinite subset  $K \in N$ , so that

$$(\nabla f_0(x^k), d^k) \leq -\varepsilon_0, \tag{62}$$

we set

$$\lim_{k \in K, k \rightarrow \infty} \inf \lambda_k = \lambda. \tag{63}$$

Then  $\lambda = 0$ , so there exists an infinite subset  $k \in K$  so that

$$\lim_{k \in K, k \rightarrow \infty} \inf \lambda_k = \lambda. \tag{64}$$

The algorithm shows

$$\sigma_1 \leq \rho_k \leq \sigma_2, \quad \psi_k = \frac{\lambda^k}{\rho^k}, \tag{65}$$

so that

$$(\nabla f_0(x^k + \psi_k d^k), d^k) > \gamma (\nabla f_0(x^k), d^k), \quad \forall k \in K. \tag{66}$$

Suppose

$$\begin{aligned} \lim_{k \in K, k \rightarrow \infty} x^k &= x^*, \\ \lim_{k \in K, k \rightarrow \infty} d^k &= d^*. \end{aligned} \tag{67}$$

There are

$$-\varepsilon_0 (1 - \gamma) \geq (\nabla f_0(x^*), d^*) (1 - \gamma) \geq 0 \tag{68}$$

which is in contradiction with

$$\gamma \in (0, 1), \quad (69)$$

and thus

$$\lim_{k \rightarrow \infty} (\nabla f_0(x^k), d^k) = 0. \quad (70)$$

Any point of  $\{x^k\}$  is the K-T point of

$$\min_{R, f, C} - \sum_{s:n=BS} U_s(x). \quad (71)$$

Then by the pseudo convexity of the

$$\min_{R, f, C} - \sum_{s:n=BS} U_{s_0}(x), \quad (72)$$

we can see  $x^* \in R^*$ .  $\{\|x^k - x^*\|\}$  decreases monotonically limit exists. Therefore,

$$\lim_{k \rightarrow \infty} (x^k - x^*) = \lim_{k \in K, k \rightarrow \infty} x^k - x^* = 0, \quad (73)$$

and so the proposed problem

$$\min_{R, f, C} - \sum_{s:n=BS} U_s(x) \quad (74)$$

is globally convergent.  $\square$

## 5. Numerical Examples

In this section, we aim to show the effectiveness of the proposed algorithm. We uniformly distribute 200 nodes in a square area of dimensions 1000 m  $\times$  1000 m, where each node can select for the transmission across each link.  $P_{\max}$  is set to be 100 mW, and the path attenuation parameter  $\tau$  is set to 4.

**5.1. The Influence of Parameter  $v_n^{(s)}$ .** We first investigate the effect of  $v_n^{(s)}$  on the optimal throughput versus links, and a specified number of links (L) is selected. Figure 1 shows the attained network throughput with  $v_n^{(s)}$  and without it, from which we can find that proposed model can get higher throughput than that without it. It is due to the adjustment of the parameter which can improve the system efficiency.

**5.2. Fairness of the System.** And we will also verify the influence of parameter  $\lambda$  for the users. We select three users randomly. Figure 2 shows the users' powers without the fairness parameter, where user 2 is allocated less power while user 3 is allocated more power; and the fairness can be assured with the parameter, as shown in Figure 3.

**5.3. Performance with Different Methods.** We also show the throughput versus links with different methods in Figure 4, which adopt the same parameter of the system. From the figures we can find that the proposed algorithm can get better performance the algorithms solved with gradient method and Quasi-Newton method.

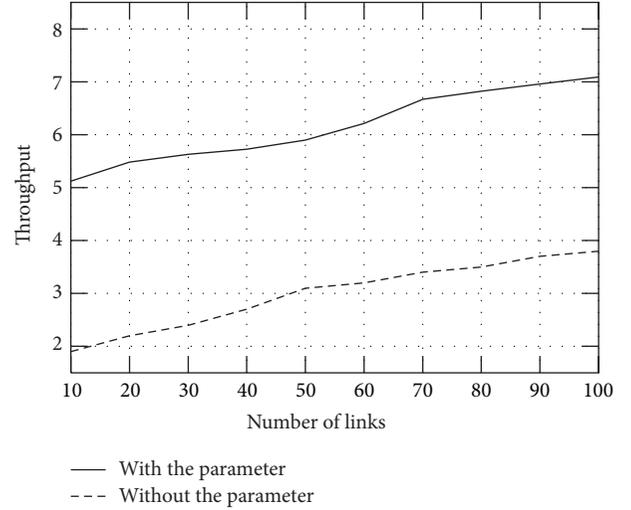


FIGURE 1: The influence of parameter  $v_n^{(s)}$ .

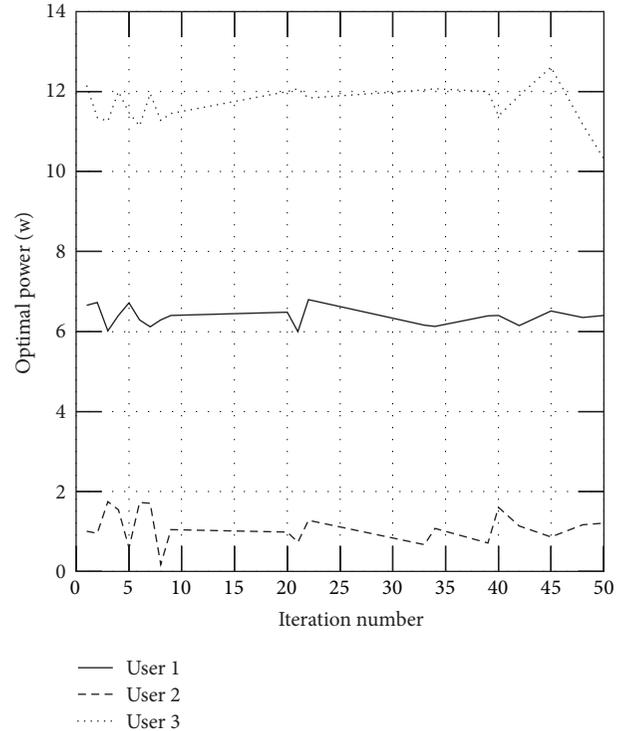


FIGURE 2: Optimal power without the parameter  $\lambda$ .

## 6. Conclusions

This work proposes to optimize joint power and multiple access control in wireless mesh network. Dual decomposition is used to decouple problem into power control and the multihop orthogonal frequency division multiple access scheduling in the transport layer and the MAC/PHY layer. Rosen projection matrix combined with the Solodov projection techniques is used to solve the proposed algorithm. Simulation studies show that the proposed algorithms are

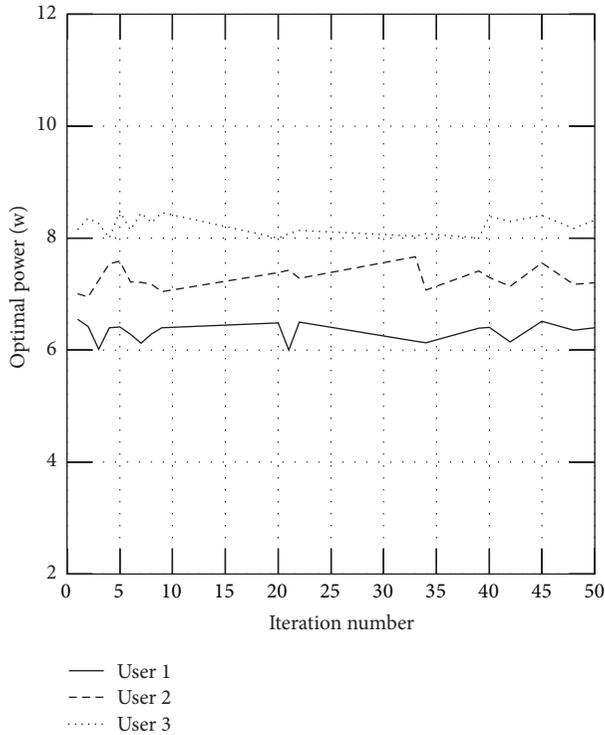


FIGURE 3: Optimal power with the parameter  $\lambda$ .

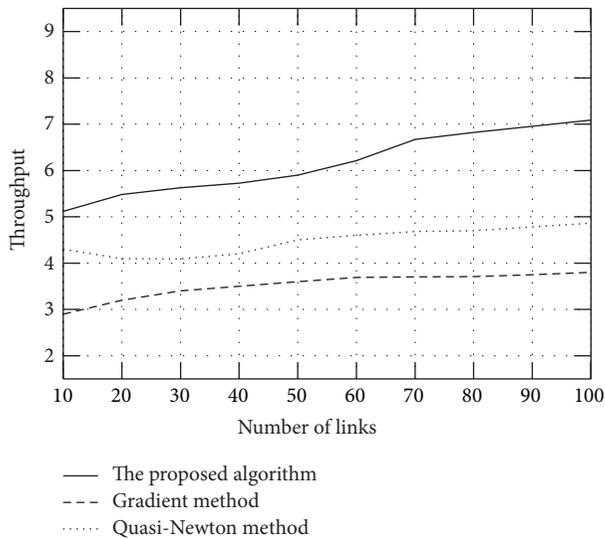


FIGURE 4: Throughput comparison with different methods.

effective to solve the optimization problem and outperform the existing approaches in terms of throughput and the fairness of the users can be assured. The nonconvexity of the resource control problem is our future research work.

### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

### Acknowledgments

This work was supported by National Science Foundation of China Grants 61174097 and 61374108 and Shandong Provincial Natural Science Foundation, China (ZR2011FM016).

### References

- [1] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability," *Journal of the Operational Research Society*, vol. 49, no. 3, pp. 237–252, 1998.
- [2] J. Luo, C. Rosenberg, and A. Girard, "Engineering wireless mesh networks: joint scheduling, routing, power control, and rate adaptation," *IEEE/ACM Transactions on Networking*, vol. 18, no. 5, pp. 1387–1400, 2010.
- [3] K. Hedayati, I. Rubin, and A. Behzad, "Integrated power controlled rate adaptation and medium access control in wireless mesh networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 7, pp. 2362–2370, 2010.
- [4] T. S. Kim, Y. Yang, J. C. Hou, and S. V. Krishnamurthy, "Resource allocation for QoS support in wireless mesh networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 2046–2053, 2013.
- [5] L. X. Cai, Y. K. Liu, T. H. Luan et al., "Sustainability analysis and resource management for wireless mesh networks with renewable energy supplies," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 2, pp. 345–355, 2014.
- [6] J. B. Rosen, "The gradient projection method for nonlinear programming. Part I. Linear constraints," *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, no. 1, pp. 182–217, 1960.
- [7] J. B. Rosen, "The gradient projection method for nonlinear programming. Part 2, nonlinear constraints," *Journal of the Society for Industrial and Applied Mathematics*, vol. 9, no. 4, pp. 514–532, 1960.
- [8] M. V. Solodov and B. F. Svaiter, "A hybrid projection-proximal point algorithm," *Journal of Convex Analysis*, vol. 6, no. 1, pp. 59–70, 1999.
- [9] M. V. Solodov and B. F. Svaiter, "A new projection method for variational inequality problems," *SIAM Journal on Control and Optimization*, vol. 37, no. 3, pp. 765–776, 1999.
- [10] M. V. Solodov and B. F. Svaiter, "A truly globally convergent Newton-type method for the monotone nonlinear complementarity problem," *SIAM Journal on Optimization*, vol. 10, no. 2, pp. 605–625, 2000.

## Research Article

# A High Performance Load Balance Strategy for Real-Time Multicore Systems

Keng-Mao Cho,<sup>1</sup> Chun-Wei Tsai,<sup>2</sup> Yi-Shiuan Chiu,<sup>1</sup> and Chu-Sing Yang<sup>1</sup>

<sup>1</sup>*Institute of Computer and Communication Engineering, Department of Electrical Engineering, National Cheng Kung University, Tainan 70101, Taiwan*

<sup>2</sup>*Department of Applied Informatics and Multimedia, Chia Nan University of Pharmacy & Science, Tainan 71710, Taiwan*

Correspondence should be addressed to Chun-Wei Tsai; [cwtsai0807@gmail.com](mailto:cwtsai0807@gmail.com)

Received 20 February 2014; Accepted 9 March 2014; Published 14 April 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Keng-Mao Cho et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Finding ways to distribute workloads to each processor core and efficiently reduce power consumption is of vital importance, especially for real-time systems. In this paper, a novel scheduling algorithm is proposed for real-time multicore systems to balance the computation loads and save power. The developed algorithm simultaneously considers multiple criteria, a novel factor, and task deadline, and is called power and deadline-aware multicore scheduling (PDAMS). Experiment results show that the proposed algorithm can greatly reduce energy consumption by up to 54.2% and the deadline times missed, as compared to the other scheduling algorithms outlined in this paper.

## 1. Introduction

To promote convenience in people's lives, "smart" has become a new requirement for various products [1], which in turn has led embedded systems to being developed. Embedded systems are now widely used in our daily life, such as digital appliances, network devices, portable devices, and diversified information products [2–8]. Various applications are employed in these devices, and multimedia applications are especially prevalent [9–11]. In order to support the plethora of applications, particularly multimedia-related signal processing, superior performance of embedded systems is required. Along with the increasing demand, the system energy consumption is also increasing. As a matter of fact, the advancement in battery technologies has been slower than the advancement of computing speed and the consequent processor energy consumption.

Due to these reasons and to enhance the performance of modern embedded systems [12–14], the system needs to (1) provide more computation power and (2) reduce power consumption while maintaining performance.

To enhance the *performance* of the embedded system, a multicore architecture is one of the possible solutions which

allow the system to process numerous jobs simultaneously by parallel computation. Keeping every processor core of the system in high utilization is an important issue to achieve high performance. In order to maximize the parallel computing of a multicore system, load balance becomes an issue that needs to be considered when scheduling. Round robin is one of the simple methods to dispatch tasks in a multicore system [15], where tasks are dispatched to processor cores in a rotated order. Shortest queue first [16] is another method that is often used. In this method, tasks are assigned to the processor core with the shortest waiting queue. To find the shortest queue, the number of tasks or the total computation time of tasks on the processor core can be used to represent the queue length. The latter is also called shortest response time first, and requires a priori knowledge about task service times. Additionally, utilization of a processor is usually considered as the criterion in load balance. To generate the maximum balanced load, tasks should be assigned to the processor core with the lowest utilization [17].

In addition to the performance of the embedded system, energy consumption is also an important issue. Over the last decade, manufacturers have been competing to improve the performance of processors by raising the clock frequency.

Under the same technology level and manufacturing processes, the higher operating frequencies of a CMOS-based processor require higher supply voltage. The dynamic power consumption ( $P_{\text{dynamic}}$ ) of a CMOS-based processor is related to the operating frequency ( $f$ ) and supply voltage ( $V_{\text{dd}}$ ) as  $P_{\text{dynamic}} \propto V_{\text{dd}}^2 \times f$ . Thus, higher operating frequency results not only in higher performance but also higher power consumption. Due to the fact that devices which use batteries carry limited energy, research on power saving has received increasing attention, where dynamic voltage and frequency scaling (DVFS) techniques are often applied to extend the battery life of portable devices. DVFS reduces the supply voltage and operating frequency of processors simultaneously to save energy when performance demand is low. Just as the human brain consumes a lot of energy, the processors of a system consume the majority of the energy too. Consequently, multicore architectures can benefit greatly from DVFS technology. In early multicore systems, all processor cores shared the same clock [18]. Under this architecture, DVFS can still be applied to save energy, but there are more limitations. The tradeoff between performance and energy consumption becomes more difficult. To support more flexible power management for multicore systems, the voltage and frequency island (VFI) technique [19, 20] has been developed, where processor cores are partitioned into groups, with processor cores belonging to the same group sharing one supply voltage and having the same processing frequency [21].

*1.1. Motivation.* In the past, most studies regarding scheduling on a multicore system [22–25] have not been designed for real-time systems. For some urgent tasks, raising the priority level of these tasks cannot satisfy the urgency completely. In this case, not only a priority but also a deadline will be used to express the character of this task. Tasks with deadlines are called real-time tasks. Nowadays, some studies have focused on scheduling for real-time multicore systems [26, 27]. However, these kinds of algorithms usually view guaranteeing the hard deadline as their main purpose, and therefore limitations arise. Also, these algorithms need more a priori knowledge of tasks. When implementing them into a real system, they must satisfy some requirements, such as fixed application, training, and specific information about the application. However, most portable devices execute nonspecific applications, which are usually not hard-real-time tasks. For example, users may download numerous applications onto their smart phones, where most are soft-real-time tasks and normal tasks. Unfortunately, it is difficult to ensure which applications will be executed on devices before users actually use them. This is why the design of algorithms for specific applications is not suitable, and the requirement of additional a priori knowledge is difficult to implement efficiently. Thus, to solve these problems and to consider the tradeoff between performance and energy consumption, this paper applies a solution to the problems of scheduling and power saving in a real-time system for a multicore platform. The proposed algorithms decrease the times of deadline missed and simultaneously consider dynamic power, static power, load balance, and practicability.

*1.2. Contribution.* The contributions of this paper are as follows.

- (1) A power and deadline-aware multicore scheduling algorithm is proposed. It is composed of two parts: a power-aware scheduling algorithm and a deadline-aware load dispatch algorithm. The proposed algorithm is simple and easy to implement and overcomes the problems related to many existing power-saving algorithms that are difficult to implement and not suitable for diverse applications.
- (2) In the frequency-scaling part of the power-aware scheduling algorithm, we propose a DVFS-based algorithm called ED<sup>3</sup>VFS. This algorithm uses task deadlines to determine when to scale the operating frequency and is able to adjust parameters dynamically to suit different task sets. Experimental results show that ED<sup>3</sup>VFS is very effective and flexible.
- (3) This paper also proposes a deadline-aware load dispatch algorithm, called the two-level deadline-aware hybrid load balancer. The proposed load dispatch algorithm includes two levels: the concept of load imbalance in the first level and a novel load balance strategy, distribution of task deadline, in the second level. We also combined the other load balance strategies in the second level and let the proposed load dispatch algorithm deal with real-time tasks and normal tasks simultaneously.
- (4) We implemented the proposed load dispatch algorithm in Linux and ported the MicroC/OS-II real-time operating system kernel to a PACDSP on a PAC Duo platform and implemented the proposed power-aware scheduling algorithm in the real-time kernel. Experimental results show that the proposed algorithms work well in a real environment.

*1.3. Organization.* The remainder of this paper is organized as follows. Section 2 gives a brief introduction to work related to scheduling in a multicore system. Section 3 discusses and defines the problems we aim to solve in this paper, as well as limitations and assumptions. Section 4 describes the proposed power and deadline-aware multicore scheduling algorithm. A performance evaluation of the proposed algorithm is presented in Section 5, with conclusion offered in Section 6.

## 2. Related Work

*2.1. DVFS-Based Power Saving Technologies.* There are two strategies for using DVFS techniques to reduce energy consumption. The first strategy is scaling voltage and frequency at task slack time. When a processor serves a task, the operating frequency is multiplied by the rate between the worst-case execution time (WCET) and the deadline of the task [28] to reduce power consumption, as shown in Figure 1. Shin et al. [29] combined offline and online components to satisfy the time constraints and reduce energy consumption. The offline component finds the lowest possible processor speed that

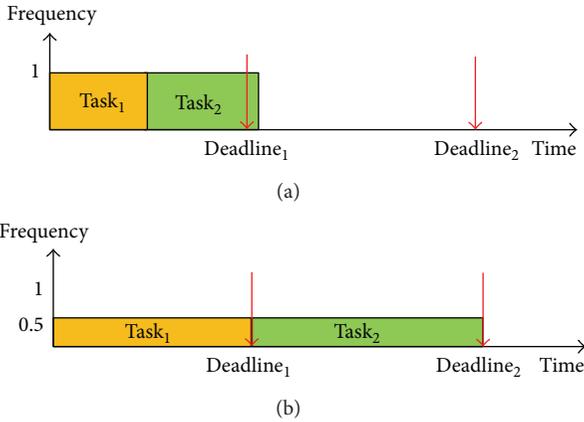


FIGURE 1: Scaling voltage and frequency at task slack time. (a) Without DVFS and (b) with DVFS.

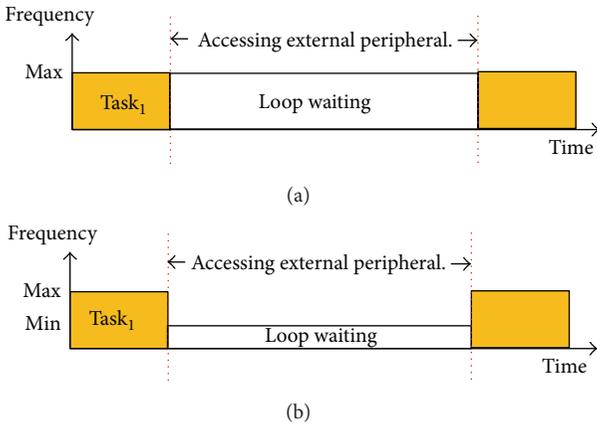


FIGURE 2: Scaling voltage and frequency when access external peripheral. (a) Without DVFS and (b) with DVFS.

satisfies all the time constraints, while the online component varies the operating speed dynamically to save more power. Since the task execution time may be changed slightly when executed, Salehi et al. [30] used an adaptive frequency update interval to follow sudden workload changes. The history data is used to predict the next workload and then, according to the prediction error, adjust the frequency update interval.

The second strategy is scaling voltage and frequency when accessing external peripherals. References [31, 32] pointed out that the operating speed of memory and peripherals is much lower than that of the processors. For tasks that are memory-bounded or I/O-bounded, the operating frequency of a processor can be decreased to save power while waiting for the external peripherals to finish their jobs, as shown in Figure 2. Liang et al. [33] proposed an approximation equation called the memory access rate-critical speed equation (MAR-CSE) and then defined and used the memory access rate (MAR) to predict its critical speed.

2.2. *Scheduling on Real-Time Multicore Systems.* Because the classical approaches need a priori knowledge of the application to achieve the target, especially when real-time

guarantees are provided, Lombardi et al. [34] developed a precedence constraint posting-based offline scheduler for uncertain task durations. This method uses the average duration of a task to replace the probability distribution and calculates an approximate completion time by this cheaper-to-obtain information. Kim et al. [35] presented two pipeline time balancing schemes, namely, workload-aware task scheduling (WATS) and applied database size control (ADSC). Because the execution time of each pipeline stage will change along with the input data, different execution times in each pipeline stage reduce the performance of a system. To achieve higher performance, the pipeline time of each pipeline stage must be in a balanced state. The basic idea of the pipeline time balance schemes is monitoring and modifying the parameter value of the function in each pipeline stage, thereby allowing the execution time of each pipeline stage to be close to the same average value. Jia et al. [36] presented a novel static mapping technique that maps a real-time application onto a multiprocessor system, which optimizes processor usage efficiency. The proposed mapping approach is composed of two algorithms: task scheduling and cluster assignment. In task scheduling, the tasks are scheduled into a set of virtual processors. Tasks that are assigned to the same virtual processors share the maximized data, while data shared among virtual processors is minimized. The goal of cluster assignment is to assign virtual processors to real processors so that the overall memory access cost is minimized.

In addition to balancing the utilization of each processor core, how to tackle the communications among tasks with performance requirements and precedence constraints is another challenge in the scheduling on real-time multicore systems. Hsiu et al. [37] considered the problem of scheduling real-time tasks with precedence constraints in multilayer bus systems and minimized the communication cost. They solved this problem via a dynamic-programming approach. First, they proposed a polynomial-time optimal algorithm for a restricted case, where one multilayer bus and the unit execution time and communication time are considered. The result was then extended as a pseudopolynomial-time optimal algorithm to consider multiple multilayer buses. To consider transition overhead and design for applications with loops, Shao et al. [38] proposed a real-time loop scheduling-algorithm called dynamic voltage loop scheduling (DVLS). In DVLS, the authors succeeded in repeatedly regrouping a loop based on rotation scheduling and decreased the energy consumed by DVS within a timing constraint.

In addition to the abovementioned studies, there are many research directions and issues regarding real-time multicore systems. For real-time applications, it is common to estimate the worst case performance early in the design process without actual hardware implementation. It is a challenge to obtain the upper bound on the worst case response time considering practical issues such as multitask applications with different task periods, precedence relations, and variable execution times. Yet, Yang et al. [39] proposed an analysis technique based on mixed integer linear programming to estimate the worst case performance of each task in a non-preemptive multitask application on a multiprocessor system.

Seo et al. [26] tackled the problem of reducing power consumption in a periodic real-time system using DVS on a multicore processor. The processor was assumed to have the limitation that all cores must run at the same performance level. And so to reduce the dynamic power, they proposed a dynamic repartitioning algorithm. The algorithm dynamically balances the task loads of multiple cores to optimize power consumption during execution. Further, they proposed a dynamic core scaling algorithm, which adjusts the number of active cores to reduce leakage power consumption under low load conditions.

Cui and Maskell [40] proposed a look-up table-based event-driven thermal estimation method. Fast event driven thermal estimation is based upon a thermal map, which is updated only when a high level event occurs. They developed a predictive future thermal map and proposed several predictive task allocation policies based on it. Differing from the utilization-based policy, they used the thermal-aware policies to reduce the peak temperature and average temperature of a system. Han et al. [27] presented synchronization-aware energy management schemes for a set of periodic real-time tasks that accesses shared resources. The mapping approach allocates tasks accessing the same resources to the same core to effectively reduce synchronization overhead. They also proposed a set of synchronization-aware slack management policies that can appropriately reclaim, preserve, release, and steal slack at runtime to slow down the execution of tasks and save more energy. Chen et al. [41] explored the online real-time task scheduling problem in heterogeneous multicore systems and considered tasks with precedence constraints and nonpreemptive task execution. In their assumption, the processor and the coprocessor have a master-slave relationship. Each task will first be executed on the processor and then dispatched to the coprocessor. During online operation, each task is tested by admission control, which ensures the schedulability. Since the coprocessor is nonpreemptive, to deal with the problem of a task having too large a blocking time, the authors inserted the preemptive points to configure the task blocking time and context switch overhead in the coprocessor.

**2.3. Summary.** To extend the system lifetime for energy-limited devices, one of the possible ways is to use DVFS-based technology [28, 29, 31, 32] to save energy. Because the requirements will change when the real-time system is being used, other studies [27, 42, 43] have combined DVFS technology and a real-time scheduler to meet the time constraint while reducing energy consumption. Now, under the environment of multicore architectures, researchers have proposed multicore schedulers, which can meet real-time constraints and consume lower energy. Along with the development of technology, the algorithms have become much more complex and have increasing restrictions when multiple issues need to be considered simultaneously. As a consequence, these algorithms become difficult to implement and work in real environments. Thus, this paper relaxed some limitations to allow the proposed algorithm to be easier to implement and work well in real environments with

simultaneous consideration to the real-time, power, and load balance issues.

### 3. Problem Definition

The DVFS-based power-aware scheduling problem is defined as finding a schedule that can satisfy all the constraints of a system while consuming less energy to execute tasks. Differing from the aforementioned traditional scheduling on a single core system, scheduling on a multi-core system needs to decide not only the execution order of tasks, but also which tasks should be executed on which processor core. A good load dispatch can improve the performance and reduce energy consumption, so load dispatch is a very important issue in multicore scheduling. In this paper, we divided the power-aware multicore scheduling problem into load dispatch and power-aware scheduling and proposed different algorithms to solve them individually. Additionally, the key points of using the DVFS technique to reduce the energy consumption of a system include deciding when the operating voltage and frequency should be adjusted and selecting the operating state. To relax the limitations and make the proposed algorithm easier to implement and more light-weight, missing deadline is allowed in this research. The problem we consider in this paper can be defined as follows and is illustrated in Figure 3.

*System Model.* There is one master processor unit and  $n$  slave processor cores in the system, and each processor core has its own operating system and can scale the operating voltage and frequency independently. When tasks are released, the master processor unit exchanges the status information with each slave processor core by IPC and dispatches them to a suitable slave processor core individually; then, the slave processor cores schedule tasks that are dispatched to them individually. Under this architecture, the proposed algorithm can apply to either homogeneous multicore systems or heterogeneous multicore systems and each processor core can manage itself. In this work, the platform that is used contains an ARM core as master processor unit and two DSPs.

*Input.* A task set  $T = \{T_{\text{real}}, T_{\text{normal}}\}$ , where  $T_{\text{real}}$  is the set of real-time tasks and  $T_{\text{normal}}$  is the set of normal tasks. Each real-time task can be represented as (*Release\_time*, *Priority*, *Related\_Deadline*) and it can be a periodic task or an aperiodic task. In a dynamic environment, it is difficult to get all the information about tasks and the overhead of using optimization method is too heavy when every task is released. Therefore, we tried to schedule real-time tasks without considering the execution time of tasks. Although hard deadline is not guaranteed, methods are proposed to decrease the missing-deadline probability. For normal tasks, we only use (*Release\_time*, *Priority*) to represent them. Generally, the execution order of real-time tasks is based on the absolute deadline of tasks. *Priority* of real-time tasks is used only when two or more absolute deadlines are identical. The details of scheduling algorithm will be described in Section 4.

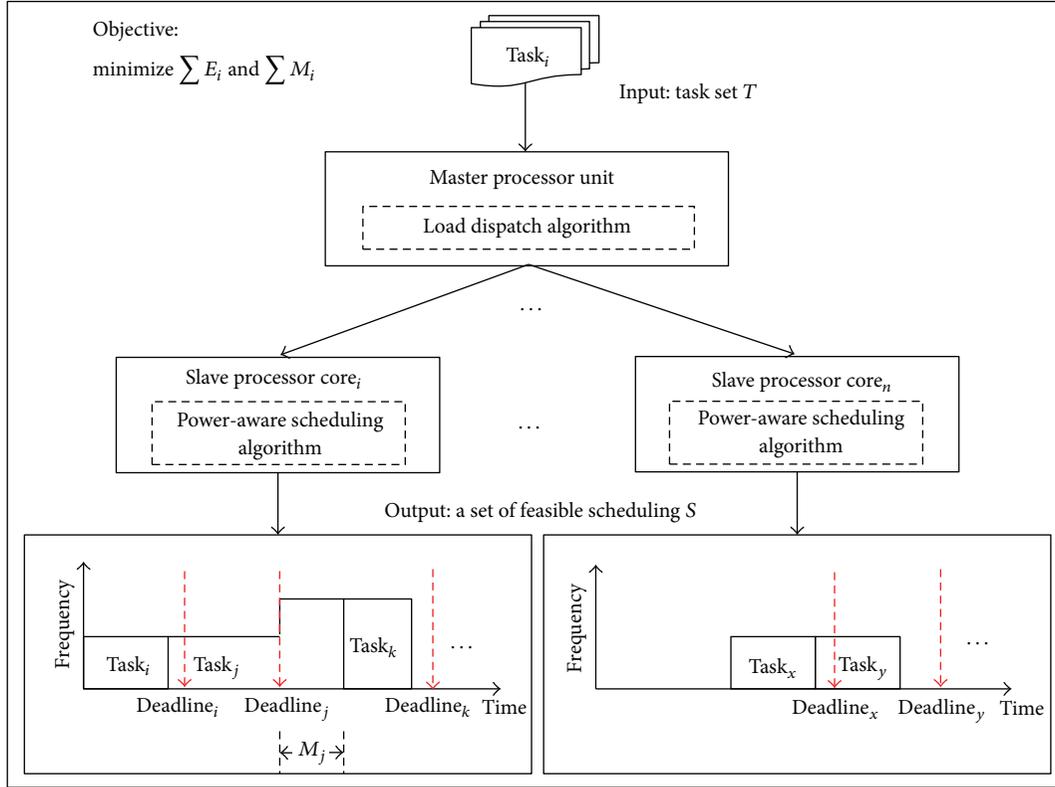


FIGURE 3: System model.

*Output.* A set of feasible scheduling,  $S = \{s_1, s_2, \dots, s_n\}$ , where  $s_i$  is the scheduling result of the  $i$ th slave processor core and scaling operating voltage and frequency produced by the proposed scaling algorithm.

*Objective.* To minimize total energy consumption  $E_{total}$ , the objective function is expressed in

$$\text{Minimize } (E_{total}), \tag{1}$$

$$E_{total} = \sum E_i, \tag{2}$$

$$E_i = \sum E_{ij}, \tag{3}$$

where  $E_i$  is the energy consumption of the  $i$ th slave processor core and  $E_{ij}$  is the energy consumption of the  $j$ th task in the  $i$ th slave processor core, as expressed in

$$E_{ij} = \sum (P_{ijk} \times t_{ijk}), \tag{4}$$

$$P_{ijk} = c \times V_{ijk}^2 \times f_{ijk}, \tag{5}$$

where  $P_{ijk}$  is the power of the  $j$ th task at the  $k$ th time slice in the  $i$ th slave processor core, expressed as (5), and  $t_{ijk}$  is the duration of the  $k$ th time slice for the  $j$ th task in the  $i$ th slave processor core. Since the operating mode will not always be the same for a given task under the proposed algorithm, the energy in each time slice must be individually calculated and then summarized. In (5),  $c$  is the load capacity,  $V_{ijk}$  is the

operating voltage, and  $f_{ijk}$  is the operating frequency of the  $j$ th task at the  $k$ th time slice in  $i$ th slave processor core.

It is allowed that tasks are finished after their deadlines in this paper. The processing speed will be increased when a deadline is missed so that the task can be finished faster. Thus, the performance constraint can be expressed as

$$\text{Minimize } (\sum M_i), \tag{6}$$

where  $M_i$  represents if the deadline of  $i$ th task is missed and is defined as

$$M_i = \begin{cases} 1, & \text{if } ft_i > d_i \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

where  $ft_i$  is the finish time and  $d_i$  is the absolute deadline of the  $i$ th task.

#### 4. Power and Deadline-Aware Multicore Scheduling

In this section, an efficient multicore scheduling algorithm is presented, called power and deadline-aware multicore scheduling, which integrates three different parts (modules): (1) mixed-earliest deadline first (MEDF) [42], (2) enhanced deadline-driven dynamic voltage and frequency scaling (ED<sup>3</sup>VFS), and (3) two-level deadline-aware hybrid load balancer (TLDHLB). Among them, MEDF is used to schedule

the tasks that have been dispatched to a processor core. ED<sup>3</sup>VFS is an enhanced version of D<sup>3</sup>VFS [42] and is used to scale the operating mode on each slave processor core. And finally, TLDHLB is used for task dispatch and is composed of two levels: the first level is the load imbalance strategy, while the second level is load balance. For example, when a new task that needs to be served by DSP arrived, TLDHLB will dispatch it to a DSP with consideration of load balance. After the DSP receives the task, MEDF will schedule the new task. At the same time, ED<sup>3</sup>VFS will be executed on DSP periodically to reduce the energy consumption.

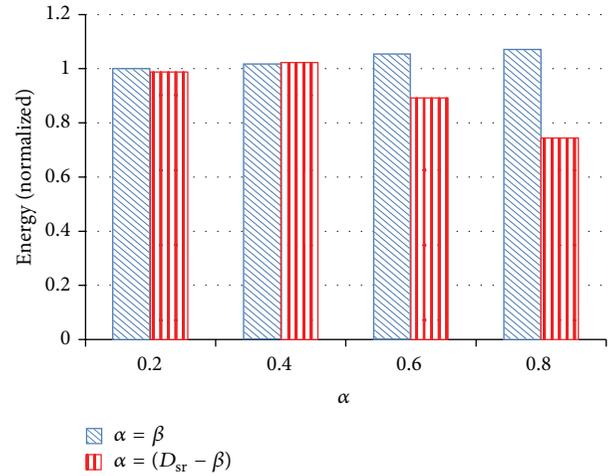
**4.1. Mixed-Earliest Deadline First.** The original scheduling algorithm used in the MicroC/OS-II kernel is a fundamental priority-based scheduling algorithm. To support real-time tasks and normal tasks in the same time, mixed-earliest deadline first is selected to replace the original scheduling algorithm. MEDF combined EDF and fixed-priority scheduling. For real-time tasks, it uses EDF to schedule the tasks. When there are two or more deadlines of real-time tasks that are identical or for normal tasks, MEDF uses fixed-priority scheduling to decide the execution order. Moreover, MEDF will always select real-time tasks first when there are real-time tasks and normal tasks in the ready queue simultaneously.

To cooperate with TLDHLB to save static power, we modified MEDF to let it turn off the processor cores while the ready task, which has the highest priority, is idle while there is no real-time task or other normal tasks. This means that when a processor core finishes all tasks, it will turn itself off to save power.

**4.2. Enhanced Deadline-Driven Dynamic Voltage and Frequency Scaling.** The D<sup>3</sup>VFS scales operating mode dynamically by the active status of a system.  $\alpha$  and  $\beta$  are two parameters used in D<sup>3</sup>VFS and are set as 10. D<sup>3</sup>VFS will scale operating modes while the system continues to be busy for  $\alpha$  time units or when no deadlines are missed for  $\beta$  time units. Inspired by observations of D<sup>3</sup>VFS, we present a better strategy to set parameters  $\alpha$  and  $\beta$  to enhance the performance of the scheduling system. First, in D<sup>3</sup>VFS, the related deadlines of tasks are not needed to be longer than 10 or a fixed threshold. When the shortest related deadline is shorter than this value of 10 (threshold),  $\alpha$  and  $\beta$  become negative, which should be amended. Second, the purposes of  $\alpha$  and  $\beta$  are not exactly the same, and so their impacts are different. The bigger value we set to  $\alpha$ , which is the longer time that the processor stays in lower speed, because the system will increase the operating frequency in a slower rate. This leads to more power savings with worse computing capacity. On the other hand,  $\beta$  is used for decreasing the operating frequency. A bigger  $\beta$  will allow the system to stay in high speed for longer, so that the performance will be better, but more power is consumed. Third, there is more than one task working in a system simultaneously. And in real environment, these tasks may not always be the same. The best setting for each task is different, and so giving different settings for different task sets is more flexible.

TABLE 1: Overall settings of  $\alpha$  and  $\beta$ .

Settings	Group 1 ( $\alpha = \beta$ )		Group 2 ( $\alpha = D_{sr} - \beta$ )	
	$\alpha$	$\beta$	$\alpha$	$\beta$
1	$0.2D_{sr}$	$0.2D_{sr}$	$0.2D_{sr}$	$0.8D_{sr}$
2	$0.4D_{sr}$	$0.4D_{sr}$	$0.4D_{sr}$	$0.6D_{sr}$
3	$0.6D_{sr}$	$0.6D_{sr}$	$0.6D_{sr}$	$0.4D_{sr}$
4	$0.8D_{sr}$	$0.8D_{sr}$	$0.8D_{sr}$	$0.2D_{sr}$

FIGURE 4: Energy comparison between original D<sup>3</sup>VFS and ED<sup>3</sup>VFS.

