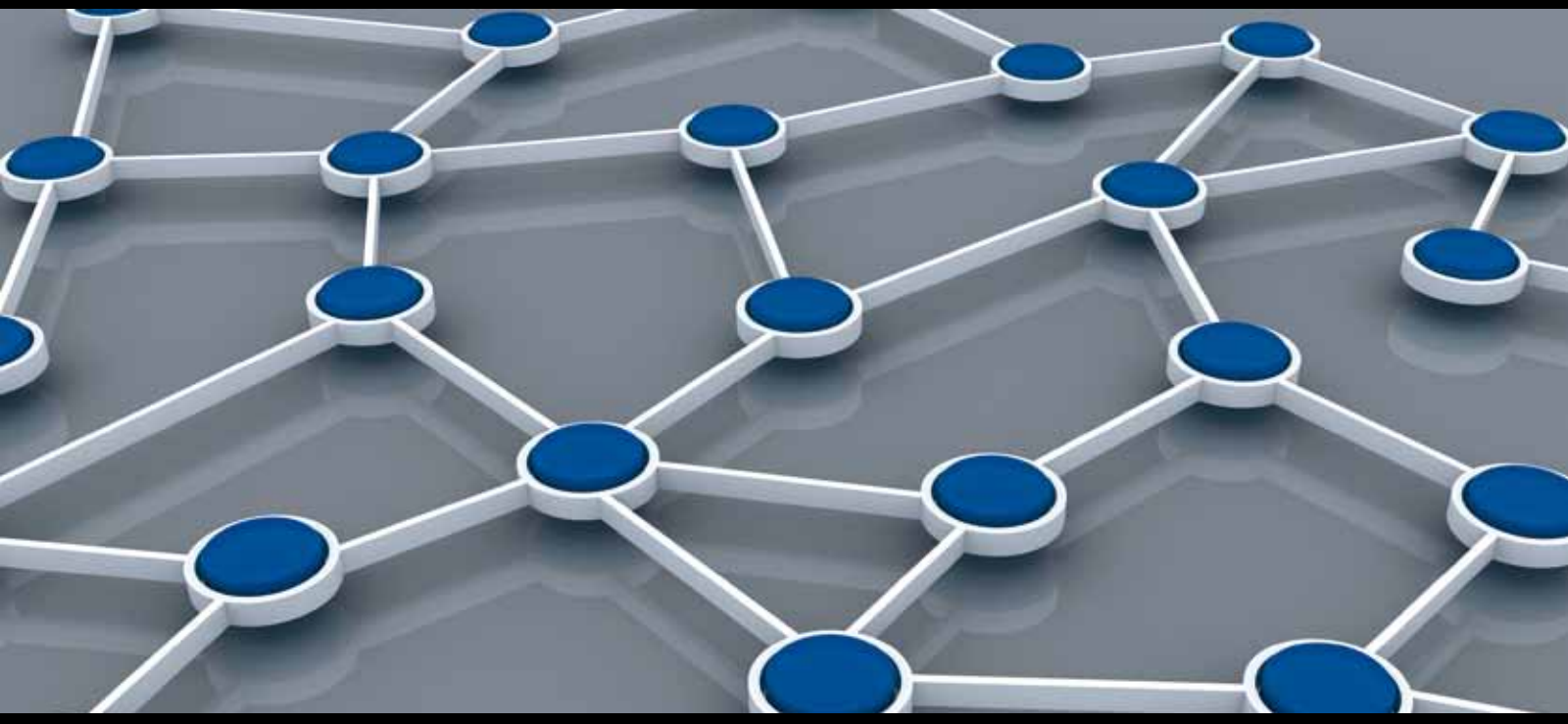
International Journal of Distributed Sensor Networks

# Wireless Ad Hoc Sensor Networks

Guest Editors: Carlos León, Julio Barbancho, Pawel Kulakowski, Adel Soudani, Pedro Marrón, and Chia-Yen Shih



# Wireless Ad Hoc Sensor Networks

International Journal of Distributed Sensor Networks

# Wireless Ad Hoc Sensor Networks

Guest Editors: Carlos León, Julio Barbancho, Pawel Kulakowski, Adel Soudani, Pedro Marrón, and Chia-Yen Shih

Copyright © 2013 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in "International Journal of Distributed Sensor Networks." 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.

# Editorial Board


Prabir Barooah, USA
R. R. Brooks, USA
P. Chatzimisios, Greece
W.-Y. Chung, Republic of Korea
George P. Efthymoglou, Greece
Frank Ehlers, Italy
Tian He, USA
Chin-Tser Huang, USA
Baoqi Huang, China
S. S. Iyengar, USA
Rajgopal Kannan, USA
Miguel A. Labrador, USA
Joo-Ho Lee, Japan
Shijian Li, China
Minglu Li, China
Shuai Li, USA
Weifa Liang, Australia
Jing Liang, China
Wen-Hwa Liao, Taiwan
Alvin S. Lim, USA
Donggang Liu, USA
Yonghe Liu, USA
Zhong Liu, China
Seng Loke, Australia
Jun Luo, Singapore
J. R. Martinez-de Dios, Spain
S. N. Merchant, India
A. Milenkovic, USA
E. F. Nakamura, Brazil
Peter C. Ölveczky, Norway
M. Palaniswami, Australia
Shashi Phoha, USA
Hairong Qi, USA
Joel Rodrigues, Portugal
Jorge Sa Silva, Portugal
Sartaj K. Sahni, USA
Weihua Sheng, USA
Sheng Wang, China
Zhi Wang, China
Qishi Wu, USA
Qin Xin, Faroe Islands
Jianliang Xu, Hong Kong
Yuan Xue, USA
Fan Ye, USA
Ning Yu, China
Tianle Zhang, China
Yanmin Zhu, China

# Contents

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 202940, 2 pages
http://dx.doi.org/10.1155/2013/202940

# *Editorial*
# Wireless Ad Hoc Sensor Networks

**Carlos León,[1] Julio Barbancho,[1] Pawel Kulakowski,[2] Adel Soudani,[3] Pedro Marrón,[4] and Chia-Yen Shih[4]**

[1] *Departamento de Tecnología Electrónica, Universidad de Sevilla, Virgen de África 7, 41011 Sevilla, Spain*
[2] *AGH University of Science and Technology, Aleja Adama Mickiewicza 30, 30-962 Krakow, Poland*
[3] *University of Monastir, Avenue de l'Environnement, 5000 Monastir, Tunisia*
[4] *University of Duisburg-Essen and Fraunhofer FKIE, Forsthausweg 2, 47057 Duisburg, Germany*

Correspondence should be addressed to Julio Barbancho; jbarbancho@us.es

Received 2 September 2013; Accepted 2 September 2013

Copyright © 2013 Carlos León 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.