To solve these aforementioned problems, this paper proposes a different concept to improve the performance of power saving algorithms and conforms to real situations while the system is working. Pseudocode 1 shows the pseudocode of ED<sup>3</sup>VFS. The basic idea of ED<sup>3</sup>VFS is that the settings of  $\alpha$  and  $\beta$  are free to change along with different task sets and these two parameters will be set to two different values. According to the basic idea of ED<sup>3</sup>VFS, we ran a series of experiments to find a better setting and to verify that the new setting is superior to the original D<sup>3</sup>VFS setting. The experimental settings include two main groups and are described as follow.

- (i) The first group is just like the original D<sup>3</sup>VFS, and the settings of  $\alpha$  and  $\beta$  are the same. There are four settings in this group, and the four settings are  $\alpha = \beta = D_{sr} \times 0.2, 0.4, 0.6,$  and  $0.8$ , where  $D_{sr}$  is the shortest related deadline.
- (ii) The second group used in ED<sup>3</sup>VFS also features four settings in. In this group,  $\alpha$  and  $\beta$  will be set into different values.  $\alpha = D_{sr} - \beta$  and  $\beta = D_{sr} \times 0.8, 0.6, 0.4,$  and  $0.2$ . Just as in the above, the effects of  $\alpha$  and  $\beta$  are opposite, so we gave them opposite values. Table 1 lists the overall settings of  $\alpha$  and  $\beta$  in these experiment series.

Figure 4 shows the energy comparison between the two different settings, namely, the original D<sup>3</sup>VFS and ED<sup>3</sup>VFS.

```

(1) Initially, setting the power mode of DSP to default level  $f_d$ .
(2) Every timer interrupt occur
(3) If (There is any real-time task.)
(4)     Setting  $\alpha = D_{sr} \times 0.2$ .
(5)     Setting  $\beta = D_{sr} - \alpha$ .
(6)     if (Deadline missed.)
(7)         Extending deadline of the task that missed deadline for  $d_e$  ticks.
(8)         Rising power mode.
(9)     else if (Utilization of DSP >  $S_r$  for  $\alpha$  ticks.)
(10)        Rising power mode.
(11)    else if (There is no deadline missed for  $\beta$  ticks.)
(12)        Falling power mode.
(13)    else
(14)        Setting the power mode of DSP to default level  $f_d$ .
(15)    End If
    
```

PSEUDOCODE 1: Pseudocode of enhanced deadline-driven dynamic voltage and frequency scaling (ED<sup>3</sup>VFS).

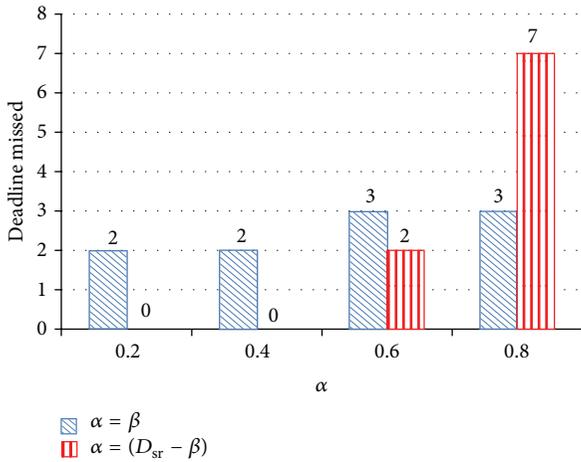


FIGURE 5: Performance comparison between original D<sup>3</sup>VFS and ED<sup>3</sup>VFS.

The vertical axis shows energy consumption while the horizontal axis shows the setting of  $\alpha$ , where  $\alpha = D_{sr} \times 0.2, 0.4, 0.6, \text{ and } 0.8$ . In this research, the first experimental result of energy is used to normalize all of the other results and make them easy to see the differences in comparison. The results show that the energy consumption of ED<sup>3</sup>VFS is lower than the original D<sup>3</sup>VFS in most cases. Figure 5 shows the performance comparison between original D<sup>3</sup>VFS and ED<sup>3</sup>VFS. We used the number of deadlines missed as the criterion to compare their performance. For a real-time system, a lower number of missed deadlines are better. The results show that ED<sup>3</sup>VFS is better than the original D<sup>3</sup>VFS in performance, except for the case when  $\alpha = D_{sr} \times 0.8$ . Although the proposed power-saving algorithm does not guarantee the hard deadline, we still try to stop missed deadlines from happening. According to the experimental results shown in Figure 5, there are two settings we can choose,  $\alpha = D_{sr} \times 0.2$  and  $0.4$  when  $\beta = D_{sr} - \alpha$ . There was no deadline missed in these two cases. Since the energy

consumption of  $\alpha = D_{sr} \times 0.2$  is less than  $\alpha = D_{sr} \times 0.4$ , we chose  $\alpha = D_{sr} \times 0.2$  and  $\beta = D_{sr} - \alpha$  as the final setting of ED<sup>3</sup>VFS. In this setting, both energy consumption and performance of ED<sup>3</sup>VFS are superior to the original D<sup>3</sup>VFS. Actually, if energy consumption is more important than performance in a given system, setting  $\alpha = D_{sr} \times 0.6$  and  $\beta = D_{sr} - \alpha$  is also a good choice. In that case, more energy can be saved and a certain level of performance maintained.

4.3. Two-Level Deadline-Aware Hybrid Load Balancer. For systems limited by battery power, letting all processor cores work continuously in active mode is not a good idea, as much energy will be consumed. In a multicore system, balancing the workload between each core can reduce the completion times of all tasks. Processor cores thus can turn to sleep mode for longer times and save more energy.

In this paper, a novel task dispatch algorithm, called two-level deadline-aware hybrid load balancer (TLDHLB), is presented. The first level is the load imbalance strategy used for saving static power, and it was inspired by [44]. The basic idea of the first level of task dispatch is dispatching tasks to the processor cores working in active mode and turning off the processor cores when all tasks are finished. For example, suppose there are one MPU and two DSPs in the system. Initially, the system will turn off two DSPs until there are tasks needing to be processed by DSP, as shown in Figure 6(a). When task<sub>1</sub> was released, MPU will check the state of the DSPs. If there is no DSP working in active mode, then turn on DSP<sub>1</sub> and dispatch task<sub>1</sub> to DSP<sub>1</sub>. According to ED<sup>3</sup>VFS, DSP<sub>1</sub> will work at the default speed, normally, and at the lowest speed in the beginning, as shown in Figure 6(b). Figure 6(c) shows that if task<sub>n</sub> is released at time<sub>n</sub> while DSP<sub>1</sub> works at full speed, then turn on DSP<sub>2</sub> and dispatch task<sub>n</sub> to DSP<sub>2</sub>. Figure 6(d) shows that, after DSP<sub>1</sub> finished all tasks assigned to it, it will turn itself off by MEDF.

The second level is used for load balance. When there are two or more DSPs working in active mode, the load balance strategies in the second level will be used to dispatch tasks that are newly released. Unlike traditional systems that do

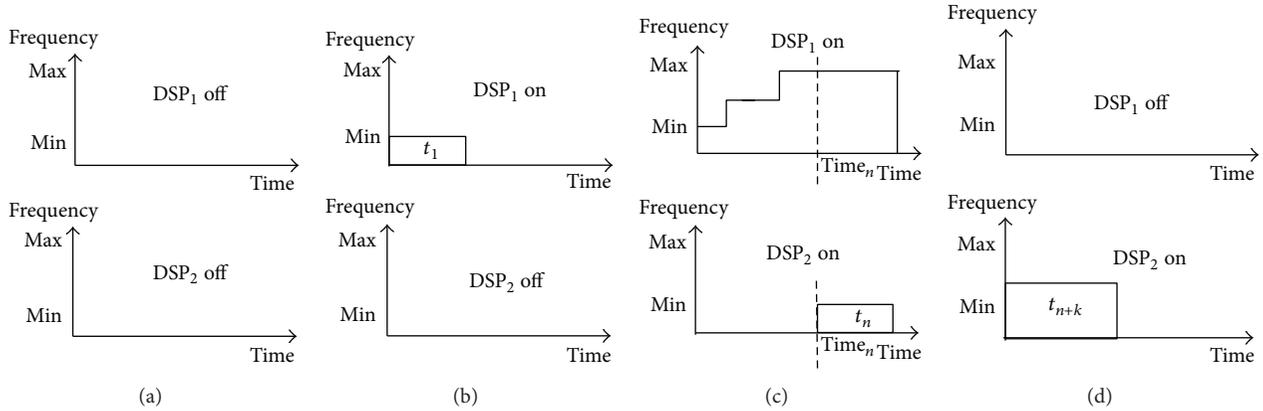


FIGURE 6: Example of load unbalance. (a) Initial state, (b) Task<sub>1</sub> released and turning on DSP<sub>1</sub>. (c) Task<sub>n</sub> released while DSP<sub>1</sub> worked in full speed and turning on DSP<sub>2</sub>. (d) DSP<sub>1</sub> finished all of tasks that dispatched to it and turning off DSP<sub>1</sub>.

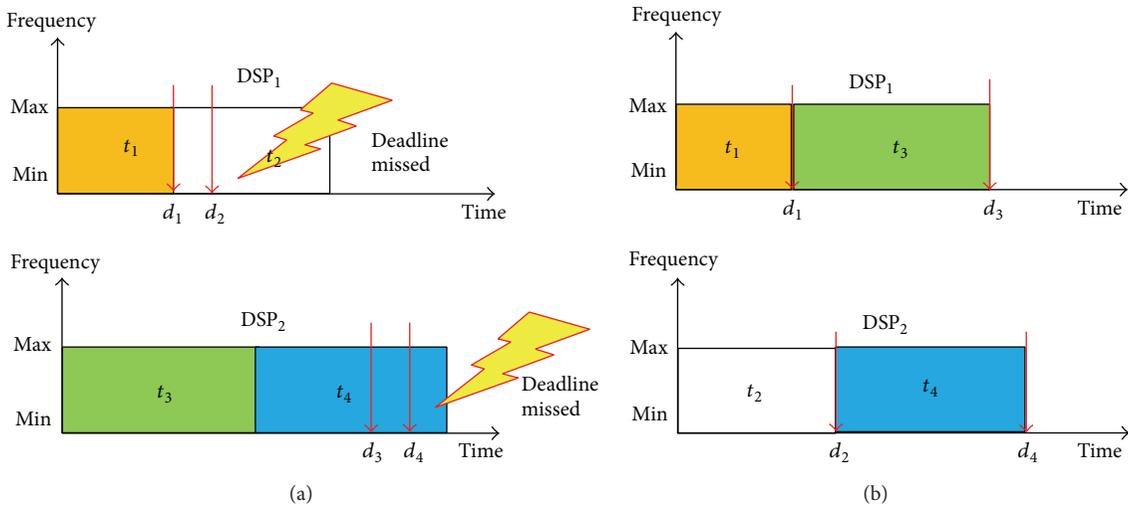


FIGURE 7: Example of real-time tasks dispatch. (a) Distribution of task deadlines is uneven. (b) Distribution of task deadlines is uniform.

not contain real-time tasks, there are simultaneously real-time tasks and normal tasks in a system in our assumption. Traditional load balance strategies were not designed for real-time systems, so we propose a new dispatch criterion to process the problem of dispatch in real-time tasks. We also combined other criteria to allow our load balance algorithm to process real-time tasks and normal tasks simultaneously and improve robustness.

The novel strategy uses the distribution of task deadlines as the criterion for load balance. According to our observation, the more uniform the distribution of task deadlines is, the lower the missing-deadline probability will be. Figure 7 is a simple example that supports our observation. There are two DSPs and four tasks; Figure 7(a) shows that one of the dispatch results' distribution of task deadlines is uneven. In this example, task<sub>1</sub> is finished at its deadline,  $d_1$ , so that there is not enough time to execute task<sub>2</sub> and a deadline missed occurs. A similar situation occurred in DSP<sub>2</sub> for task<sub>4</sub>. Figure 7(b) shows that a different dispatch result features a more uniform distribution of task deadlines. In this situation,

the time slot between two deadlines is longer, which means that there is more time to execute the next task when a task is finished, leading to the probability of a deadline missed being lower. Now, the problem is how to express the distribution of task deadlines.

In this paper, the variance of task deadlines was used as the feature of deadline distribution. Since variance expresses how far a set of numbers is spread out, variance can be used to express the density of data distribution, which is what we need. A smaller variance of task deadlines implies that the time slot between two task deadlines is shorter. Equation (8) is the formula of variance, where  $N$  is the number of data,  $x_i$  expresses  $i$ th data, and  $\bar{x}$  is the data mean:

$$\text{Var}(X) = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2. \tag{8}$$

Except for the distribution of task deadlines, the other three strategies were combined to dispatch normal tasks. These three strategies can also be used to dispatch real-time

```

(1) Initially, turning off all of DSPs,  $DSP_{off} = DSP_{all}$ .
(2) When a task that needs be processed by DSP is released
(3)   getStatus( $DSP_{all}$ )
(4) % First level for load unbalance%
(5)   If ( $DSP_{off} == DSP_{all}$ )
(6)      $DSP_t = minSerial(DSP_{off})$ 
(7)     Turn on minSerial( $DSP_{off}$ ).
(8)   Else if ( $DSP_{fs} == DSP_{on}$ )
(9)     if ( $DSP_{on} == DSP_{all}$ )
(10)      go to line 21
(11)    else
(12)       $DSP_t = minSerial(DSP_{off})$ 
(13)      Turn on minSerial( $DSP_{off}$ ).
(14)   Else if ( $Num(DSP_{on} - DSP_{fs}) == 1$ )
(15)      $DSP_t = DSP_{on} - DSP_{fs}$ 
(16)   Else
(17)     go to line 21
(18)   Dispatch the task to  $DSP_t$ .
(19)   End
(20) % Second level for load balance%
(21)   If ( $Num(maxUniformity(DSP_{on})) == 1$ )
(22)      $DSP_t = maxUniformity(DSP_{on})$ 
(23)   Else if ( $Num(minOrder(DSP_{on})) == 1$ )
(24)      $DSP_t = minOrder(DSP_{on})$ 
(25)   Else if ( $Num(minTaskNum(DSP_{on})) == 1$ )
(26)      $DSP_t = minTaskNum(DSP_{on})$ 
(27)   Else
(28)      $DSP_t = minSerial(DSP_{on})$ 
(29)   Dispatch the task to  $DSP_t$ .
(30)   End

```

PSEUDOCODE 2: Pseudocode of two levels deadline-aware hybrid load balancer.

tasks when the uniformity of deadline distributions is equal. The first strategy is the execution order of the task. This means that the dispatcher will dispatch a task to the DSP that provides higher priority. For example, assume there are two DSPs and two tasks on each DSP. When a new task is released, if the execution order of the task is second in  $DSP_1$  and third in  $DSP_2$ , this task will be dispatched to  $DSP_1$ . The second strategy is the number of tasks. The dispatcher will dispatch a task to the DSP that has the fewest number of tasks on it. When the dispatcher cannot make the decision via the three strategies in the second level mentioned above, then the last strategy will be used. The last strategy is very simple, it simply chooses the DSP whose serial number is the minimum and is working in active mode. The pseudocode of two levels deadline-aware hybrid load balancer is shown in Pseudocode 2, where  $DSP_{all}$  is the set of all DSPs,  $DSP_{off}$  is the set of DSPs that are not staying in active mode,  $DSP_t$  is the target that the task will be dispatched to,  $DSP_{on}$  is the set of DSPs that are staying in active mode,  $DSP_{fs}$  is the set of DSPs that are working at full speed, function *getStatus*() is used to obtain the status information of the DSPs, function *minSerial*() returns the DSP whose serial number is the minimum, function *Num*() returns the number of input data, function *maxUniformity*() returns the DSP whose uniformity of deadline distribution is the maximum, function

*minOrder*() returns the DSP that can provide the highest priority to new released task and function, *minTaskNum*() returns the DSP whose number of tasks is the fewest.

What is worth noticing is that when calculating the uniformity of deadline distribution, we should take the newly released task into account because we want to find a DSP whose deadline distribution is still uniform after inserting the newly released task. Furthermore, *maxUniformity*(), *minOrder*(), and *minTaskNum*() may return more than one DSP, when there are two or more DSPs with the same status. In that case, TLDHLB will use another strategy to dispatch the task and is why we combined four criteria to become a hybrid strategy.

## 5. Experiments

In this section, we describe the experimental environment and the setting of parameters. Experimental results and analyses are then shown.

**5.1. Experimental Environment.** In this study the PAC Duo platform was used for the experiments and includes an ARM926 processor and two PACDSPs. The operating system kernel running on ARM is Linux 2.6.27. We ported the

TABLE 2: Operating voltage and frequency of PACDSP.

Power Mode (operating mode)	Voltage (V)	Frequency (MHz)
7	1.0	204
6	1.0	136
5	0.9	102
4	0.9	68
3	0.9	51
2	0.8	34
1	0.8	24

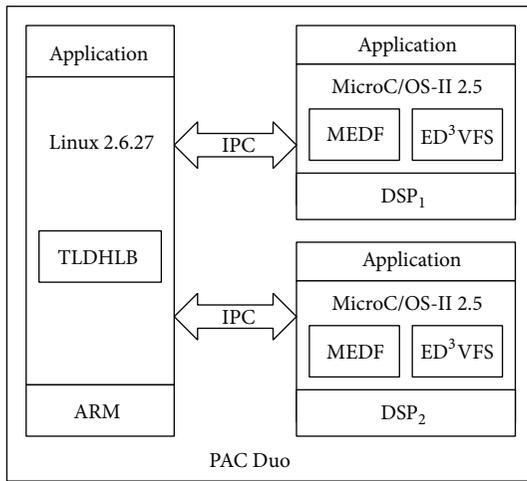


FIGURE 8: System architecture.

MicroC/OS-II kernel (version 2.5) to the PACDSPs and implemented the proposed power-aware scheduling algorithm on the MicroC/OS-II kernel and the proposed load dispatch algorithm on ARM. Figure 8 shows the experimental system architecture, while Table 2 shows the operating frequencies used in the experiments and the corresponding voltages. In the experiments, we used a digital multimeter (FLUKE 8846A) to measure the voltage and current of the PACDSPs, the data of which was used to calculate energy consumption.

**5.2. Experimental Settings.** In the experiments, we used matrix multiplication,  $\pi$  calculation, quick sort, jpeg decoder, and histogram equalization as the workload. Other than the proposed algorithms, we also implemented two load balance algorithms and three frequency scaling strategies as the comparisons. Table 3 shows the algorithm usage in the experiments.

- (i) Seven sets of settings were used. The first set is the proposed algorithms and combines the proposed load dispatch algorithm and power-aware scheduling algorithms. Worthy of note is that the DSPs on our experimental platform cannot turn off and then turn on, so we scaled the operating frequency to the lowest frequency to represent the DSP as being turned off

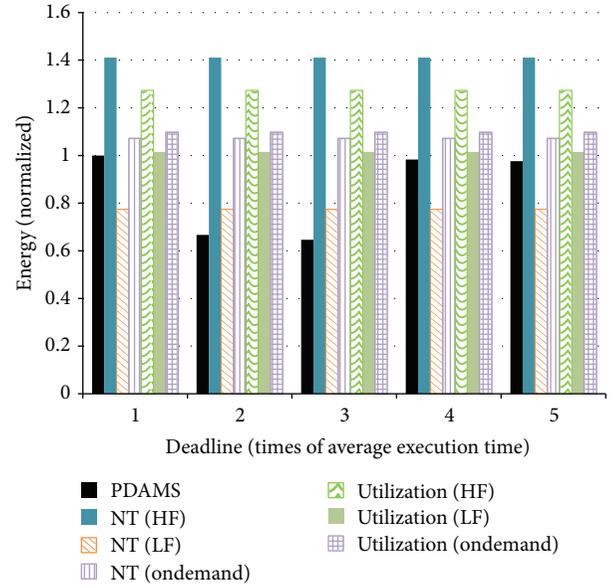


FIGURE 9: Comparison of energy consumption.

and assumed the energy consumption was zero until the proposed algorithms turn it on again.

- (ii) The second set to the fourth set used the same load balance algorithm. In these three sets, the load balancer used the number of tasks as the criterion to dispatch tasks. The frequency scaling strategies used in the second and third sets were two static settings: one is the highest operating frequency and the other is the lowest operating frequency. The fourth set used Linux-ondemand [45] as the frequency scaling strategy. Linux-ondemand is a dynamic frequency scaling algorithm; it is used in Linux kernel and dynamically scales the operating frequency according to the utilization of the processor.
- (iii) The fifth set to seventh set used the utilization of processor as the criterion to dispatch tasks. The frequency scaling strategies used in these three sets were the highest frequency, the lowest frequency, and Linux-ondemand in ordering. Except for the first set, the other settings use original scheduling used in MicroC/OS-II to schedule tasks.

For each task, we used the times of the average execution time of the task as its deadline, from one time to five times. There are five settings of task deadline for each set of settings in the experiments.

### 5.3. Experimental Results

**5.3.1. Comparison of Energy Consumption.** Figure 9 shows the comparison of energy consumption. The vertical axis shows the energy consumption while the horizontal axis shows the setting of the task deadline. The results show that the energy consumption of the proposed algorithms is lower than that of other algorithms in almost every case. Compared

TABLE 3: Usage of algorithms for experiments.

Set	Load dispatch			Frequency scaling		Linux-ondemand
	TLDHLB	The number of tasks	Utilization	ED <sup>3</sup> VFS + MEDF	Always in the highest frequency	
1 PDAMS	✓			✓		
2 NT (HF)		✓			✓	
3 NT (LF)		✓				✓
4 NT (Ondemand)		✓				✓
5 Utilization (HF)			✓		✓	
6 Utilization (LF)			✓			✓
7 Utilization (Ondemand)			✓			✓

with other algorithms, the proposed algorithm can reduce energy consumption by up to 54.2%. Experimental results also show that using the number of tasks as the criterion of load balance and always working in the lowest operating frequency can reduce energy consumption the most. The consequence is obvious and predictable, but the computing capacity under this condition is not satisfying. Although the proposed algorithm considered saving static power, due to hardware constrains, the static power consumption could barely be measured independently. As a result, static power consumption was not taken in count in the experiments. Besides, the overhead of scaling voltage and frequency has not been considered either. These will be added to our next work.

5.3.2. *Comparison of Performance.* Other than energy consumption, performance is a very important criterion to evaluate the effect of an algorithm. How to deal with the tradeoff between energy and performance is a difficult issue. Differing from traditional systems, the finish time of an overall system cannot represent the performance completely in a real-time system. For a real-time task, there is no difference between a system finishing the task very quickly and finishing the task just at deadline. Before the deadline of a real-time task, no matter what time is required to finish it, the effects of the real-time task are the same. Therefore, the number of missed deadlines is a better criterion to represent the performance of a real-time system. Figure 10 offers the performance comparisons. The vertical axis shows the number of deadlines missed while the horizontal axis shows the setting of task deadlines. The results show that the performance of the proposed algorithms is the best. Except for setting the deadline of each task into one time of an average execution time, there is no deadline missed while using the proposed algorithms. Since there are only two slave processor cores in our experimental platform, when there are more than three real-time tasks and their deadlines are just

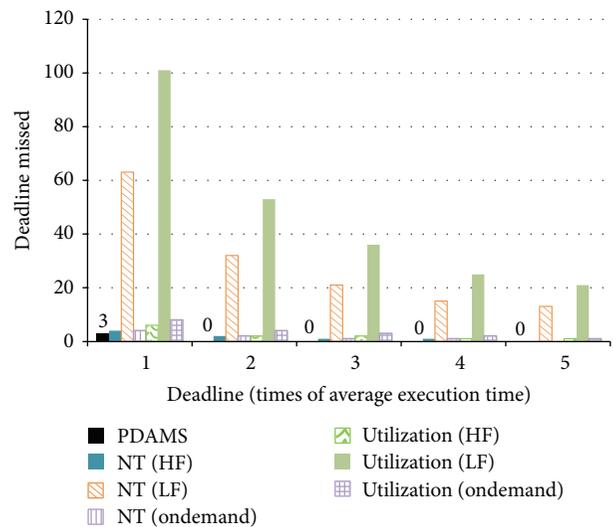


FIGURE 10: Comparison of performance.

one time of their average execution time, missing deadlines cannot be avoided. Worth to note is that even DSPs always run at the fastest frequency, using less appropriate dispatch method and scheduling algorithm may produce more missed deadlines. The proposed algorithm tries to decrease the probability of missing deadline not only in load dispatch, but also in scheduling. That is why the proposed algorithm has the chance to use less energy and get higher performance.

Although always using the lowest operating frequency can reduce energy consumption the most, the performance is not acceptable. The number of deadlines missed is much more than other algorithms, regardless of which load balance algorithm is used. Experimental results show that the proposed algorithms found a good balance point between energy consumption and performance. By considering the deadline

of tasks, the distribution of task deadlines is more uniform on each slave processor core. This can reduce the probability of deadlines being missed. A lower probability of deadline missed not only represents higher performance, but also reduces more energy consumption because there is a longer time that the system will work in lower speed when using ED<sup>3</sup>VFS. Moreover, the concept of load imbalance made the proposed algorithms reduce more energy consumption. Differing from dynamic voltage and frequency scaling technology that only reduces the dynamic power, load imbalance also reduces static power.

Both the number of tasks and processor utilization are the most popular strategies to dispatch tasks in real systems, which is why we chose them as the comparisons. The aim of this paper is to develop not only a novel and effective load balance algorithm, but also an algorithm that can be applied in a real environment. Although the performance of some state-of-the-art algorithms may be better than the proposed algorithm, it is very difficult to satisfy their assumptions. Hence, these kinds of algorithms are hard to make work in a real environment. The other reason is the similarity of assumptions that all the algorithms used in this paper do not need the worst case execution time. Nowadays, portable devices that feature the ability to connect to internet are very popular, such as smart phones and tablet PCs. When a user downloads an application from the internet, there is no way for the system to get the worst case execution time of this application immediately. Although not using the worst case execution time of the tasks made the proposed algorithms unable to guarantee the hard-deadline, the proposed algorithm still tries to avoid missing deadline and become more flexible. We implemented the proposed algorithm in a real platform and the experimental results show that the proposed algorithms work well and are superior to others in general performance.

## 6. Conclusion

This paper applied a solution to the problems of load dispatch and power saving in a real-time system on a multicore platform, called power and deadline-aware multicore scheduling. The proposed algorithm simultaneously considers dynamic power, static power, and load balance. To reduce the dynamic power, we implemented MEDF and fine-tuned the parameters of D<sup>3</sup>VFS to save more power and improve performance. The concept of load imbalance was introduced in saving static power. Instead of dispatching the workload to every processor core equally, the proposed algorithm turns power on only in parts of processor cores and lets other unnecessary cores turn to sleep mode or turn-off. Finally, deadline is used as a novel strategy for load balance between processor cores in active mode. Combining load imbalance and load balance, this paper proposed a two-level task dispatch algorithm called two-level deadline-aware hybrid load balancer.

To verify that the proposed algorithms are useful, we implemented them on a multicore platform, PACDuo. We also implemented some load balance algorithms and frequency scaling algorithms for comparison. Experimental results show that compared to six combinations of load

balance algorithms and frequency scaling algorithms, the proposed algorithms can reduce energy consumption by up to 54.2% and the performance of the proposed algorithms is superior to others. However, much work still needs to be completed in the future. Some areas for future study include (1) adding theoretical analysis to support the proposed algorithm, (2) modelling the energy consumption more detailed, (3) considering the demands of hard-real-time and task migration while keeping the algorithm light-weight, and (4) introducing the concept of heuristic algorithms and improving the proposed algorithms.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

The authors would like to thank the editors and anonymous reviewers for their valuable comments and suggestions on the paper that greatly improve the quality of the paper. This work was supported in part by the National Science Council of Taiwan, under Grants NSC102-2221-E-041-006 and NSC100-2218-E-006-028-MY3 and in part by the Research Center of Energy Technology and Strategy of National Cheng Kung University under Grant D102-23015.

## References

- [1] W. Elmenreich and D. Egarter, "Design guidelines for smart appliances," in *Proceedings of the 10th Workshop on Intelligent Solutions in Embedded Systems (WISES '12)*, pp. 76–82, Klagenfurt, Austria, July 2012.
- [2] Y.-W. Bai, L.-S. Shen, and Z.-H. Li, "Design and implementation of an embedded home surveillance system by use of multiple ultrasonic sensors," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 1, pp. 119–124, 2010.
- [3] H. Kim, Y. Won, and S. Kang, "Embedded NAND flash file system for mobile multimedia devices," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 2, pp. 545–552, 2009.
- [4] I. Kramberger, M. Grasic, and T. Rotovnik, "Door phone embedded system for voice based user identification and verification platform," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 3, pp. 1212–1217, 2011.
- [5] H. Jo, H. Kim, J. Jeong, J. Lee, and S. Maeng, "Optimizing the startup time of embedded systems: a case study of digital TV," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 4, pp. 2242–2247, 2009.
- [6] N. Barsoum, "Speed control of the induction drive by temperature and light sensors via PIC," *Global Journal of Technology and Optimization*, vol. 1, pp. 53–59, 2010.
- [7] U. R. Shikoska, D. Davcev, R. Reckoski, G. Petrovska, C. Andreevski, and J. Sikoski, "Space and time in wireless sensor network," *Global Journal of Technology and Optimization*, vol. 2, pp. 73–83, 2011.
- [8] C.-F. Lai, M. Chen, M. Qiu, A. V. Vasilakos, and J. H. Park, "A RF4CE-based remote controller with interactive graphical user

- interface applied to home automation system,” *ACM Transactions on Embedded Computing Systems*, vol. 12, no. 2, article 30, 2013.
- [9] A. Gotchev, G. B. Akar, T. Capin, D. Strohmeier, and A. Boev, “Three-dimensional media for mobile devices,” *Proceedings of the IEEE*, vol. 99, no. 4, pp. 708–741, 2011.
- [10] G. Correa, P. Assuncao, L. Agostini, and L. A. D. S. Cruz, “Complexity control of high efficiency video encoders for power-constrained devices,” *IEEE Transactions on Consumer Electronics*, vol. 57, no. 4, pp. 1866–1874, 2011.
- [11] C. A. Martinez, J. C. C. San Adrian, and M. V. Cortes, “Dynamic tolerance region computing for multimedia,” *IEEE Transactions on Computers*, vol. 61, no. 5, pp. 650–665, 2012.
- [12] T. Simunic, L. Beniani, and G. de Micheli, “Event-driven power management of portable systems,” in *Proceedings of the 12th International Symposium on System Synthesis*, pp. 18–23, San Jose, Calif, USA, November 1999.
- [13] T. Simunic, L. Benini, A. Acquaviva, P. Glynn, and G. de Micheli, “Dynamic voltage scaling and power management for portable systems,” in *Proceedings of the 38th Design Automation Conference (DAC '01)*, pp. 524–529, Las Vegas, Nev, USA, June 2001.
- [14] Y.-S. Hwang, S.-K. Ku, and K.-S. Chung, “A predictive dynamic power management technique for embedded mobile devices,” *IEEE Transactions on Consumer Electronics*, vol. 56, no. 2, pp. 713–719, 2010.
- [15] M. C. Cera, G. P. Pezzi, E. N. Mathias, N. Maillard, and P. O. A. Navaux, “Improving the dynamic creation of processes in MPI-2,” in *Recent Advances in Parallel Virtual Machine and Message Passing Interface*, vol. 4192 of *Lecture Notes in Computer Science*, pp. 247–255, 2006.
- [16] S. Dandamudi, “Performance implications of task routing and task scheduling strategies for multiprocessor systems,” in *Proceedings of the 1st International Conference on Massively Parallel Computing Systems*, pp. 348–353, Ischia, Italy, May 1994.
- [17] H. Aydin and Q. Yang, “Energy-aware partitioning for multiprocessor real-time systems,” in *Proceedings of the 17th International Parallel and Distributed Processing Symposium (IPDPS '03)*, p. 113, IEEE, April 2003.
- [18] G. Magklis, G. Semeraro, D. H. Albonesi, S. G. Dropsho, S. Dwarkadas, and M. L. Scott, “Dynamic frequency and voltage scaling for a multiple-clock-domain microprocessor,” *IEEE Micro*, vol. 23, no. 6, pp. 62–68, 2003.
- [19] P. Choudhary and D. Marculescu, “Power management of voltage/frequency island-based systems using hardware-based methods,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 17, no. 3, pp. 427–438, 2009.
- [20] W. Jang and D. Z. Pan, “A voltage-frequency island aware energy optimization framework for networks-on-chip,” *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 1, no. 3, pp. 420–432, 2011.
- [21] L.-F. Leung and C.-Y. Tsui, “Energy-aware synthesis of networks-on-chip implemented with voltage islands,” in *Proceedings of the 44th ACM/IEEE Design Automation Conference (DAC '07)*, pp. 128–131, San Diego, Calif, USA, June 2007.
- [22] T. Liu, Y. Sun, Z. Zhang, and L. Guo, “Load balancing for flow-based parallel processing systems in CMP architecture,” in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '09)*, pp. 1–7, Honolulu, Hawaii, USA, December 2009.
- [23] Y. Ahn, W.-J. Kim, K.-S. Chung, S.-H. Kim, H.-S. Kim, and T. H. Han, “A novel load balancing method for multi-core with non-uniform memory architecture,” in *Proceedings of the International SoC Design Conference (ISOCC '10)*, pp. 412–415, Seoul, Korea, November 2010.
- [24] X. Geng, G. Xu, D. Wang, and Y. Shi, “A task scheduling algorithm based on multi-core processors,” in *Proceedings of the International Conference on Mechatronic Science, Electric Engineering and Computer (MEC '11)*, pp. 942–945, Jilin, China, August 2011.
- [25] X. Kavousianos, K. Chakrabarty, A. Jain, and R. Parekhji, “Test schedule optimization for multicore SoCs: handling dynamic voltage scaling and multiple voltage islands,” *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 31, no. 11, pp. 1754–1766, 2012.
- [26] E. Seo, J. Jeong, S. Park, and J. Lee, “Energy efficient scheduling of real-time tasks on multicore processors,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 19, no. 11, pp. 1540–1552, 2008.
- [27] J.-J. Han, X. Wu, D. Zhu, H. Jin, L. Yang, and J. L. Gaudiot, “Synchronization-aware energy management for VFI-based multicore real-time systems,” *IEEE Transactions on Computers*, vol. 61, no. 12, pp. 1682–1696, 2012.
- [28] C. Tianzhou, H. Jiangwei, X. Liangxiang, and Z. Zhenwei, “A practical dynamic frequency scaling scheduling algorithm for general purpose embedded operating system,” in *Proceedings of the 2nd International Conference on Future Generation Communication and Networking (FGCN '08)*, vol. 2, pp. 213–216, Hainan, China, December 2008.
- [29] Y. Shin, K. Choi, and T. Sakurai, “Power optimization of real-time embedded systems on variable speed processors,” in *Proceedings of the IEEE/ACM International Conference on Computer Aided Design (ICCAD '00)*, pp. 365–368, San Jose, Calif, USA, November 2000.
- [30] M. E. Salehi, M. Samadi, M. Najibi, A. Afzali-Kusha, M. Pedram, and S. M. Fakhraie, “Dynamic voltage and frequency scheduling for embedded processors considering power/performance tradeoffs,” *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 19, no. 10, pp. 1931–1935, 2011.
- [31] M. Spiga, A. Alimonda, S. Carta, F. Aymerich, and A. Acquaviva, “Exploiting memory-boundedness in energy-efficient hard real-time scheduling,” in *Proceedings of the International Symposium on Industrial Embedded Systems (IES '06)*, pp. 1–10, Antibes, France, October 2006.
- [32] W.-Y. Liang, S.-C. Chen, Y.-L. Chang, and J.-P. Fang, “Memory-aware dynamic voltage and frequency prediction for portable devices,” in *Proceedings of the 14th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA '08)*, pp. 229–236, Kaohsiung, Taiwan, August 2008.
- [33] W.-Y. Liang, Y.-L. Chen, and M.-F. Chang, “A memory-aware energy saving algorithm with performance consideration for battery-enabled embedded systems,” in *Proceedings of the 15th IEEE International Symposium on Consumer Electronics (ISCE '11)*, pp. 547–551, Singapore, June 2011.
- [34] M. Lombardi, M. Milano, and L. Benini, “Robust scheduling of task graphs under execution time uncertainty,” *IEEE Transactions on Computers*, vol. 62, no. 1, pp. 98–111, 2013.
- [35] J.-Y. Kim, M. Kim, S. Lee, J. Oh, S. Oh, and H.-J. Yoo, “Real-time object recognition with neuro-fuzzy controlled workload-aware task pipelining,” *IEEE Micro*, vol. 29, no. 6, pp. 28–43, 2009.

- [36] Z. J. Jia, T. Bautista, and A. Nunez, "Real-time application to multiprocessor-system-on-chip mapping strategy for system-level design tool," *Electronics Letters*, vol. 45, no. 12, pp. 613–615, 2009.
- [37] P.-C. Hsiu, C.-K. Hsieh, D. N. Lee, and T.-W. Kuo, "Multilayer bus optimization for real-time embedded systems," *IEEE Transactions on Computers*, vol. 61, no. 11, pp. 1638–1650, 2012.
- [38] Z. Shao, M. Wang, Y. Chen et al., "Real-time dynamic voltage loop scheduling for multi-core embedded systems," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 54, no. 5, pp. 445–449, 2007.
- [39] H. Yang, S. Kim, and S. Ha, "An MILP-based performance analysis technique for non-preemptive multitasking MPSoC," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 29, no. 10, pp. 1600–1613, 2010.
- [40] J. Cui and D. L. Maskell, "A fast high-level event-driven thermal estimator for dynamic thermal aware scheduling," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 31, no. 6, pp. 904–917, 2012.
- [41] Y.-S. Chen, H. C. Liao, and T.-H. Tsai, "Online real-time task scheduling in heterogeneous multicore system-on-a-chip," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 1, pp. 118–130, 2013.
- [42] K.-M. Cho, C.-H. Liang, J.-Y. Huang, and C.-S. Yang, "Design and implementation of a general purpose power-saving scheduling algorithm for embedded systems," in *Proceedings of the IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC '11)*, pp. 1–5, Xi'an, China, September 2011.
- [43] S. Liu, J. Lu, Q. Wu, and Q. Qiu, "Harvesting-aware power management for real-time systems with renewable energy," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 20, no. 8, pp. 1473–1486, 2012.
- [44] H. Jeon, W. H. Lee, and S. W. Chung, "Load unbalancing strategy for multicore embedded processors," *IEEE Transactions on Computers*, vol. 59, no. 10, pp. 1434–1440, 2010.
- [45] V. Pallipadi and A. Starikovskiy, "The ondemand governor," in *Proceedings of the Linux Symposium*, pp. 215–230, Ottawa, Canada, 2006.

## Research Article

# A Game-Theoretical Approach to Multimedia Social Networks Security

Enqiang Liu,<sup>1</sup> Zengliang Liu,<sup>1</sup> Fei Shao,<sup>2</sup> and Zhiyong Zhang<sup>3</sup>

<sup>1</sup> University of Science and Technology Beijing, Beijing 100083, China

<sup>2</sup> Xidian University, Xi'an 710126, China

<sup>3</sup> Information Engineering College, Henan University of Science and Technology, Luoyang 471023, China

Correspondence should be addressed to Enqiang Liu; [chinyung@126.com](mailto:chinyung@126.com)

Received 27 February 2014; Accepted 19 March 2014; Published 13 April 2014

Academic Editors: N. Barsoum, V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 Enqiang Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The contents access and sharing in multimedia social networks (MSNs) mainly rely on access control models and mechanisms. Simple adoptions of security policies in the traditional access control model cannot effectively establish a trust relationship among parties. This paper proposed a novel two-party trust architecture (TPTA) to apply in a generic MSN scenario. According to the architecture, security policies are adopted through game-theoretic analyses and decisions. Based on formalized utilities of security policies and security rules, the choice of security policies in content access is described as a game between the content provider and the content requester. By the game method for the combination of security policies utility and its influences on each party's benefits, the Nash equilibrium is achieved, that is, an optimal and stable combination of security policies, to establish and enhance trust among stakeholders.

## 1. Introduction

Multimedia social networks (MSNs) are currently in the wave of popularity. It allows users to share music, pictures, home movies, blogs, and other digital contents with friends, family, colleagues, and students quickly and easily. In the past few years, MSNs, such as MySpace, Facebook, LinkedIn, Flickr, and YouTube, have become the most convenient online sharing method in sharing of images, videos, audios, and other multimedia contents. Although MSNs make communication between people easier and faster and enhance information dissemination among people, there are also security issues, such as privacy disclosure and copyright disputes. This has, undoubtedly, brought serious harm to the dissemination and development of the Internet information. In response to these security issues, the access control mechanism provides a method that allows selective media contents sharing in MSNs. The access control mechanism determines which users can access what resources and how to use these resources and allows users to selectively share their digital contents. Using access control mechanism on digital rights management, content providers can choose to accept or reject access

requests after verifying the access conditions of the digital contents [1, 2].

In the existing MSNs, the access control includes two main types: the relationship-based access control and the trust-based access control.

(1) *Relationship-Based Access Control*. Gates [3] described a new, relationship-based access control security paradigm to meet the needs of Web 2.0. Hart et al. [4] proposed a content- and relationship-based access control system using relationship information in web based social network (WBSN) to represent the authorized agent, which satisfied the key requirements for protecting WBSN resources. However, the system did not achieve the enhanced privacy needs in access control, considered only the direct relationship, and did not take consideration of the node trust in access authorization. In terms of privacy concerns, it focused on privacy protection and data mining techniques and allowed social network analysis for potential sensitive information that had no public disclosure possibility. Park et al. [5] proposed a user-behavior-centric access control framework and identified four core control behaviors: attributes, policies, relationships,

and sessions. The proposed online social network (OSN) had the following characteristics. First, in personalized policies, the OSN users had their own security and privacy policies and attributes. Second, the proposed OSN separated the users from resource policies. Third, the proposed OSN supported access control that was independent of the user relationship and sessions that represented actions. It also took into account the enhanced control, which is not referred to in the existing OSN services. Many of the latest literatures on the OSN access control cannot distinguish between sessions and users.

(2) *Trust-Based Access Control*. Ali et al. [6] applied a multi-level security approach, in which trust was the only parameter that was used to determine the security levels of the users and resources. More precisely, each user was assigned a reputation value. The reputation value was a user's average trust level that was specified by other users. However, Ali and his colleagues only considered direct trust relationship without taking into account the indirect trust relationship. Kruk et al. [7], then, proposed a distributed authentication management system based on the second round "friend" relationship to bring out the management of access rights and trust authorization. Wang and Sun [8] proposed a trust-related management framework that included access control policies and a privacy protection mechanism. This mechanism administers the access policies on the data that contain the provable information, enhances the support to the highly complex privacy related policies, and takes consideration of the purpose and obligations. Under this mechanism, the agent can perform access rights on the objects based on relationships, trusts, purposes, and obligations. This mechanism also introduced strategic operations and the concept of policy conflicts and proposed a purpose related access control policy framework. Sachan et al. [1] pointed out that the traditional access control cannot meet the fine-grained access control requirements and the large number of users. To solve this problem, they proposed an efficient bit-vector transform based access control mechanism suitable for MSNs. They converted the content related certificate into an efficient architecture and, then, verified the security, storage, and execution efficiency of the proposed mechanism rough simulations. Villegas [2] proposed a personal data access control (PDAC) scheme. PDAC computes a "trusted distance" measure between users that is composed of the hop distance on the social network and an affine distance derived from experiential data. Zhang and Wang [9] proposed a trust model for social networks. Based on deep analysis of the characteristics of social networks, they developed a computational model for calculating trust in social networks. Carminati et al. [10] proposed a rule-based access control model and used certificate chain as a parameter for calculating trust, so as to realize effective control of content access in social networks.

These studies focused mainly on relationship and trust-based user access control and realized the controllable, safe transmission of digital content in the MSNs. However, in the relationship-based user access control, only the direct relationship is considered, while the indirect relationship between users, the type of relationship, and closeness of

the relationships are not considered. In the study of trust-based access control, there was no unified understanding of the trust threshold. The setting of the trust threshold can directly affect the security of digital content and controllable dissemination.

In order to solve these issues in social network access control and to prevent excessive denial to normal access or access to much malicious contents, the approach of adopting security policies through game-theoretic analyses is proposed. Tian and Lin [11] proposed a trust prediction-based game control mechanism for trustworthy networks. This mechanism could not only predict behavior trust level with single trust attribute but also could predict trust level with the multiple trust attributes, so as to help participants to achieve the maximization of utility. Wen et al. [12] proposed game-theoretic model for information dissemination in social networks. This model reflected the influence of human behavior on information dissemination and conceptualized participants' utility function based on different parties' interests. An empirical study indicated that information dissemination can be divided into several stages, and the dissemination speed is limited by the characteristics of each person in the network. Zhang et al. [13] proposed game-based social network access control. For the "nonfriend" type of access users, on the basis of defining user trust and its calculation method, this study conducted game-theoretic analyses by integrating the payoff matrix of both the content provider and the content requester, calculated the hybrid Nash equilibrium, provided decision-making criteria for access control, and finally analyzed the utility of the access control method with examples. None of these three models consider the personalization problem of the content providers' security policy and only conceptualize it as accepting or denying access. In addition, in the participants' utility function, none of the models consider the inherent cost, such as the cost of implementing security policies for the content providers and the cost of malicious access for the access requesters. Zhang et al. [14, 15], for a general digital rights management (DRM) value chain system, proposed a layered analysis of multiparty trust architecture by using game-theoretic analyses of adoption of security policies. Based on formalized utilities of security policies and services, the adoption of security policies with external relativity is described as a game between the content provider, the digital services/providers, and the content requester. Based on the utility of the security policies and their influence on each party's benefits, the Nash equilibrium value was achieved, which is an optimal and stable combination of security policies, thus establishing and strengthening multiparty trust. In order to effectively select and deploy security policies in content sharing scenarios, Zhang et al. [16] introduced the game theory to analyze the influence of security policies that use trusted-computing-enhanced security policy stakeholders. At last, Zhang et al. conducted game-theoretic analyses and swarm simulation. The results indicated that the obtained digital content and security cost had direct impact on the content provider's choice of security policies. In addition, different basic-sharing models, including local, intermediate, and extensive sharing models, will further affect the choice of the content providers.

The mixed-sharing model was much more similar to the real content sharing situations. Due to limited power, sharing, and higher security cost, the dynamic security policy is better than the fully enhanced security policies; but with the reduction of more power and enhanced security cost, the latter strategy would be the best and the most stable Nash equilibrium [17, 18].

As noted above, there are a lot of studies about access control issues in MSNs; however, a successful access of multimedia digital content (MMDC) should have the following three factors: security, trust, and benefits. So far, because of the lack of access control in MSNs, the MMDC access is only based on security policies and the related mechanisms. Therefore, how to make a rational use of security policies to maximize the benefits of the participants is worth considering. This paper proposes a game-based security policies adoption approach for MSNs. This system is benefits-centric that enables the participants to find an optimal and stable security policy in MSNs.

## 2. Formalized Game of Security Policies

*2.1. Two-Party Trust Architecture.* Recently, game theory is widely applied in economics, biology evolution, and information technology, especially for the decision-making on information security policies, when multiple stakeholders have their own benefits and strategies moving. The MSN scenario has such characteristics as needed by game theory.

A general MSN is composed of different stakeholders, such as  $P$  (content providers) and  $R$  (content requester). Based on the basic analysis of the trust relationship, two-party trust architecture (TPTA) is the trust architecture between  $P$  and  $R$ . This system includes a set of security rules, namely, the basic security rules and the optional security rules. The specific security policies can be achieved by using these security rules. As shown in Figure 1, participants are rational agent (RA), who can, rationally, select and deploy a security policy based on the game theory.

### 2.2. Basic Components

*Definition 1* (party). A symbol  $\wp$  denotes personal player participating in content provider and content access; these two roles are interchangeable in MSNs. The  $\wp$  can take different roles in MMDC sharing; the content providers can assume the role of the content requester, while the content requester can also take the content provider role. The formalized  $\wp$  participant is as follows:

$$\wp = \{\alpha \mid \text{stakeholder accessing to contents}\} \quad (1)$$

$$\text{MMSN\_VauleChain}_{\text{MPTA}} = \{P, R, \text{MMDC}\}.$$

*Definition 2* (security rules). In response to the participants' security requirements, one security rule corresponds to one user attribute constraint, which ensures the security of MMDC. A symbol  $\text{SR}^*$  denotes basic security rules; the other SR denotes optional security rules. Notation of  $f$ ,  $w$ ,  $u$  denotes an effective factor from factor set  $F$  influencing

benefit of  $\wp$ , the weight value of factor, and a positive/negative utility, respectively. Here, the normalized weight is based on the weight of all of the factors of SR:

$$\text{security rule} = \{\text{SR}_1^*, \text{SR}_2^*, \dots, \text{SR}_i^*, \text{SR}_1, \text{SR}_2, \dots, \text{SR}_j\}$$

$$F(\text{sr}_s) = \{f_{\text{sr}_1}, f_{\text{sr}_2}, \dots, f_{\text{sr}_l}\} \quad (1 \leq s \leq l), \quad (2)$$

$$\mu(\text{sr}_s) = \sum_{i=1}^l u_i \left( \frac{w_i}{\sum_{k=1}^l w_k} \right).$$

*Property 1* (external relativity of optional security rules). If two or multiple optional security rules are from different parties, choose to adopt simultaneously or adopt only one of them according to the needs of participants. The external relativity of these rules is described as follows, in which  $\mathbb{C}(\wp)$  denotes the base set of  $\wp$ .

- (1) If  $P$  has some strict requirements for MMDC access ( $R$  must meet all the security rules before accessing the MMDC),

$$\text{Relative\_Components} = \{\text{sr}_1, \text{sr}_2, \dots, \text{sr}_p\}$$

$$\forall i, j (1 \leq i, j \leq p, 2 \leq p \in \mathbb{C}(\wp)) \exists s, \quad (3)$$

$$t(s, t \in \{P, R\}) (\text{sr}_i \in \text{SR}_s, \text{sr}_j \in \text{SR}_t, i \neq j \rightarrow s \neq t).$$

- (2) If  $P$  has relaxed requirements for MMDC access ( $R$  only needs to meet any one of the security rules to access MMDC),

$$\text{Relative\_Components} = \{\text{sr}_1, \text{sr}_2, \dots, \text{sr}_p\}$$

$$\forall i, j (1 \leq i, j \leq p, 2 \leq p \in \mathbb{C}(\wp)) \exists s, \quad (4)$$

$$t(s, t \in \{P, R\}) (\text{sr}_i \in \text{SR}_s \vee \text{sr}_j \in \text{SR}_t).$$

*Definition 3* (security rules).  $\text{sp}$  includes the  $P$  and  $R$ 's  $\text{sp}$ , denoted, respectively, as  $\text{sp}_P$  and  $\text{sp}_R$ .  $\text{sp}_P$  is considered as a set of security rules and services;  $\text{sp}_R$  includes normal access and malicious access:

$$\text{sp}_P = \{\text{sr}_1^* \dots, \text{sr}_i^*, \text{sr}_1, \text{sr}_2, \dots, \text{sr}_s\} \quad (0 \leq s \leq j),$$

$$\text{SP}_{P_i} = \{\text{sp}_i^1, \text{sp}_i^2, \dots, \text{sp}_i^{\mathbb{C}(\text{SP}_i)}\} \quad (\mathbb{C}(\text{SP}_i) = 2^j, i \in \{P, R\}),$$

$$\text{sp}_R = \{\text{normal, malicious}\}. \quad (5)$$

*Definition 4* (utility of  $\text{sp}$ ). Utility  $U_P$  of  $\text{sp}_P$  is a sum of utilities of all rules or services involved in  $\text{sp}_P$ ; utility  $U_R$  of  $\text{sp}_R$  is a sum of utilities of all rules or services involved in  $\text{sp}_R$ :

$$U(\text{sp}_P) = \sum_{p=0}^i \mu(\text{sr}_p^*) + \sum_{p=0}^j \mu(\text{sr}_p) + \sum_{p=0}^k \mu(\text{MMDC}) \quad (6)$$

$$U(\text{sp}_R) = \sum_{p=0}^i \mu(\text{sr}_R) + \sum_{p=0}^j \mu(\text{MMDC}).$$

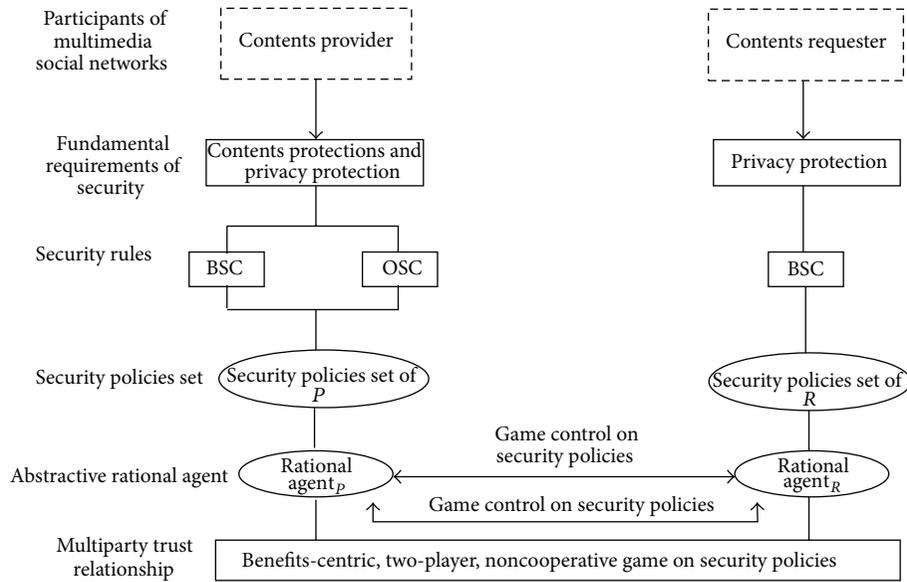


FIGURE 1: Two-party trust architecture in multimedia social networks.

### 2.3. Formalized Game of Security Policies

**Definition 5** (rational agent). A symbol RA denotes a rational actor aiming at a maximization of benefit and makes a decision on adopting a certain security policy. In TPTA, there are two RAs with respect to two parties, namely,  $RA_P$  and  $RA_R$ .

**Definition 6** (payoff of RA). In TPTA, a payoff RA denotes the acquired benefits from security policies set. It is the carry for RA adoption of security policies. Benefits include two aspects: RA or changes of RA.

**Definition 7** (two-party game). Two-party game  $G$  of security polices denotes a process of making decision on effective and rational adoption of security policies that have effect on benefit of the opposing parties. To achieve utility maximization and balance, the game is depicted by a set of three tuples as  $\langle \rho, sp, payoff \rangle$ . SP represents the security policies set:

$$G = \{ \langle RA_i, SP_i, Payoff(RA_i, RA_{-i}) \rangle \mid i = \{P, R\} \}. \quad (7)$$

**Definition 8** (Nash equilibrium under policies combination). For any RA, when adopting a security policy,  $sp^*$  acquires greater benefit than the benefit acquired by choosing any other sp; the combination of each RA's  $sp^*$  is considered as a balance of payoffs by adopting relatively dominant security policies:

$$\begin{aligned} Payoff(RA_i^{sp^*}, RA_{-i}^{sp^*}) &\geq Payoff(RA_i^{sp^j}, RA_{-i}^{sp^*}) \\ j \in SP_i, j \neq sp^*, \quad i \in \{P, R\} \quad (-i \in \{P, R\}, -i \neq i), \end{aligned} \quad (8)$$

where  $(sp_P^*, sp_R^*)$  is a relatively dominant pure policies combination.

### 2.4. Game of Security Policies in Two Scenarios

**Theorem 9** (two parties both change game in content access). Content access is a general scenario in MSNs. In this scenario, the adoption of security policies is considered to be a particular game process in which both  $P$  and  $R$  change simultaneously.

*Proof.* In TPTA, according to  $RA_P$  and  $RA_R$  in Definition 5, denote their security policies combinations as  $SP_P$  and  $SP_R$ , respectively. Game was further formalized as  $G_{\text{acquisition}} = \{ \langle RA_i, SP_i, Payoff(RA_i, RA_{-i}) \rangle \}$ , in which  $i = \{P, R\}$ . For MMDC access,  $P$  needs to set up security rules for  $R$ 's MMDC access, that is, choosing a particular sp from SP. Under normal circumstances, the process of content access has timing characteristics; after  $RA_R$  requests MMDC access to  $RA_P$ ,  $RA_R$  should meet the access control policies. However, when each RA adopts and initializes SP, they do not know other RA's changes of sps. In addition, during the content transaction, the setting of the MMDC security polices in MSNs cannot be changed. Therefore, the change process of RA in security policies is a simultaneous change of the game, rather than a continuous change of the game.  $\square$

**Theorem 10.** The trust values of the content providers' benefits and those of the content requesters' benefits are proportional.

*Proof.* Based on the utilities of the content provider and the content requester in Definition 4, the trust values of content requesters  $R_i$  and  $R_j$  are  $i, j$  assuming that  $i < j$ . Because the larger the trust value, the larger the  $i, j$  values; therefore, the trust value of  $i$  is larger than that of  $j$ . The larger the user trust value, the larger the  $\mu(\text{MMDC})$  value, the greater the utility and therefore the larger the  $P$  and  $R$  benefits.  $\square$

**Deduction 1** (repeated game in content access scenario). When several content access sessions are carried out, the participants in MSNs will choose to reactivate a game in order

to select a security policy. The new game can be seen as a repetitive game, which is based on the process and results of the previous game, and get a new equilibrium.

*Proof.* In a given scenario, as the access to content increases, the adoption of security policies will change accordingly. When  $RA_P$  and  $RA_R$  select security policies again, a repeated game will happen, combined with sessions of the previous game and transaction to obtain a new security policies combination, which is called a new Nash equilibrium.  $\square$

### 3. Game-Theoretic Analysis of Typical Security Strategy

In an access control model of multimedia social network which has universal significance, each party has a security strategy set and practical choice set representing moving in content access. Some typical  $P$  and  $R$  security strategies are listed in Section 2.1. The following two sections cite the security benefits of all security rules, effective strategy combinations, and participant benefits, respectively. A strategy selection example is finally analyzed.

**3.1. Typical Security Strategy.** In this study, some typical security strategies are presented. A real access control for multimedia social network may include but is not limited to these strategies. In Definition 2 in Section 2.1, some security rules that can meet the security demand of any party are mentioned first, and then security strategy set can be easily derived.

The security rules of two participants include relationship type (RelT), depth (Dep), compactness (C), and trust (T).

Similarly, since the SR set of  $P$  can be denoted as  $\{G^*, Dep^*, C^*, T^*\}$ , the security strategy set is {general security strategy, enhanced security strategy}. The enhanced security strategy is  $k_1G^* + k_2Dep^* + k_3C^* + k_4T^*$ , where  $k_i$  ( $i = 1, 2, 3, 4$ )  $\in \{0, 1\}$ ,  $\sum_{i=1}^4 k_i \neq 0$ , denoted by  $sp_p$ .

For the access into MMDC, there are two types of  $R$ , normal access (NA) and malicious access (MA). Therefore, the security strategy set is  $\{MA^*, NA^*\}$ .

By the typical security strategies and related SR analysis above, the utility impact factor, weight, and the utility of SR will be introduced in this section. As  $SR^*$  cannot change the utility of  $sp$ , only the utility of SR is considered here.

**3.2. Effective Strategy Combination and Its Utility.** Since there are two security strategies for each party, there are 4 possible strategy combinations in the game. Figure 2 describes the security strategy combinations of participants, where  $sp_i$  ( $i \in SP, RP$ ) denotes the strategy mentioned in the subgraph.

The benefits of content provider and content requester are defined as follows.

The symbol  $U_P^{Benefit.NA}$  denotes the fact that when the content provider implements the general safety strategy and the content requester adopts the normal access, the content provider may obtain normal average benefit, such as the rise of the number of friends, increase of attention degree, and acquisition of the information of content requester.

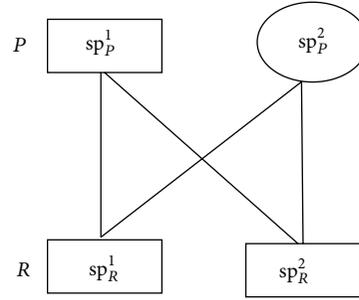


FIGURE 2: Security strategy combinations.

The symbol  $U_P^{Damage.MASuccess} > 0$  denotes the possible average amount of loss after the content provider implements general safety strategy and content requester adopts the malicious access, such as the multimedia digital content of the content provider being forwarded casually. Another maliciousness includes impersonating user identity using the content provider's information.

The symbol  $U_P^{Damage.NA} > 0$  denotes the possible average amount of loss of content provider when the content provider implements the enforced security strategy and the content requester accesses normally, such as declining normal user accessing so that the social network resource is not fully used and the loss of no cooperation caused by distrust between the two parties.

The symbol  $U_P^{Cost} > 0$  denotes the cost of deploying security strategy for content provider, such as the increase of time expenditure. Consider  $U_P^{Cost} = C_0(k_1G^* + k_2Dep^* + k_3C^* + k_4T^*)$ .

The symbol  $U_R^{Benefit.NA}$  denotes the average benefit obtained by the content requester when the content requester accesses normally and content provider implements the enforced security strategy, such as rise of the number of friends, increase of attention degree, and promotion of digital content.

The symbol  $U_R^{Benefit.MASuccess}$  denotes the excess benefit obtained by content requester when the content requester adopts malicious behaviors and content provider implements general security strategy, such as distributing the multimedia digital content casually and disclosing private information without permission.

The symbol  $U_R^{Cost} > 0$  denotes the cost of attacking the security strategy platform by the content requester.

The symbol  $U_R^{Punish} > 0$  denotes the punishment that may be given to the content requester adopting malicious behaviors, such as decreasing the trustworthiness of content requester, suspending the access right to social network for the content requester, or suing the content requester, where  $U_R^{Benefit.MASuccess} > U_P^{Cost}$ .

First, we analyze the gain and loss of the benefits of both content provider and requester. If the content requester accesses normally and the content provider implements the enforced security strategy, then the content requester and provider will both benefit and their benefits are denoted as  $U_P^{Benefit.NA}$  and  $U_R^{Benefit.NA}$ , respectively. If the content requester

accesses maliciously and the content provider implements the general security strategy, then the loss of content provider is  $U_R^{\text{Benefit.NA}}$ , while the benefit of content requester contains an excess benefit  $U_P^{\text{Cost}}$  obtained by malicious access, in addition to the normal average benefit  $U_R^{\text{Benefit.NA}}$ . However,

the content requester may be subjected to a punishment  $U_P^{\text{Cost}}$  if adopting malicious access. If the content provider implements the enforced security strategy, then there will be neither benefit nor loss but only the cost of implementing the enforced security strategy  $U_P^{\text{Cost}}$ . Based on Definition 4 and Figure 2, the payoff matrix of participants under multiple combinations is as follows:

$$A = \begin{bmatrix} U_P^{\text{Benefit.NA}} \cdot \alpha_1^{i-1} & -U_P^{\text{Damage.MASuccess}} \cdot \alpha_2^{i-1} \\ -U_P^{\text{Damage.NA}} \cdot \alpha_3^{i-1} - U_P^{\text{Cost}} \alpha_4^{i-1} & -U_P^{\text{Cost}} \alpha_4^{i-1} \end{bmatrix}$$

$$B = \begin{bmatrix} U_R^{\text{Benefit.NA}} \cdot \alpha_5^{i-1} + U_R^{\text{Benefit.MASuccess}} \cdot \alpha_6^{i-1} - U_R^{\text{Punish}} \cdot \alpha_7^{i-1} - U_R^{\text{Punish}} \cdot \alpha_8^{i-1} & -U_R^{\text{Punish}} \cdot \alpha_7^{i-1} - U_R^{\text{Cost}} \alpha_8^{i-1} \\ U_R^{\text{Benefit.NA}} \cdot \alpha_5^{i-1} & 0 \end{bmatrix}, \tag{9}$$

where  $\alpha_i$  ( $i = 1, 2, \dots, 8$ )  $> 1$  is the parameter factor, mainly used to adjust the ratio of user benefit to punishment. The setting of this value is based on requirements of the decision makers. The benefit matrices  $A$  and  $B$  denote that the user's benefit or loss is closely related to his/her attributes and is proportional to the trustworthiness. The reason why the content requester chooses malicious access is that it is believed that the benefit obtained by malicious access is larger than that by normal access; that is, the user is rational. However, the content provider in social networks increases his or her attention degree and maximizes the

benefit by making more friends, which means that the content provider is also rational. Supposing that the probability of content provider implementing the general security strategy is  $x$ , then the probability of implementing enforced security strategy is  $1 - x$  and the mixed strategy for the content provider is  $P = (x, 1 - x)$ . Similarly, supposing that the malicious access probability of content requester is  $y$ , then the probability of normal access is  $1 - y$  and the mixed strategy of content requester is  $R = (y, 1 - y)$ . Based on Definition 4 and Figure 1, the benefit obtained by the participants under multiple combinations is as follows:

$$E_{SR} = \text{Payoff}_R \cdot B \cdot \text{Payoff}_P^T = (y, 1 - y) \cdot \begin{bmatrix} U_R^{\text{Benefit.NA}} \cdot \alpha_5^{i-1} + U_R^{\text{Benefit.MASuccess}} \cdot \alpha_6^{i-1} - U_R^{\text{Punish}} \cdot \alpha_7^{i-1} - U_R^{\text{Punish}} \cdot \alpha_8^{i-1} & -U_R^{\text{Punish}} \cdot \alpha_7^{i-1} - U_R^{\text{Cost}} \alpha_8^{i-1} \\ U_R^{\text{Benefit.NA}} \cdot \alpha_5^{i-1} & 0 \end{bmatrix} \cdot \begin{pmatrix} x \\ 1 - x \end{pmatrix} \tag{10}$$

$$= x \cdot y \cdot U_R^{\text{Benefit.MASuccess}} \cdot \alpha_6^{i-1} + x \cdot U_R^{\text{Benefit.NA}} \cdot \alpha_5^{i-1} - y \cdot (U_R^{\text{Punish}} \cdot \alpha_7^{i-1} + U_R^{\text{Cost}} \alpha_8^{i-1}).$$

By taking the partial derivative of the above equation with respect to  $y$ , the condition for the content provider getting the optimal strategy is

$$\frac{\partial E_R}{\partial y} = x \cdot U_R^{\text{Benefit.MASuccess}} \cdot \alpha_6^{i-1} - (U_R^{\text{Punish}} \cdot \alpha_7^{i-1} + U_R^{\text{Cost}} \alpha_8^{i-1}) = 0. \tag{11}$$

Hence, there is

$$x^* = \frac{U_R^{\text{Punish}} \cdot \alpha_7^{i-1} + U_R^{\text{Cost}} \alpha_8^{i-1}}{U_R^{\text{Benefit.MASuccess}} \cdot \alpha_6^{i-1}}; \tag{12}$$

that is to say,  $P^* = (x^*, 1 - x^*)$  is the optimal strategy for the content provider.

It can be seen from the result of observation and analysis that the accepting probability of content provider is only

related to the benefit and payment of the user. By increasing the punishment for malicious access from the content requester, increasing the cost of attacking security strategy by requester, and decreasing the benefit obtained by successful malicious accesses from the content requester, the probability of content provider adopting the general security strategy can be improved and the normal operation of the social network can be promoted. When the content provider adopts general security strategy in the probability of  $x > x^*$ , the content requester can obtain benefit by normal access; otherwise, the optimal strategy for the content requester is to adopt the normal access strategy. The strategy can only be used to determine at what probability the content provider should accept the access and to select the parameters for decision makers in a macroscopic way. It still depends on the attributes of the content requester and relevant history when it comes to a specific access. A rational content requester

seeks a method to maximize his/her own payment to play the game. Therefore, the one that can meet the demand and enable both parties to keep a stable state is the mixed strategy Nash equilibrium, which is the lowest condition acceptable for the content provider. The benefit function of the content provider is expressed as

$$\begin{aligned}
 E_P &= \text{Payoff}_P \cdot A \cdot \text{Payoff}_R^T = (x, 1 - x) \\
 &\cdot \begin{bmatrix} U_P^{\text{Benefit.NA}} \cdot \alpha_1^{i-1} & -U_P^{\text{Damage.MASuccess}} \cdot \alpha_2^{i-1} \\ -U_P^{\text{Damage.NA}} \cdot \alpha_3^{i-1} - U_P^{\text{Cost}} \alpha_4^{i-1} & -U_P^{\text{Cost}} \alpha_4^{i-1} \end{bmatrix} \\
 &\cdot \begin{pmatrix} y \\ 1 - y \end{pmatrix} \\
 &= x \cdot y \left( U_P^{\text{Benefit.NA}} \cdot \alpha_1^{i-1} + U_P^{\text{Damage.MASuccess}} \cdot \alpha_2^{i-1} \right. \\
 &\quad \left. + U_P^{\text{Damage.NA}} \cdot \alpha_3^{i-1} \right) \\
 &\quad - x \left( U_P^{\text{Damage.MASuccess}} \cdot \alpha_2^{i-1} - U_P^{\text{Cost}} \alpha_4^{i-1} \right) \\
 &\quad - y U_P^{\text{Damage.NA}} \cdot \alpha_3^{i-1} - U_P^{\text{Cost}} \alpha_4^{i-1}. \tag{13}
 \end{aligned}$$

By taking the partial derivative of the above equation with respect to  $x$ , the condition for the content requester getting the optimal strategy is

$$\begin{aligned}
 \frac{\partial E_P}{\partial x} &= y \left( U_P^{\text{Benefit.NA}} \cdot \alpha_1^{i-1} + U_P^{\text{Damage.MASuccess}} \cdot \alpha_2^{i-1} \right. \\
 &\quad \left. + U_P^{\text{Damage.NA}} \cdot \alpha_3^{i-1} \right) \\
 &\quad - \left( U_P^{\text{Damage.MASuccess}} \cdot \alpha_2^{i-1} - U_P^{\text{Cost}} \alpha_4^{i-1} \right) = 0. \tag{14}
 \end{aligned}$$

Hence

$$\begin{aligned}
 y^* &= \left( U_P^{\text{Damage.MASuccess}} \cdot \alpha_2^{i-1} - U_P^{\text{Cost}} \alpha_4^{i-1} \right) \\
 &\times \left( U_P^{\text{Benefit.NA}} \cdot \alpha_1^{i-1} + U_P^{\text{Damage.MASuccess}} \cdot \alpha_2^{i-1} \right. \\
 &\quad \left. + U_P^{\text{Damage.NA}} \cdot \alpha_3^{i-1} \right)^{-1}, \tag{15}
 \end{aligned}$$

where  $R^* = (y^*, 1 - y^*)$  is the optimal strategy for the content requester.

It can be seen from the observation result that the mixed strategy Nash equilibrium for content requester gives an uncertain game-theoretic result to the user. Illegal user is not able to get the payoff matrix and decision probability and therefore is unable to judge how the content provider will process the request. These users can obtain the payoff matrix and decision probability by illegal means, but how the content provider will make decision is not certain.

**3.3. Dynamic Strategy Control Based on Mixed Strategy Nash Equilibrium.** In the above section, the mixed strategy Nash equilibrium for the content provider and requester is calculated, and the issue of user controlling strategy probability is presented. However, it is not certain what the decision will be

TABLE 1: Parameters settings of the example on Scene 1.

	Parameters			
	$U_P^{\text{Benefit.NA}}$	$U_P^{\text{Damage.NA}}$	$U_P^{\text{Damage.MASuccess}}$	$U_P^{\text{Cost}}$
1	100	100	600	70
2	150	150	800	100

each time. Besides, it is necessary to decide by combining with the strategy selected by the content requester. This is due to the fact that the attributes and decision probabilities of different content requesters are different and the game controlling strategy depends on the game-theoretic analysis of the two parties, instead of the strategy inference of one party. Hence, the content provider needs to adjust the strategies according to the decision probability of himself/herself and that of the content requester and the requirement of his/her decision probability.

$P^* = (x^*, 1 - x^*)$ ,  $R^* = (y^*, 1 - y^*)$ , while the requirement by the content provider on the strategy probability of the content requester is  $R_0 = (y_0, 1 - y_0)$ .

- (1) The strategy requirement by the content provider is strict; that is,  $y_0 \geq y^*$ :  $P$  adopts enforced security strategy to increase  $x^*$ .
- (2) The strategy requirement by the content provider is strict; that is,  $y_0 < y^*$ :  $P$  does not need to increase  $x^*$  and the general security strategy can be adopted.

## 4. Use Cases Analyses

**4.1. Background.** In multimedia social network, the content provider distributes the multimedia digital content and content requester can ask to access the multimedia digital content. When all attributes of the content requester satisfy the requirement of the access control model of multimedia social network platform, the requester can access the digital content. However, after some content providers access the digital content, they casually distribute the multimedia digital content and disclose the private information without permission to seek illegal benefits. In order to prevent such malicious access behaviour, the multimedia social network platform will adopt certain punishment methods, such as declining users to access digital contents. However, mistakenly refusing normal users to access multimedia digital content is not beneficial to the promotion of digital content and drawing attention, while no access control will not achieve the purpose of preventing malicious access, which will damage the interests of the content provider. By using the proposed mixed strategy, the content provider can avoid malicious access from the users and accept normal access.

**4.2. Use Cases Game Decision on Security Policies.** The parameter factors of game-theoretic analysis  $\alpha_i$  ( $i = 1, 2, \dots, 8$ ) are 1, 1.1, 1, 1.1, 1.2, 1.1, 1.2, and 1.1, respectively. The assumed values of other parameters are shown as the second to eighth columns in Tables 1 and 2. By substituting the above parameters into (6) and (3), the probability of content requester adopting malicious access,  $y^*$ , and that of content

TABLE 2: Parameters settings of the example on Scene 2.

	Parameters			
	$U_R^{\text{Benefit\_NA}}$	$U_R^{\text{Benefit\_MASuccess}}$	$U_R^{\text{Cost}}$	$U_R^{\text{Punish}}$
1	100	500	50	300
2	180	700	50	350

provider adopting the general security strategy,  $x^*$ , can be calculated. The  $y^*$  and  $x^*$  can be calculated according to the data in the example. The content provider can make decision based on the dynamic strategy control rule of mixed strategy Nash equilibrium. For multimedia social network application platform, two typical scenes are set up.

*Scene 1.* The level of the relationship between content requester and content provider is 1, set to be 1.  $C_0$  is set to be 1,  $k_i$  as 1, depth as 1, closeness as 68, and trustworthiness as 1.

*Scene 2.* The level of the relationship between content requester and content provider is 1.  $C_0$  is set to be 1,  $k_i$  as 1, depth as 1, closeness as 98, and trustworthiness as 1.

Based on the above mentioned scenes and use cases, the value requirement of  $R$  from  $P$  is given as (0.66, 0.34). According to the value of each parameter in Table 1, it can be calculated that  $P$  (general security strategy, enforced security strategy) of Scene 1 = (0.75, 0.25),  $R$  (malicious access, normal access) = (0.68, 0.32). At this time,  $P$  needs to adopt the enforced security strategy. In Scene 2,  $P$  (general security strategy, enforced security strategy) = (0.68, 0.32),  $R$  (malicious access, normal access) = (0.65, 0.32). At this time,  $P$  only needs to adopt the general security strategy.

The benefit and punishment obtained by the content requester increase with the increase of trustworthiness and closeness and decrease with the deepening of the relationship. With the increase of trustworthiness and closeness of content provider and the decrease of relationship depth, the probability of content requester adopting malicious access is decreasing, while the probability of content provider adopting the general security strategy is increasing. This is in accordance with the actual practice on the social network. The content provider can implement the access control based on the mixed strategy Nash equilibrium between the two parties to further adjust the probability of adopting the corresponding strategy. However, the content requester does not know which strategy the content provider will adopt, and the cost of adopting malicious access and the received punishment is far greater than the benefit obtained from successful malicious access. Hence, the provider requester will not adopt the malicious access strategy easily.

*4.3. Discussions.* The decision-making model and method for adopting of security policies are firmly based on the game theory and its applications on information security, so it is complete and robust. Besides, it has also flexibility due to an ability to represent the game on multiparticipant and multisecurity policies, not only two parties and two strategies. The proposed approach to decision has significant

advantages, including effectiveness on MSNs security policies combination realization and deployment, convenience on the least overhead of security management, and benefits and productivity for contents providers owing to wider contents access and sharing in MSNs.

## 5. Conclusions

This paper proposed a game-based analysis on security policies to obtain an optimal combination of security policies for content access in MSNs, thus achieving utility maximization between users. For this reason, this study created the TPTA between the content provider and the content requester. And then, we proposed a typical game-theoretic control of security strategy, obtained the mixed strategy Nash equilibrium based on security attribute of the user, and analyzed a practical example. In this study, the strategy selection under the existing access control mechanism of social network is addressed. A game-theoretic analysis method is provided for the selection of security strategy by the content provider and for the protection of multimedia digital content. In the future, the research challenge will focus on an in-depth consideration to effectively and rationally deploy security policies by the MSNs game-theoretic analysis of security strategy under content sharing conditions, so as to improve the security, credibility, and flexibility of the real MSNs applications and services. In general, the novel game-theoretical model for MSNs is also suitable for the same scenarios and services where multiple stakeholders have their own benefits and strategies choices, including general social media network and applications.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work was sponsored by the National Natural Science Foundation of China Grant no. 61370220, Plan For Scientific Innovation Talent of Henan Province Grant no. 134100510006, Key Program for Basic Research of the Education Department of Henan Province Grant nos. 13A520240 and 14A520048, and the Joint-Research Project titled "Key technologies research on a novel network security defense system" Grant no. 61440144. The authors also thank Qingli Chen for her works on related experiments and analyses.