Wireless Ad Hoc Sensor Networks (WAdSNs) have become an important research area in the last ten years. Main contributions have focused on the development of new hardware, software, and protocols supporting distributed applications. In these kinds of networks, cooperation is a key issue, especially when nodes collaborate with their neighbors and with the whole network, considering it as a single entity. This paradigm can be used to face complex problems with high requirements: ease of deployment, self-configuration, and self-repair, among others. This special issue is among these sorts of applications.

Considering the point of view of network cooperation, the papers that form this special issue could be classified in four important categories.

The first category deals with routing purposes. Four papers could be included in this category.

The paper titled "*Energy-efficient routing algorithms based on OVSF code and priority in clustered wireless sensor networks,*" by X. Wu et al., describes an improved protocol based on the classic routing protocol LEACH. Different studies are done to analyze several environments where wireless sensor networks are deployed.

Another routing algorithm is proposed in the paper titled "*Energy-aware routing in wireless sensor networks using local betweenness centrality*", by X.-H. Li et al. This proposal uses local betweenness centrality to estimate the energy consumption of the neighborhood. The main goal is to provide balanced energy consumption in wireless sensor networks. The results provided demonstrate that there exist several advantages using this algorithm.

The paper titled "*Towards efficient and secure geographic routing protocol for hostile wireless sensor network*," by C. Lyu et al., proposes a new routing protocol, named ESGR, that exploits the geographic location, cryptography mechanisms, and broadcast wireless channel. The authors examine the impact of a wide variety of attacks in malicious wireless sensor network scenarios and demonstrate that ESGR avoids a specific variety of attacks and ensures high packet delivery rate in malicious sensor network environment.

The paper "*A credible routing based on a novel trust mechanism in ad hoc networks*," by R. Feng et al., provides a novel routing algorithm for Mobile Ad hoc Networks (MANETs) focusing on finding paths in dynamic networks and considering security.

The second category deals with network formation and maintenance.

The paper "*A nonuniform sensor distribution strategy for avoiding energy holes in wireless sensor networks*," by G. Ma and Z. Tao, presents a nonuniform sensor distribution strategy based on unequal cluster for WSNs whose main objective is to resolve the energy hole problem to improve the performance of multihop communications.

Another approximation to sensor network formation is described in "*A design approach for controlled self-organization-based sensor networks focused on control timescale*," by D. Kominami and M. Murata. The authors propose and

evaluate a design for the network formation and maintenance supervised by control timescale.

The third category deals with network cooperation for data transport.

The paper "*Implementing a distributed WSN based on IPv6 for ambient monitoring,*" by D. F. Larios et al., evaluates different communication protocols distributed and centralized, in order to determine the best trade-off for environmental monitoring in different migratory areas of waterbirds. The results demonstrate that the use of fully distributed algorithms, such as IPv6 over WSNs, could be suitable for certain kind of cooperative applications.

A reliable collection protocol for aggregating data packets from all the sensor nodes to the sink in a large-scale WSN is presented in "*A reliable data collection protocol based on erasure-resilient code in asymmetric wireless sensor networks*," by J.-J. Lei et al.

The paper "*Improved reliable trust-based and energy-efficient data aggregation for wireless sensor networks*", by C.-x. Liu et al., presents an improved, reliable, trust-based, and energy-efficient data-aggregation protocol for wireless sensor networks.

The fourth and last category deals with the security of the information. In this sense the last two papers included in the first category and the last one included in the third category could be understood also under the point of view of the security.

Another paper focused on security is presented in this special issue as an example of networking cooperation. This paper is titled "*On the security of certificateless signature schemes,*" by G. Sharma et al. The authors examine the use of certificateless public key cryptography in wireless sensor networks. The results prove that this security scheme has some vulnerabilities in front of certain kind of malicious attacks.

The papers included in this special issue deal with four important topics about Wireless Ad Hoc Sensor Networks. We hope that they can improve the design and development of this kind of networks.

## Acknowledgments

The guest editors would like to thank the authors for the great level of the contributions included in this special issue and the work developed by many experts who participated in the review process providing constructive comments to the authors to improve the quality of the papers.

*Carlos León*
*Julio Barbancho*
*Pawel Kulakowski*
*Adel Soudani*
*Pedro Marrón*
*Chia-Yen Shih*

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 792315, 7 pages
http://dx.doi.org/10.1155/2013/792315

*Research Article*

# A Supermodular Game Framework for Power Control of Wireless Sensor Networks

**Zhide Chen,[1] Yanqing Zeng,[2] and Li Xu[1]**

[1] *School of Mathematics and Computer Science, Fujian Normal University, Key Lab of Network Security and Cryptography, Fuzhou, Fujian 350007, China*
[2] *Department of Information Management and Engineering, Fujian Commercial College, Fuzhou, Fujian 350012, China*

Correspondence should be addressed to Zhide Chen; zhidechen@fjnu.edu.cn

Received 21 December 2012; Revised 15 July 2013; Accepted 16 July 2013

Academic Editor: Julio Barbancho

Copyright © 2013 Zhide Chen 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 consider a distributed power control scheme for wireless sensor networks. To derive decentralized solutions that do not require complete information and any cooperation among the users, we formulate this problem as a supermodular game, which each user maximizes its utility function provided transmission rate constraints. Through analyzing the supermodular property of the game, the existence and uniqueness of the Nash equilibrium (NE) are established. Furthermore, we propose a distributed price and power update algorithm (DPPA) to compute the solution of the game which is based on myopic best response. Performance evaluations via numerical simulations verify the existence of theNE and the convergence property of the DPPA algorithm.