## References

- [1] A. Sachan, S. Emmanuel, and M. Kankanhalli, "An efficient access control method for multimedia social networks," in *Proceedings of the 2nd ACM SIGMM Workshop on Social Media (WSM '10)*, pp. 33–38, Firenze, Italy, October 2010.
- [2] W. Villegas, *A trust-based access control scheme for social networks [M.S. thesis]*, School of Computer Science, McGill University, Montreal, Canada, 2008.
- [3] B. Carminati, "Access control and privacy in web-based social networks," *International Journal of Web Information Systems*, vol. 4, no. 4, pp. 395–415, 2008.

- [4] M. Hart, R. Johnson, and A. Stent, "More content-less control: access control in the web 2.0," in *Proceedings of the Workshop on Web 2.0 Security and Privacy at the IEEE Symposium on Security and Privacy*, pp. 1-3, Oakland, Calif, USA, May 2007.
- [5] J. Park, R. Sandhu, and Y. Cheng, "A user-activity-centric framework for access control in online social networks," *IEEE Internet Computing*, vol. 15, no. 5, pp. 62-65, 2011.
- [6] B. Ali, W. Villegas, and M. Maheswaran, "A trust based approach for protecting user data in social networks," in *Proceedings of the Conference of the Center for Advanced Studies on Collaborative Research (CASCON '07)*, pp. 288-293, Richmond Hill, Montreal, Canada, October 2007.
- [7] S. R. Kruk, S. Grzonkowski, A. Gzella et al., "D-FOAF: distributed identity management with access rights delegation," in *The Semantic Web*, vol. 4185 of *Lecture Notes in Computer Science*, pp. 140-154, 2006.
- [8] H. Wang and L. Sun, "Trust-involved access control in collaborative open social networks," in *Proceedings of the 4th International Conference on Network and System Security (NSS '10)*, pp. 239-246, Melbourne, Australia, September 2010.
- [9] Z. Y. Zhang and K. L. Wang, "A trust model for multimedia social networks," *Social Networks Analysis and Mining*, vol. 3, no. 4, pp. 969-979, 2012.
- [10] B. Carminati, E. Ferrari, and A. Perego, "Rule-based access control for social networks," in *Proceedings of the Move to Meaningful Internet Systems 2006: OTM, 2006 Workshops*, pp. 1734-1744, Montpellier, France, 2006.
- [11] L.-Q. Tian and C. Lin, "A kind of game-theoretic control mechanism of user behavior trust based on prediction in trustworthy network," *Chinese Journal of Computers*, vol. 30, no. 11, pp. 1930-1938, 2007.
- [12] Q. Wen, Y. Z. Wang, and J. Y. Yu, "A game theoretical model of information dissemination in social network," in *Proceedings of International Conference on Complex Systems*, Agadir, Morocco, November 2012.
- [13] S. B. Zhang, W. D. Cai, and Y. J. Li, "A game-theory based access control method suitable for social network," *Journal of Northwestern Polytechnical University*, vol. 29, no. 4, pp. 652-657, 2011.
- [14] Z. Zhang, Q. Pei, J. Ma, and L. Yang, "Establishing multi-party trust architecture for drm by using game-theoretic analysis of security policies," *Chinese Journal of Electronics*, vol. 18, no. 3, pp. 519-524, 2009.
- [15] Z. Zhang, Q. Pei, J. Ma, L. Yang, and K. Fan, "Cooperative and non-cooperative game-theoretic analyses of adoptions of security policies for DRM," in *Proceedings of the 6th IEEE Consumer Communications and Networking Conference (CCNC '09)*, pp. 1-5, Las Vegas, Nev, USA, January 2009.
- [16] Z. Zhang, Q. Pei, J. Ma, and L. Yang, "Game-theoretic analyses and simulations of adoptions of security policies for DRM in contents sharing scenario," *Intelligent Automation & Soft Computing*, vol. 17, no. 2, pp. 191-203, 2011.
- [17] Z. Zhang, S. Lian, Q. Pei, and J. Pu, "Fuzzy risk assessments on security policies for digital rights management," *Neural Network World*, vol. 20, no. 3, pp. 265-284, 2010.
- [18] Z. Y. Zhang, *Risk Assessment and Management*, Academy Publish, 2012.

## Research Article

# Energy Efficiency of Task Allocation for Embedded JPEG Systems

Yang-Hsin Fan,<sup>1</sup> Jan-Ou Wu,<sup>2</sup> and San-Fu Wang<sup>3</sup>

<sup>1</sup> Department of Computer Science and Information Engineering, National Taitung University, Taitung 95002, Taiwan

<sup>2</sup> Department of Electronic Engineering, De Lin Institute of Technology, New Taipei 23654, Taiwan

<sup>3</sup> Department of Electronic Engineering, Ming Chi University of Technology, New Taipei 24301, Taiwan

Correspondence should be addressed to Yang-Hsin Fan; yhfan@nttu.edu.tw

Received 28 February 2014; Accepted 17 March 2014; Published 10 April 2014

Academic Editors: N. Barsoum, V. N. Dieu, and P. Vasant

Copyright © 2014 Yang-Hsin Fan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Embedded system works everywhere for repeatedly performing a few particular functionalities. Well-known products include consumer electronics, smart home applications, and telematics device, and so forth. Recently, developing methodology of embedded systems is applied to conduct the design of cloud embedded system resulting in the applications of embedded system being more diverse. However, the more energy consumes result from the more embedded system works. This study presents hyperrectangle technology (HT) to embedded system for obtaining energy saving. The HT adopts drift effect to construct embedded systems with more hardware circuits than software components or vice versa. It can fast construct embedded system with a set of hardware circuits and software components. Moreover, it has a great benefit to fast explore energy consumption for various embedded systems. The effects are presented by assessing a JPEG benchmarks. Experimental results demonstrate that the HT, respectively, achieves the energy saving by 29.84%, 2.07%, and 68.80% on average to GA, GHO, and Lin.

## 1. Introduction

Smarter, smaller, and portable characteristics make embedded systems to serve diverse functionalities. Nowadays, embedded systems have rapidly increasing requirements in the applications as in automobiles, avionics, and mobile devices. According to the IDC [1] reports, nearly 1 billion smart connected devices were shipped in 2011 and it would be double by 2016. Likewise, the Intel [2] predicts that there will be more than 30 billion devices constantly linked and another 150 billion fitfully connected in the end of this decade. It will greatly increase the demand of energy while these devices are served. However, the more embedded systems that are served, the more energy they consume. As a result, Bernd [3] summarizes the market and technology analysis which are towards energy efficiency for mobile devices, cloud computing, and storage services.

From an architectural perspective, all embedded systems have hardware and software components. These hardware components, such as *application specific integrated circuits*

(ASIC) or standard logic, offer specific functionalities or programmable gates when developing hardware circuits. Conversely, the software components, such as microcontrollers or programmable *digital signal processing* (DSPs), provide an environment for various application programs. By assembling these programmable hardware and software, any embedded system can be developed depending on system specifications.

Inside the embedded systems there are a few tasks that are designed by programmable hardware or software components. Each task consumes energy regardless of the forms in either programmable hardware or software components. Energy consumption is classified as dynamic and static energy consumption based on its mode of state. The dynamic energy consumption is defined while the task is working for providing functionalities. On the other hand, task consumes static energy dissipation when its state is idle. For an embedded system with  $n$  tasks, executing one task consumes dynamic energy and the other tasks arise static energy consumption. For example, when task 1 runs, task

2, task 3 till task  $n$  occur static energy consumption. In consideration of task 2 runs, it consumes dynamic energy consumption, at the same time, task 1, task 3, task 4 till task  $n$  occur static energy consumption. To iterate the process for every task execution, the energy consumption of embedded system can be assessed.

The aforementioned statements manifest embedded systems to incessantly consume energy for repeatedly performing a few particular tasks. In order to improve energy consumption, this study proposes *hyperrectangle technology* (HT) to embedded system target to obtain energy saving. This paper is organized as five sections. In Section 2, we investigate some previous works with respect to energy saving for embedded system. Section 3 describes HT for embedded systems to achieve energy saving. Section 4 demonstrates experimental results of *joint photographic experts group* (JPEG) encoding system that is performed by HT. We conclude this work in Section 5.

## 2. Preliminary Works

Researchers pay much attention to energy consumption of embedded systems with respect to the fields of processor energy consumption, real-time power consumption, dynamic power consumption, and scheduling power consumption. From the processors viewpoints, Vilcu [4] aims at real time embedded system to minimize the CPU power consumption. First, he studies task execution in the power consumption of processor(s). Then, he finds the effects of optimal configuration processor(s) for energy consumption. Finally, he defines globally optimal scheduling which gains minimal energy consumption for homogeneous multiprocessor system. Gao et al. [5] present *energy-efficient architecture for embedded software* (EAES) and dynamic energy-saving method to solve energy-saving problem. The former uses a processor with dynamic voltage scaling capability, FPGA modules, and extends directed acyclic graph to embedded system. The latter adopts preassignment to achieve dynamic runtime scheduling and minimize energy consumption. Qiu et al. [6] discuss the execution time of tasks with conditional instructions or operations problem. They adopt probabilistic random variable approach to model execution time of tasks. Then, they propose practical algorithm VACP to minimize energy consumption for uniprocessor embedded systems. Silva-Filho and Lima [7] state that memory hierarchy consumes power up to 50% in microprocessor system. Consequently, they propose an automated architecture exploration mechanism to NIOS II processor and memory hierarchy with parameter variation. The experimental results show the reduction of energy consumption is approximately 27%. In 2008, Zeng et al. [8] present generalized *dynamic energy performance scaling* (DEPS) framework to hard real-time embedded systems for exploring application-specific energy-saving potential issue. Three energy performance tradeoff technologies, DHRC, DVFS, and DPM, are integrated into DEPS. Their experimental results of simulation show the static DEPS has been improved, respectively, 13.6% and 13.7% in DVFS and DHRC. Also, dynamic DEPS improves 5.7% when comparing to static DEPS.

Real-time power information is a valuable data for software designer for battery-powered embedded systems. Genser et al. [9] propose power profiling approach to collect real-time power information at early designing stages. Moreover, they present an emulation-based power profiling approach to achieve real-time power analysis for embedded systems. Because the power information is collected at early designing stages, the development efficiency and time to market are improved. In 2008, Elewi et al. [10] first discuss the real-time scheduling of dependent tasks problem and then present enhanced *multispeed* (MS) algorithm for energy saving. With energy consumption problem of battery-powered embedded systems, Casares et al. [11] aim at embedded smart camera to analyze the power consumption and performance. Not only graph of energy consumption but also instructions of collections are presented. They conclude the importance of lightweight algorithm, the time of transfer data, and transferred data type.

Dynamic power consumption of *field programmable gate array* (FPGA) is discussed in [12, 13]. In 2009, Tsang and So [12] adopt precomputation approach to reduce dynamic power consumption in commercial off-the-shelf FPGAs. The experimental results of comparator show that 83% of dynamic power in logic or 43.1% of total dynamic power is reduced if the increased resource consumption can be negligible. In 2010, Bhandari et al. [13] present fly partial reconfiguration as well as scaling the clock on FPGA for reducing dynamic power consumption in embedded system. They conclude that the factors of dynamic power dissipation consist of application, architecture, and reconfigurable time.

The application of different algorithms to arrange scheduling issues for reducing power consumption is discussed in [14, 15]. In 2010, Bashiri and Miremadi [14] investigate *earliest-deadline-first* (EDF) and *rate-monotonic* (RM) algorithms on power efficiency of task scheduling. The results show that the BF-EDF and FF-EDF have the best power efficiency. In 2011, Cho et al. [15] propose power-saving scheduling algorithm and use soft-deadline to reduce energy consumption by about 40%. However, it is a trade off on performance and power savings for embedded systems. Kan et al. [16] present a heuristic algorithm called TGPM-ALL with interior point method to handle the frequency assignment on multiple soft-deadlines embedded systems. Their empirical results show the effectiveness in comparing TGPM-ALL with TGPM-I and BEST algorithms.

## 3. Hyperrectangle Approach

*Task graph* (TG) is a conceptual graph that facilitates to describe operation for embedded system. TG comprises of vertices ( $V$ ), edges ( $A$ ), and levels ( $L$ ) that can be represented by a 3-tuple set,  $G(V, A, L)$ . The  $V$  is a unit of work which may take dependencies one or more antecedents. The  $A$  is used to exhibit the flow among  $V$ . The  $L$  indicates the order of works for  $V$ . Based on TG, we propose *energy-consumed task graph* (ETG) as system model of HT that adds a factor of energy consumption on TG.

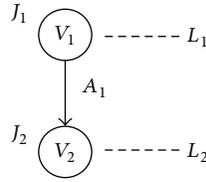


FIGURE 1: A simple ETG with two tasks.

**3.1. System Model.** ETG comprises of vertices ( $V$ ), edges ( $A$ ), levels ( $L$ ), and energy consumption ( $J$ ) that can be represented by a 4-tuple set,  $G(V, A, L, J)$ . Symbol  $V$  stands for task that is a working unit on embedded system. For example, one task is represented as  $V_1$  and a number of tasks are labeled as  $V_2, V_3$ , and  $V_4$ , and so forth. Hence, application program inside embedded system can be defined as a set of tasks as  $V_1, V_2$ , and  $V_3$  to  $V_n$ . Another symbol  $A$  is used to direct work flow of applications among  $V$ s. For instance, symbol  $A_1$  guides the working flow from task  $V_1$  to another task  $V_2$ . Moreover, two tasks are connected by  $A_1$  that implies their location on different levels. Label  $L$  defines the state of  $V$  and the height of ETG. The state is organized into two categories. One state is named working ( $W$ ) and the other idle ( $I$ ). Both states simultaneously incur when a symbol  $A_1$  activates. For instance, working state  $W_1$  on  $V_2$  and idle state  $I_1$  on  $V_1$  separately form when  $A_1$  activates. Sign  $J$  denotes the energy consumption of  $V$ . Each task consumes energy depending on either  $W$  or  $I$ . Figure 1 displays a ETG with 2 vertices (i.e.,  $V_1$  and  $V_2$ ), 1 edge, 2 energy consumption (i.e.,  $J_1$  and  $J_2$ ), and 2 levels (i.e.,  $L_1$  and  $L_2$ ).

**3.2. Energy Consumption Definition.** Power consumption of the task is classified into dynamic or static power consumption according to their state of work. Dynamic power consumption  $D$  occurs while the task locates  $W$ . On the contrary, the task in  $I$  consumes static power consumption  $S$ . Take Figure 1 as an example, the  $V_1$  and  $V_2$  first consume  $D$  and  $S$  separately in  $L_1$  because the former locates at  $W$  and the latter places on  $I$ . After that, the  $A_1$  directs the work flow to  $L_2$ . In  $L_2$ , the  $V_2$  and  $V_1$  individually consume  $D$  on  $W$  and  $S$  on  $I$ . In summary, both  $D$  and  $S$  are consumed by  $V$  depending on  $W$  or  $I$ . It should be noticed that each task must consume energy at any time even though it is idle.

Low power dissipation model and analysis for embedded systems are discussed by Fan et al. [17]. They derive power dissipation with dynamic and static power dissipation from TG. The expression of the sum of power consumption for embedded system is calculated by using

$$P = (L - 1) \times (P_{s,t_1} + P_{s,t_2} + \dots + P_{s,t_n}) + (P_{d,t_1} + P_{d,t_2} + \dots + P_{d,t_n}), \quad (1)$$

where  $L$  is the height of TG,  $P_s$  is static power consumption,  $P_d$  is dynamic power consumption, and  $t_1, t_2, \dots, t_n$  is a set of task. In consideration of energy consumption, the energy consumption is formulated as follows:

$$E = P \times T, \quad (2)$$

where  $P$  is power dissipation and  $T$  represents execution time. Owing to power consumption which is divided into  $D$  and  $S$ , the energy consumption is categorized into dynamic ( $E_d$ ) and static ( $E_s$ ) energy dissipation. Moreover, each task can be separately implemented as two forms of  $f$  as *hardware circuit* (HC) and *software component* (SC) so that the energy consumption of embedded system can be formulated as follows:

$$E_{d,v_i}^f = P_{d,v_i}^f \times T_{d,v_i}^f, \quad (3)$$

$$E_{s,v_i}^f = P_{s,v_i}^f \times T_{s,v_i}^f,$$

where  $d$  is dynamic energy consumption,  $s$  is static energy consumption,  $f$  is a form of hardware circuit or software component, and  $v$  is task,  $i = 1, 2, \dots, n$ . In summarizing equations from (1) to (3), the total energy consumption can be derived as follows:

$$E = (L - 1) \times (E_{s,v_1}^f + E_{s,v_2}^f + \dots + E_{s,v_n}^f) + (E_{d,v_1}^f + E_{d,v_2}^f + \dots + E_{d,v_n}^f), \quad (4)$$

where  $L$  is the height of ETG,  $E_s$  is static energy consumption, and  $E_d$  is dynamic energy consumption.

**3.3. Hyperrectangle Model.** To construct hyperrectangle model of energy consumption for embedded system, we first analyze the energy consumption of ETG with two tasks which is shown in Figure 1. Then derive complicated model from it. By holding the principle of one task that has two forms (i.e., HC and SC), a ETG with two tasks (i.e.,  $v_1$  and  $v_2$ ) can be constructed four embedded systems namely HC-HC ( $E_1$ ), HC-SC ( $E_2$ ), SC-HC ( $E_3$ ), and SC-SC ( $E_4$ ). According to (3), the energy consumption of  $E_1, E_2, E_3$ , and  $E_4$  can be defined in the following:

$$E_1 = E_{d,v_1}^{HC} + E_{s,v_2}^{HC} + E_{d,v_2}^{HC} + E_{s,v_1}^{HC}, \quad (5)$$

$$E_2 = E_{d,v_1}^{HC} + E_{s,v_2}^{SC} + E_{d,v_2}^{SC} + E_{s,v_1}^{HC}, \quad (6)$$

$$E_3 = E_{d,v_1}^{SC} + E_{s,v_2}^{HC} + E_{d,v_2}^{HC} + E_{s,v_1}^{SC}, \quad (7)$$

$$E_4 = E_{d,v_1}^{SC} + E_{s,v_2}^{SC} + E_{d,v_2}^{SC} + E_{s,v_1}^{SC}. \quad (8)$$

The first and the second terms or the third and the fourth terms from (5) to (8) prove the description in Section 3.1, which indicates that  $W$  and  $I$  simultaneously incur while a symbol  $A_1$  activates. Moreover, the third and the fourth terms can be regarded as mutual functions as the second and the first terms. Consequently, the third and the fourth terms from (5) to (8) can be formulated as the third term in the following:

$$E_1 = E_{d,v_1}^{HC} + E_{s,v_2}^{HC} + f(E_{d,v_1}^{HC}, E_{s,v_2}^{HC}), \quad (9)$$

$$E_2 = E_{d,v_1}^{HC} + E_{s,v_2}^{SC} + f(E_{d,v_1}^{HC}, E_{s,v_2}^{SC}), \quad (10)$$

$$E_3 = E_{d,v_1}^{SC} + E_{s,v_2}^{HC} + f(E_{d,v_1}^{SC}, E_{s,v_2}^{HC}), \quad (11)$$

$$E_4 = E_{d,v_1}^{SC} + E_{s,v_2}^{SC} + f(E_{d,v_1}^{SC}, E_{s,v_2}^{SC}). \quad (12)$$

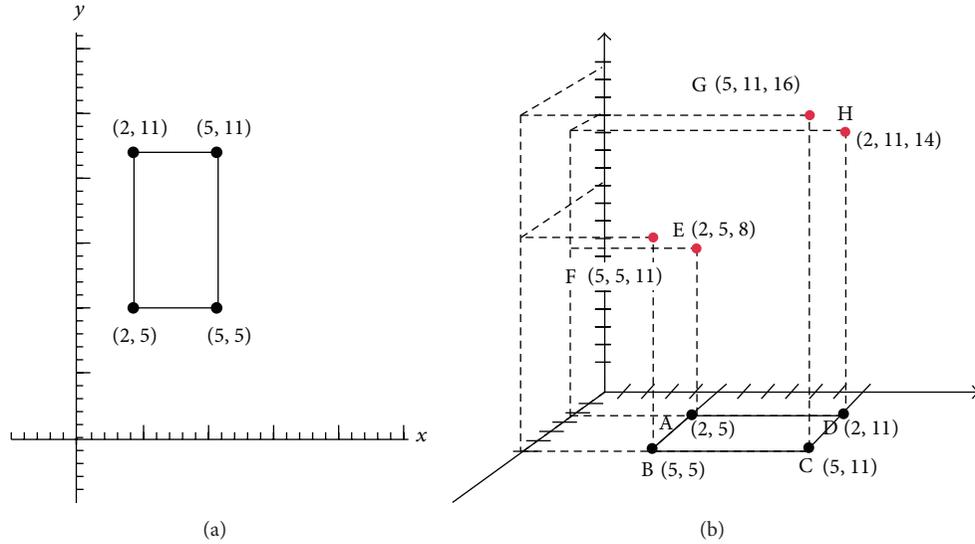


FIGURE 2: Example of a rectangle  $R^2$  in hyperrectangle model.

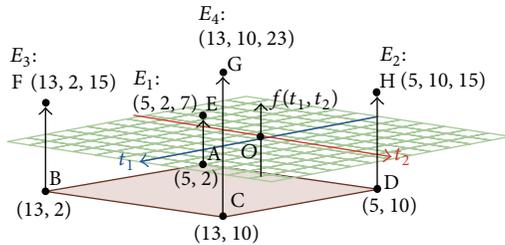


FIGURE 3: Hyperrectangle schema of embedded system with two tasks.

From the axial coordination's perspective, the first and the second terms from (9) to (12) can be represented as four points of a rectangle  $R^1$ . Moreover, the third term from (9) to (12) forms four points in another rectangle  $R^2$  in the tridimensionality. As a result, the hyperrectangle model of energy consumption for embedded system with two tasks is constructed and transferred to the three-dimensional space.

Based on the previous description, the hyperrectangle model of energy consumption for embedded system with two tasks can be defined in the following.

For a rectangle  $R^2$ , it has vertex  $(x_i, y_j)$ ,  $i, j = 1, 2$ , and a solution set  $D \subseteq R^2$ . A given function  $f : D \rightarrow R$ ,  $f(x, y) = u$  and  $b \in R$ , the solution can be obtained as follows:

$$x_i + y_j + f(x_i, y_j) \leq b, \quad \text{where } i, j = 1, 2. \quad (13)$$

*Example 1.* Figure 2(a) shows a set of vertex  $D = \{(2, 5), (5, 5), (2, 11), (5, 11)\}$ . Figure 2(b) displays  $f : D \rightarrow R$  where  $f(2, 5) = 8$ ,  $f(5, 5) = 11$ ,  $f(2, 11) = 14$ , and  $f(5, 11) = 16$ . For a given  $b = 20$ , the solution comprises  $\{(2, 5), (5, 5)\}$  since it meets  $x_i + y_j + f(x_i, y_j) \leq b$ .

*Example 2.* Table 1 displays an energy consumption example of embedded system with two tasks. The evaluating factors of

TABLE 1: Energy consumption of embedded system with 2 tasks.

Tasks	Energy consumption				
	$E_d^{HC}$	HC	$E_s^{HW}$	SC	$E_s^{SC}$
$v_1$	5		1	13	9
$v_2$	6		2	14	10

energy consumption include the name of task, dynamic, and static energy consumption of HC and SC. Figure 3 shows a hyperrectangle schema of embedded system with two tasks (i.e.,  $v_1$  and  $v_2$ ). According to (9) to (12), the first and the second term are transferred to XY axes for  $E_1, E_2, E_3$ , and  $E_4$ , where locates at points A, D, B, and C. Points E, H, F, and G are marked for the third term from (9) to (12). We observe the most energy consumption occurring at point G (i.e.,  $13 + 10 + 23 = 46$ ), which comprises SC-SC ( $E_4$ ). On the other hand, the HC-HC ( $E_1$ ) consumes the fewest energy consumption. Eventually, the energy consumption can be improved if one task is substituted from SC to HC where less energy is consumed.

Similarly, an embedded system with three tasks can be defined as follows. For a cuboid  $R^3$ , it has vertex  $(x_i, y_j, z_k)$ ,  $i, j, k = 1, 2$ , and a solution set  $D \subseteq R^3$ . A given function  $f : D \rightarrow R$ ,  $f(x, y, z) = v$  and  $b \in R$ , the solution can be obtained from the following:

$$x_i + y_j + z_k + f(x_i, y_j, z_k) \leq b, \quad \text{where } i, j, k = 1, 2. \quad (14)$$

According to (13) and (14), we summarize the general expression for embedded system with  $n$  tasks as follows. For a hyperrectangle  $R^n$ , it has vertex,  $(x_{(1,i_1)}, x_{(2,i_2)}, \dots, x_{(n,i_n)})$ ,  $i_j = 1, 2$ , and a solution set  $D \subseteq R^n$  where  $D$  is shown in

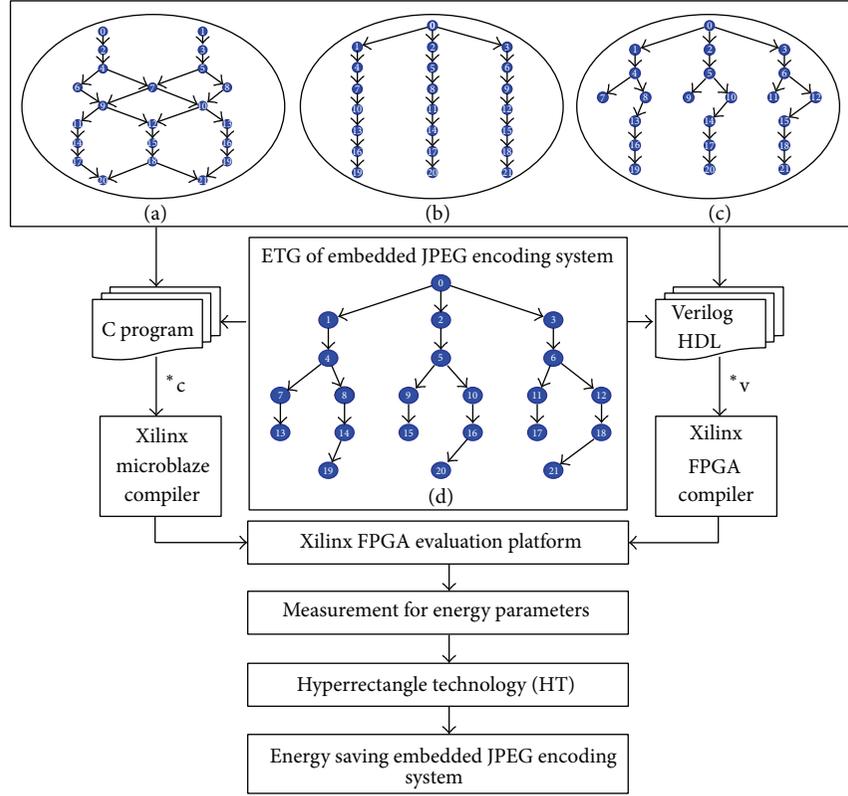


FIGURE 4: Flow chart of experimental setup.

TABLE 2: Technology of system parameters.

	GA	GHO
Number of chromosomes	100000	100000
Number of populations	500	500
Probability of crossover	1	1
Probability of mutation	1	1

(15). Given a function  $f : D \rightarrow R$  and  $b \in R$ , the solution can be obtained as follows:

$$D = \left\{ (x_{(1,i_1)}, x_{(2,i_2)}, \dots, x_{(n,i_n)}) \mid i_j = 1, 2 \right\}, \quad (15)$$

$$x_{(1,i_1)} + x_{(2,i_2)} + \dots + x_{(n,i_n)} + f(x_{(1,i_1)}, x_{(2,i_2)}, \dots, x_{(n,i_n)}) \leq b, \quad (16)$$

where  $1 \leq j \leq n$ .

Applying hyperrectangle approach to embedded systems for gaining energy consumption consists of the following steps. First, the number of tasks  $U$  of embedded system must be defined. Next, separately constructing the number of tasks  $U_{hc,i}$  with HC and  $U_{sc,j}$  with SC being the same as  $U$ . Third, constructing sets of  $f(U_{hc,i})$  and  $f(U_{sc,j})$  depicts energy consumption for each task. Therefore, energy consumption of each task can be fast evaluated. Fourth, constructing the first embedded system  $ES_a$  that comprises a set of tasks with  $U_{sc}$ . Fifth, the task with the most energy consumption in  $f(U_{sc,a})$  is swapped with  $f(U_{hc,a})$ . After the swapping process

is iterated until each task is made of  $U_{hc}$ , the HT exploits a set of embedded systems  $HT_x$  according to the number of  $U_{hc}$  and  $U_{sc}$ . The first embedded system  $ES_1$  is assembled by one of two  $U_{hc}$  and  $U_{sc}$ . If  $U$  is odd, extra SC is set preceding privilege to deploy to the  $ES_1$ . The second embedded system  $ES_2$  is set where the number of tasks with SC is more than HC. Alternatively, the number of tasks with HC that is more than SC is set to the third embedded system  $ES_3$ . In setting  $ES_2$  and  $ES_3$  until  $ES_m$ , the above process is executed repeatedly. We observe that  $ES_{2i}$  and  $ES_{(2i+1)}$  form the drift effect with SC and HC, respectively.

#### 4. Experimental Results and Analysis

The experimental platform is Xilinx FPGA ML507 [18]. Table 2 presents the technology of system parameters. The tested example is *joint photographic experts group* (JPEG) encoding system that consists of 22 tasks and 9 levels. From level 1 to 9 in Figure 4(a), the number of tasks is 2, 2, 2, 3, 2, 3, 3, 3, and 2, respectively. Each task is individually implemented as hardware circuit and software component form, which are designed by Verilog programming language and C programming language. Figure 4 demonstrates the flow chart of experimental setup.

The measured data of energy consumption is shown in Table 3. In the Task column, it shows the name of task that works in the JPEG encoding system. For the dynamic and static energy consumption of hardware circuit, it is displayed

TABLE 3: Measured data of tasks for JPEG encoding system.

Tasks	Energy Consumption			
	$E_d^{HC}$	HC ( $10^{-6}$ )	$E_s^{HC}$	SC ( $10^{-3}$ )
$v_0$ (LevelOffset)	0.0180504		0.0119968	0.591286362
$v_1$ (DCT)	1.414320718		0.631151123	181.6586949
$v_2$ (DCT)	1.414320718		0.631151123	181.6586949
$v_3$ (DCT)	1.414320718		0.631151123	181.6586949
$v_4$ (Quant.)	0.629914608		0.123049296	0.93692231
$v_5$ (Quant.)	0.629914608		0.123049296	0.93692231
$v_6$ (Quant.)	0.629914608		0.123049296	0.93692231
$v_7$ (DPCM)	0.000575667		0.000336501	0.040158143
$v_8$ (ZigZag)	0.02977856		0.003429459	0.561685068
$v_9$ (DPCM)	0.000575667		0.000336501	0.040158143
$v_{10}$ (ZigZag)	0.02977856		0.003429459	0.561685068
$v_{11}$ (DPCM)	0.000575667		0.000336501	0.040158143
$v_{12}$ (ZigZag)	0.02977856		0.003429459	0.561685068
$v_{13}$ (VLC)	0.086972024		0.028369015	0.054136441
$v_{14}$ (RLE)	0.046884314		0.023175988	0.838115122
$v_{15}$ (VLC)	0.093019754		0.030341698	0.054136441
$v_{16}$ (RLE)	0.046884314		0.023175988	0.838115122
$v_{17}$ (VLC)	0.093019754		0.030341698	0.054136441
$v_{18}$ (RLE)	0.046884314		0.023175988	0.838115122
$v_{19}$ (VLC)	0.116959532		0.022825184	0.996227367
$v_{20}$ (VLC)	0.116959532		0.022825184	0.976457285
$v_{21}$ (VLC)	0.116959532		0.022825184	0.976457285

TABLE 4: Energy consumption of HT.

Embedded systems		HT (ours)			
$v_1$	$v_{22}$	Energy dissipation (J)			
		Structure 1: Figure 4(a)	Structure 2: Figure 4(b)	Structure 3: Figure 4(c)	Structure 4: Figure 4(d)
$ES_a(00000000000000000000)$		3.9713661	3.544421533	3.117476966	2.69053
$ES_b(01000000000000000000)$		2.680031235	2.391796366	2.103561497	1.81533
$ES_c(01100000000000000000)$		1.38869637	1.239171199	1.089646028	0.94012
$ES_d(01110000000000000000)$		0.097361505	0.086546032	0.075730559	0.06492
$ES_e(011100000000000000100)$		0.088413403	0.078591929	0.068770455	0.05895
$ES_f(011100000000000000010)$		0.079642539	0.070795381	0.061948222	0.05310
$ES_g(011100000000000000011)$		0.070871676	0.062998833	0.05512599	0.04725
$ES_h(011110000000000000011)$		0.0624566	0.05518605	0.04858061	0.04164
$ES_i(011111000000000000011)$		0.054041524	0.048038377	0.04203523	0.03603
$ES_j(011111100000000000011)$		0.045626448	0.040558149	0.03548985	0.03042
$ES_k(0111111100000001000011)$		0.038096141	0.033864372	0.029632602	0.02540
$ES_l(0111111100000001010011)$		0.030565834	0.027170595	0.023775355	0.02038
$ES_m(0111111100000001010111)$		0.023035527	0.020476818	0.017918108	0.01536
$ES_n(1111111100000001010111)$		0.017720291	0.015752077	0.013783864	0.01182
$ES_o(1111111101000001010111)$		0.012670802	0.011263568	0.009856333	0.00845
$ES_p(1111111101010001010111)$		0.007621314	0.006775058	0.005928803	0.00508
$ES_q(11111111010101010111)$		0.002571825	0.002286549	0.002001273	0.00172
$ES_r(111111110101011010111)$		0.002084963	0.001853789	0.001622614	0.00139
$ES_s(111111110101011110111)$		0.001598124	0.001421048	0.001243973	0.00107
$ES_t(111111110101011111111)$		0.001111284	0.000988308	0.000865333	0.00074
$ES_u(111111111010111111111)$		0.000749892	0.000667071	0.00058425	0.00050
$ES_v(111111111101111111111)$		0.000388501	0.000345834	0.000303167	0.00026
$ES_w(111111111111111111111)$		2.711E - 05	2.4597E - 05	2.20841E - 05	1.95711E - 05

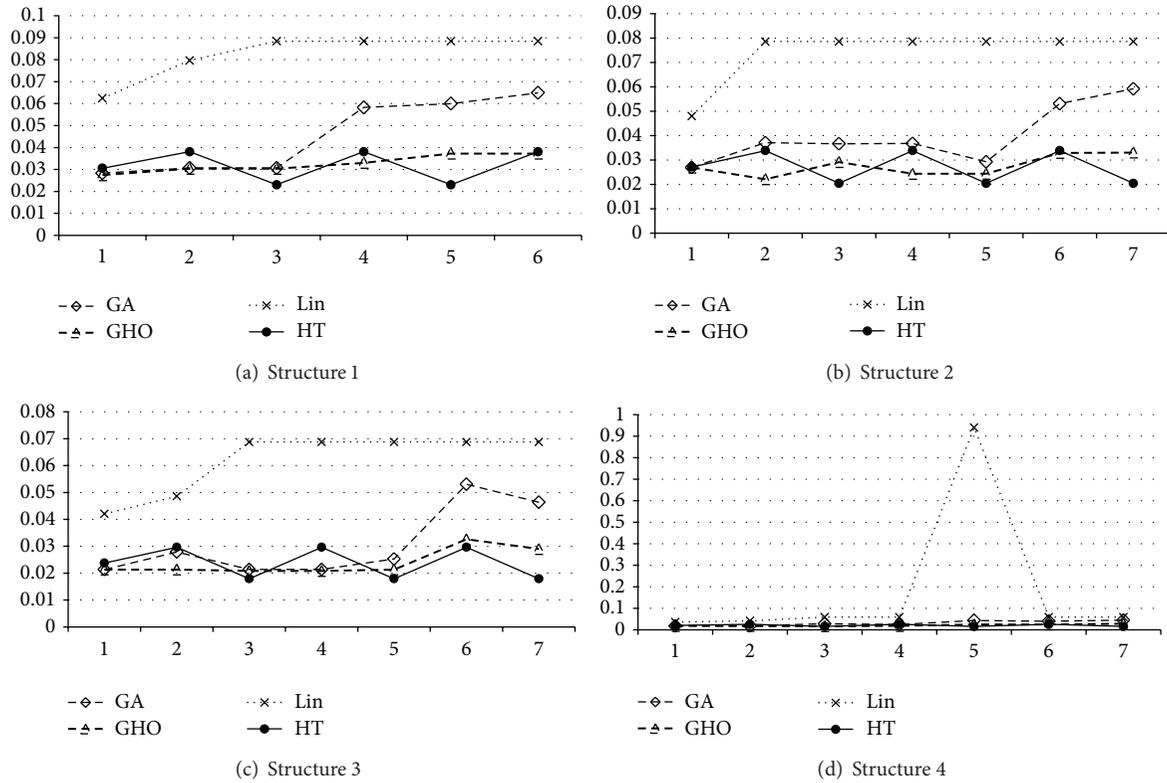


FIGURE 5: Energy consumption of embedded JPEG encoding system.

in column 2 and 3. The software tasks with dynamic and static energy consumption are illustrated in column 4 and 5.

Table 4 lists the experimental results of the proposed approach. Embedded systems column displays the results according to (16). In the energy dissipation column, it is calculated via (4). The  $ES_1$  is set to  $ES_1$ ,  $ES_k$ , and  $ES_m$  is set to  $ES_2$  and  $ES_3$ , respectively. On one hand, the SC drift effect, respectively, diffuses from  $ES_k$  to  $ES_a$ . Similarly, the HC drift effect diffuses from  $ES_m$  to  $ES_w$ , respectively. All designs by HT to embedded systems can be fast explored for energy consumption.

To present the efficiency of the proposed HT, we compare HT to *genetic algorithm* (GA) [19], GHO [20], and Lin [21] via JPEG benchmarks. Four structures shown in Figure 4 of energy consumption are used to demonstrate the effects of HT. The energy consumption of each structure is set to 0.1, 0.09, 0.08, and 1 joule, respectively. The structure 1 is shown in Figure 4(a) and the experimental result is depicted in Figure 5(a). The HT gains the energy saving in comparison with GA [19], GHO [20], and Lin [21]. Moreover, the HT improves the energy consumption by 30.00%, 2.38%, and 61.49% on average to GA [19], GHO [20], and Lin [21], respectively. Figure 5(b) displays the results of structure 2 (Figure 4(b)), in which the HT separately improves the energy consumption by 31.89%, 1.52%, and 63.40% on average to GA [19], GHO [20], and Lin [21]. Figure 5(c) shows the results of structure 3 (Figure 4(c)), in which the HT individually improves the energy consumption by 23.18%, 0.44%

and, 61.69% on average to GA [19], GHO [20], and Lin [21]. Figure 5(d) exhibits the results of structure 4 (Figure 4(d)), in which the HT separately improves the energy consumption by 34.28%, 3.93%, and 88.62% on average to GA [19], GHO [20], and Lin [21]. In summary, the HT achieves the energy saving by 29.84%, 2.07%, and 68.80% on average to GA [19], GHO [20], and Lin [21], respectively.