## 1. Introduction

Wireless sensor networks (WSN) have broad application prospects, that is, widely used in health care, military, environmental monitoring, and forecasting, as well as intelligent home, building condition monitoring, and so on. The power control problem with interference management, throughput maximization, or energy efficient target, is an interesting issue and attracts tremendous attention [1–6]. Since different power levels result in a different performance and fundamentally affect many aspects of the operation of the networks, the power control problem becomes a complex and intriguing problem. Since the sensor networks are totally distributed and have no fundamental infrastructures, the traditional centralized power control mechanisms are not feasible. It is well known that game theory is a good tool to handle the distributed problems. Game theory techniques have widely been applied to engineering problems in which the action of one component has impact on and perhaps conflicts with that of any other component [7]. More particularly, supermodular games are interesting since they have several desirable properties, such as they encompass many applied models, have the remarkable property that many solution concepts yield the same predictions and have nice comparative statics properties and behave well under various learning rules [8].

In this paper, we study the power control problem, focus on distributed algorithms with no centralized control using a supermodular game framework, and investigate if any optimality is achievable. Our work is motivated by the following points.

(i) The networks have no central infrastructures; a distributed power control mechanism is desired.

(ii) The game theoretic formulation is a useful tool to design distributed algorithms. And it is amenable to prove the convergence for the distributed algorithm.

(iii) Supermodular game model has several desirable properties.

To formulate the power control design, we take the rate constraints into account and propose a strategic game, where all rational users maximize their own utilities by choosing appropriate power levels. We refer this game to distributed power control game (DPCG). By building the supermodular property of the DPCG, the existence of the Nash equilibrium solution is established. Furthermore, we

present an analysis for users' myopic best response (MBR) dynamics. Based on the MBR dynamics, a distributed price and power update algorithm (DPPA) is proposed. Then we prove the convergence of the algorithm and show the uniqueness of the NE.

The rest of this paper is organized as follows. In Section 2, the related works are discussed. We depict the system model in Section 3. In Section 4, the proof of the existence of the NE is provided. The proposed pricing mechanism and the distributed convergence algorithm are presented in Section 5. The simulation results are given in Section 6. Finally, Section 7 draws the conclusions.

## 2. Related Work

There has been a rich literature to study power control in wireless networks. In [9], the authors proposed a joint power and channel resource allocation iterative optimization algorithm. References [1–6] illustrate power allocation problems using traditional game theory and take different factors into account. Evolutionary game theory was used to depict the power problem in [10, 11]. In particular [10] established an evolutionary power mechanisms for W-CDMA and WIM wireless systems. With the analytical limitations of evolutionary game, we can just analyze low-dimension strategy space. Shamik et al. proposed a game theoretic framework for power control in wireless sensor network. They showed the performance differences between continuous levels and discrete levels and gave a framework to find the best transmit power span. However, they did not show the convergence of the power control mechanism. In [12], they used potential game theory and replicator dynamics learning scheme to analyze the distributed power allocation problem in parallel multiple-access channels, and they depicted the sufficient conditions for the unique NE and showed the convergence property. The interaction of several radio devices aiming to obtain wireless connectivity by using a set of base stations was modeled as a noncooperative game in [13]. They showed the existence of the Nash equilibrium of two situations, that is, BS selection and BS sharing. The authors of [14] considered the power control for open-loop overlaid network MIMO systems in a game theoretical perspective. They viewed the problem as a noncooperative game, and the numerical simulations verified the significant performance advantages of the proposed scheme.

## 3. System Model

*3.1. Link Capacity.* The links between nodes are modeled by a set $\iota = \{1, 2, \ldots, L\}$. The channels are assumed to be additive white Gaussian noise (AWGN) channel with noise power spectral density $N_0$ over the bandwidth of operation $B$. The channel gain from link $i$'s transmitter to the link $j$'s receiver is denoted by $G_{ij}$. The interference from link $i$'s transmitter to link $j$'s receiver is denoted as $G_{ij}p_i$. Then, the total interference and noise power at the link $j$'s receiver is as follows:

$$\iota_j = \sum_{i \neq j} G_{ij} p_i + N_0 B. \tag{1}$$

Therefore, the instantaneous signal to interference plus noise ratio (SINR) of node $j$ is

$$\gamma_j = \frac{G_{jj} p_j}{\iota_j}. \tag{2}$$

Denote the link capacity of link $j$ by $R_j$; we can get the link capacity that can be supported over link $j$ as follows:

$$R_j = B \log\left(1 + \gamma_j\right). \tag{3}$$

To guarantee the communication quality, we assume that

$$R_j \geq R_j^*, \tag{4}$$

where $R_j^*$ is the rate constraint of link $j$, that is, the minimum transmission rate required by each link.

*3.2. Distributed Power Control Game.* The power control game is modeled as a strategic game, in which the players are the links and the payoff is the difference between their transmission gain and the power consumption cost. To maximize its own utility, each player chooses transmission power with a given constraint on the minimum achievable information rate to compete against others. The strategy of each player is one-dimension subset of $R$. The strategy of player $i$ is denoted by $p_i$, and the strategy of player $i$'s opponents is defined as $p_{-i} = (p_1, \ldots, p_{i-1}, p_{i+1}, \ldots, p_L)$. Then the game has the following structure:

$$\vartheta = \left\{\iota, \{P_i\}_{i \in \iota}, \{u_i(p_i)\}_{i \in \iota}\right\}, \tag{5}$$

where $\iota = \{1, 2, \ldots, L\}$ is the set of links and $p_i$ is the power allocation strategy of player $i$, which is defined as

$$p_i \triangleq \left\{p_i \in P_i : R_i(p_i, p_{-i}) \geq R_i^*\right\}, \tag{6}$$