## 5. Conclusion

Energy saving issue is always discussed and concerned in electronic devices. Nowadays, nearly any electronic device either already has existed or will embed computing systems resulting in the applications of embedded systems that are more diverse. It reveals that embedded systems are growing exponentially. While more and more embedded systems are repeated day by day in order to provide various functionalities, the speed of energy consumption is greatly increased.

This study presents *hyperrectangle technology* (HT) to embedded systems target to achieve energy saving. The drift effect on HT facilitates the designer to fast explore energy consumption of embedded systems. The effectiveness of the proposed approach is demonstrated by assessing a JPEG benchmarks. Experimental results demonstrate that the HT achieves the energy saving by 29.84%, 2.07%, and 68.80% on average to GA, GHO, and Lin, respectively. Consequently, this work is valuable for developing energy-saving embedded systems.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

The authors would like to thank the National Science Council, Taiwan, for financially supporting this research under Contract nos. NSC 101-2221-E-143-002 and NSC 102-2221-E-143-002.

## References

- [1] Businesswire web site, [http://www.Businesswire.com/news/home/20120328005370/en/1-Billion-Smart-Connected-Devices-Shipped-2011#.U0Qsr\\_mSx8E](http://www.Businesswire.com/news/home/20120328005370/en/1-Billion-Smart-Connected-Devices-Shipped-2011#.U0Qsr_mSx8E).
- [2] Intel web site, February 2013, <http://www.intel.com/content/dam/www/public/us/en/documents/white-papers/consumerization-of-it-trends-whitepaper.pdf>.
- [3] P. S. Bernd, "Technology, market and cost trends 2012," CERN IT Report, 2012, [https://espace.cern.ch/WLCG-document-repository/Technical\\_Documents/Technology\\_Market\\_Cost\\_Trends\\_2012\\_v23.pdf](https://espace.cern.ch/WLCG-document-repository/Technical_Documents/Technology_Market_Cost_Trends_2012_v23.pdf).
- [4] D. Vilcu, "Real time scheduling and CPU power consumption in embedded systems," in *Proceedings of the 2008 IEEE International Conference on Automation, Quality and Testing, Robotics*, pp. 261–266, May 2008.
- [5] Z. Gao, G. Dai, P. Liu, and P. Zhang, "Energy-efficient architecture for embedded software with hard real-time requirements in partial reconfigurable systems," in *Proceedings of the 2009 IEEE International Conference on Embedded Computing, Scalable Computing, and Communications*, pp. 387–392, September 2009.
- [6] M. Qiu, J. Wu, F. Hu, S. Liu, and L. Wang, "Voltage assignment for soft real-time embedded systems with continuous probability distribution," in *Proceedings of the 15th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications*, pp. 413–418, August 2009.
- [7] A. G. Silva-Filho and S. M. L. Lima, "Energy consumption reduction mechanism by tuning cache configuration using nios II processor," in *Proceedings of the 2008 IEEE International SOC Conference*, pp. 291–294, September 2008.
- [8] G. Zeng, H. Tomiyama, H. Takada, and T. Ishihara, "A generalized framework for system-wide energy savings in hard real-time embedded systems," in *Proceedings of the IEEE/IFIP International Conference on Embedded and Ubiquitous Computing (EUC '08)*, pp. 206–213, December 2008.
- [9] A. Genser, C. Bachmann, J. Haid, C. Steger, and R. Weiss, "An emulation-based real-time power profiling unit for embedded software," in *Proceedings of the IEEE International Symposium on Systems, Architectures, Modelling, and Simulation*, pp. 67–73, July 2009.
- [10] A. M. Elewi, M. H. A. Awadalla, and M. I. Eladawy, "Energy-efficient multi-speed algorithm for scheduling dependent real-time tasks," in *Proceedings of the International Conference on Computer Engineering and Systems (ICCES '08)*, pp. 237–242, November 2008.
- [11] M. Casares, A. Pinto, Y. Wang, and S. Velipasalar, "Power consumption and performance analysis of object tracking and event detection with wireless embedded smart cameras," in *Proceedings of the 3rd International Conference on Signal Processing and Communication Systems (ICSPCS '09)*, pp. 1–8, September 2009.
- [12] C. C. Tsang and H. K.-H. So, "Reducing dynamic power consumption in FPGAs using precomputation," in *Proceedings of the IEEE International Conference on Field-Programmable Technology (FPT '09)*, pp. 407–410, December 2009.
- [13] S. U. Bhandari, S. Subbaraman, and S. Pujari, "Power reduction in embedded system on FPGA using the fly partial reconfiguration," in *Proceedings of the 2010 IEEE International Symposium on Electronic System Design*, pp. 77–80, December 2010.
- [14] M. Bashiri and S. G. Miremadi, "Investigating the effects of schedulability conditions on the power efficiency of task scheduling in an embedded system," in *Proceedings of the 13th IEEE International Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing*, pp. 102–106, May 2010.
- [15] K. Cho, C. Liang, J. Huang, and C. Yang, "Design and implementation of a general purpose power-saving scheduling algorithm for embedded systems," in *Proceedings of the IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC '11)*, pp. 1–5, September 2011.
- [16] E. Y. Y. Kan, W. K. Chan, and T. H. Tse, "Leveraging performance and power savings for embedded systems using multiple target deadlines," in *Proceedings of the IEEE International Conference on Quality Software*, pp. 473–480, July 2010.
- [17] Y.-H. Fan, J.-O. Wu, and S.-F. Wang, "Low power dissipation model analysis for embedded systems," *Journal of Research Notes in Information Science*, vol. 13, pp. 184–188, 2013.
- [18] Xilinx ML507 website, <http://www.xilinx.com>.
- [19] Y. Zou, Z. Zhuang, and H. Chen, "HW-SW partitioning based on genetic algorithm," in *Proceedings of the 2004 Congress on Evolutionary Computation (CEC '04)*, pp. 628–633, Portland, Ore, USA, June 2004.
- [20] T.-Y. Lee, Y.-H. Fan, Y.-M. Cheng, and C.-C. Tsai, "Hardware-software partitioning for embedded multiprocessor FPGA systems," *Journal of Innovative Computing, Information and Control A*, vol. 5, no. 10, pp. 3071–3083, 2009.
- [21] T. Y. Lin, Y. T. Hung, and R. G. Chang, "Efficient hardware/software partitioning approach for embedded multiprocessor systems," in *Proceedings of the International Symposium on VLSI Design, Automation and Test (VLSI-DAT '06)*, pp. 231–234, Hsinchu, Taiwan, April 2007.

## Research Article

# Comparative Study of Popular Objective Functions for Damping Power System Oscillations in Multimachine System

Naz Niamul Islam,<sup>1</sup> M. A. Hannan,<sup>1</sup> Hussain Shareef,<sup>1</sup> Azah Mohamed,<sup>1</sup> and M. A. Salam<sup>2</sup>

<sup>1</sup> Department of Electrical, Electronic and Systems Engineering, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, 43600 Bnagi, Selangor, Malaysia

<sup>2</sup> Electrical and Electronic Engineering Program, Faculty of Engineering, Institut Teknologi Brunei (Engineering and Technology University), Jalan Tungku Link, Gadong BE1410, Brunei Darussalam

Correspondence should be addressed to Naz Niamul Islam; [niamul@eng.ukm.my](mailto:niamul@eng.ukm.my)

Received 24 February 2014; Accepted 19 March 2014; Published 6 April 2014

Academic Editors: N. Barsoum, P. Vasant, and G.-W. Weber

Copyright © 2014 Naz Niamul Islam et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Power oscillation damping controller is designed in linearized model with heuristic optimization techniques. Selection of the objective function is very crucial for damping controller design by optimization algorithms. In this research, comparative analysis has been carried out to evaluate the effectiveness of popular objective functions used in power system oscillation damping. Two-stage lead-lag damping controller by means of power system stabilizers is optimized using differential search algorithm for different objective functions. Linearized model simulations are performed to compare the dominant mode's performance and then the nonlinear model is continued to evaluate the damping performance over power system oscillations. All the simulations are conducted in two-area four-machine power system to bring a detailed analysis. Investigated results proved that multiobjective D-shaped function is an effective objective function in terms of moving unstable and lightly damped electromechanical modes into stable region. Thus, D-shape function ultimately improves overall system damping and concurrently enhances power system reliability.

## 1. Introduction

Reliability and security of multimachine power system are achieved through numerous controllers. Design of robust damping controllers for the interconnected power system is very sensitive in order to suppress oscillations [1]. Power system oscillation is very harmful to the overall interconnected system. Therefore, proper and adequate damping is required to prevent damages caused by oscillations [1]. Damping power system oscillations are performed by means of power system stabilizers (PSSs) installed in the generator's excitation system and flexible AC transmission system (FACTS) devices installed in transmission line for fast acting damping performance [2–4]. The damping performance by PSS and FACTS depends on proper selection of its controller parameters [3]. The design of PSS and FACTS controllers usually is achieved in the linearized model by minimizing or maximizing an objective function using metaheuristic

optimization algorithms [5–9]. Improving damping performance has been proposed by many superior metaheuristic algorithms using inconsistent objective functions [5, 9–13]. However, metaheuristic algorithms are used to minimize or maximize an objective function that correspond optimized parameters under several system operating conditions [11, 14]. Therefore, selection of objective function may influence the overall controller optimization problem and that consequently affects existing system reliability.

In the last few years, many approaches to formulate the objective functions have been used in the damping controller design problem. That may be categorized into two types: (i) single objective functions and (ii) multiobjective functions. Application of single objective functions is presented either in terms of damping factors or damping ratios of electromechanical modes. In many literatures, the worst damping factor (largest value) has been selected and then the aim of optimization was set to minimize it [5–8]. On

the other hand, damping ratio is also important to limit the overshoot of oscillations which ultimately improves damping performance of controllers. Therefore, the lowest damping ratio has been determined and controller parameters were optimized in such a way to maximize it in [9]. However, system damping depends on the performance of dominant electromechanical modes rather than a single mode only. So, a single objective function known as comprehensive damping index (CDI) has been proposed in [10] that includes dominant modes. As both of the damping factor and damping ratio contribute to enhance damping performance, significance of multiobjective functions is realized in numerous researches [11–19]. In multiobjective techniques, the formulation to place electromechanical modes into a D-shaped region is very popular. However, different approaches have been applied to form D-shaped stability area. In [11, 14–18], a technique to form D-shaped stability zone is discussed based on the expected damping factor and ratio for dominant modes. In that objective function, only selected dominant modes are considered into optimization that do not have expected damping factor or damping ratio. Another approach to set up D-shaped region has been claimed in [12, 19] by the algebraic sum of the worst damping factor and ratio. Moreover, a different arrangement is reported in [13] in order to form D-shaped sector.

Application of the objective function with different formulations has been proposed for the design of robust damping controllers. As an objective function formulation is quite sensitive to damping controller parameter optimization, it is very significant to evaluate their comparative performance under a common base. Therefore, this research comes forward to conduct a comparative and comprehensive investigation to evaluate the existing objective functions for damping controller design in multimachine power system. Popular objective functions are considered to use with a metaheuristic optimization algorithm under a base setting. Later on, the analysis is extended to assess performances in linearized model and nonlinear model of the power system.

## 2. Power System and Damping Controller

Two-area four-machine power system shown in Figure 1 is a benchmark system to study power system oscillations [1, 2]. The performance of local and interarea modes of oscillations can be observed easily in this system [2]. Therefore, it is adopted to design damping controllers by means of PSSs. The overall system has 13 buses, two load areas, and four synchronous generators equipped with similar components. The excitation system is the IEEE type ST1 exciter. The detail system data are taken from [2].

A PSS is added with each synchronous generator to provide an additional damping torque if system tempts to become unstable for oscillations. All PSSs are conventional PSS (CPSS) which consists of two-stage lead-lag compensation blocks with one gain and one washout block. The rotor speed deviation ( $\Delta\omega$ ) is the input of PSS. IEEE type ST1 excitation system with PSS is presented in Figure 2, where  $T_w$  is the washout time constant and it is set to 10 s for all PSSs.

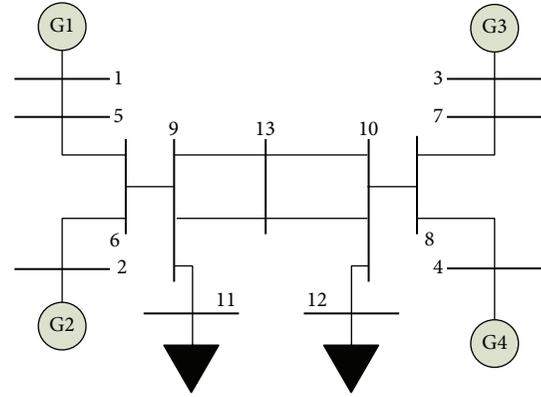


FIGURE 1: Single line diagram of two-area four-machine power system.

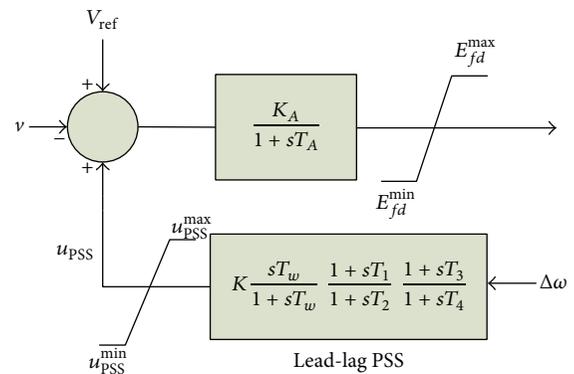


FIGURE 2: IEEE type ST1 excitation system with power system stabilizer.

On the other hand, optimizing parameters are gain constant ( $K$ ) and lead-lag time constants ( $T_1$ ,  $T_2$ ,  $T_3$ , and  $T_4$ ).

The entire systems are modelled based on linear-time-invariant (LTI) state-space (SS) theory around an operating point from a set of nonlinear differential equations. The linearized system model is represented in terms of the state matrix ( $A$ ), input matrix ( $B$ ), output matrix ( $C$ ), and feed-forward matrix ( $D$ ) shown in

$$\begin{aligned} \dot{x} &= A \cdot x + B \cdot u, \\ y &= C \cdot x + D \cdot u, \end{aligned} \quad (1)$$

where  $x$  is the vector of state variables and  $u$  is the vector of input variables. The eigenvalues determined from state matrix  $A$  reflect the overall system stability. Therefore, objective functions are formulated to move eigenvalue forcefully from unstable region (right side of  $s$ -plane) to stable region (left side of  $s$ -plane). The key properties of an eigenvalue ( $\lambda = \sigma \pm j\omega$ ) are damping factor ( $\sigma$ ) and the damping ratio ( $\zeta$ ) that are determined from

$$\begin{aligned} \sigma_i &= \text{real}(\lambda_i), \\ \zeta_i &= -\frac{\sigma_i}{\sqrt{\sigma_i^2 + \omega_i^2}}, \end{aligned} \quad (2)$$

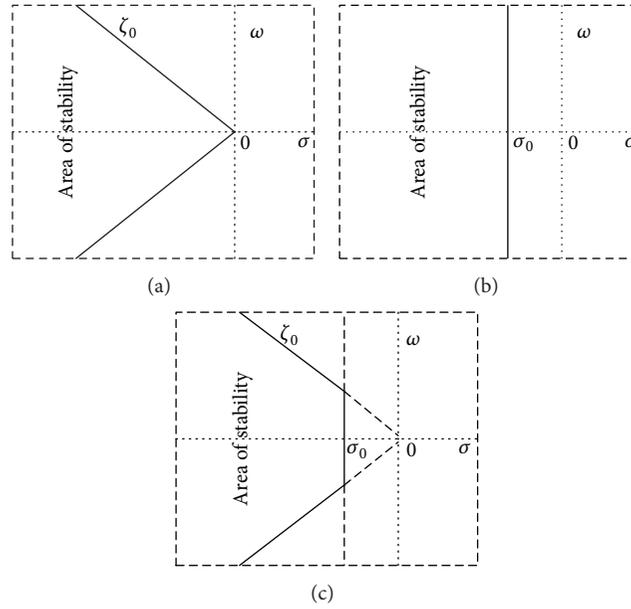


FIGURE 3: Stability regions defined by different objective functions in s-plane.

where  $i = 1, 2, 3, \dots, n$  and  $n$  is the total number of eigenvalues in the power system. The objective functions are maximized or minimized in order to achieve system stability under constraints of damping controller parameters shown

$$\begin{aligned}
 K_{\min} &\leq K \leq K_{\max}, \\
 T_{1,\min} &\leq T_1 \leq T_{1,\max}, \\
 T_{2,\min} &\leq T_2 \leq T_{2,\max}, \\
 T_{3,\min} &\leq T_3 \leq T_{3,\max}, \\
 T_{4,\min} &\leq T_4 \leq T_{4,\max}.
 \end{aligned} \tag{3}$$

### 3. Objective Functions

The main purpose of objective function formulation is to shift eigenvalues from unstable region to stable region shown in Figure 3 to achieve sufficient damping over power system oscillations. The popular objective functions are described in (4)–(8).

**3.1. Single Objective Functions.** These types of objective functions are formulated by focusing one property of eigenvalues (either  $\sigma$  or  $\zeta$ ) that partially ensure system stability for total  $j$  operating points. From research paper [5–8], the objective function is constructed to improve only damping factor shown in (4). In this case, the objective function  $J_1$  is set to minimize

$$J_1 = \max \{ \text{real}(\lambda_{ij}) \}. \tag{4}$$

Damping ratio is a measure of damping oscillations. In [9], the eigenvalue with poorest damping ratio is selected to maximize  $J_2$  shown in

$$J_2 = \min(\zeta_{ij}). \tag{5}$$

In other literature [10], CDI is considered as the objective function that considered all dominant modes to achieve maximum damping ratios instead of one mode only. CDI is described by (6). Here,  $J_3$  is minimized in order to optimize controller parameters.

$$J_3 = \sum_{i=1}^n (1 - \zeta_{ij}). \tag{6}$$

**3.2. Multiobjective Functions.** Different approaches are adopted to move eigenvalues into D-shaped stability regions. Multiobjective functions in (7) [11, 14–18] and in (8) [12, 19] are reported to minimize ( $J_4$ ) and maximize ( $J_5$ ), respectively, in order to achieve D-shaped stability criteria:

$$J_4 = \sum_{j=1}^{np} \sum_{\sigma_{ij} \geq \sigma_0} (\sigma_0 - \sigma_{ij})^2 + a \cdot \sum_{j=1}^{np} \sum_{\zeta_{ij} \leq \zeta_0} (\zeta_0 - \zeta_{ij})^2, \tag{7}$$

$$J_5 = -\max(\sigma_{ij}) + \min(\zeta_{ij}). \tag{8}$$

### 4. Differential Search Algorithm

Differential search algorithm (DSA), as a metaheuristic optimization algorithm, is selected to assess the objective functions (4)–(8) for improving damping performances by multimachine PSSs. Tuning the parameters of PSSs damping

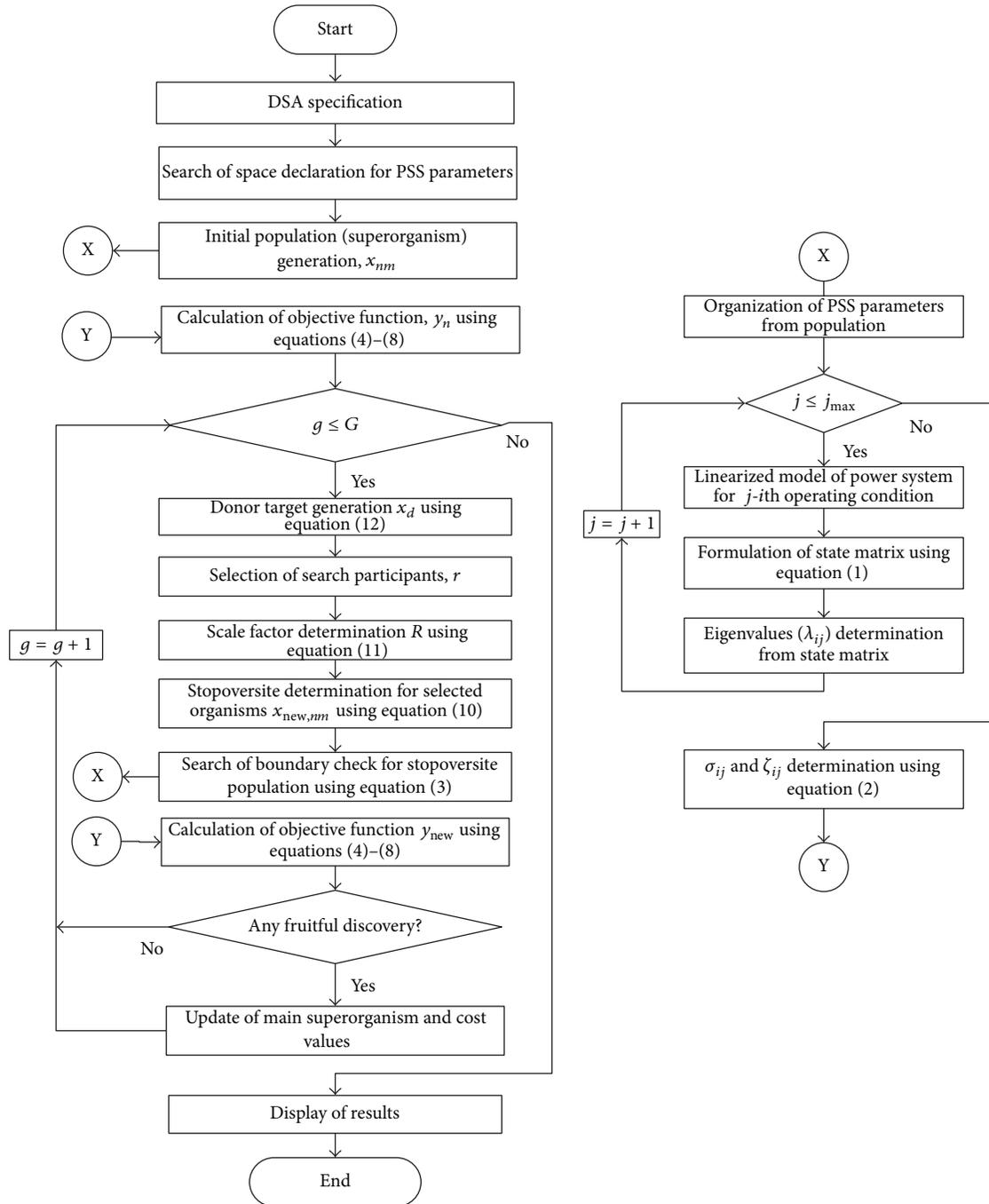


FIGURE 4: Flow chart of PSS parameters optimization with different objective functions using DSA.

controllers is a multimodal optimization problem and DSA is recommended specially to solve this type of problem [20]. In DSA, the initial population of organisms is generated using (9) and their movements are characterized by (10):

$$x_{nm} = low_m + (up_m - low_m) \cdot rand, \quad (9)$$

$$x_{new, nm} = x_{nm} + R \cdot (x_d - x_{nm}), \quad (10)$$

where  $x_{nm}$  is the initial population and  $low_m$  and  $up_m$  are the lower and upper constraint vectors of PSS parameters,

respectively. The scale factor  $R$  is generated using a gamma random function as shown in

$$R = randg \cdot [2 \cdot rand] \cdot (rand - rand), \quad (11)$$

where  $randg$  and  $rand$  are the gamma random function and uniformly distributed pseudorandom number, respectively. In DSA, donors are considered to be located in fruitful area. Therefore, generated populations are moved towards the

TABLE 1: System operating conditions considered in damping controller design.

Case	Operating points	Faults at bus	Between buses
1	Base case (all line in services)		
2	Three phase fault	9	9 and 13
3	Line to line fault	13	13 and 10

location of donors and donors are made up by reshuffling the original populations as shown in

$$x_d = (x_{nm})_{\text{random\_reshuffling}} \quad (12)$$

For a fair comparison, the simulation settings of DSA are kept the same for each objective function. The simulation settings are mentioned in the appendix. To understand the overall procedures of this research, the step by step works are depicted in flow chart diagram shown in Figure 4. Interested authors are referred to the original paper of this optimization algorithm in [20].

## 5. Results and Discussion

In order to evaluate the effectiveness of selected objective functions in (4)–(8), simulations are performed in linearized model and nonlinear model of the power system. The linear model simulations are performed using DSA technique to minimize or maximize the objective functions that concurrently bring unstable and poorly damped electromechanical modes into stable region. The common base setting for optimization simulations using DSA is the same (mentioned in the appendix) for all objective functions. All the simulations are performed on a desktop computer with 3.14 GHz processor and 4 GB DDR3 RAM. The system operating points are considered as mentioned in Table 1. The test system has four PSSs and total 20 parameters are to be optimized.

Simulations for each objective function have been performed 10 times and data are recorded. Then the best one (according to the objective function value) is selected for comparison. The same process is executed for all five objective functions. The best optimized parameters of four PSSs are listed in Table 2. From the best optimized value, the major dominant modes are determined as shown in Table 3 with corresponding damping ratios and frequencies. Sensitive modes are highlighted (bold) for easy understanding.

For objective function  $J_1$ , although dominant modes have achieved expected damping ratios, (0.15) it may lead to system instability. The reason is its interarea mode's damping factor, which is below expected value (−1.5). In addition, the frequency of interarea modes is very low; therefore it is enough to show poor damping. For remaining objective functions, it is noticeable that dominant modes for only  $J_4$  are achieved expecting damping factor and damping ratio. Moreover,  $J_4$  has no modes with very low frequency. The improvement is based on the damping factors and damping ratios. Single objective functions  $J_1$ ,  $J_2$ , and  $J_3$  do not have both expected values, which is a sign of poor damping performance. Whereas multiobjective functions  $J_4$ ,  $J_5$  have

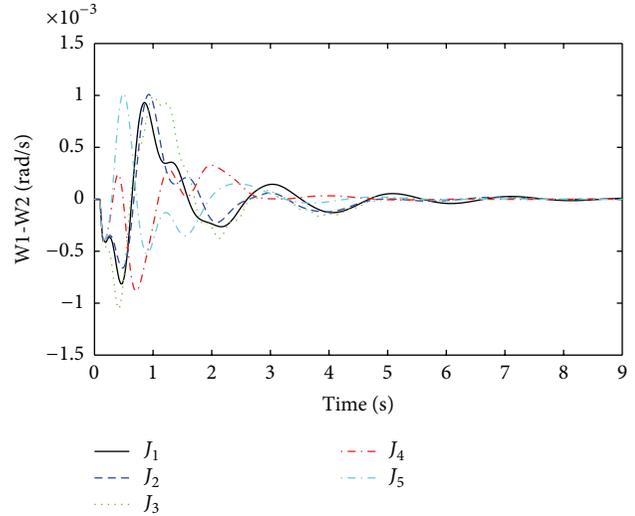


FIGURE 5: Local mode of oscillations between G1 and G2 (W1-W2) for three-phase fault at bus 9.

better performance compared with single objective functions. Even though  $J_5$  do not have expected damping factor for its low frequency mode, which may lead to poor damping.

The objective functions performances are assessed further in nonlinear time domain simulations by comparing damping performance. In order to check the robustness of damping controller, it is important to evaluate damping performances under considered (Table 1) and not considered operating points during linear model optimization. As a considered operating condition, three-phase fault near bus 9 (between lines 9 to 13) and, as a not considered operating point, three-phase fault near bus 13 (between lines 13–9) are executed individually. For both conditions, faults are applied for 100 ms duration and cleared by opening remote end of the faulted line at 150 ms. After that the system is operating in only one tie line and oscillations between generators to generators start growing. Optimized PSSs sense the oscillations and provide damping to suppress it. The quality of a good damping controller is measured on how fast the settling time is achieved and how smoothly the overshoot it can handle to avoid major damages by oscillations. Figures 5 and 6 are the responses of local mode and interarea mode oscillations obtained for fault at bus 9. In Figure 5, it is observed that only  $J_4$  has lower overshoot and very fast settling time than others. For interarea mode of oscillations in Figure 6,  $J_4$  and  $J_5$  have almost the same settling time except high overshoot for  $J_5$ .

Rotor oscillation responses for the fault at bus 13 are obtained as shown in Figures 7, 8, 9, and 10. From Figures 7–10, it is observed that single objective functions  $J_1$  and  $J_2$  have very poor settling times and overshoot than other. Objective function CDI has quite a satisfactory damping performance except in Figure 7. For  $J_4$ , the best settling times are achieved in all modes and the overshoot values are also quite reasonable. Multiobjective function  $J_5$  has damping performance just like single objective functions  $J_1$ ,  $J_2$ . Thus, from Figures 7–10, the time domain simulations are used to

TABLE 2: Best optimized parameters for each objective function using DSA.

Objective functions	$K$	$T_1$	$T_2$	$T_3$	$T_4$	
$J_1$	PSS1	5.1903	1.1099	0.2358	1.6706	1.3299
	PSS2	7.2073	1.1549	1.2313	0.7452	1.5635
	PSS3	6.3461	0.8985	0.5746	0.8453	0.3337
	PSS4	10.7097	0.0906	0.7335	1.4030	1.6083
$J_2$	PSS1	5.0422	0.9573	1.8312	1.6562	0.2857
	PSS2	29.3894	0.2130	1.7013	0.0165	1.9849
	PSS3	21.8602	0.2569	0.1395	1.3798	1.1226
	PSS4	25.9223	0.0634	1.7467	0.9554	0.8587
$J_3$	PSS1	15.0137	1.3111	1.0455	0.4944	0.0597
	PSS2	40.2497	0.2833	1.8740	0.5664	1.9577
	PSS3	31.1375	0.7925	0.9937	0.1303	0.0108
	PSS4	22.8605	0.2659	1.6025	1.0201	1.6150
$J_4$	PSS1	25.5890	1.3868	0.9446	0.0788	0.7359
	PSS2	3.10292	0.7043	0.7753	1.3975	0.1023
	PSS3	12.5965	0.0191	0.5753	1.1611	0.4992
	PSS4	16.5014	0.6357	0.3429	1.0108	0.9893
$J_5$	PSS1	33.794	0.0411	1.9849	0.7937	1.1574
	PSS2	31.8907	0.3248	0.0561	0.3886	1.8845
	PSS3	28.9673	1.1705	0.7381	0.1170	0.0134
	PSS4	35.8091	1.1773	0.8512	0.0769	0.6321

TABLE 3: Comparative improvement in dominant modes for different objective functions.

Objective functions	Dominant modes eigenvalues	Damping ratio ( $\zeta$ )	Frequency ( $\omega$ )
$J_1$	$-0.7180 - 3.8405i$	0.1838	<b>0.6112</b>
	$-2.0831 - 10.7434i$	0.1904	1.7099
	$-3.2354 - 17.3224i$	0.1836	2.7570
$J_2$	$-0.7376 + 4.0867i$	0.1776	<b>0.6504</b>
	$-2.2613 + 12.3786i$	0.1797	1.9701
	$-3.0685 + 17.0267i$	0.1774	2.7099
$J_3$	$-1.6484 - 4.8225i$	0.3234	0.7675
	$-2.6479 + 18.0621i$	<b>0.1451</b>	2.8747
	$-0.7421 + 18.5663i$	<b>0.0399</b>	2.9549
$J_4$	$-2.7588 - 11.0712i$	0.2418	1.7620
	$-4.4346 + 13.9203i$	0.3035	2.2155
	$-2.7934 + 17.2958i$	0.1594	2.7527
$J_5$	$-1.2487 - 3.3262i$	0.3515	<b>0.5294</b>
	$-2.5901 - 6.5999i$	0.3653	<b>1.0504</b>
	$-3.8229 - 16.0199i$	0.2321	2.5496

validate the optimization performance in order to check the system stability for the three-phase fault at bus 13.

## 6. Conclusion

In this research, comparative analyses of different objective functions used for power system oscillation damping controller design have been performed. Damping power

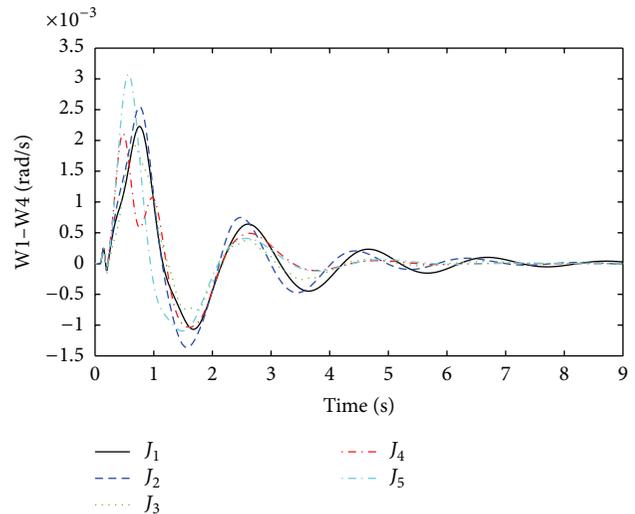


FIGURE 6: Interarea mode of oscillations between G1 and G4 (W1-W4) for three-phase fault at bus 9.

system oscillations by optimal tuning of PSS is approached. Metaheuristic algorithm DSA is kept the same in settings for all objective functions. Simulations for each objective function are conducted 10 times and best results are selected for comparison. The comparisons are performed in terms of dominant eigenvalues (damping factor, damping ratio) and the damping performance in nonlinear time domain simulations. The comparative results show that the multiobjective D-shaped stability function is a more efficient objective function than that of others. The performance of damping

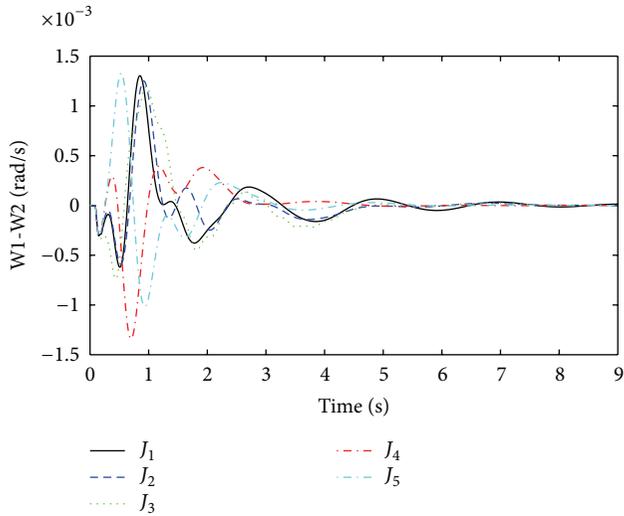


FIGURE 7: Local mode of oscillations between G1 and G2 (W1-W2) for three-phase fault at bus 13.

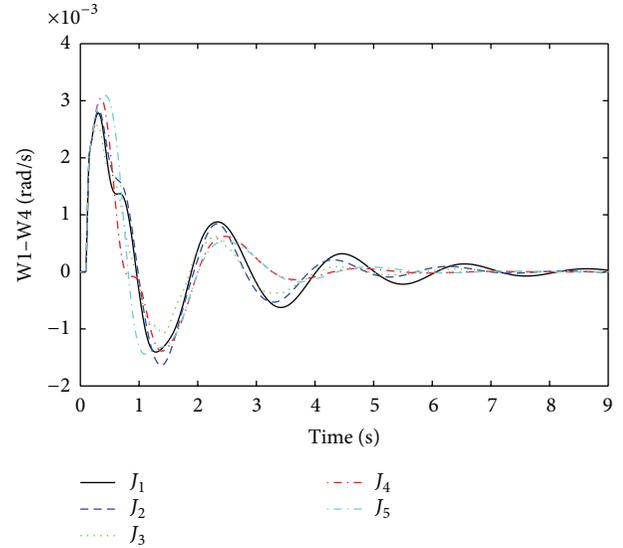


FIGURE 9: Interarea mode of oscillations between G1 and G4 (W1-W4) for three-phase fault at bus 13.

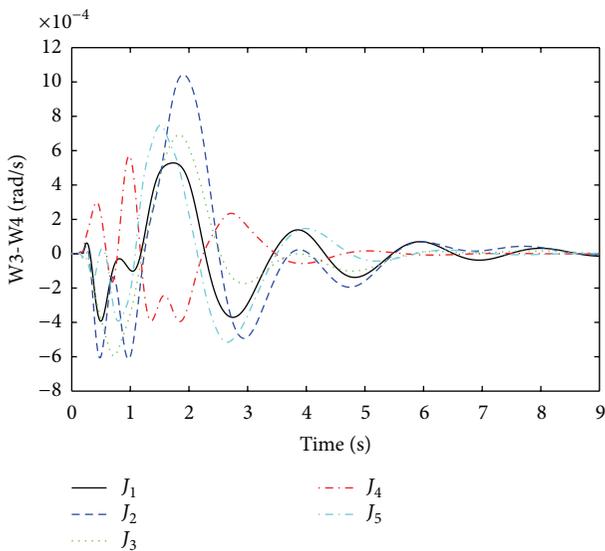


FIGURE 8: Local mode of oscillations between G3 and G4 (W3-W4) for three-phase fault at bus 13.

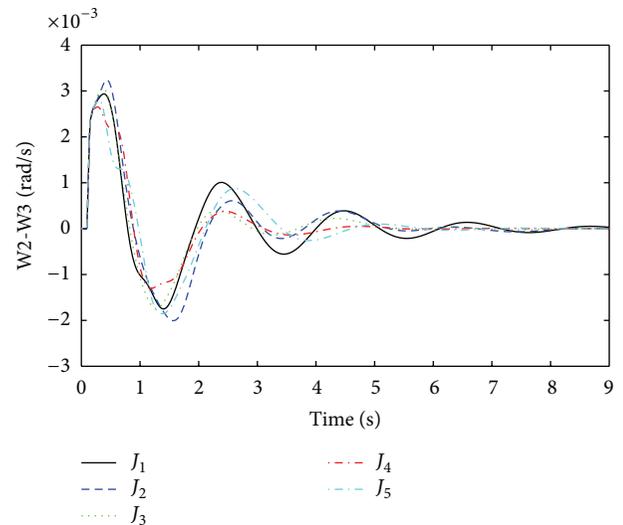


FIGURE 10: Interarea mode of oscillations between G2 and G3 (W2-W3) for three-phase fault at bus 13.

oscillation is further verified in nonlinear time domain simulations. Therefore, problem formulation into D-shaped objective function is recommended in order to design robust damping controller. Further researches should be conducted to develop more efficient objective functions reflecting better performance of damping controllers. Formulation of new and improved objective functions should be challenged based on more extensive comparative analysis with existing uses.