where $R_i(p_i, p_{-i}) = B \log(1 + (G_{ii}p_i/\iota)_i)$. As defined above, $R_i^*$ is the transmission rate constraint. We assume that $R_i^* \geq B$ for all links. $P_i$ denotes the action set for player $i$ as follows

$$P_i = \left[p_{i,\min}, p_{i,\max}\right], \tag{7}$$

where $p_{i,\min} = 0$ means that user $i$ decides not to transmit. We denote the vector $R^* \triangleq (R_i^*)_{i \in \iota}$ by the rate constraint profile. The payoff function of user $i$ is defined as follows:

$$u_i(p_i) = B \log(\gamma_i) - c_i p_i, \tag{8}$$

where $c_i$ is the pricing parameter for user $i$. We use $\log(\gamma_i)$ to replace $\log(\gamma_i + 1)$ due to the following reasons: in the high SINR regime, logarithmic utility approximates the Shannon capacity $B\log(1 + \gamma_i)$ and for low SINR, a user's rate is approximately linear in SINR, and so this utility is proportional to the logarithm of the rate. Each user's feasible strategies depend not only on their own action, but also on their opponents'.

Given the power allocation of the others $p_{-i}$, the optimal strategy for the $i$th user is the solution of the following maximization problem:

$$\max_{p_i} u_i(p_i, p_{-i}), \quad i \in \iota, \quad \text{subject to } p_i \in P_i. \tag{9}$$

We refer to the game $\vartheta$ with maximization problem (9) by distributed power control game (DPCG). The strategy space of the game is $P = P_1 \times \cdots \times P_L$.

Now we define the Nash equilibrium of this game. An NE of the DPCG game $\vartheta$ is defined as follows.

*Definition 1.* Consider an $L$-player game, each player maximizing its individual cost function $u_i : P_i \to R_+$, subject to coupled inequality constraint $R_i^* - R_i(p_i, p_{-i}) \le 0$, a vector $p^* = (p_i^*, p_{-i}^*)$ is called an NE solution of DPCG $\vartheta$ if every given $p_{-i}^*$ is as follows:

$$u_i\left(p_i^*\right) \ge u_i\left(p_i\right), \quad \forall p_i \in P_i, \ \forall i \in \iota. \tag{10}$$

In the following sections, we will discuss whether the NE solution exists. If it exists, is it unique and how to reach it in a totally distributed way?

## 4. Existence of the NE

Herein, we prove that the DPCG game has the NE solution by establishing the supermodular property for it. By verifying the Hessian matrix of the game, we prove the uniqueness of the NE. Following [8], we have the following.

*Definition 2* (supermodular game). The strategic form game $\langle I; (S_i); (u_i)\rangle$ is a supermodular game if for all $i$ one has the following:

(1) $S_i$ is a compact subset of $R$;

(2) $u_i$ is upper semicontinuous in $s_i$, continuous in $s_{-i}$;

(3) $u_i$ has increasing differences in $(s_i; s_{-i})$.

With the above definition, we can get the following.

**Theorem 3.** *The game $\vartheta$ defined by (5)–(9) is supermodular and admits a compact sublattice NE solution.*

*Proof.* Apparently, the strategy set $P_i$ of user $i$ is a sublattice of $R$; that is, for all $p_i, p_j \in P_i$, $\max(p_i, p_j) \in P_i$ and $\min(p_i, p_j) \in P_i$. For the game DPCG's utility,

$$u_i\left(p_i\right) = B\log\left(\frac{G_{ii}p_i}{\sum_{j\ne i} G_{ji}p_j + N_0 B}\right) - c_i p_i. \tag{11}$$

Let $f(\gamma_i) = B\log(\gamma_i)$; then we have

$$\frac{-\gamma_i f_i''\left(\gamma_i\right)}{f_i'\left(\gamma_i\right)} = 1 \ge 1. \tag{12}$$

For all $j \ne i$,

$$\frac{\partial^2 u_i}{\partial p_i \partial p_j}\left(p_i, p_{-i}\right) = -\frac{\gamma_i^2 G_{ji}}{p_i^2 G_{ii}}\left[\gamma_i f''\left(\gamma_i\right) + f'\left(\gamma_i\right)\right]. \tag{13}$$

Obviously, $(\partial^2 u_i/\partial p_i \partial p_j)(p_i, p_{-i}) \ge 0$. It indicates that the utility function of the DPCG meets increasing differences. According to Definition 2, we conclude that the DPCG is a supermodular game. Based on Theorem 1 in [15], the set of the NEs of game $\vartheta$ is a nonempty and compact sublattice, and it admits an NE solution. □

## 5. Distributed Convergence Algorithm and Uniqueness of the NE

*5.1. Pricing Mechanism.* In the above section, we did not depict the pricing parameter precisely. Herein, we define the price charged to other users for generating interference to user $i$ as follows:

$$\pi_i\left(p_i, p_{-i}\right) = -\frac{\partial u_i\left(p_i, p_{-i}\right)}{\partial I_i}, \tag{14}$$

where $I_i = \sum_{j\ne i} p_j G_{ji}$ is the total interference that is created by opponents received by user $i$. Obviously, $\pi_i(p_i, p_{-i})$ is always nonnegative and represents user $i$'s marginal increase in utility per unit and decrease in total opponents' interference. User $i$ therefore maximizes the difference between its gain minus its cost to the other users in the network due to the interference it creates. The cost is its transmit power times a weighted sum of other users' prices, where the weights equal the channel gains between user $i$'s transmitter and the other links' receivers. Then the maximization problem becomes that each user $i$ specifies a power level $p_i \in P_i$ to maximize the following utility function:

$$u_i\left(p_i, p_{-i}\right) = B\log\left(\gamma_i\right) - p_i \sum_{j\ne i} \pi_j G_{ij}. \tag{15}$$