**Appendix**

Simulation settings of DSA: DSA control parameter,  $P1 = P2 = 0.3 \times \text{rand}$ , population size,  $n = 40$ ; problem dimension,  $m = 20$ , and epoch size = 300.

Objective functions settings: expected damping factor,  $\sigma_0 = -1.5$ , expected damping ratio,  $\zeta_0 = 0.15$ .

**Conflict of Interests**

The authors declare that there is no conflict of interests regarding the publication of this paper.

**Acknowledgment**

The authors are grateful to Universiti Kebangsaan Malaysia for supporting this research financially under Grants DIP-2012-03 and 06-01-02-SF1060.

## References

- [1] P. Kundur, *Power System Stability and Control*, McGraw-Hill, 1994.
- [2] G. Rogers, *Power System Oscillations*, Kluwer Academic, 2000.
- [3] P. Kundur, M. Klein, G. J. Rogers, and M. S. Zywno, "Application of power system stabilizers for enhancement of overall system stability," *IEEE Transactions on Power Systems*, vol. 4, no. 2, pp. 614–626, 1989.
- [4] M. A. Hannan and K. W. Chan, "Transient analysis of FACTS and custom power devices using phasor dynamics," *Journal of Applied Sciences*, vol. 6, no. 5, pp. 1074–1081, 2006.
- [5] M. A. Abido, "Parameter optimization of multimachine power system stabilizers using genetic local search," *International Journal of Electrical Power & Energy Systems*, vol. 23, no. 8, pp. 785–794, 2001.
- [6] M. A. Abido, "Novel approach to conventional power system stabilizer design using tabu search," *International Journal of Electrical Power & Energy Systems*, vol. 21, no. 6, pp. 443–454, 1999.
- [7] Y. L. Abdel-Magid, M. A. Abido, and A. H. Mantawy, "Robust tuning of power system stabilizers in multimachine power systems," *IEEE Transactions on Power Systems*, vol. 15, no. 2, pp. 735–740, 2000.
- [8] H. E. Mostafa, M. A. El-Sharkawy, A. A. Emary, and K. Yassin, "Design and allocation of power system stabilizers using the particle swarm optimization technique for an interconnected power system," *International Journal of Electrical Power & Energy Systems*, vol. 34, no. 1, pp. 57–65, 2012.
- [9] M. M. Linda and N. K. Nair, "A new-fangled adaptive mutation breeder genetic optimization of global multi-machine power system stabilizer," *International Journal of Electrical Power & Energy Systems*, vol. 44, no. 1, pp. 249–258, 2013.
- [10] L.-J. Cai and I. Erlich, "Simultaneous coordinated tuning of PSS and FACTS damping controllers in large power systems," *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 294–300, 2005.
- [11] Y. L. Abdel-Magid and M. A. Abido, "Optimal multiobjective design of robust power system stabilizers using genetic algorithms," *IEEE Transactions on Power Systems*, vol. 18, no. 3, pp. 1125–1132, 2003.
- [12] H. Alkhatib and J. Duveau, "Dynamic genetic algorithms for robust design of multimachine power system stabilizers," *International Journal of Electrical Power & Energy Systems*, vol. 45, no. 1, pp. 242–251, 2013.
- [13] M. M. Farsangi, S. Kyanzadeh, S. Haidari, and H. Nezamabadi-pour, "Coordinated control of low-frequency oscillations using real immune algorithm with population management," *Energy Conversion and Management*, vol. 51, no. 2, pp. 271–276, 2010.
- [14] S. M. Abd-Elazim and E. S. Ali, "A hybrid particle swarm optimization and bacterial foraging for optimal power system stabilizers design," *International Journal of Electrical Power & Energy Systems*, vol. 46, pp. 334–341, 2013.
- [15] H. Shayeghi, H. A. Shayanfar, S. Jalilzadeh, and A. Safari, "Multi-machine power system stabilizers design using chaotic optimization algorithm," *Energy Conversion and Management*, vol. 51, no. 7, pp. 1572–1580, 2010.
- [16] H. Yassami, A. Darabi, and S. M. R. Rafiei, "Power system stabilizer design using Strength Pareto multi-objective optimization approach," *Electric Power Systems Research*, vol. 80, no. 7, pp. 838–846, 2010.
- [17] A. Khodabakhshian and R. Hemmati, "Multi-machine power system stabilizer design by using cultural algorithms," *International Journal of Electrical Power & Energy Systems*, vol. 44, no. 1, pp. 571–580, 2013.
- [18] M. Eslami, H. Shareef, M. R. Taha, and M. Khajezadeh, "Adaptive particle swarm optimization for simultaneous design of UPFC damping controllers," *International Journal of Electrical Power & Energy Systems*, vol. 57, pp. 116–128, 2014.
- [19] M. Khaleghi, M. M. Farsangi, H. Nezamabadi-Pour, and K. Y. Lee, "Pareto-optimal design of damping controllers using modified artificial immune algorithm," *IEEE Transactions on Systems, Man and Cybernetics C: Applications and Reviews*, vol. 41, no. 2, pp. 240–250, 2011.
- [20] P. Civicioglu, "Transforming geocentric cartesian coordinates to geodetic coordinates by using differential search algorithm," *Computers & Geosciences*, vol. 46, pp. 229–247, 2012.

## Review Article

# Fault Location Based on Synchronized Measurements: A Comprehensive Survey

A. H. Al-Mohammed<sup>1</sup> and M. A. Abido<sup>2,3</sup>

<sup>1</sup> Saudi Electricity Company, Dammam, Saudi Arabia

<sup>2</sup> Electrical Engineering Department, King Fahd University of Petroleum & Minerals, KFUPM, Dhahran, Saudi Arabia

<sup>3</sup> Electrical Engineering Department, Faculty of Engineering, Menoufia University, Shebin El-Kom, Egypt

Correspondence should be addressed to M. A. Abido; [mabido@kfupm.edu.sa](mailto:mabido@kfupm.edu.sa)

Received 6 December 2013; Accepted 25 December 2013; Published 16 February 2014

Academic Editors: V. N. Dieu, P. Vasant, and G.-W. Weber

Copyright © 2014 A. H. Al-Mohammed and M. A. Abido. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a comprehensive survey on transmission and distribution fault location algorithms that utilize synchronized measurements. Algorithms based on two-end synchronized measurements and fault location algorithms on three-terminal and multiterminal lines are reviewed. Series capacitors equipped with metal oxide varistors (MOVs), when set on a transmission line, create certain problems for line fault locators and, therefore, fault location on series-compensated lines is discussed. The paper reports the work carried out on adaptive fault location algorithms aiming at achieving better fault location accuracy. Work associated with fault location on power system networks, although limited, is also summarized. Additionally, the nonstandard high-frequency-related fault location techniques based on wavelet transform are discussed. Finally, the paper highlights the area for future research.

## 1. Introduction

Transmission and distribution lines are exposed to faults that are caused by different reasons such as short circuits, birds, and storms. Most of these faults result in mechanical damage of power lines which must be repaired before returning the line to service. Power line faults must be located accurately to allow maintenance crews to arrive at the scene and repair the faulted section as soon as possible. Rugged terrain and geographical layout cause some sections of power transmission lines to be difficult to reach. Therefore, robustness of the accurate fault location determination under a variety of power system operating constraints and fault conditions is an important requirement [1, 2].

Generally, fast and accurate fault location will expedite supply restoration and enhance the supply quality and reliability. In addition, this will minimize the customer inconvenience. Therefore, fault location can be considered as one of the first functions to be integrated into modern substation control system [2, 3].

Varieties of fault-location algorithms have been developed and presented in the literature. The majority of them

are based on an impedance principle, making use of the fundamental frequency voltages and currents. Fault-location algorithms based on traveling-wave phenomenon, high-frequency components of currents and voltages generated by faults, and artificial intelligence such as fuzzy neural network [4] have also been developed. Depending on the availability of the fault-locator input signals, fault location algorithms can be categorized as one-end, two-end, and multiend. This survey is concerned with the two-end and multiend algorithms utilizing synchronized measurements. A fault location algorithm based on aforesaid measurements proves to be robust under power swing and out-of-step conditions [5] and has been proposed [6] as part of a strategy that aimed at preventing or mitigating the cascading blackouts that involve relay misoperations or inadequate local diagnostic support.

Accuracy evaluation of various fault location algorithms, as reported in the literature, takes into account different combination of the following main factors [2]:

- (i) fault position (location),
- (ii) fault type,

- (iii) fault resistance,
- (iv) prefault power flow and its direction,
- (v) strength of equivalent sources behind the line terminals,
- (vi) line imbalance due to lack of transposition,
- (vii) inaccurate impedance data of the line,
- (viii) presence and status of series and shunt compensating devices with metal oxide varistors (MOVs),
- (ix) fault inception angle,
- (x) VTs and CTs transient and steady errors.

In this survey, the issue of fault location utilizing two-end synchronized measurements is addressed in Section 2. Fault location on three-terminal and multiterminal lines is then presented in Section 3. A special attention shall be paid while developing a fault location algorithm for a series-compensated transmission line in order to locate any fault accurately. Therefore, Section 4 is devoted to discuss this matter. In order to overcome the limitations of the conventional fault location algorithms, some work has been carried out to locate power system faults in an adaptive manner and this subject is reported in Section 5. In comparison with the work pertaining to fault location on transmission lines, there is a quite limited published work related to fault location on power system networks as highlighted in Section 6. A short review of the nonstandard high-frequency-related fault location techniques based on wavelet transform is given in Section 7. Area for future research is highlighted in Section 8.

## 2. Fault Location Utilizing Two-End Synchronized Measurements

At present, phasor measurement units (PMUs) have come out of their academic infancy with commercial viability. They represent a revolution in power systems monitoring and control. A PMU can measure current and voltage and calculate the phase angle. Therefore, real-time calculation of phase angles around the system can be achieved. This can be attributed to time stamping and synchronization that are not available with traditional measurement. With the satellite GPS (global positioning system) availability, digital measurements at different line terminals can be performed synchronously. A synchronized measurement system includes the phase-angle data in addition to the magnitudes. The phase angles are measured with respect to an arbitrary but common reference. Having been known the absolute time of the measurements, phase information can be obtained. The time for all measurements must be synchronized with a time reference that must be the same for all local systems. This time reference is obtained from the GPS [7].

The potential uses of the subsecond GPS-synchronized phasor data collected from various locations within an electric power system promise endless benefits for the applications targeting reliable operation of electric power system [8]. Various PMU applications in power systems including fault location [9–15] have been reported. Merging of time correlated information from PMU, SCADA, and nonoperational

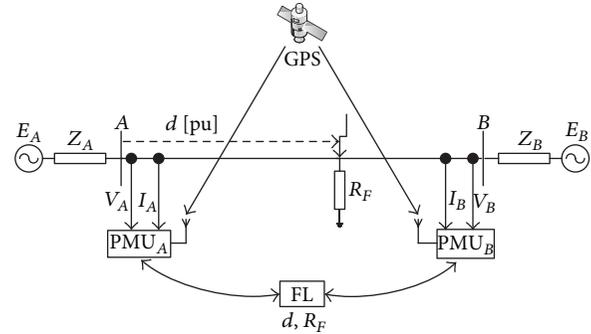


FIGURE 1: A schematic diagram for two-end synchronized fault-location arrangement using PMUs.

data has also been suggested [16] to improve the effectiveness of alarm processing, accuracy of fault location, and ability to detect cascades.

Two-end fault location algorithms have been proposed with the aim of overcoming the limitations of the one-end fault location techniques and improving fault-location accuracy. A schematic diagram for two-end synchronized fault-location arrangement using PMUs is shown in Figure 1 [2].

Fault location algorithms based on two-end synchronized measurements have been developed using either complete or incomplete two-end measurements. With the use of complete two-end measurements the three-phase voltages and currents are utilized. In case of incomplete two-end measurements, the following options are of interest [2]:

- (i) three-phase voltages from both ends with three-phase current from only one end,
- (ii) three-phase currents from both ends with three-phase voltage from only one end,
- (iii) three-phase voltages from both ends.

Various fault location algorithms utilizing complete two-end synchronized measurements [1, 17–39] have been developed. Both iterative [17] and noniterative [24, 26] methods have been proposed to locate faults on a single transmission line. In the noniterative method, an analytical synchronization of the unsynchronized voltage and current measurements from two ends of a line is performed with use of the determined synchronization operator. Then, the synchronized measurements are used to calculate the distance to fault. Simultaneous usage of two kinds of symmetrical components for determining the synchronization operator makes the calculations simple and highly accurate.

Fault location taking into account the arc faults has been addressed in [21, 25, 29–32, 35]. The proposed fault detection/location technique [21, 35] for both arcing and permanent faults is achieved by a combination of a fault detection index and a fault location index which are obtained by processing synchronized fundamental phasors. In order to discriminate between arcing and permanent faults, the proposed technique estimates the amplitude of arc voltage by least error squares method through the measured synchronized harmonic phasors caused by the nonlinear arc behavior.

The discrimination is then achieved by comparing the estimated amplitude of arc voltage to a given threshold value. In order to eliminate the error caused by exponentially decaying dc offset on the computations of fundamental and harmonic phasors, an extended discrete Fourier transform algorithm is also proposed. In [25, 29, 30, 32], the electric arc is considered as a source of higher harmonics and is included in the complete fault model accordingly. The developed algorithm can determine both the arc and the fault resistance. In [31], the fault location algorithm is derived in the time domain. The faulted phase voltage is modeled as a serial connection of fault resistance and arc voltage. The algorithm does not require the line zero sequence resistance as an input datum. The influences of remote infeed, fault resistance, higher order harmonics, and network topology are investigated.

For short and medium-length lines using the lumped model is usually sufficient. In order to improve fault-location accuracy, especially in the case of long-length lines, the distributed nature of overhead-line parameters has to be considered [2]. Algorithms presented in [20, 33, 37, 38] take such requirement into account by representing long lines with distributed parameters where shunt capacitance is included in the line model.

Transmission systems may sometimes consist of an overhead line in combination with an underground power cable. The fault location scheme developed in [22, 36] for such systems requires synchronized phasor measurement data at one end of the transmission line and the most far end of the power cable. The algorithm is derived using distributed line model, modal transformation theory, and discrete Fourier transform. In [39], a fault location technique for two-terminal multisection compound transmission line is presented.

The fault location algorithm presented in [34] takes the three-phase unbalance into consideration. The algorithm models the line with its distributed parameters and uses the theory of mode transformation. In [40], a fault location algorithm in joint parallel lines is proposed using six-sequence fault components in fault location. Although the algorithm is not influenced by factors such as the load current, the operating mode of the power system, or the fault resistance, the associated percentage error can reach up to 2%. A fault detection/location algorithm on transmission line based on linear state estimation is presented in [41] where fault location and voltage of fault point are added as the new state variables in a linear state estimator based on PMU data.

Fault location algorithms based on incomplete two-end synchronized measurements utilizing only three-phase voltages from both line ends [42–45] have also been developed. Algorithm proposed in [43] can be applied for both transposed and untransposed lines and algorithms presented in [44, 45] suit single or double transmission lines. However, algorithm suggested in [44] utilizes, in addition to the synchronized voltage measurements at both ends of the faulted line, the synchronized voltages at neighboring nodes. Although the method is highly accurate, the number of utilized PMUs is not optimal. Fault location algorithms utilizing only synchronized measurements of two-end voltages have

the advantage of being immune to CT saturation as they completely reject the currents from the input signals.

### 3. Fault Location on Three-Terminal and Multiterminal Lines

Generally, multiterminal lines are those having three or more terminals with substantial generation behind each. Similarly, tapped lines are those having three or more terminals with substantial power generation behind at maximum two of them. Multiterminal and tapped lines are used for economical or environmental reasons. The taps feed only loads, that is, passive networks, while at the remaining terminals they are terminated by active networks [2].

Various fault location algorithms on three-terminal lines are presented in [46–52]. In [46], a fault location method is developed based on synchronized measurements of three-phase current from all three terminals and additionally three-phase voltage from the terminal at which a fault locator is installed. The delivered fault-location algorithm consists of three subroutines designated for locating faults within particular line sections and a procedure for indicating the faulted line section. An approach for fault location on EHV teed feeders [47] utilizes synchronized voltages and currents at all three ends of a teed feeder. Measurements are then digitally filtered to accurately extract the power frequency phasors. In this approach, use is made of superimposed modal components of signals so as to minimize errors arising in accuracy due to line loading or source impedances. Algorithms presented in [48, 49] use synchronized voltage and current data from two terminals only. They are not influenced by fault resistance, fault location, prefault loading conditions, source impedance, and fault types.

Algorithms discussed in [53–59] are related to multiterminal lines. The iterative method presented in [53] uses synchronized voltage and current measurements from all terminals. Current measurements, however, were avoided in [54] to overcome current-transformer errors in the current measurements that can be as high as 10%. In [56], a universal noniterative fault location technique for N-terminal transmission lines based on two-terminal fault location technique is presented. The method discussed in [57] is also noniterative and it is based on distributed line model and synchronized positive sequence voltage and current phasors. In [55], a fault location algorithm for transmission line with tapped legs is developed. The algorithm only uses the synchronized phasors measured on two terminals of the original line to calculate the fault location. The algorithm does not need the model of tapped leg and, therefore, can be applied to any type of tapped leg such as generators, loads, or combined system.

### 4. Fault Location on Series-Compensated Lines

The one-line diagrams of a series-compensated transmission line with series capacitors (SCs) and MOVs installed at midpoint and at both ends of the line are shown, respectively, in Figures 2 and 3. MOVs are installed to protect SCs against overvoltages.

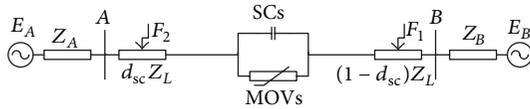


FIGURE 2: Single transmission line compensated with SCs and MOVs installed at midpoint.

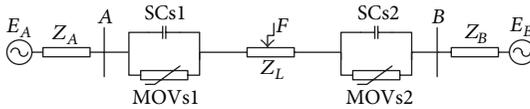


FIGURE 3: Transmission line compensated with SCs and MOVs installed at both ends.

Various fault location algorithms on series-compensated lines are presented in [60–65]. A fault location algorithm, presented in [60–62], does not need the series device model and information of the protection function of series device to predict the voltage drop. Instead, two iteration steps, prelocation step, and correction step are used to calculate the voltage drop and fault location. The algorithm can be applied to any series FACTS system with a very high accuracy.

Use of instantaneous values for fault location of series-compensated transmission lines while avoiding the accuracy limitation caused by the operation of MOV is discussed in [63]. The method requires only a short duration of fault measurement data to estimate the location and can be applied with minimum filtering of high frequencies. It is independent of the fault type and does not require the fault to be pure resistive. However, it requires knowledge of the source impedance at both ends of the line.

In [64], a fault location algorithm based on distributed time domain line model for a transmission line with a FACTS device connected in series is presented. The algorithm can be applied to any series FACTS compensated line since the series device model and knowledge of the operating mode of the compensating device are not utilized to compute the voltage drop across the series device during the fault period. Filtering of DC and other frequency components is not needed. The algorithm is not sensitive to fault resistance and fault inception angle and does not require knowledge of source impedance.

## 5. Adaptive Fault Location

The prime aim of adaptive fault location algorithms is to achieve a better fault location accuracy. The idea of adaptive fault location on transmission lines boils down to proper estimation of line parameters and system impedance. The environmental conditions and operation history of the transmission line affect its sag. As the conductor current increases, its temperature increases and, consequently, its sag. The line resistance changes with the line temperature. In addition, the line reactance will change since it depends on the distance between the phase conductors which is affected by the line sag. Therefore, the uncertainty of the line parameters could

affect substantially the accuracy of the fault location. The effect of such uncertainty can reach up to 6-7% if the parameters vary 20% of the practical parameters. It is worth mentioning that the power utility usually provides very ideal parameters of the line that do not take into consideration its operation history [66].

With normal operation, PMU can measure the voltage at both ends of the line along with its current. The line parameters can be calculated using the pre-fault phase and amplitude measurements of the voltages and currents. Four methods are presented in [67] to identify transmission line impedance parameters from synchronized measurements.

Algorithms discussed in [66, 68–81] are related to adaptive fault location on transmission lines. All adaptive fault location algorithms presented in the above-mentioned references utilize synchronized voltage and current measurements at both ends of a transmission line. Adaptive fault location for aged power cables is presented in [70, 79]. The algorithm is incorporated with distributed line model, modal transformation theory, and discrete Fourier transform. It solves the problem of cable changing parameters especially the change of the relative permittivity over its age and, thus, for the operating positive, negative, and zero-sequence capacitance changes.

Algorithm presented in [71] utilizes synchronized measurements for online estimation of line parameters. A fault location index in terms of Clarke components of the synchronized voltage and current phasors is proposed to calculate the fault location. Also, a discrete Fourier transform-based algorithm is proposed to eliminate system noise and measurement errors. This work has been extended in [76, 77] by adding a fault detection index to the algorithm. A similar adaptive relaying scheme has been developed in [74, 75].

An adaptive fault protection scheme for transmission lines is discussed in [68, 69]. The work includes fault detection, direction discrimination, classification, and location. Both fault detection and fault location indices are derived by using two-terminal synchronized measurements incorporated with distributed line model and modal transformation theory. The fault detection index is composed of two complex phasors and the angle difference between the two phasors determines whether the fault is internal or external to the protected zone. The fault types can be classified by the modal fault detection index. The proposed scheme also combines online parameter estimation to assure protection scheme performance and to achieve adaptive protection. Simulation studies show that fault location accuracy is high under various system and fault conditions.

An adaptive protection scheme is presented in [72, 73] for both transposed and untransposed parallel transmission lines based on the distributed line model. The fault detection and location indices are derived using the eigenvalue/eigenvector theory to decouple the mutual coupling effects between parallel lines. The two proposed indices are used in coordination such that the internal and external fault events can be distinguished. By online estimating of line parameters under the actual system conditions, the proposed scheme responds more accurately to power system faults.

In [80], an adaptive fault location algorithm for transmission line tapped with a source of generation using the concept of superimposed voltage and current phasors is discussed. A discrimination index is proposed to identify the faulted section while taking the effects caused by tapped lines into account. The equivalent source impedance outside the considered transmission lines is estimated online.

Adaptive fault location for single, double, and teed transmission lines is addressed in [66]. In the proposed algorithm, line parameters are calculated online. In addition, suddenly changed voltage and current are utilized to obtain suddenly changed positive voltage and current components to solve the system's impedance at the fault time that exactly reflects the generation mode of the power system. The subject of adaptive fault location for double-circuit transmission lines has also been discussed in [81] where a six-sequence fault component method is employed to implement fault location. Line parameters are estimated online and the line is represented with its distributed parameters.

### 6. Fault Location in Power System Networks

Fault location in distribution networks creates new problems compared with the same task in HV and EHV transmission lines. In HV and EHV networks each transmission line may be equipped with a dedicated fault locator (FL). In such a case, the FL algorithm is a numerical procedure that converts voltage and current into a single number being a distance to fault. The distribution networks, in contrast to the transmission lines, are usually nonhomogeneous with branches and loads along the line which make the fault location difficult [2].

Various fault location algorithms in power system networks are presented in [82–99]. A method for location of single phase to ground faults in distribution network based on wide-area synchronizing information is discussed in [82]. Network lines are modeled with their distributed parameters and a fault location function is constructed according to the relations between fault current and fault distance. The transverse fault current corresponding to the hypothetical fault distance is obtained using synchronized voltage and currents from two ends of a line. The maximum transverse fault current is obtained using the interactive search algorithm on the whole line and the distance corresponding to the maximum of the transverse fault current is the fault distance.

In [84], a fault location algorithm for urban distribution system with distributed generation is presented. The method uses only currents and thus avoids the need for installation of voltage transformers. A differential current ratio (DCR) contains the differential current information to describe the feature of the faulted segment and differential current information of normal one. The faulted segment would have a DCR of a value less than one while the normal one would have a DCR greater than one. The largest fault current would be sensed by one FTU in the faulted segment.

Fault-location observability with minimum installation of PMUs in a power system network is an optimization problem that has been solved by different methods such as GA

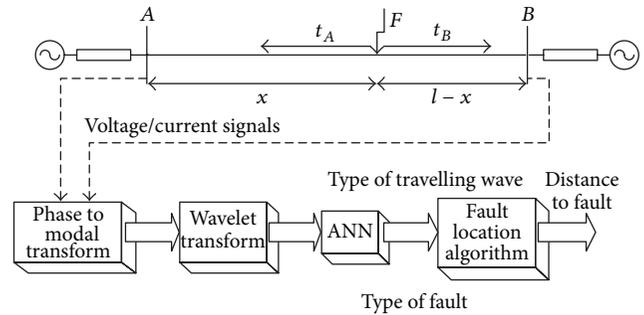


FIGURE 4: Illustration of a distorted voltage analysis using wavelet transform.

[88], branch and bound [89], and tabu search [94]. The aim here is to utilize synchronized measurements of optimally installed PMUs in a suitable algorithm to locate faults that occur anywhere in the network.

Some attempts to use discriminant analysis theory of multivariate statistical analysis theory [90], cluster analysis theory [91], power flow fingerprint [92], depth first search of graph theory [93], neural networks [86, 96], structural analysis [95], and principal component analysis (PCA) [97] have also been reported for fault location in power system networks. Fault location methods for large transmission networks are proposed in [98, 99] utilizing PMU voltage measurements.

### 7. Fault Location Using Wavelet Transform

The idea of the wavelet transform- (WT-) based fault location method is illustrated in Figure 4. Measurement of arriving times of waves at the terminal buses, together with knowing the velocity of these waves, allows determination of a distance to fault. Thus, the problem in fact boils down to proper detection of arriving waves, to distinguish the type of these waves (direct waves or reflected waves) and to capture times of their arrival at terminal buses. All these tasks can be directly realized by using wavelet transforms (time location of wave) combined with ANN (distinguishing kinds of wave) [2].

Various fault location algorithms using wavelet transform are presented in [85, 100–105]. A fault-location scheme for multiend aged cable system utilizing synchronized measurements at the two terminals of each cable is presented in [101]. The developed scheme is applied on the modal coordinates instead of the phase coordinates. It can eliminate the impact of the change in the propagation velocity of the travelling waves on the fault-location calculations. This solves the problem of cable changing parameters especially the change of the relative permittivity of the cable with age.

The application of discrete wavelet transform (DWT) to fault location on power distribution lines with tapped loads is investigated in [102]. When a fault occurs, a transient wave travels from the fault point to substation busbars and load terminals. The time taken for the fault generated transient wave to arrive at busbars or load terminals depends on

the distance traveled and the velocity of the travelling wave. Fault transient detectors are installed at substation busbars and load terminals to capture the time taken for the transient to arrive and be synchronized with a GPS clock. From the recorded time and the topological structure of the network, fault location is deduced.

Fault location for teed circuits with mutually coupled lines and series capacitors is proposed in [104], based on DWT, by processing of travelling waves to extract the arrival times of fault initiated waves reflected from the discontinuities. The method is not influenced by series compensation rate and its location, fault resistance, fault type, and any existing mutual coupling between the lines.

Fault location for three-terminal transmission lines is addressed in [103, 105] where WT is used to analyze high and low frequency components generated by a fault. The proposed fault location scheme is based on travelling waves and on calculation of the fundamental frequency component. The method is independent of fault impedance, fault type, fault inception angle, fault position, and mutual coupling effects.

## 8. Future Work

In case two-end synchronized measurements are utilized for fault location, it is possible to formulate more equations than the number of unknowns. The resulting redundancy may be explored to obtain certain improvements in fault location. Also, values of line parameters may not be always available in reality and, therefore, fault location techniques not requiring line parameters may be developed. Moreover, the work of designing a fault location approach capable of detecting, identifying, and removing the possible bad measurements seems interesting as it helps achieve better fault-location accuracy. Finally, application of expert systems for fault location in power systems may also be an area for future research.

## 9. Conclusion

This paper presented a comprehensive survey on fault location algorithms utilizing synchronized measurements. Algorithms based on two-end synchronized measurements have been reported and discussed. Published work related to fault location algorithms on three- and multiterminal lines has been summarized. Additionally, the paper presented the fault location algorithms on series-compensated transmission lines. Some light was shed on adaptive fault location algorithms. Fault location in power system networks and fault location based on wavelet transform have also been highlighted and discussed. Area for future research has been suggested.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

The authors acknowledge the support of Deanship of Scientific Research, King Fahd University of Petroleum and Minerals, through the Electrical Power and Energy Systems Research Group funded Project no. RG1007-1&2.

## References

- [1] M. Kezunovic and B. Perunicic, "Synchronized sampling improves fault location," *IEEE Computer Applications in Power*, vol. 8, no. 2, pp. 30–33, 1995.
- [2] M. Saha, J. Izykowski, and E. Rosolowski, *Fault Location on Power Network*, Springer, New York, NY, USA, 2010.
- [3] V. Madani, "Western interconnection experience with phasor measurements," in *Proceedings of the IEEE PES Power Systems Conference and Exposition (PSCE '06)*, pp. 343–352, November 2006.
- [4] C.-W. Liu, M.-C. Su, and S.-S. Tsay, "Application of a novel fuzzy neural network to real-time transient stability swings prediction based on synchronized phasor measurements," *IEEE Transactions on Power Systems*, vol. 14, no. 2, pp. 685–692, 1999.
- [5] N. Zhang and M. Kezunovic, "A study of synchronized sampling based fault location algorithm performance under power swing and out-of-step conditions," in *Proceedings of the IEEE Russia Power Tech*, St. Petersburg, Russia, June 2005.
- [6] N. Zhang and M. Kezunovic, "Improving real-time fault analysis and validating relay operations to prevent or mitigate cascading blackouts," in *Proceedings of the IEEE PES Transmission and Distribution Conference and Exhibition*, pp. 847–852, May 2006.
- [7] J. Altman, *A practical comprehensive approach to PMU placement for full observability [M.S. thesis]*, Faculty of the Virginia Polytechnic Institute and State University, Blacksburg, Va, USA, 2007.
- [8] Y. Makarov, C. Miller, T. Nguen, and M. Jian, "Characteristic ellipsoid method for monitoring power system dynamic behavior using phasor measurements," in *Proceedings of the Bulk Power System Dynamics and Control-VII, Revitalizing Operational Reliability, iREP Symposium*, pp. 1–5, August 2007.
- [9] M. A. El-Hadidy, D. H. Helmi, H. G. Negm, and H. M. El-Shaer, "Starting synchrophasor measurements in Egypt: a pilot project using fault recorders," in *Proceedings of the 12th International Middle East Power System Conference (MEPCON '08)*, pp. 157–161, March 2008.
- [10] K. E. Martin, "Precise timing in electric power systems," in *Proceedings of the IEEE Annual International Frequency Control Symposium*, pp. 15–22, June 1993.
- [11] R. O. J. Burnett, M. M. Butts, and P. S. Sterlina, "Power system applications for phasor measurement units," *IEEE Computer Applications in Power*, vol. 7, no. 1, pp. 8–13, 1994.
- [12] R. E. Wilson, "PMUs," *IEEE Potentials*, vol. 13, no. 2, pp. 26–28, 1994.
- [13] J. Tlustý, A. Kasembe, Z. Muller et al., "The monitoring of power system events on transmission and distribution level by the use of phasor measurement units (PMU)," in *Proceedings of the 20th International Conference and Exhibition on Electricity Distribution (CIRED '09)*, pp. 1–4, June 2009.
- [14] S. López, J. Gómez, R. Cimadevilla, and O. Bolado, "Synchrophasor applications of the national electric system operator of Spain," in *Proceedings of the 61st Annual Conference for Protective Relay Engineers*, pp. 436–456, April 2008.

- [15] A. G. Phadke and J. S. Thorp, *Synchronized Phasor Measurements and Their Applications*, Springer, New York, NY, USA, 2008.
- [16] M. Kezunovic, C. Zheng, and C. Pang, "Merging PMU, operational, and non-operational data for interpreting alarms, locating faults and preventing cascades," in *Proceedings of the 43rd Annual Hawaii International Conference on System Sciences (HICSS '10)*, pp. 1–4, January 2010.
- [17] S. M. Brahma, "New fault location scheme for a two-terminal transmission line using synchronized phasor measurements," in *Proceedings of the IEEE PES Transmission and Distribution Conference and Exhibition*, pp. 853–857, May 2006.
- [18] S. Samantaray, L. Tripathy, and P. Dash, "Differential equation-based fault locator for unified power flow controller-based transmission line using synchronised phasor measurements," *IET Generation, Transmission & Distribution*, vol. 3, no. 1, pp. 86–98, 2009.
- [19] S. El Safty, M. M. A. El Nasr, S. F. Mekhemer, and M. M. Mansour, "New technique for fault location in interconnected networks using phasor measurement unit," in *Proceedings of the 12th International Middle East Power System Conference (MEP-CON '08)*, pp. 6–10, March 2008.
- [20] C. Zheng, L. Chengmu, S. Jinxi, and W. Xinrong, "A fault location algorithm for transmission line based on distributed parameter," in *Proceedings of the IEE International Conference on Developments in Power System Protection*, pp. 411–413, 2001.
- [21] Y.-H. Lin, C.-W. Liu, and C.-S. Chen, "A new PMU-based fault detection/location technique for transmission lines with consideration of arcing fault discrimination—part I: theory and algorithms," *IEEE Transactions on Power Delivery*, vol. 19, no. 4, pp. 1587–1593, 2004.
- [22] M. Gilany, E. S. T. El Din, M. M. Abdel Aziz, and D. K. Ibrahim, "An accurate scheme for fault location in combined overhead line with underground power cable," in *Proceedings of the IEEE Power Engineering Society General Meeting*, vol. 3, pp. 2521–2527, June 2005.
- [23] M. Kezunovic and B. Perunicic, "Automated transmission line fault analysis using synchronized sampling at two ends," *IEEE Transactions on Power Systems*, vol. 11, no. 1, pp. 441–447, 1996.
- [24] J. Izykowski, E. Rosolowski, P. Balcerek, M. Fulczyk, and M. M. Saha, "Accurate noniterative fault location algorithm utilizing two-end unsynchronized measurements," *IEEE Transactions on Power Delivery*, vol. 25, no. 1, pp. 72–80, 2010.
- [25] Z. Radojević and V. Terzija, "Two terminals numerical algorithm for fault distance calculation and fault analysis," in *Proceedings of the IEEE PES Power Systems Conference and Exposition (PSCE '06)*, pp. 1037–1042, November 2006.
- [26] J. Izykowski, E. Rosolowski, P. Balcerek, M. Fulczyk, and M. M. Saha, "Accurate noniterative fault location algorithm utilizing two-end unsynchronized measurements," *IEEE Transactions on Power Delivery*, vol. 25, no. 1, pp. 72–80, 2010.
- [27] M. Kezunovic and B. Perunicic, "Automated transmission line fault analysis using synchronized sampling at two ends," in *Proceedings of the IEEE Power Industry Computer Application Conference*, pp. 407–413, May 1995.
- [28] A. Gopalakrishnan, M. Kezunovic, S. M. McKenna, and D. M. Hamai, "Fault location using the distributed parameter transmission line model," *IEEE Transactions on Power Delivery*, vol. 15, no. 4, pp. 1169–1174, 2000.
- [29] Z. Radojević and V. Terzija, "Numerical algorithm for overhead lines protection and disturbance records analysis," *IET Generation, Transmission and Distribution*, vol. 1, no. 2, pp. 357–363, 2007.
- [30] V. Terzija, N. I. Elkalashy, G. Preston, V. Stanojević, and G. Štrbac, "Detection of arcing faults: modelling, simulation, testing and algorithms aspects," in *Proceedings of the IEEE Lausanne Power Tech*, pp. 1147–1152, July 2007.
- [31] M. B. Djurić, V. V. Terzija, and Z. M. Radojević, "Overhead lines fault location and arc voltage estimation numerical algorithm derived in time domain," *Electrical Engineering*, vol. 81, no. 1, pp. 45–53, 1998.
- [32] C. J. Lee, J. B. Park, J. R. Shin, and Z. M. Radojević, "A new two-terminal numerical algorithm for fault location, distance protection, and arcing fault recognition," *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1460–1462, 2006.
- [33] H. H. Kim, C. J. Lee, Z. M. Radojević, J. B. Park, and J. R. Shin, "An improved two-terminal numerical algorithm for fault location estimation considering shunt capacitance," in *Proceedings of the IEEE Power Engineering Society General Meeting (PES '06)*, June 2006.
- [34] S. Urano, T. Yamada, Y. Ooura, Y. Xu, and Y. Yamaguchi, "Development of the high accuracy impedance type fault locator using a mode transformation," in *Proceedings of the IEEE/PES Transmission and Distribution Conference and Exhibition: Asia and Pacific*, pp. 1–6, August 2005.
- [35] Y.-H. Lin, C.-W. Liu, and C.-S. Chen, "A new PMU-based fault detection/location technique for transmission lines with consideration of arcing fault discrimination—part II: performance evaluation," *IEEE Transactions on Power Delivery*, vol. 19, no. 4, pp. 1594–1601, 2004.
- [36] E. Sayed Tag El Din, M. M. Abdel Aziz, D. K. Ibrahim, and M. Gilany, "Fault location scheme for combined overhead line with underground power cable," *Electric Power Systems Research*, vol. 76, no. 11, pp. 928–935, 2006.
- [37] C.-J. Lee, Z. Radojević, H.-H. Kim, J.-B. Park, and J.-R. Shin, "A new numerical algorithm for fault location estimation using two-terminal synchronized voltage and current phasors," in *Proceedings of the 5th IFAC Symposium on Power Plants and Power Systems Control (PPS '06)*, pp. 131–136, June 2006.
- [38] N. Zhang and M. Kezunovic, "Complete fault analysis for long transmission line using synchronized sampling," in *Proceedings of the 5th IFAC Symposium on Power Plants and Power Systems Control (PPS '06)*, pp. 137–142, June 2006.
- [39] C.-W. Liu, T.-C. Lin, C.-S. Yu, and J.-Z. Yang, "A fault location technique for two-terminal multisection compound transmission lines using synchronized phasor measurements," *IEEE Transactions on Smart Grid*, vol. 3, no. 1, pp. 113–121, 2012.
- [40] C. Fan, H. Cai, and W. Yu, "Application of six-sequence fault components in fault location for joint parallel transmission line," *Tsinghua Science and Technology*, vol. 10, no. 2, pp. 247–253, 2005.
- [41] M. Shiroei, S. Daniar, and M. Akhbari, "A new algorithm for fault location on transmission lines," in *Proceedings of the IEEE Power and Energy Society General Meeting (PES '09)*, July 2009.
- [42] K. G. Firozjahan and A. Sheikholeslami, "A current independent synchronized phasor measurement based method for fault location on transmission lines," in *Proceedings of the International Conference on Electrical Engineering (ICEE '07)*, April 2007.
- [43] S. M. Brahma and A. A. Girgis, "Fault location on a transmission line using synchronized voltage measurements," *IEEE Transactions on Power Delivery*, vol. 19, no. 4, pp. 1619–1622, 2004.

- [44] C. Wang, C.-X. Dou, X.-B. Li, and Q.-Q. Jia, "A WAMS/PMU-based fault location technique," *Electric Power Systems Research*, vol. 77, no. 8, pp. 936–945, 2007.
- [45] C. Wang, Q.-Q. Jia, X.-B. Li, and C.-X. Dou, "Fault location using synchronized sequence measurements," *International Journal of Electrical Power and Energy Systems*, vol. 30, no. 2, pp. 134–139, 2008.
- [46] J. Izykowski, E. Rosolowski, M. M. Saha, M. Fulczyk, and P. Balcerek, "A fault-location method for application with current differential relays of three-terminal lines," *IEEE Transactions on Power Delivery*, vol. 22, no. 4, pp. 2099–2107, 2007.
- [47] R. K. Aggarwal, D. V. Coury, A. T. Johns, and A. Kalam, "Computer aided design and testing of an accurate fault locator for EHV teed feeders," in *Proceedings of the 5th International Conference on Developments in Power System Protection*, pp. 60–64, April 1993.
- [48] Y.-H. Lin, C.-W. Liu, and C.-S. Yu, "A new fault locator for three-terminal transmission lines—using two-terminal synchronized voltage and current phasors," *IEEE Transactions on Power Delivery*, vol. 17, no. 2, pp. 452–459, 2002.
- [49] J. Xin and X. Lin, "A novel method to improve the location accuracy of a category of fault locators," in *Proceedings of the Asia-Pacific Power and Energy Engineering Conference (APPEEC '11)*, pp. 1–4, 2011.
- [50] K. G. Firouzjah and A. Sheikholeslami, "A current independent method based on synchronized voltage measurement for fault location on transmission lines," *Simulation Modelling Practice and Theory*, vol. 17, no. 4, pp. 692–707, 2009.
- [51] T. Funabashi, H. Otoguro, Y. Mizuma et al., "Digital fault location for high resistance grounded transmission lines," *IEEE Transactions on Power Delivery*, vol. 14, no. 1, pp. 80–85, 1999.
- [52] A. Esmailian, M. Mohseninezhad, M. Doostizadeh, and M. Khanabadi, "A precise PMU based fault location method for multi terminal transmission line using voltage and current measurement," in *Proceedings of the 10th International Conference on Environment and Electrical Engineering*, pp. 1–4, May 2011.
- [53] S. M. Brahma, "New fault-location method for a single multi-terminal transmission line using synchronized phasor measurements," *IEEE Transactions on Power Delivery*, vol. 21, no. 3, pp. 1148–1153, 2006.
- [54] S. M. Brahma, "Fault location scheme for a multi-terminal transmission line using synchronized voltage measurements," *IEEE Transactions on Power Delivery*, vol. 20, no. 2, pp. 1325–1331, 2005.
- [55] C.-S. Yu, C.-W. Liu, and Y.-H. Lin, "A fault location algorithm for transmission lines with tapped leg—PMU based approach," in *Proceedings of the IEEE Power Engineering Society Summer Meeting*, vol. 2, pp. 915–920, July 2001.
- [56] C.-W. Liu, K.-P. Lien, C.-S. Chen, and J.-A. Jiang, "A universal fault location technique for N-terminal ( $N \geq 3$ ) transmission lines," *IEEE Transactions on Power Delivery*, vol. 23, no. 3, pp. 1366–1373, 2008.
- [57] K.-P. Lien, C.-W. Liu, J.-A. Jiang, C.-S. Chen, and C.-S. Yu, "A novel fault location algorithm for multi-terminal lines using phasor measurement units," in *Proceedings of the 37th Annual North American Power Symposium*, pp. 576–581, October 2005.
- [58] M. Abe, N. Otsuzuki, T. Emura, and M. Takeuchi, "Development of a new fault location system for multi-terminal single transmission lines," in *Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference*, pp. 259–268, April 1994.
- [59] M. Abe, N. Otsuzuki, T. Emura, and M. Takeuchi, "Development of a new fault location system for multi-terminal single transmission lines," *IEEE Transactions on Power Delivery*, vol. 10, no. 1, pp. 159–168, 1995.
- [60] C.-S. Yu, C.-W. Liu, and J.-A. Jiang, "A new fault location algorithm for series compensated lines using synchronized phasor measurements," in *Proceedings of the Power Engineering Society Summer Meeting*, pp. 1350–1354, July 2000.
- [61] C. Yu, C. Liu, S. Yu, and J. Jiang, "A new PMU-based fault location algorithm for series compensated lines," *IEEE Power Engineering Review*, vol. 21, no. 11, pp. 58–58, 2001.
- [62] C.-S. Yu, C.-W. Liu, S.-L. Yu, and J.-A. Jiang, "A new PMU-based fault location algorithm for series compensated lines," *IEEE Transactions on Power Delivery*, vol. 17, no. 1, pp. 33–46, 2002.
- [63] M. Al-Dabbagh and S. K. Kapuduwage, "Using instantaneous values for estimating fault locations on series compensated transmission lines," *Electric Power Systems Research*, vol. 76, no. 1–3, pp. 25–32, 2005.
- [64] J. Sadeh and A. Adinehzadeh, "Accurate fault location algorithm for transmission line in the presence of series connected FACTS devices," *International Journal of Electrical Power and Energy Systems*, vol. 32, no. 4, pp. 323–328, 2010.
- [65] J.-C. Gu, K.-Y. Shen, S.-L. Yu, and C.-S. Yu, "Removal of dc offset and subsynchronous resonance in current signals for series compensated transmission lines using a novel Fourier filter algorithm," *Electric Power Systems Research*, vol. 76, no. 5, pp. 327–335, 2006.
- [66] C. Fan, X. Du, S. Li, and W. Yu, "An adaptive fault location technique based on PMU for transmission line," in *Proceedings of the IEEE Power Engineering Society General Meeting (PES '07)*, pp. 1–6, June 2007.
- [67] D. Shi, D. J. Tylavsky, N. Logic, and K. M. Koellner, "Identification of short transmission-line parameters from synchrophasor measurements," in *Proceedings of the 40th North American Power Symposium (NAPS '09)*, pp. 1–8, September 2009.
- [68] J.-A. Jiang, C.-S. Chen, and C.-W. Liu, "A new protection scheme for fault detection, direction discrimination, classification, and location in transmission lines," *IEEE Transactions on Power Delivery*, vol. 18, no. 1, pp. 34–42, 2003.
- [69] J.-A. Jiang, C.-S. Chen, and C.-W. Liu, "A new protection scheme for fault detection, direction discrimination, classification, and location in transmission lines," *IEEE Transactions on Power Delivery*, vol. 18, no. 1, pp. 34–42, 2003.
- [70] E. S. Tag El Din, M. Gilany, M. M. Abdel Aziz, and D. K. Ibrahim, "An PMU double ended fault location scheme for aged power cables," in *Proceedings of the IEEE Power Engineering Society General Meeting*, vol. 1, pp. 80–86, June 2005.
- [71] J. Joe-Air, L. Ying-Hong, L. Chih-Wen, Y. Jun-Zhe, and T. Tong-Ming, "An adaptive fault locator system for transmission lines," in *Proceedings of the IEEE Power Engineering Society Summer Meeting*, vol. 2, pp. 930–936, 1999.
- [72] C.-S. Chen, C.-W. Liu, and J.-A. Jiang, "A new adaptive PMU based protection scheme for transposed/untransposed parallel transmission lines," *IEEE Transactions on Power Delivery*, vol. 17, no. 2, pp. 395–404, 2002.
- [73] C.-S. Chen, C.-W. Liu, and J.-A. Jiang, "A new adaptive PMU based protection scheme for transposed/untransposed parallel transmission lines," *IEEE Transactions on Power Delivery*, vol. 17, no. 2, pp. 395–404, 2002.
- [74] J.-A. Jiang, C.-W. Liu, and C.-S. Chen, "A novel adaptive PMU-based transmission-line relay—design and EMTP simulation

- results," *IEEE Transactions on Power Delivery*, vol. 17, no. 4, pp. 930–937, 2002.
- [75] J.-A. Jiang, C.-W. Liu, and C.-S. Chen, "A novel adaptive PMU-based transmission-line relay—design and EMTP simulation results," *IEEE Transactions on Power Delivery*, vol. 17, no. 4, pp. 930–937, 2002.
- [76] J.-A. Jiang, Y.-H. Lin, J.-Z. Yang, T.-M. Too, and C.-W. Liu, "An adaptive PMU based fault detection/location technique for transmission lines—part II: PMU implementation and performance evaluation," *IEEE Transactions on Power Delivery*, vol. 15, no. 4, pp. 1136–1146, 2000.
- [77] J.-A. Jiang, J.-Z. Yang, Y.-H. Lin, C.-W. Liu, and J.-C. Ma, "An adaptive PMU based fault detection/location technique for transmission lines part I: theory and algorithms," *IEEE Transactions on Power Delivery*, vol. 15, no. 2, pp. 486–493, 2000.
- [78] H. Khorashadi-Zadeh and Z. Li, "A novel PMU-based transmission line protection scheme design," in *Proceedings of the 39th North American Power Symposium (NAPS '07)*, pp. 13–19, October 2007.
- [79] E. M. T. Eldin, M. I. Gilany, M. M. Abdelaziz, and D. K. Ibrahim, "An accurate fault location scheme for connected aged cable lines in double-fed systems," *Electrical Engineering*, vol. 88, no. 5, pp. 431–439, 2006.
- [80] Y.-H. Lin, C.-W. Liu, J.-A. Jiang, and J.-Z. Yang, "An adaptive fault locator for transmission lines tapped with a source of generation—using synchronized voltage and current phasors," in *Proceedings of the Power Engineering Society Summer Meeting*, pp. 1379–1383, July 2000.
- [81] L. Shengfang, F. Chunju, Y. Weiyong, C. Huarong, and K. K. Li, "A new phase measurement unit (PMU) based fault location algorithm for double circuit lines," in *Proceedings of the 8th IEE International Conference on Developments in Power System Protection*, vol. 1, pp. 188–191, April 2004.
- [82] Z. Liang, Y. Mu, Q. Wu, and W. Lu, "A new method about location for single-phase grounding fault in distribution network," in *Proceedings of the 20th International Conference and Exhibition on Electricity Distribution (CIRED '09)*, pp. 1–4, June 2009.
- [83] K.-P. Lien, C.-W. Liu, C.-S. Yu, and J.-A. Jiang, "Transmission network fault location observability with minimal PMU placement," *IEEE Transactions on Power Delivery*, vol. 21, no. 3, pp. 1128–1136, 2006.
- [84] G.-F. Zhu and Y.-P. Lu, "Development of fault location algorithm for distribution networks with DG," in *Proceedings of the IEEE International Conference on Sustainable Energy Technologies (ICSET '08)*, pp. 164–168, November 2008.
- [85] S. A. Gafoor and P. V. Ramana Rao, "Wavelet based fault detection, classification and location in transmission lines," in *Proceedings of the 1st International Power and Energy Conference (PECon '06)*, pp. 114–118, November 2006.
- [86] S. A. M. Javadian, A. M. Nasrabadi, M.-R. Haghifam, and J. Rezvantalab, "Determining fault's type and accurate location in distribution systems with DG using MLP neural networks," in *Proceedings of the International Conference on Clean Electrical Power (ICCEP '09)*, pp. 284–289, June 2009.
- [87] W. Bo, Q. Jiang, and Y. Cao, "Transmission network fault location using sparse PMU measurements," in *Proceedings of the 1st International Conference on Sustainable Power Generation and Supply (SUPERGEN '09)*, pp. 1–6, April 2009.
- [88] S. S. Geramian, H. A. Abyane, and K. Mazlumi, "Determination of optimal pmu placement for fault location using genetic algorithm," in *Proceedings of the 13th International Conference on Harmonics and Quality of Power (ICHQP '08)*, pp. 1–5, October 2008.
- [89] K. Mazlumi, H. A. Abyaneh, S. H. H. Sadeghi, and S. S. Geramian, "Determination of optimal PMU placement for fault-location observability," in *Proceedings of the 3rd International Conference on Deregulation and Restructuring and Power Technologies (DRPT '08)*, pp. 1938–1942, April 2008.
- [90] Y. Zhang, J. Zhang, J. Ma, and Z. Wang, "Fault detection based on discriminant analysis theory in electric power system," in *Proceedings of the 1st International Conference on Sustainable Power Generation and Supply (SUPERGEN '09)*, pp. 1–5, April 2009.
- [91] Y. Zhang, J. Ma, J. Zhang, and Z. Wang, "Fault diagnosis based on cluster analysis theory in wide area backup protection system," in *Proceedings of the Asia-Pacific Power and Energy Engineering Conference (APPEEC '09)*, pp. 1–4, March 2009.
- [92] W. Zhao, X. G. Chen, Y. J. Cao, and M. W. Peng, "A novel method of fault diagnosis based on synchronized phasor measuring and flow fingerprint identification technology," in *Proceedings of the 1st International Conference on Sustainable Power Generation and Supply (SUPERGEN '09)*, pp. 1–5, April 2009.
- [93] Y. Zhang, J. Zhang, J. Ma, and Z. Wang, "Fault detection and identification based on DFS in electric power network," in *Proceedings of the IEEE International Symposium on Knowledge Acquisition and Modeling Workshop (KAM '08)*, pp. 742–745.
- [94] R. A. F. Pereira, L. G. W. Da Silva, and J. R. S. Mantovani, "PMUs optimized allocation using a tabu search algorithm for fault location in electric power distribution system," in *Proceedings of the IEEE/PES Transmission and Distribution Conference and Exposition: Latin America*, pp. 143–148, November 2004.
- [95] D. Düstegör, S. V. Poroseva, M. Y. Hussaini, and S. Woodruff, "Automated graph-based methodology for fault detection and location in power systems," *IEEE Transactions on Power Delivery*, vol. 25, no. 2, pp. 638–646, 2010.
- [96] N. Rezaei and M.-R. Haghifam, "Protection scheme for a distribution system with distributed generation using neural networks," *International Journal of Electrical Power and Energy Systems*, vol. 30, no. 4, pp. 235–241, 2008.
- [97] Z. Wang, Y. Zhang, and J. Zhang, "Principal components fault location based on WAMS/PMU measure system," in *Proceedings of the IEEE PES General Meeting: The Electrification of Transportation and the Grid of the Future*, July 2011.
- [98] J. Quanyuan, L. Xingpeng, W. Bo, and W. Haijiao, "PMU-based fault location using voltage measurements in large transmission networks," *IEEE Transactions on Power Delivery*, vol. 27, no. 3, pp. 1644–1652, 2012.
- [99] E. Nashawati, R. Garcia, and T. Rosenberger, "Using synchrophasor for fault location identification," in *Proceedings of the 65th Annual Conference for Protective Relay Engineers*, pp. 14–21, 2012.
- [100] G.-T. Kim, H.-S. Kim, and J.-Y. Choi, "Wavelet transform based power transmission line fault location using GPS for accurate time synchronization," in *Proceedings of the 27th Annual Conference of the IEEE Industrial Electronics Society (IECON '01)*, vol. 1, pp. 495–499, December 2001.
- [101] M. Gilany, D. K. Ibrahim, and E. S. Tag Eldin, "Traveling-wave-based fault-location scheme for multiend-aged underground cable system," *IEEE Transactions on Power Delivery*, vol. 22, no. 1, pp. 82–89, 2007.
- [102] H. Nouri, W. Chun, and T. Davies, "An accurate fault location technique for distribution lines with tapped loads using wavelet

- transform,” in *Proceedings of the IEEE Porto Power Tech Proceedings*, vol. 3, 2001.
- [103] M. Da Silva, M. Oleskovicz, and D. V. Coury, “A fault locator for three-terminal lines based on wavelet transform applied to synchronized current and voltage signals,” in *Proceedings of the IEEE PES Transmission and Distribution Conference and Exposition: Latin America (TDC '06)*, pp. 1–6, August 2006.
- [104] C. Evrenosoglu and A. Abur, “Fault location for teed circuits with mutually coupled lines and series capacitors,” in *Proceedings of the IEEE Bologna Power Tech Conference*, 2003.
- [105] M. Da Silva, D. V. Coury, M. Oleskovicz, and Ê. C. Segatto, “An Alternative fault location algorithm based on wavelet transforms for three-terminal lines,” in *Proceedings of the IEEE Power and Energy Society General Meeting: Conversion and Delivery of Electrical Energy in the 21st Century*, July 2008.

## Research Article

# Design of an Optimal Preview Controller for Linear Discrete-Time Descriptor Noncausal Multirate Systems

Mengjuan Cao<sup>1,2</sup> and Fucheng Liao<sup>1</sup>

<sup>1</sup> School of Mathematics and Physics, University of Science and Technology Beijing, Beijing 100083, China

<sup>2</sup> School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

Correspondence should be addressed to Fucheng Liao; [fcliao@ustb.edu.cn](mailto:fcliao@ustb.edu.cn)

Received 4 December 2013; Accepted 17 December 2013; Published 23 January 2014

Academic Editors: N. Barsoum, P. Vasant, and G.-W. Weber

Copyright © 2014 M. Cao and F. Liao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The linear discrete-time descriptor noncausal multirate system is considered for the presentation of a new design approach for optimal preview control. First, according to the characteristics of causal controllability and causal observability, the descriptor noncausal system is constructed into a descriptor causal closed-loop system. Second, by using the characteristics of the causal system and elementary transformation, the descriptor causal closed-loop system is transformed into a normal system. Then, taking advantage of the discrete lifting technique, the normal multirate system is converted to a single-rate system. By making use of the standard preview control method, we construct the descriptor augmented error system. The quadratic performance index for the multirate system is given, which can be changed into one for the single-rate system. In addition, a new single-rate system is obtained, the optimal control law of which is given. Returning to the original system, the optimal preview controller for linear discrete-time descriptor noncausal multirate systems is derived. The stabilizability and detectability of the lifted single-rate system are discussed in detail. The optimal preview control design techniques are illustrated by simulation results for a simple example.

## 1. Introduction

Descriptor system theory has obtained many excellent results in the control areas; the main scholarly reports can be seen in [1, 2]. In recent years, the literature [3] considered the optimal fusion problem for the state estimation of discrete-time stochastic singular systems with multiple sensors and correlated measurement noise and obtained the optimal full-order filters and smoothers for the original system. The literature [4] proposed a novel suboptimal control method for a class of nonlinear singularly perturbed systems based on adaptive dynamic programming; the literature [5] discussed finite-time robust dissipative control for a class of descriptor systems, and the control system was effectively confined within the desired state-space ellipsoid. The literature [6] provided a necessary and sufficient condition to guarantee admissibility for positive continuous-time descriptor systems. Notably, the literature [7] combined descriptor system theory with preview control theory and successfully obtained the optimal preview controller with preview action for the linear discrete-time descriptor causal system; the literature [8]

derived the optimal preview controller for discrete-time descriptor causal systems in a multirate setting. The literature [9] obtained the optimal preview controller with preview feedforward compensation for linear discrete-time descriptor systems with state delay. In addition, linear quadratic optimal regulator theory for the continuous and discrete descriptor system tends to be complete as discussed in [10–12].

In recent years, the multirate digital control system has also obtained many new results as discussed in [13–16]. The characteristics of multirate systems are as follows. First of all, the systems are multiinput and multioutput systems. Second, the sampler and retainer of input channels and output channels have different sampling periods as discussed by Xiao [13]. For such systems, if the designed regulator satisfies appropriate multirate characteristics, it should have a better performance than that of the single-rate regulator.

The previous multirate systems have been basically studied for normal systems; however, this paper successfully constructs the optimal preview controller on the basis of the literature [8] for linear discrete-time descriptor noncausal

multirate systems. The effectiveness of the proposed method is shown by simulation.

## 2. Description of the Problem and Basic Assumptions

Consider the regular linear discrete-time descriptor non-causal system described by

$$\begin{aligned} Ex(k+1) &= Ax(k) + Bu(k), \\ y(k) &= Cx(k), \end{aligned} \tag{1}$$

where  $x(k) \in R^n$ ,  $u(k) \in R^r$ , and  $y(k) \in R^m$  are its state, control input, and measure output, respectively;  $E, A \in R^{n \times n}$ ,  $B \in R^{n \times r}$ , and  $C \in R^{m \times n}$  are constant matrices; here,  $E$  is a singular matrix with  $\text{rank}(E) = q < n$ .

As [8], we need to make the following basic assumptions:

Assumption 1 (A1): system (1) is stabilizable.

Assumption 2 (A2): system (1) is detectable.

Assumption 3 (A3): the system (1) is both causally controllable and causally observable.

Assumption 4 (A4): the state vector  $x(k)$  and output vector  $y(k)$  can only be measured at  $k = iN$  ( $i = 0, 1, 2, \dots$ ), where  $N$  is a positive integer.

Assumption 5 (A5): the preview length of the reference signal  $R(k)$  is  $M_R$ ; that is, at each time  $k$ , the  $M_R$  future values  $R(k+1), R(k+2), \dots, R(k+M_R)$ , and the present and past values of the reference signal are available where  $M_R = SN$  and  $S$  is a nonnegative integer.

The future values of the desired signal are assumed not to change beyond the  $k + M_R$ ; namely

$$R(k+j) = R(k+M_R), \quad j = M_R + 1, M_R + 2, \dots \tag{2}$$

*Remark 1.* (A1)–(A3) and (A5) are the basic assumptions, and (A4) makes the system multirate.

By (A3), there must exist a static output feedback

$$\begin{aligned} u(k) &= My(k) + v(k) \\ &= MCx(k) + v(k) = Kx(k) + v(k), \end{aligned} \tag{3}$$

where  $K = MC$ ,  $v(k) \in R^r$ , and  $M \in R^{r \times m}$  such that the closed-loop system

$$Ex(k+1) = (A + BK)x(k) + Bv(k) \tag{4}$$

is causal as discussed by [2]; that is,

$$\text{deg} \{ \det [sE - (A + BK)] \} = \text{rank}(E). \tag{5}$$

Obviously, taking advantage of the characteristic that any matrix can be transformed to a canonical form by

elementary transformation, there always exist nonsingular matrices  $Q_1, P_1$ , such that  $QEP = \begin{bmatrix} I_q & 0 \\ 0 & 0 \end{bmatrix}$ . Denote

$$\begin{aligned} x(k) &= P_1 \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}, \quad Q_1(A + BK)P_1 = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \\ Q_1B &= \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad CP_1 = [C_1 \ C_2], \end{aligned} \tag{6}$$

where  $x_1(k) \in R^q$  and  $x_2(k) \in R^{n-q}$ .

As [8], the system (4) is restricted equivalent to

$$\begin{aligned} x_1(k+1) &= A_{11}x_1(k) + A_{12}x_2(k) + B_1v(k), \\ 0 &= A_{21}x_1(k) + A_{22}x_2(k) + B_2v(k), \\ y(k) &= C_1x_1(k) + C_2x_2(k). \end{aligned} \tag{7}$$

Because elementary transformation does not change the causality of the system, the system (7) is also a causal system. As a result, matrix  $A_{22}$  is nonsingular as discussed by [2]. Then the optimal preview problem for the descriptor non-causal system is transformed into the one for the descriptor causal system.

As [8], let the error signal

$$e(k) = y(k) - R(k). \tag{8}$$

We want to get

$$\lim_{k \rightarrow \infty} e(k) = \lim_{k \rightarrow \infty} [y(k) - R(k)] = 0. \tag{9}$$

The quadratic performance index function for the system (1) is defined as

$$J = \sum_{k=1}^{\infty} [e^T(k) Q_e e(k) + \Delta u^T(k) H_u \Delta u(k)], \tag{10}$$

where the weight matrices satisfy  $Q_e > 0$  and  $H_u > 0$ .  $\Delta$  is the first-order forward difference operator; that is,  $\Delta u(k) = u(k+1) - u(k)$ .

In order to smooth the conduct of the study, we will also make the following assumptions.

Assumption 6 (A6): the matrix

$$\begin{bmatrix} \bar{K}_2 & 0 & 0 & \cdots & 0 \\ \bar{K}_1 \bar{B}_1 & \bar{K}_2 & 0 & \cdots & 0 \\ \bar{K}_1 \bar{A}_1 \bar{B}_1 & \bar{K}_1 \bar{B}_1 & \bar{K}_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \bar{K}_1 \bar{A}_1^{N-2} \bar{B}_1 & \bar{K}_1 \bar{A}_1^{N-3} \bar{B}_1 & \bar{K}_1 \bar{A}_1^{N-4} \bar{B}_1 & \cdots & \bar{K}_2 \end{bmatrix} \tag{11}$$

is nonsingular, where the meaning of the various symbols is given in the following discussion.

Assumption 7(A7): the matrix

$$\Psi = \left[ \begin{array}{c|cccc} \overline{A_1^N} - I & \overline{A_1^{N-1}}\overline{B_1} & \overline{A_1^{N-2}}\overline{B_1} & \cdots & \overline{A_1}\overline{B_1} & \overline{B_1} \\ \hline \overline{C_1} & \overline{C_2} & 0 & \cdots & \cdots & 0 \\ \overline{C_1}\overline{A_1} & \overline{C_1}\overline{B_1} & \overline{C_2} & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \ddots & \vdots \\ \overline{C_1}\overline{A_1}^{N-2} & \overline{C_1}\overline{A_1}^{N-3}\overline{B_1} & \overline{C_1}\overline{A_1}^{N-4}\overline{B_1} & \cdots & \overline{C_2} & 0 \\ \overline{C_1}\overline{A_1}^{N-1} & \overline{C_1}\overline{A_1}^{N-2}\overline{B_1} & \overline{C_1}\overline{A_1}^{N-3}\overline{B_1} & \cdots & \overline{C_1}\overline{B_1} & \overline{C_2} \end{array} \right] \quad (12)$$

is of full row rank, where the meaning of the various symbols is given in the following discussion.

### 3. The Derivation of the Single-Rate System

The system (7) is a multirate system according to the above discussion. We adopt the discrete lifting technique to convert (7) to a single-rate system.

By the obtained results in [8], (7) can be lifted as

$$\begin{aligned} \tilde{x}_1(i+1) &= \overline{A_1^N} \tilde{x}_1(i) + \tilde{B}_1 \tilde{v}(i), \\ \tilde{y}(i) &= \overline{C_1} \tilde{x}_1(i) + \overline{C_2} \tilde{v}(i), \end{aligned} \quad (13)$$

where

$$\begin{aligned} \overline{A_1} &= A_{11} - A_{12}A_{22}^{-1}A_{21}, \\ \overline{B_1} &= B_1 - A_{12}A_{22}^{-1}B_2, \quad \tilde{x}_1(i) = x_1(iN) \in R^q, \\ \tilde{x}_2(i) &= x_2(iN) \in R^{n-q}, \quad \tilde{v}(i) = \begin{bmatrix} v(iN) \\ \vdots \\ v(iN+N-2) \\ v(iN+N-1) \end{bmatrix}, \\ \tilde{B}_1 &= [\overline{A_1}^{N-1}\overline{B_1} \ \cdots \ \overline{A_1}\overline{B_1} \ \overline{B_1}], \\ \overline{C_1} &= C_1 - C_2A_{22}^{-1}A_{21}, \\ \overline{C_2} &= -C_2A_{22}^{-1}B_2, \quad \tilde{y}(i) = \begin{bmatrix} y(iN) \\ y(iN+1) \\ \vdots \\ y(iN+N-2) \\ y(iN+N-1) \end{bmatrix}, \\ \overline{C_1} &= \begin{bmatrix} \overline{C_1} \\ \overline{C_1}\overline{A_1} \\ \vdots \\ \overline{C_1}\overline{A_1}^{N-2} \\ \overline{C_1}\overline{A_1}^{N-1} \end{bmatrix}, \end{aligned}$$

$$\overline{C_2} = \begin{bmatrix} \overline{C_2} & 0 & \cdots & 0 & 0 \\ \overline{C_1}\overline{B_1} & \overline{C_2} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \overline{C_1}\overline{A_1}^{N-3}\overline{B_1} & \overline{C_1}\overline{A_1}^{N-4}\overline{B_1} & \cdots & \overline{C_2} & 0 \\ \overline{C_1}\overline{A_1}^{N-2}\overline{B_1} & \overline{C_1}\overline{A_1}^{N-3}\overline{B_1} & \cdots & \overline{C_1}\overline{B_1} & \overline{C_2} \end{bmatrix}. \quad (14)$$

In order to design the optimal preview controller for linear discrete-time descriptor noncausal multirate systems (1), continue to lift the static output feedback (3).

First, denoting  $KP_1 = [K_1 \ K_2]$ , where  $K_1 \in R^{r \times q}$  and  $K_2 \in R^{r \times (n-q)}$ , we have

$$\begin{aligned} u(k) &= KP_1 \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + v(k) = [K_1 \ K_2] \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + v(k) \\ &= K_1x_1(k) + K_2x_2(k) + v(k). \end{aligned} \quad (15)$$

We know that matrix  $A_{22}$  is nonsingular. Then we can derive

$$x_2(k) = -A_{22}^{-1}A_{21}x_1(k) - A_{22}^{-1}B_2v(k), \quad (16)$$

from the second equation of (7).

By (16), (15) will become

$$\begin{aligned} u(k) &= K_1x_1(k) + K_2(-A_{22}^{-1}A_{21}x_1(k) - A_{22}^{-1}B_2v(k)) \\ &\quad + v(k) \\ &= (K_1 - K_2A_{22}^{-1}A_{21})x_1(k) \\ &\quad + (I - K_2A_{22}^{-1}B_2)v(k) = \overline{K_1}x_1(k) + \overline{K_2}v(k), \end{aligned} \quad (17)$$

where

$$\overline{K_1} = K_1 - K_2A_{22}^{-1}A_{21}, \quad \overline{K_2} = I - K_2A_{22}^{-1}B_2. \quad (18)$$

Substituting (16) into the first equation in (7), we get

$$x_1(k+1) = \overline{A_1}x_1(k) + \overline{B_1}v(k). \quad (19)$$

Using (19) repeatedly, (17) will become

$$\begin{aligned} u(iN) &= \overline{K_1}x_1(iN) + \overline{K_2}v(iN), \\ u(iN+1) &= \overline{K_1}x_1(iN+1) + \overline{K_2}v(iN+1) \\ &= \overline{K_1}(\overline{A_1}x_1(iN) + \overline{B_1}v(iN)) + \overline{K_2}v(iN+1) \\ &= \overline{K_1}\overline{A_1}x_1(iN) + \overline{K_1}\overline{B_1}v(iN) + \overline{K_2}v(iN+1), \end{aligned}$$

$$\begin{aligned}
 u(iN + 2) &= \bar{K}_1 \bar{A}_1 x_1(iN + 1) + \bar{K}_1 \bar{B}_1 v(iN + 1) \\
 &\quad + \bar{K}_2 v(iN + 2) \\
 &= \bar{K}_1 \bar{A}_1 (\bar{A}_1 x_1(iN) + \bar{B}_1 v(iN)) \\
 &\quad + \bar{K}_1 \bar{B}_1 v(iN + 1) + \bar{K}_2 v(iN + 2) \\
 &= \bar{K}_1 \bar{A}_1^2 x_1(iN) + \bar{K}_1 \bar{A}_1 \bar{B}_1 v(iN) \\
 &\quad + \bar{K}_1 \bar{B}_1 v(iN + 1) + \bar{K}_2 v(iN + 2), \\
 &\quad \vdots \\
 u(iN + N - 1) &= \bar{K}_1 \bar{A}_1^{N-1} x_1(iN) + \bar{K}_1 \bar{A}_1^{N-2} \bar{B}_1 v(iN) \\
 &\quad + \dots + \bar{K}_1 \bar{B}_1 v(iN + N - 2) \\
 &\quad + \bar{K}_2 v(iN + N - 1).
 \end{aligned} \tag{20}$$

The above equations may be represented in the matrix form:

$$\bar{u}(i) = \bar{K}_1 \bar{x}_1(i) + \bar{K}_2 \bar{v}(i), \tag{21}$$

where

$$\bar{K}_1 = \begin{bmatrix} \bar{K}_1 \\ \bar{K}_1 \bar{A}_1 \\ \bar{K}_1 \bar{A}_1^2 \\ \vdots \\ \bar{K}_1 \bar{A}_1^{N-1} \end{bmatrix},$$

$$\bar{K}_2 = \begin{bmatrix} \bar{K}_2 & 0 & 0 & \dots & 0 \\ \bar{K}_1 \bar{B}_1 & \bar{K}_2 & 0 & \dots & 0 \\ \bar{K}_1 \bar{A}_1 \bar{B}_1 & \bar{K}_1 \bar{B}_1 & \bar{K}_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \bar{K}_1 \bar{A}_1^{N-2} \bar{B}_1 & \bar{K}_1 \bar{A}_1^{N-3} \bar{B}_1 & \bar{K}_1 \bar{A}_1^{N-4} \bar{B}_1 & \dots & \bar{K}_2 \end{bmatrix},$$

$$\bar{u}(i) = \begin{bmatrix} u(iN) \\ u(iN + 1) \\ u(iN + 2) \\ \vdots \\ u(iN + N - 1) \end{bmatrix}. \tag{22}$$

Remark 2.  $\bar{K}_2$  is exactly the matrix in (A6).

### 4. Construction of the Descriptor Augmented Error System

As [8], we take advantage of the first-order forward difference operator  $\Delta$ :

$$\Delta \bar{x}_1(i) = \bar{x}_1(i + 1) - \bar{x}_1(i). \tag{23}$$

Construct the vector

$$\bar{R}(i) = \begin{bmatrix} R(iN) \\ R(iN + 1) \\ \vdots \\ R(iN + N - 2) \\ R(iN + N - 1) \end{bmatrix}. \tag{24}$$

Then we can obtain the error vector

$$\begin{aligned}
 \bar{e}(i) = \bar{y}(i) - \bar{R}(i) &= \begin{bmatrix} y(iN) \\ y(iN + 1) \\ \vdots \\ y(iN + N - 2) \\ y(iN + N - 1) \end{bmatrix} - \begin{bmatrix} R(iN) \\ R(iN + 1) \\ \vdots \\ R(iN + N - 2) \\ R(iN + N - 1) \end{bmatrix} \\
 &= \begin{bmatrix} e(iN) \\ e(iN + 1) \\ \vdots \\ e(iN + N - 2) \\ e(iN + N - 1) \end{bmatrix}.
 \end{aligned} \tag{25}$$

By the second equation in (13), we derive

$$\bar{e}(i) = \bar{C}_1 \bar{x}_1(i) - \bar{R}(i) + \bar{C}_2 \bar{v}(i). \tag{26}$$

Using  $\Delta$  on both sides of (26) and noticing  $\Delta \bar{e}(i) = \bar{e}(i + 1) - \bar{e}(i)$ , we obtain

$$\bar{e}(i + 1) = \bar{e}(i) + \bar{C}_1 \Delta \bar{x}_1(i) - \Delta \bar{R}(i) + \bar{C}_2 \Delta \bar{v}(i). \tag{27}$$

Using  $\Delta$  on both sides of the first equation of (13), we can derive

$$\Delta \bar{x}_1(i + 1) = \bar{A}_1^N \Delta \bar{x}_1(i) + \bar{B}_1 \Delta \bar{v}(i). \tag{28}$$

Combine (27) and (28) to produce

$$\begin{aligned}
 \begin{bmatrix} \bar{e}(i + 1) \\ \Delta \bar{x}_1(i + 1) \end{bmatrix} &= \begin{bmatrix} I & \bar{C}_1 \\ 0 & \bar{A}_1^N \end{bmatrix} \begin{bmatrix} \bar{e}(i) \\ \Delta \bar{x}_1(i) \end{bmatrix} \\
 &\quad + \begin{bmatrix} -I_{mN} \\ 0 \end{bmatrix} \Delta \bar{R}(i) + \begin{bmatrix} \bar{C}_2 \\ \bar{B}_1 \end{bmatrix} \Delta \bar{v}(i).
 \end{aligned} \tag{29}$$

Contrasting (1), the observed vector can be taken as  $\bar{e}(i)$ . Letting

$$\begin{aligned}
 X_0(i) &= \begin{bmatrix} \bar{e}(i) \\ \Delta \bar{x}_1(i) \end{bmatrix}, & \Phi &= \begin{bmatrix} I & \bar{C}_1 \\ 0 & \bar{A}_1^N \end{bmatrix}, \\
 G &= \begin{bmatrix} \bar{C}_2 \\ \bar{B}_1 \end{bmatrix}, & G_R &= \begin{bmatrix} -I_{mN} \\ 0 \end{bmatrix}, & C_0 &= [I \ 0].
 \end{aligned} \tag{30}$$

As [8], we have

$$\begin{aligned} X_0(i+1) &= \Phi X_0(i) + G\Delta\tilde{v}(i) + G_R\Delta\tilde{R}(i), \\ \tilde{e}(i) &= C_0X_0(i). \end{aligned} \tag{31}$$

Equation (31) is the error system, which is a normal system. For (31), the previewed desired signal is  $\Delta\tilde{R}(i)$ ; that is, at each time  $i$ ,  $\Delta\tilde{R}(i), \Delta\tilde{R}(i+1), \dots, \Delta\tilde{R}(i+S-1)$  are available, and

$$\Delta\tilde{R}(i+l) = 0 \quad (l = S, S+1, \dots). \tag{32}$$

Then we continue to construct the descriptor augmented error system and denote

$$\begin{aligned} X_R(i) &= \begin{bmatrix} \Delta\tilde{R}(i) \\ \Delta\tilde{R}(i+1) \\ \vdots \\ \Delta\tilde{R}(i+S-1) \end{bmatrix}, \\ A_R &= \begin{bmatrix} 0 & I_{mN} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & I_{mN} \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix}, \end{aligned} \tag{33}$$

where  $A_R$  is a  $mNS \times mNS$  matrix; notice the identity  $X_R(i+1) = A_RX_R(i)$ . Using the identity and (31), we obtain

$$\begin{aligned} X_{R0}(i+1) &= \Phi_{R0}X_{R0}(i) + G_{R0}\Delta\tilde{v}(i), \\ \tilde{e}(i) &= C_{R0}X_{R0}(i). \end{aligned} \tag{34}$$

This is the constructed descriptor augmented error system. The dimension of the system (34) is  $mNS + mN + q$ , and

$$\begin{aligned} X_{R0}(i) &= \begin{bmatrix} X_R(i) \\ X_0(i) \end{bmatrix}, \quad \Phi_{R0} = \begin{bmatrix} A_R & 0 \\ G_{PR} & \Phi \end{bmatrix}, \quad G_{R0} = \begin{bmatrix} 0 \\ G \end{bmatrix}, \\ G_{PR} &= [G_R \ 0 \ \cdots \ 0], \quad C_{R0} = [0 \ C_0]. \end{aligned} \tag{35}$$

### 5. Design of an Optimal Regulator for Descriptor Augmented Error Systems

As [8], we convert the performance index (10) as follows

$$\begin{aligned} J &= \sum_{k=1}^{\infty} [e^T(k) Q_e e(k) + \Delta u^T(k) H_u \Delta u(k)] \\ &= \sum_{i=1}^{\infty} \sum_{j=0}^{N-1} [e^T(iN+j) Q_e e(iN+j) \\ &\quad + \Delta u^T(iN+j) H_u \Delta u(iN+j)] \\ &= \sum_{i=1}^{\infty} [e^T(i) \bar{Q} \tilde{e}(i) + \Delta \tilde{u}^T(i) H \Delta \tilde{u}(i)], \end{aligned} \tag{36}$$

where

$$\begin{aligned} \bar{Q} &= \text{diag}(\underbrace{Q_e \ Q_e \ \cdots \ Q_e}_N) > 0, \\ H &= \text{diag}(\underbrace{H_u \ H_u \ \cdots \ H_u}_N) > 0, \\ \bar{Q} &\in R^{Nm \times Nm}, \quad H \in R^{Nr \times Nr}. \end{aligned} \tag{37}$$

By (A6),  $\bar{K}_2^T H \bar{K}_2 > 0$ . Then, adopting the first-order forward difference operator on both sides of (21), the performance index (36) can be written as

$$\begin{aligned} J &= \sum_{i=1}^{\infty} [e^T(i) \bar{Q} \tilde{e}(i) + \Delta \tilde{u}^T(i) H \Delta \tilde{u}(i)] \\ &= \sum_{i=1}^{\infty} [e^T(i) \bar{Q} \tilde{e}(i) + (\bar{K}_1 \Delta \tilde{x}_1(i) + \bar{K}_2 \Delta \tilde{v}(i))^T \\ &\quad \times H (\bar{K}_1 \Delta \tilde{x}_1(i) + \bar{K}_2 \Delta \tilde{v}(i))] \\ &= \sum_{i=1}^{\infty} [e^T(i) \bar{Q} \tilde{e}(i) + \Delta \tilde{v}^T(i) \bar{K}_2^T H \bar{K}_2 \Delta \tilde{v}(i) \\ &\quad + 2\Delta \tilde{v}^T(i) \bar{K}_2^T H \bar{K}_1 \Delta \tilde{x}_1(i) + \Delta \tilde{x}_1^T(i) \bar{K}_1^T H \bar{K}_1 \Delta \tilde{x}_1(i)] \\ &= \sum_{i=1}^{\infty} [e^T(i) \bar{Q} \tilde{e}(i) + \Delta \tilde{v}^T(i) \bar{K}_2^T H \bar{K}_2 \Delta \tilde{v}(i) \\ &\quad + 2\Delta \tilde{v}^T(i) \bar{K}_2^T H \bar{K}_1 \Delta \tilde{x}_1(i) \\ &\quad + \Delta \tilde{x}_1^T(i) \bar{K}_1^T H \bar{K}_2 [\bar{K}_2^T H \bar{K}_2]^{-1} \bar{K}_2^T H \bar{K}_1 \Delta \tilde{x}_1(i) \\ &\quad + \Delta \tilde{x}_1^T(i) \\ &\quad \times [\bar{K}_1^T H \bar{K}_1 - \bar{K}_1^T H \bar{K}_2 [\bar{K}_2^T H \bar{K}_2]^{-1} \bar{K}_2^T H \bar{K}_1] \Delta \tilde{x}_1(i)] \\ &= \sum_{i=1}^{\infty} \left[ [\Delta \tilde{v}(i) + [\bar{K}_2^T H \bar{K}_2]^{-1} \bar{K}_2^T H \bar{K}_1 \Delta \tilde{x}_1(i)]^T \right. \\ &\quad \times (\bar{K}_2^T H \bar{K}_2) [\Delta \tilde{v}(i) + [\bar{K}_2^T H \bar{K}_2]^{-1} \bar{K}_2^T H \bar{K}_1 \Delta \tilde{x}_1(i)] \\ &\quad + e^T(i) \bar{Q} \tilde{e}(i) + \Delta \tilde{x}_1^T(i) \\ &\quad \times [\bar{K}_1^T H \bar{K}_1 - \bar{K}_1^T H \bar{K}_2 [\bar{K}_2^T H \bar{K}_2]^{-1} \bar{K}_2^T H \bar{K}_1] \Delta \tilde{x}_1(i) \left. \right]. \end{aligned} \tag{38}$$

If we denote  $\bar{Q} = \bar{K}_1^T H \bar{K}_1 - \bar{K}_1^T H \bar{K}_2 [\bar{K}_2^T H \bar{K}_2]^{-1} \bar{K}_2^T H \bar{K}_1$ , it is easy to see  $\bar{Q} \geq 0$ .