*5.2. Distributed Update Algorithm.* In Section 4, the DPCG game has been shown to admit an NE. Now we focus on distributed algorithm to compute the NE solutions. In patricular we employ asynchronous *myopic best response* (MBR) update rules; that is, the users update their strategies according to their best response assuming that other player's strategies are fixed. Given $p_{-i}$ at the user $i$'s update time epoch, we can express the MBR updates as follows:

$$\begin{aligned}\overline{p}_i &= \arg\max_{p_i\in P_i} u_i\left(p_i\right) \\ &= \arg\max_{p_i\in P_i}\left[B\log\left(\frac{G_{ii}p_i}{\iota_i}\right) - p_i\sum_{j\ne i}\pi_j G_{ij}\right].\end{aligned} \tag{16}$$

**Theorem 4.** *As the other player's strategies are fixed, the users' MBR is single-valued. Moreover, the user's MBR dynamics can be given as follows:*

$$\overline{p}_i^{(t+1)} = \max\left[\min\left[p_{i,\max}, \frac{e^{R^*/B}}{G_{ii}}\iota_i\right], p_{i,\min}\right]. \tag{17}$$

*Proof.* The MBR updates (16) are as follows:

$$\max_{p_i\in P_i} u_i = \max_{p_i\in P_i}\left[B\log\left(\frac{G_{ii}p_i}{\iota_i}\right) - p_i\sum_{j\ne i}\pi_j G_{ij}\right], \tag{18}$$

subject to

$$\begin{aligned} p_{i,\min} \le p_i \le p_{i,\max}, \\ g\left(p_i\right) = R^* - B\log\left(1 + \frac{G_{ii}p_i}{\iota_i}\right) \le 0.\end{aligned} \tag{19}$$

(1) INITIALIZATION: For each user $i \in L$ choose some power $p_i \in P_i$;
(2) POWER UPDATE: At each $t \in T_i^{\text{power}}$, $\forall i \in L$, each user $i$ measures the total interference and update its power level according to (17);
(3) PRICE UPDATE: In each $t \in T_i^{\text{price}}$, each user $i$ updates its price according to (14).

Algorithm 1: Distributed power control algorithm.

The Lagrangian function associated with problem (18) can be given as follows:

$$\begin{aligned}\Psi(\lambda_i, p_i) &= \log\left(\frac{G_{ii}p_i}{\iota_i}\right) - p_i\sum_{j\neq i}\pi_j G_{ij} - \lambda_{i1}(p_{i,\min} - p_i) \\ &\quad - \lambda_{i2}(p_i - p_{i,\max}) - \lambda_{i3}\left(R^* - B\log\left(1 + \frac{G_{ii}p_i}{\iota_i}\right)\right),\end{aligned} \tag{20}$$

where $\lambda_{i1}, \lambda_{i2}, \lambda_{i3}$ are the Lagrange multipliers on link $i$. The Karush-Kuhn-Tucker (KKT) conditions for user $i$ are given by

$$\begin{aligned}\frac{B}{p_i} - \sum_{j\neq i} G_{ij}\pi_j - \lambda_{i1} - \lambda_{i2} + \frac{B}{p_i}\lambda_{i3} = 0,\\ \lambda_{i1}(p_{i,\min} - p_i) = 0,\\ \lambda_{i2}(p_i - p_{i,\max}) = 0,\\ \lambda_{i3}\left(R^* - B\log\left(\frac{p_i G_{ii}}{\sum_{j\neq i} G_{ji}p_j + N_0 B}\right)\right) = 0.\end{aligned} \tag{21}$$

Consider these two situations:

(1) $\lambda_{i1} = 0, \lambda_{i2} = 0, \lambda_{i3} = 0, \overline{p}_i^{(t+1)} = B/\sum_{j\neq i}\pi_j\, G_{ij}$;

(2) $\lambda_{i1} = 0, \lambda_{i2} = 0, \lambda_{i3} = 1, \overline{p}_i^{(t+1)} = (e^{R^*/B}/G_{ii})\iota_i$.

In situation 1, all the constraints are not active and we cannot get a desirable solution. It is necessary to let the constraints to be active, and thus we can easily get the solution of problem (18) as follows:

$$\begin{aligned}\lambda_{i1} = 0, \qquad \lambda_{i2} = 0, \qquad \lambda_{i3} = 1,\\ \overline{p}_i^{(t+1)} = \frac{e^{R^*/B}}{G_{ii}}\iota_i.\end{aligned} \tag{22}$$

Based on the above analysis, the MBR dynamics can be simplified as (17), and thus the proof is completed. □

From these expressions, we can conclude that to implement the update process, each user only needs to know the following information:

(i) its own utility $u_i$, the current SINR $\gamma_i$ (this can be obtained over observations, that is, the SINR before current update time epoch), and its own channel gain $G_{ii}$;

(ii) the channel gains $G_{ij}$ for $j \in L$ and $j \neq i$;

(iii) the price profile $\pi$.

These pieces of information are not difficult to get due to the SINR $\gamma_i$, and channel game $G_{ii}$ can be measured at the receiver and fed back to the transmitter. Also measuring the adjacent channel gains $G_{ij}$ can be accomplished by letting each receiver periodically broadcast a control beacon as we assumed channel reciprocity. The price information can also be broadcast through this control beacon.

Based on Theorem 4, the detailed implementation of the asynchronous distributed power control algorithm can be shown in Algorithm 1.

Note that the power and price need not to update simultaneously. And for each user, the two update processed also need not have to be at the same time. We referes to this distributed power control algorithm as DPPA.

### 5.3. Uniqueness of the NE

**Theorem 5.** *The power control game DPCG $\vartheta$ has a unique Nash equilibrium.*

*Proof.* The Hessian matrix of each user, $H(p) = \nabla u(p)$, consists of diagonal elements as follows:

$$H_{ii}(p) = \frac{1}{p_i^2}, \tag{23}$$

for all $i \in L$, and off-diagonal elements as follows:

$$H_{ij}(p) = 0 \tag{24}$$

for all $j \neq i$. Then, it is easy to verify that $H(p)$ is positive definite as desired. It follows that it has a unique global optimum, which is the only solution to the KKT conditions. □

*5.4. Convergence of the Distributed Algorithm.* We characterize the convergence of the distributed algorithm in this section. We consider it in a game theoretic framework. Each user $i$ specifies a power $p_i$ and a price $\pi_i$ to maximize its utility function (15). It is easy to notice that the best response for each user is to choose a large enough price to force all other users transmit at $P^{\min}$ since there is no penalty for users to announce a high price. This is not a desirable result from the system's perspective. To improve this situation, we consider an externally procedure to determine the price parameter. Let each user split to two fictitious players, and we consider the following External Power-Price (EPP) control game as follows:

$$G_{\text{EPP}} = \left[\text{EW} \cup \text{EC}, \left\{P_i^{\text{EW}}, P_i^{\text{EC}}\right\}, \left\{s_i^{\text{EW}}, s_i^{\text{EC}}\right\}\right], \tag{25}$$

where the players are from the union of set EW and set EC, which are both copies of $L$. EW is a fictitious power player set; each player $i \in$ EW chooses a power $p_i$ from the strategy set $P_i^{\mathrm{EW}} = P_i$ and receives the following payoff:

$$u_i^{\mathrm{EW}}\left(p_i; p_{-i}, \pi_{-i}\right) = \log\left(\gamma_i\right) - p_i \sum_{j \neq i} \pi_j G_{ij}. \tag{26}$$

EC is a fictitious price player set; each player $i \in$ EC chooses a price $\pi_i$ from the strategy set $P_i^{\mathrm{EC}} = [0, \overline{\pi}_i]$ and receives the following payoff:

$$u_i^{\mathrm{EC}}\left(\pi_i; p\right) = -\left(\pi_i - C_i\left(p\right)\right)^2. \tag{27}$$

Here, $\overline{\pi}_i = \sup_p C_i(p)$, which could be infinite for some utility functions. The players in the $G_{\mathrm{EPP}}$ are selfish and maximize their own payoff function.

In $G_{\mathrm{EPP}}$ the players' best responses are given by $B_i^{\mathrm{EW}}(p_{-i}, \pi_{-i}) = W_i(p_{-i}, \pi_i)$ for $i \in$ EW and by $B_i^{\mathrm{EC}}(p) = C_i(p)$ for $i \in$ EC, where $W_i$ and $C_i$ are the update rules for the distributed algorithm. In other words, the distributed algorithm can be interpreted as if the players in $G_{\mathrm{EPP}}$ employ asynchronous myopic best response (MBR) updates; that is, the players update their strategies according to their best responses to the given other players' strategies. It is known that the set of fixed points of MBR updates is the same as the set of the NEs of a game.

**Proposition 6.** *$G_{EPP}$ is supermodular in the transformed strategies $(p, -\pi)$.*

*Proof.* The proof of the proposition is straightforward. For the player in EW, the utility function is as follows:

$$u_i^{\mathrm{EW}}\left(p_i; p_{-i}, \pi_{-i}\right) = \log\left(\gamma_i\right) - p_i \sum_{j \neq i} \pi_j h_{ij}. \tag{28}$$

Let $\pi_j' = -\pi_j$; then, we get

$$\begin{aligned}
u_i^{\mathrm{EW}}\left(p_i; p_{-i}, \pi_{-i}'\right) &= \log\left(\gamma_i\right) + p_i \sum_{j \neq i} \pi_j' h_{ij}, \\
\frac{\partial^2 u_i^{\mathrm{EW}}\left(p_i; p_{-i}, \pi_{-i}'\right)}{\partial p_i \partial p_j} &= \frac{\partial^2 u_i^{\mathrm{EW}}\left(p_i; p_{-i}, \pi_{-i}'\right)}{\partial p_j \partial p_i} = 0 \geq 0, \\
\frac{\partial^2 u_i^{\mathrm{EW}}\left(p_i; p_{-i}, \pi_{-i}'\right)}{\partial p_i \partial \pi_j'} &= \frac{1}{p_i} + \sum_{i \neq j} h_{ij} > 0, \\
\frac{\partial^2 u_i^{\mathrm{EW}}\left(p_i; p_{-i}, \pi_{-i}'\right)}{\partial \pi_j' \partial p_i} &= \sum_{i \neq j} h_{ij} > 0, \\
\frac{\partial^2 u_i^{\mathrm{EW}}\left(p_i; p_{-i}, \pi_{-i}'\right)}{\partial p_j \partial \pi_j'} &= \frac{\partial^2 u_i^{\mathrm{EW}}\left(p_i; p_{-i}, \pi_{-i}'\right)}{\partial \pi_j' \partial p_j} = 0 \geq 0.
\end{aligned} \tag{29}$$

Similarly, the player's utility in FC is as follows:

$$\begin{aligned}
u_i^{\mathrm{EC}}\left(\pi_{-i}'; p\right) &= -\left(-\pi_{-i}' - C_i\left(p\right)\right)^2, \\
\frac{\partial^2 u_i^{\mathrm{EC}}\left(\pi_{-i}'; p\right)}{\partial p_i \partial p_j} &= \frac{\partial^2 u_i^{\mathrm{EC}}\left(\pi_{-i}'; p\right)}{\partial p_j \partial p_i} = 0 \geq 0, \\
\frac{\partial^2 u_i^{\mathrm{EC}}\left(\pi_{-i}'; p\right)}{\partial p_i \partial \pi_i'} &= \frac{\partial^2 u_i^{\mathrm{EC}}\left(\pi_{-i}'; p\right)}{\partial \pi_i' \partial p_i} = 0 \geq 0, \\
\frac{\partial^2 u_i^{\mathrm{EC}}\left(\pi_{-i}'; p\right)}{\partial \pi_i' \partial p_j} &= \frac{\partial^2 u_i^{\mathrm{EC}}\left(\pi_{-i}'; p\right)}{\partial p_j \partial \pi_i'} = 0.
\end{aligned} \tag{30}$$

Then the $G_{\mathrm{EPP}}$ is supermodular game, and the proof is completed. □

With Proposition 6, we can conclude that the fictitious game $G_{\mathrm{EPP}}$ is a supermodular game. We can also conclude that the game holds the following properties.