Let

$$w(i) = \Delta \tilde{v}(i) + [\bar{K}_2^T H \bar{K}_2]^{-1} \bar{K}_2^T H \bar{K}_1 \Delta \tilde{x}_1(i), \tag{39}$$

$$R = [\bar{K}_2^T H \bar{K}_2]^{-1} \bar{K}_2^T H \bar{K}_1. \tag{40}$$

From  $\Delta\tilde{x}_1(i) = [0 \ 0 \ I_q] X_{R0}(i)$ , (39) can be written as

$$\begin{aligned} w(i) &= \Delta\tilde{v}(i) + R\Delta\tilde{x}_1(i) \\ &= \Delta\tilde{v}(i) + R[0 \ 0 \ I_q] X_{R0}(i) \\ &= \Delta\tilde{v}(i) + \widehat{K}_1 X_{R0}(i), \end{aligned} \tag{41}$$

where  $\widehat{K}_1 = R[0 \ 0 \ I_q]$ .

The performance index (38) can continue to be written as

$$\begin{aligned} J &= \sum_{i=1}^{\infty} [w^T(i) \widehat{K}_2^T H \widehat{K}_2 w(i) + \tilde{e}^T(i) \overline{Q} \tilde{e}(i) + \Delta\tilde{x}_1^T(i) \overline{Q} \Delta\tilde{x}_1(i)] \\ &= \sum_{i=1}^{\infty} \left[ w^T(i) \widehat{K}_2^T H \widehat{K}_2 w(i) \right. \\ &\quad \left. + \begin{bmatrix} X_R(i) \\ \tilde{e}(i) \\ \Delta\tilde{x}_1(i) \end{bmatrix}^T \begin{bmatrix} 0 & 0 & 0 \\ 0 & \overline{Q} & 0 \\ 0 & 0 & \overline{Q} \end{bmatrix} \begin{bmatrix} X_R(i) \\ \tilde{e}(i) \\ \Delta\tilde{x}_1(i) \end{bmatrix} \right] \\ &= \sum_{i=1}^{\infty} \left[ w^T(i) \widehat{K}_2^T H \widehat{K}_2 w(i) + X_{R0}^T(i) \begin{bmatrix} 0 & 0 & 0 \\ 0 & \overline{Q} & 0 \\ 0 & 0 & \overline{Q} \end{bmatrix} X_{R0}(i) \right] \\ &= \sum_{i=1}^{\infty} [X_{R0}^T(i) \widehat{Q} X_{R0}(i) + w^T(i) \widehat{K}_2^T H \widehat{K}_2 w(i)], \end{aligned} \tag{42}$$

where  $\widehat{Q} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \overline{Q} & 0 \\ 0 & 0 & \overline{Q} \end{bmatrix}$ .

From (41), we derive

$$\Delta\tilde{v}(i) = w(i) - \widehat{K}_1 X_{R0}(i). \tag{43}$$

Substituting (43) into (34), we get

$$\begin{aligned} X_{R0}(i+1) &= (\Phi_{R0} - G_{R0} \widehat{K}_1) X_{R0}(i) + G_{R0} w(i), \\ \tilde{e}(i) &= C_{R0} X_{R0}(i). \end{aligned} \tag{44}$$

Then, the problem becomes an optimal control problem for a normal system (44) under the performance index (42). According to the results in Duan [17], we immediately get the following.

**Theorem 3.** *If  $(\Phi_{R0} - G_{R0} \widehat{K}_1 \mid G_{R0})$  is stabilizable and  $(\widehat{Q}^{1/2} \mid \Phi_{R0} - G_{R0} \widehat{K}_1)$  is detectable, the optimal regulator of the system (44) minimizing the performance index (42) is given by*

$$\begin{aligned} w(i) &= -[\widehat{K}_2^T H \widehat{K}_2 + G_{R0}^T P G_{R0}]^{-1} \\ &\quad \times G_{R0}^T P (\Phi_{R0} - G_{R0} \widehat{K}_1) X_{R0}(i), \end{aligned} \tag{45}$$

where  $P$  is the unique symmetric semipositive definite solution of the algebraic Riccati equation:

$$\begin{aligned} P &= (\Phi_{R0} - G_{R0} \widehat{K}_1)^T P (\Phi_{R0} - G_{R0} \widehat{K}_1) \\ &\quad - (\Phi_{R0} - G_{R0} \widehat{K}_1)^T P G_{R0} [\widehat{K}_2^T H \widehat{K}_2 + G_{R0}^T P G_{R0}]^{-1} \\ &\quad \times G_{R0}^T P (\Phi_{R0} - G_{R0} \widehat{K}_1) + \widehat{Q}. \end{aligned} \tag{46}$$

## 6. The Existence Conditions of the Optimal Regulator

We will verify the existence conditions of the optimal regulator for (44).

**Theorem 4.**  *$(\Phi_{R0} - G_{R0} \widehat{K}_1 \mid G_{R0})$  is stabilizable if and only if  $(\Phi_{R0} \mid G_{R0})$  is stabilizable.*

*Proof.* Notice that the system (44) is derived from the system (34) under the state feedback (43). We know that the state feedback does not change the stabilizability of the system as discussed by [17], so the system (44) is stabilizable if and only if the system (34) is stabilizable; that is,  $(\Phi_{R0} - G_{R0} \widehat{K}_1 \mid G_{R0})$  is stabilizable if and only if  $(\Phi_{R0} \mid G_{R0})$  is stabilizable. This completes the proof.  $\square$

**Theorem 5.**  *$(\Phi_{R0} \mid G_{R0})$  is stabilizable if and only if  $(\overline{A}_1^N \mid \overline{B}_1)$  is stabilizable and*

$$\begin{bmatrix} \overline{A}_1^N - I & \overline{B}_1 \\ \overline{C}_1 & \overline{C}_2 \end{bmatrix} \tag{47}$$

*is of full row rank.*

*Proof.* First, we have

$$\begin{aligned} \text{rank} [\lambda I - \Phi_{R0} \ G_{R0}] &= \text{rank} \begin{bmatrix} \lambda I - A_R & 0 & 0 \\ -G_{PR} & \lambda I - \Phi & G \end{bmatrix} \\ &= mS + \text{rank} [\lambda I - \Phi \ G]. \end{aligned} \tag{48}$$

Noticing the structure of  $\Phi$  and  $G$ , Theorem 5 can be proved by Lemma 1(a) in Liao et al. [14]. Here we omit the proof.  $\square$

Note that the matrix in (47) is  $\Psi$  in (A7).

**Theorem 6.**  *$(\overline{A}_1^N \ \overline{B}_1)$  is stabilizable if and only if (A1) holds.*

*Proof.* First, from [8], we know that  $(\overline{A}_1^N \ \overline{B}_1)$  is stabilizable if and only if the system (7) is stabilizable.

By using formula (6) and the nonsingularity of  $\begin{bmatrix} P_1 & 0 \\ 0 & I \end{bmatrix}$  and  $Q_1$ , we have

$$\begin{aligned} \text{rank} [\lambda E - (A + BK) \ B] \\ = \text{rank} \left( Q_1 [\lambda E - (A + BK) \ B] \begin{bmatrix} P_1 & 0 \\ 0 & I \end{bmatrix} \right) \end{aligned}$$

$$\begin{aligned}
 &= \text{rank} [\lambda Q_1 E P_1 - Q_1 (A + BK) P_1 \quad Q_1 B] \\
 &= \text{rank} \left[ \lambda \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \right].
 \end{aligned}
 \tag{49}$$

So, the system (7) is stabilizable if and only if the system (4) is stabilizable.

By rank  $[\lambda E - (A + BK) B] = \text{rank} [\lambda E - A \quad B] \begin{bmatrix} I & 0 \\ -K & I \end{bmatrix} = \text{rank} [\lambda E - A \quad B]$ , the system (4) is stabilizable if and only if the system (1) is stabilizable; that is, (A1) holds.

In summary, this completes the proof.  $\square$

*Remark 7.* This theorem also proves that the systems (7) and (1) have the same stabilizability.

Combining Theorems 4, 5, and 6, if the original system (1) is stabilizable and  $\Psi$  in (A7) is of full row rank, the final formal system (44) is also stabilizable. Furthermore, the condition is both necessary and sufficient. These conditions ensure that the state feedback gain in Theorem 3 exists.

Next, we examine the detectability of  $(\widehat{Q}^{1/2} \mid \Phi_{R0} - G_{R0} \widehat{K}_1)$ .

**Theorem 8.** *If (A2) holds, the system (4) is detectable.*

*Proof.* Since the output feedback does not change the detectability of the system as discussed by [2], this completes the proof.  $\square$

**Theorem 9.** *The system (4) is detectable if and only if  $(\overline{C}_1 \quad \overline{A}_1)$  is detectable.*

*Proof.* First, by the Popov-Belevitch-Hautus (PBH) rank test as discussed by [17], the system (4) is detectable if and only if, for any complex  $\lambda$  satisfying  $|\lambda| \geq 1$ ,

$$\text{rank} \begin{bmatrix} \lambda E - (A + BK) \\ C \end{bmatrix} = n \text{ (full column rank)}. \tag{50}$$

By using formula (6) and the nonsingularity of  $\begin{bmatrix} Q_1 & 0 \\ 0 & I \end{bmatrix}$  and  $P_1$ , we have

$$\begin{aligned}
 &\text{rank} \begin{bmatrix} \lambda E - (A + BK) \\ C \end{bmatrix} \\
 &= \text{rank} \left( \begin{bmatrix} Q_1 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \lambda E - (A + BK) \\ C \end{bmatrix} P_1 \right) \\
 &= \text{rank} \begin{bmatrix} \lambda Q_1 E P_1 - Q_1 (A + BK) P_1 \\ C P_1 \end{bmatrix} \\
 &= \text{rank} \left[ \lambda \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \right].
 \end{aligned}
 \tag{51}$$

This shows that the systems (4) and (7) have the same detectability.

Again from [8], the system (7) is detectable if and only if  $(\overline{C}_1 \quad \overline{A}_1)$  is detectable.

In summary, this completes the proof.  $\square$

**Theorem 10.**  *$(\overline{C}_1 \quad \overline{A}_1)$  is detectable if and only if  $(\overline{C}_1 \quad \overline{A}_1^N)$  is detectable.*

This theorem is a proven lemma in [8, 14].

**Theorem 11.** *If  $(\overline{C}_1 \quad \overline{A}_1^N)$  is detectable,  $(\widehat{Q}^{1/2} \mid \Phi_{R0} - G_{R0} \widehat{K}_1)$  is detectable.*

*Proof.* First, we have

$$\begin{aligned}
 &\text{rank} \begin{bmatrix} \lambda I - (\Phi_{R0} - G_{R0} \widehat{K}_1) \\ \widehat{Q}^{1/2} \end{bmatrix} \\
 &= \text{rank} \begin{bmatrix} \lambda I - A_R & 0 & 0 \\ I_m & (\lambda - 1) I_m & -\overline{C}_1 + \overline{C}_2 R \\ 0 & 0 & \lambda I - \overline{A}_1^N + \overline{B}_1 R \\ 0 & 0 & 0 \\ 0 & \overline{Q}^{1/2} & 0 \\ 0 & 0 & \overline{Q}^{1/2} \end{bmatrix} \\
 &= \text{rank} \begin{bmatrix} -\overline{C}_1 + \overline{C}_2 R \\ \lambda I - \overline{A}_1^N + \overline{B}_1 R \\ \overline{Q}^{1/2} \end{bmatrix}.
 \end{aligned}
 \tag{52}$$

Assuming  $V = \begin{bmatrix} H^{1/2} \overline{K}_1 \\ j \{ [H \overline{K}_2]^{-1} \}^{1/2} \overline{K}_2^T H \overline{K}_1 \end{bmatrix}$ , we have  $\overline{Q} = V^T V$ . So

$$\begin{aligned}
 &\text{rank} \begin{bmatrix} -\overline{C}_1 + \overline{C}_2 R \\ \lambda I - \overline{A}_1^N + \overline{B}_1 R \\ \overline{Q}^{1/2} \end{bmatrix} \\
 &= \text{rank} \begin{bmatrix} -\overline{C}_1 + \overline{C}_2 R \\ \lambda I - \overline{A}_1^N + \overline{B}_1 R \\ V \end{bmatrix} \\
 &= \text{rank} \begin{bmatrix} -\overline{C}_1 + \overline{C}_2 R \\ \lambda I - \overline{A}_1^N + \overline{B}_1 R \\ H^{1/2} \overline{K}_1 \\ j \{ [ \overline{K}_2^T H \overline{K}_2 ]^{-1} \}^{1/2} \overline{K}_2^T H \overline{K}_1 \end{bmatrix}
 \end{aligned}$$

$$= \text{rank} \begin{bmatrix} -\tilde{C}_1 \\ \lambda I - \tilde{A}_1^N \\ \tilde{K}_1 \\ 0 \end{bmatrix} = \text{rank} \begin{bmatrix} \lambda I - \tilde{A}_1^N \\ \tilde{C}_1 \\ \tilde{K}_1 \end{bmatrix}. \tag{53}$$

If  $(\tilde{C}_1 \ \tilde{A}_1^N)$  is detectable,  $\begin{bmatrix} \lambda I - \tilde{A}_1^N \\ \tilde{C}_1 \\ \tilde{K}_1 \end{bmatrix}$  is of full column rank.

This completes the proof.  $\square$

Combining Theorems 8 and 11, if the original system (1) is detectable,  $(\tilde{Q}^{1/2} \mid \Phi_{R0} - G_{R0}\tilde{K}_1)$  is also detectable. Furthermore, the condition is just sufficient.

### 7. The Optimal Preview Controller for the Original System

Returning to the optimal control input (45) of the descriptor augmented error system and the related formula (43), we get

$$\begin{aligned} \Delta \tilde{v}(i) &= w(i) - \tilde{K}_1 X_{R0}(i) \\ &= -[\tilde{K}_2^T H \tilde{K}_2 + G_{R0}^T P G_{R0}]^{-1} \\ &\quad \times G_{R0}^T P (\Phi_{R0} - G_{R0} \tilde{K}_1) X_{R0}(i) - \tilde{K}_1 X_{R0}(i) \\ &= \left\{ -[\tilde{K}_2^T H \tilde{K}_2 + G_{R0}^T P G_{R0}]^{-1} G_{R0}^T P (\Phi_{R0} - G_{R0} \tilde{K}_1) - \tilde{K}_1 \right\} \\ &\quad \times X_{R0}(i) = T X_{R0}(i), \end{aligned} \tag{54}$$

where  $T = -[\tilde{K}_2^T H \tilde{K}_2 + G_{R0}^T P G_{R0}]^{-1} G_{R0}^T P (\Phi_{R0} - G_{R0} \tilde{K}_1) - \tilde{K}_1$ .  
From (21) and (54), we continue to get

$$\begin{aligned} \Delta \tilde{u}(i) &= \tilde{K}_1 \Delta \tilde{x}_1(i) + \tilde{K}_2 \Delta \tilde{v}(i) \\ &= \tilde{K}_1 \Delta \tilde{x}_1(i) + \tilde{K}_2 T X_{R0}(i) \\ &= \tilde{K}_1 [0 \ 0 \ I_q] X_{R0}(i) + \tilde{K}_2 T X_{R0}(i) \\ &= \{ \tilde{K}_1 [0 \ 0 \ I_q] + \tilde{K}_2 T \} X_{R0}(i) \\ &= \{ \tilde{K}_1 [0 \ 0 \ I_q] + \tilde{K}_2 T \} \begin{bmatrix} X_R(i) \\ \tilde{e}(i) \\ \Delta \tilde{x}_1(i) \end{bmatrix} \\ &= \hat{T} \begin{bmatrix} X_R(i) \\ \tilde{e}(i) \\ \Delta \tilde{x}_1(i) \end{bmatrix}, \end{aligned} \tag{55}$$

where  $\hat{T} = \tilde{K}_1 [0 \ 0 \ I_q] + \tilde{K}_2 T$ .

Noticing

$$X_R(i) = \begin{bmatrix} \Delta \tilde{R}(i) \\ \Delta \tilde{R}(i+1) \\ \vdots \\ \Delta \tilde{R}(i+S-1) \end{bmatrix}, \tag{56}$$

$\hat{T}$  is partitioned into

$$\hat{T} = [T_R(0) \ T_R(1) \ \cdots \ T_R(S-1) \mid T_e \mid T_x]. \tag{57}$$

Equation (55) can be written as

$$\Delta \tilde{u}(i) = \sum_{l=0}^{S-1} T_R(l) \Delta \tilde{R}(i+l) + T_e \tilde{e}(i) + T_x \Delta \tilde{x}_1(i). \tag{58}$$

Noticing

$$\Delta \tilde{u}(i) = \begin{bmatrix} \Delta u(iN) \\ \Delta u(iN+1) \\ \vdots \\ \Delta u(iN+N-1) \end{bmatrix}, \tag{59}$$

$T_R(l)$ ,  $T_e$ , and  $T_x$  are decomposed into

$$T_R(l) = \begin{bmatrix} T_R^{(0)}(l) \\ T_R^{(1)}(l) \\ \vdots \\ T_R^{(N-1)}(l) \end{bmatrix}, \quad T_e = \begin{bmatrix} T_e^{(0)} \\ T_e^{(1)} \\ \vdots \\ T_e^{(N-1)} \end{bmatrix}, \quad T_x = \begin{bmatrix} T_x^{(0)} \\ T_x^{(1)} \\ \vdots \\ T_x^{(N-1)} \end{bmatrix} \tag{60}$$

$(l = 0, 1, \dots, S-1).$

Then (58) can be written as

$$\begin{aligned} \begin{bmatrix} \Delta u(iN) \\ \Delta u(iN+1) \\ \vdots \\ \Delta u(iN+N-1) \end{bmatrix} &= \sum_{l=0}^{S-1} \begin{bmatrix} T_R^{(0)}(l) \\ T_R^{(1)}(l) \\ \vdots \\ T_R^{(N-1)}(l) \end{bmatrix} \Delta \tilde{R}(i+l) \\ &+ \begin{bmatrix} T_e^{(0)} \\ T_e^{(1)} \\ \vdots \\ T_e^{(N-1)} \end{bmatrix} \tilde{e}(i) \\ &+ \begin{bmatrix} T_x^{(0)} \\ T_x^{(1)} \\ \vdots \\ T_x^{(N-1)} \end{bmatrix} \Delta \tilde{x}_1(i). \end{aligned} \tag{61}$$

The above equation can be further written as

$$\Delta u(iN + j) = \sum_{l=0}^{S-1} T_R^{(j)}(l) \Delta \tilde{R}(i+l) + T_e^{(j)} \tilde{e}(i) + T_x^{(j)} \Delta \tilde{x}_1(i). \tag{62}$$

That is,

$$u((i+1)N + j) = u(iN + j) + \sum_{l=0}^{S-1} T_R^{(j)}(l) \Delta \tilde{R}(i+l) + T_e^{(j)} \tilde{e}(i) + T_x^{(j)} \Delta \tilde{x}_1(i). \tag{63}$$

If  $i$  is substituted by  $i - 1$ , we obtain the control input of the most important theorem.

**Theorem 12.** *If (A1)–(A7) hold and  $Q_e > 0$  and  $H_u > 0$ , then the Riccati equation (46) has a unique symmetric semipositive definite solution, and the optimal control input of the system (1) is*

$$u(iN + j) = u((i-1)N + j) + \sum_{l=0}^{S-1} T_R^{(j)}(l) \Delta \tilde{R}(i+l-1) + T_e^{(j)} \tilde{e}(i-1) + T_x^{(j)} [x_1(iN) - x_1((i-1)N)]$$

$$i = 1, 2, \dots; j = 0, 1, 2, \dots, N-1, \tag{64}$$

where

$$\Delta \tilde{R}(i-1) = \begin{bmatrix} R(iN) - R((i-1)N) \\ R(iN+1) - R((i-1)N+1) \\ \vdots \\ R(iN+N-2) - R((i-1)N+N-2) \\ R(iN+N-1) - R((i-1)N+N-1) \end{bmatrix},$$

$$\tilde{e}(i-1) = \begin{bmatrix} e((i-1)N) \\ e((i-1)N+1) \\ \vdots \\ e((i-1)N+N-2) \\ e((i-1)N+N-1) \end{bmatrix} \tag{65}$$

are determined by

$$\tilde{e}(i-1) = \tilde{C}_1 \tilde{x}_1(i-1) - \tilde{R}(i-1) + \tilde{C}_2 \tilde{v}(i-1), \tag{66}$$

where

$$\tilde{x}_1(i-1) = x_1((i-1)N),$$

$$\tilde{R}(i-1) = \begin{bmatrix} R((i-1)N) \\ R((i-1)N+1) \\ \vdots \\ R((i-1)N+N-2) \\ R((i-1)N+N-1) \end{bmatrix}. \tag{67}$$

In addition,  $\tilde{v}(i-1)$  can be derived from (21) and (A6) as follows:

$$\tilde{v}(i-1) = \tilde{K}_2^{-1} \tilde{u}(i-1) - \tilde{K}_2^{-1} \tilde{K}_1 \tilde{x}_1(i-1), \tag{68}$$

where

$$\tilde{u}(i-1) = \begin{bmatrix} u((i-1)N) \\ u((i-1)N+1) \\ \vdots \\ u((i-1)N+N-2) \\ u((i-1)N+N-1) \end{bmatrix}. \tag{69}$$

### 8. Numerical Example

Consider the following regular linear discrete-time descriptor noncausal system in the form of (1):

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} x(k+1) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -2 & 2 \\ 0 & 0 & -1 \end{bmatrix} x(k) + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} u(k),$$

$$y(k) = [1 \ -1 \ 1] x(k). \tag{70}$$

In this case, the coefficient matrices are

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -2 & 2 \\ 0 & 0 & -1 \end{bmatrix},$$

$$B = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad C = [1 \ 0 \ 1], \tag{71}$$

respectively.

Through calculating, the above system satisfies all conditions required in the paper. By MATLAB simulation, the gain matrix in output feedback is taken as  $M = 2$ , and the coefficient matrices in (7) are

$$A_{11} = \begin{bmatrix} 2 & -1 \\ 1 & -2 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad A_{21} = [2 \ -2],$$

$$A_{22} = 3, \quad B_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad B_2 = 1, \tag{72}$$

$$C_1 = [1 \ -1], \quad C_2 = 2.$$

We assume that  $N = 3$  in (A4). To calculate  $\bar{A}_1, \bar{B}_1, \bar{C}_1, \bar{C}_2$ , and  $\Psi$  give

$$\bar{A}_1 = \begin{bmatrix} 0 & 1 \\ -1.6667 & 0.6667 \end{bmatrix}, \quad \bar{B}_1 = \begin{bmatrix} 0 \\ -1.3333 \end{bmatrix},$$

$$\bar{C}_1 = [-0.3333 \ 0.3333], \quad \bar{C}_2 = -0.6667,$$

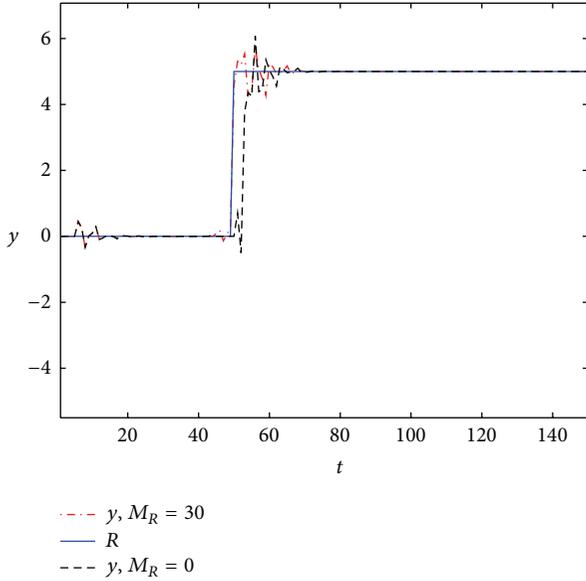


FIGURE 1: The output response to step function.

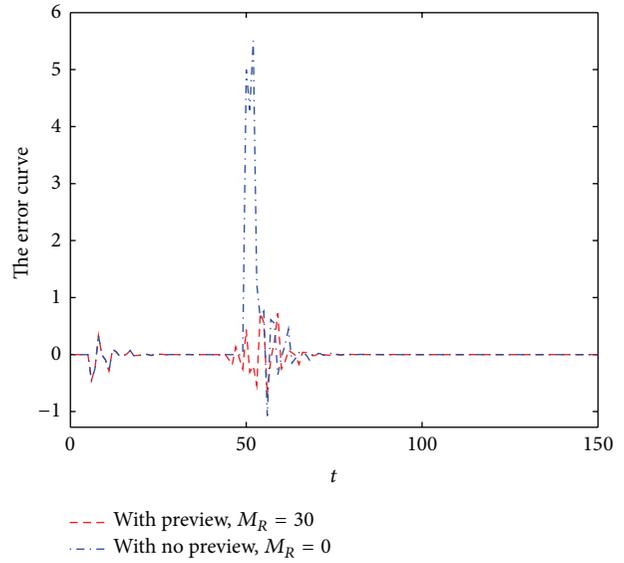


FIGURE 2: The error of step function.

$$\Psi = \left[ \begin{array}{c|ccc} \overline{A}_1^3 - I & \overline{A}_1^2 \overline{B}_1 & \overline{A}_1 \overline{B}_1 & \overline{B}_1 \\ \hline \overline{C}_1 & \overline{C}_2 & 0 & 0 \\ \overline{C}_1 \overline{A}_1 & \overline{C}_1 \overline{B}_1 & \overline{C}_2 & 0 \\ \hline \overline{C}_1 \overline{A}_1^2 & \overline{C}_1 \overline{A}_1 \overline{B}_1 & \overline{C}_1 \overline{B}_1 & \overline{C}_2 \end{array} \right]$$

$$= \left[ \begin{array}{cc|ccc} -2.1111 & -1.2222 & -0.8889 & -1.3333 & 0 \\ 2.0370 & -2.9259 & 1.6296 & -0.8889 & -1.3333 \\ \hline -0.3333 & 0.3333 & -0.6667 & 0 & 0 \\ -0.5556 & -0.1111 & -0.4444 & -0.6667 & 0 \\ 0.1852 & -0.6296 & 0.1481 & -0.4444 & -0.6667 \end{array} \right]. \tag{73}$$

Let the initial state vector  $\tilde{x}_1(0) = \begin{bmatrix} -2 \\ -1 \end{bmatrix}$ . In addition, take the weight matrices  $Q_e = 100$  and  $H_u = 10$ . Let preview length be  $M_R = 30$ ; that is,  $S = 10$ . We present MATLAB simulation results for two cases.

(1) *Step Function.* Let the desired signal be

$$R(k) = \begin{cases} 0, & k \leq 50 \\ 5, & k > 50. \end{cases} \tag{74}$$

By MATLAB simulation, the output response of the linear discrete-time descriptor noncausal multirate system (with preview action and no preview action) is shown in Figure 1. The error signals are shown in Figure 2. Note that the preview action significantly reduces the error. In particular, the error signal is asymptotically zero.

(2) *Ramp Function.* Let the desired signal be

$$R(k) = \begin{cases} 0, & k \leq 30 \\ 0.25(k - 30), & 30 < k \leq 50 \\ 5, & k > 50. \end{cases} \tag{75}$$

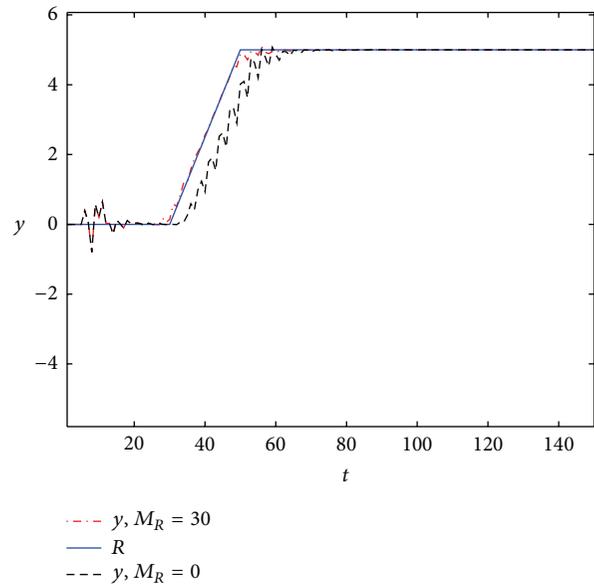


FIGURE 3: The output response to ramp function.

The output responses are shown in Figure 3. The error signals are shown in Figure 4.

From Figures 1–4, we can easily see the effectiveness of the present controller of this paper. On the one hand, when using preview control, the output curve can track the desired signal faster; on the other hand, the overshoot is smaller.

### 9. Conclusion

This paper studied the optimal preview controller for linear discrete-time descriptor noncausal multirate systems. By making use of the characteristics of causal controllability and causal observability, the original system was converted into

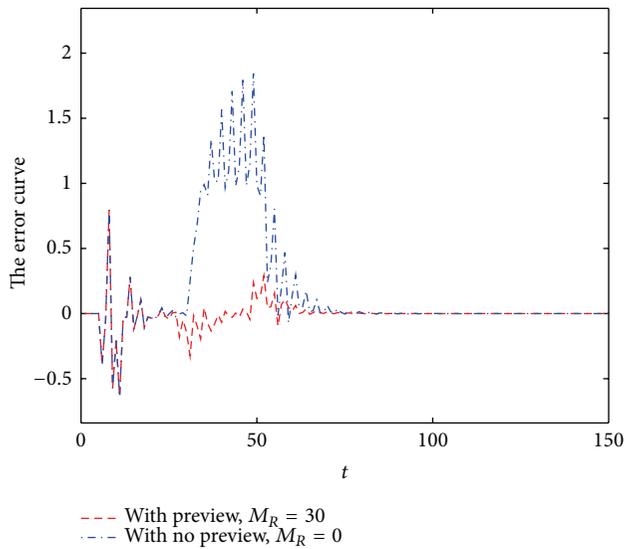


FIGURE 4: The error of ramp function.

a descriptor causal closed-loop system. Then, using the characteristics of a causal system and a discrete lifting technique, the descriptor causal closed-loop multirate system was changed into a single-rate normal system. Taking advantage of the conventional method of the error system in preview control theory, a descriptor augmented error system is constructed, and the problem is transformed into a regulator problem. Finally, the optimal preview controller is designed according to the related theory of preview control. From preview control theory, the obtained closed-loop system contains an integrator so that the response of the system does not have static error. The numerical simulation showed the effectiveness of the proposed preview control system.

## Acknowledgments

This work is supported by the National Natural Science Foundation of China (no. 61174209) and the Oriented Award Foundation for Science and Technological Innovation, Inner Mongolia Autonomous Region, China (no. 2012).

## References

- [1] L. Y. Dai, *Singular Control Systems*, Lecture Notes in Control and Information Science, Springer, New York, NY, USA, 1989.
- [2] D. M. Yang, Q. L. Zhang, and B. Yao, *The Descriptor Systems*, Science Press, Beijing, China, 2004.
- [3] J.-B. Sun, C.-J. Zhang, and B. Guo, "Optimal fusion full-order estimators for discrete-time stochastic singular systems," *Control and Decision*, vol. 25, no. 2, pp. 263–268, 2010.
- [4] N. Cao, H.-G. Zhang, Y.-H. Luo, D.-Z. Feng, and Y. Liu, "Suboptimal control of a class of nonlinear singularly perturbed systems," *Control Theory & Applications*, vol. 28, no. 5, pp. 688–692, 2011.
- [5] Z. Su, Q. L. Zhang, and J. Ai, "Finite-time robust dissipative control for a class of descriptor systems," *Journal of Northeastern University*, vol. 33, no. 9, pp. 1217–1221, 2012.
- [6] Y. M. Zhang, Q. L. Zhang, T. Tanaka, and M. Cai, "Admissibility for positive continuous-time descriptor systems," *International Journal of Systems Science*, vol. 44, no. 11, pp. 2158–2165, 2013.
- [7] F. C. Liao, M. J. Cao, Z. X. Hu, and P. D. An, "Design of an optimal preview controller for linear discrete-time causal descriptor systems," *International Journal of Control*, vol. 85, no. 10, pp. 1616–1624, 2012.
- [8] F. C. Liao, M. Tomizuka, M. J. Cao, and D. Wang, "Optimal preview control for discrete-time descriptor causal systems in a multirate setting," *International Journal of Control*, vol. 86, no. 5, pp. 844–854, 2013.
- [9] M. J. Cao and F. C. Liao, "Design of an optimal preview controller for linear discrete-time descriptor systems with state delay," *International Journal of Systems Science*, 2013.
- [10] D. J. Bender and A. J. Laub, "The linear-quadratic optimal regulator for descriptor systems: discrete-time case," *Automatica*, vol. 23, no. 1, pp. 71–85, 1987.
- [11] G.-S. Zhang and L. Liu, "Linear quadratic optimal control based on dynamic compensation for rectangular descriptor systems," *Acta Automatica Sinica*, vol. 36, no. 12, pp. 1752–1757, 2010.
- [12] C. Zhaolin, H. Huimin, and Z. Jifeng, "The optimal regulation of generalized state-space systems with quadratic cost," *Automatica*, vol. 24, no. 5, pp. 707–710, 1988.
- [13] J. Xiao, *Multirate Digital Control Systems*, Science Press, Beijing, China, 2003.
- [14] F. C. Liao, K. Takaba, and T. Katayama, "Design of an optimal preview servomechanism for discrete-time systems in a multirate setting," *Dynamics of Continuous, Discrete and Impulsive Systems B*, vol. 10, no. 5, pp. 727–744, 2003.
- [15] F. C. Liao and H. P. Liu, "Design of an optimal preview controller for a kind of discrete-time systems," *Journal of University of Science and Technology Beijing*, vol. 29, no. 5, pp. 542–547, 2007.
- [16] Q.-S. Shi and F.-C. Liao, "Design of an optimal preview controller for linear discrete-time multirate systems with state-delay," *Journal of University of Science and Technology Beijing*, vol. 33, no. 3, pp. 363–375, 2011.
- [17] G. R. Duan, *Linear System Theory*, Harbin Institute of Technology Press, Harbin, China, 1998.