(i) The NE set of the game is a nonempty and compact sublattice, and there exists a component-wise smallest and largest NE. (Followed by Lemmas 4.2.1 and 4.2.2 in [16].)

(ii) If each user starts from feasible strategy and uses MBR update rule, the strategy profile will eventually locate in the set bounded by the smallest NE and the largest NE. And if the NE is unique then the MBR update rule globally converge to the NE (followed by Theorem 8 in [17]).

In this case, Theorem 1 in [15] can again be used to characterize the structure of $G_{\mathrm{EPP}}$ as well as the convergence of the DPPA. Hence, the fixpoint set of the distributed algorithm is a singleton set containing only the global optimum point. Therefore, the distributed algorithm globally converges to this point. Together with Theorem 5, we conclude the following.

**Proposition 7.** *The DPPA algorithm is globally convergent to the unique NE of the DPCG.*

## 6. Numerical Results

We provide some numerical results to verify the performance of the distributed update algorithm. The simple sensor network example considered here consists of four pairs of links. The maximum power value $p_{i,\max}$, $p_{i,\min}$, for all $i \in L$, band width $B$, and noise power spectral density $N_0$ are chosen as 100 mW, 0 mW, 0.1, and 100 Mbps, respectively. The channel gains of links are determined by independent exponential random variables, where the expected value of the gain value of the gain matrix is

$$E\left[G_{ij}\right] = \begin{pmatrix} 1 & 0.05 & 0.03 & 0.02 \\ 0.06 & 1 & 0.04 & 0.04 \\ 0.06 & 0.04 & 1 & 0.05 \\ 0.03 & 0.04 & 0.05 & 1 \end{pmatrix}. \tag{31}$$

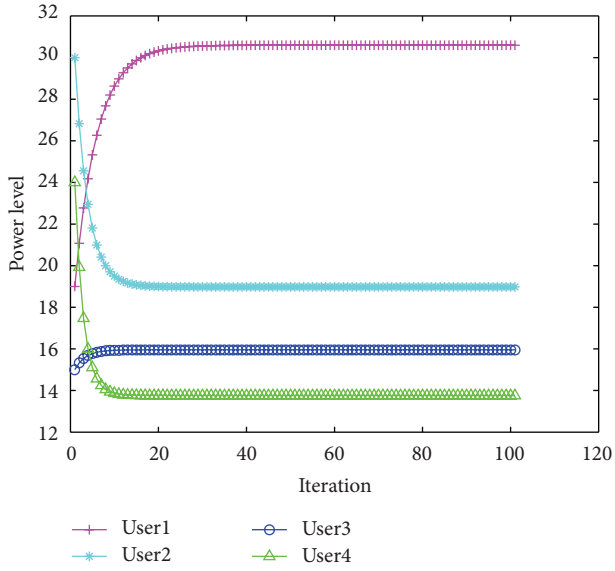



Figure 1: Convergence of distributed update algorithm. The rate constraints are the same.

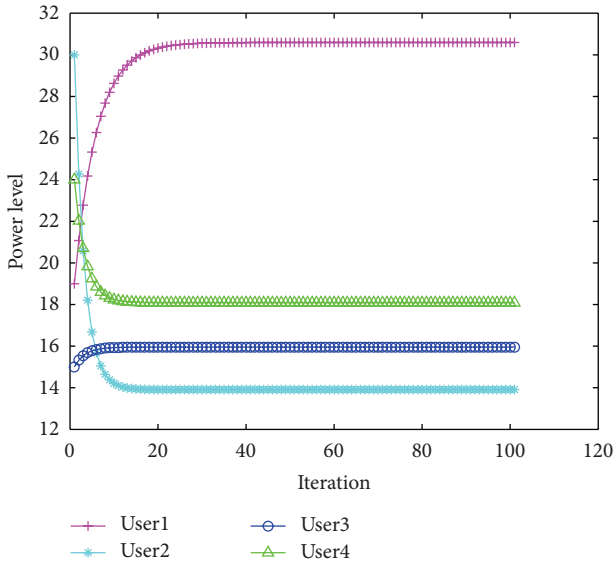



Figure 2: Convergence of distributed update algorithm. The rate constraints are different.

We evaluate the convergence properties of the proposed distributed update algorithm with the same and heterogeneous rate constraints, respectively. Figure 1 illustrates the situation that the four users have the same rate requirement, which means $R_1^* = R_2^* = R_3^* = R_4^* = 17$ Mbps. We see that each of them converges to a reasonable power level after iterations. The power values are different due to the various channel gains.

The results in Figure 2 are nearly the same but with different rate constraints for users, that is, $R_1^* = 17$, $R_2^* = 16$, $R_3^* = 17$, $R_4^* = 18$ Mbps. Note that although user 3 in Figures 1 and 2 has the same rate requirement, the power values

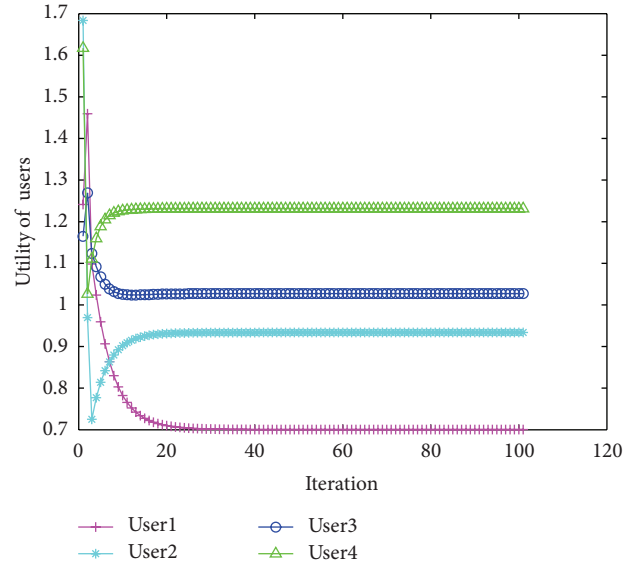



Figure 3: The utility of users. The rate constraints are the same.

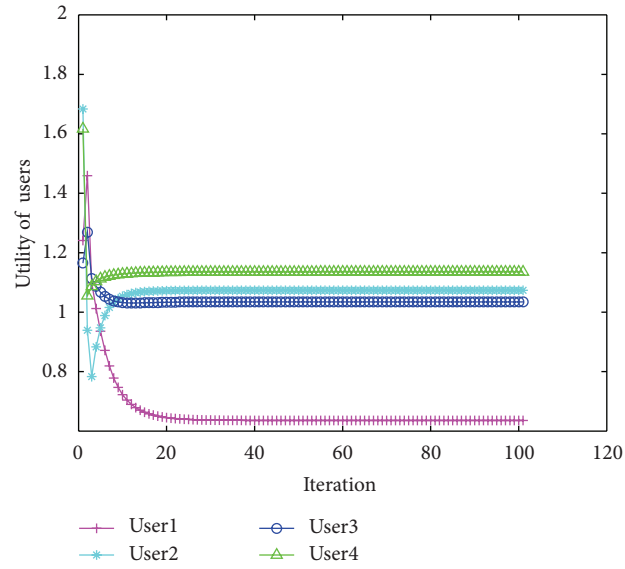



Figure 4: The utility of users. The rate constraints are different.

assigned to it are not equal because of the different network scenario.

The utility of each user is depicted in Figure 3. In the same rate requirement situation, we can see that the payoff of the users converges to a stable state, and the system reaches the NE state after iterations. Similarly, the utility of the users with different rate target is shown in Figure 4. We can conclude that after the system reaches the NE, the users cannot get extra profit by changing their strategies unilaterally.

## 7. Conclusion

In this paper, we use game theory to address the issue of power control in wireless sensor networks. We formulate

the distributed power control game as a noncooperative game, where each user maximizes its own utility by choosing a power level from the feasible area. The existence and uniqueness of the NE are established by building the game's supermodular properties. A distributed power and price update algorithm is proposed which is based on MBR update rules. By correlating to another supermodular game, we show that the proposed distributed update algorithm converges to the unique NE. Numerical experiments have been done to evaluate the algorithm. The results indicate that after relatively little iteration, the game converges to the NE, and the users cannot get extra profit by changing their strategies unilaterally.

## Acknowledgments

This work was supported by Natural Science Foundation of Fujian Province with Grant no. 2012J01252; National Natural Science Foundation of China Grant no. 61072080; the development project of Fujian provincial strategic emerging industries technologies: key technologies in development of next generation integrated high performance Gateway; Fujian development and reform commission high-technical [2013]266; and Fujian Province University-Industry Cooperation of Major Science and Technology Project (2011H6008).

## References

[1] 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.

[2] Z. Gengzhong, L. Sanyang, and Q. Xiaogang, "A power control algorithm based on non-cooperative game for wireless CDMA sensor networks," *International Journal of Digital Content Technology and Its Applications*, vol. 4, no. 3, pp. 137–145, 2010.

[3] H. Cui, G. Wei, Q. Huang, and Y. Yu, "A game theoretic approach for power allocation with QoS constraints in wireless multimedia sensor networks," *Multimedia Tools and Applications*, vol. 51, no. 3, pp. 983–996, 2011.

[4] E.-V. Belmega, S. Lasaulce, M. Debbah, M. Jungers, and J. Dumont, "Power allocation games in wireless networks of multi-antenna terminals," *Telecommunication Systems*, vol. 47, no. 1-2, pp. 109–122, 2011.

[5] R. Valli and P. Dananjayan, "A non-cooperative game theoretical approach for power control in virtual MIMO wireless sensor network," *International Journal Of UbiComps*, vol. 1, no. 3, p. 44, 2010.

[6] 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.

[7] D. E. Charilas and A. D. Panagopoulos, "A survey on game theory applications in wireless networks," *Computer Networks*, vol. 54, no. 18, pp. 3421–3430, 2010.

[8] J. Levin, *Supermodular Games*, Lectures Notes, Department of Economics, Stanford University, 2003.

[9] Y. Sun, Y. Xiao, M. Zhao, X. Zhong, S. Zhou, and N. B. Shroff, "Joint power and channel resource allocation for F/TDMA decode and forward relay networks," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '09)*, pp. 1–6, December 2009.

[10] E. Altman, R. El-Azouzi, Y. Hayel, and H. Tembine, "Evolutionary power control games in wireless networks," in *Networking 2008 Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet*, vol. 4982 of *Lecture Notes in Computer Science*, pp. 930–942, 2008.

[11] G. He, S. Lasaulce, and Y. Hayel, "Stackelberg games for energy-efficient power control in wireless networks," in *Proceedings of IEEE International Conference on Computer Communications (INFOCOM '11)*, pp. 591–595, April 2011.

[12] J. Xu, K. Li, and G. Min, "Layered multi-path power control in underwater sensor networks," in *Proceedings of the 53rd IEEE Global Communications Conference (GLOBECOM '10)*, pp. 1–5, December 2010.

[13] S. M. Perlaza, E. V. Belmega, S. Lasaulce, and M. Debbah, "On the base station selection and base station sharing in self-configuring networks," in *Proceedings of the 4th International Conference on Performance Evaluation Methodologies and Tools (VALUETOOLS '09)*, article 71, 2009.

[14] H. Yu, S. Zhang, and V. K. N. Lau, "Game theoretical power control for open-loop overlaid network MIMO systems with partial cooperation," *IEEE Transactions on Wireless Communications*, vol. 10, no. 1, pp. 135–141, 2011.

[15] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 5, pp. 1074–1084, 2006.

[16] D. M. Topkis, *Supermodularity and Complementarity*, Princeton University Press, Princeton, NJ, USA, 1998.

[17] P. Milgrom and J. Roberts, "Rationalizability, learning and equilibrium in games with strategic complementarities," *Econometrica*, vol. 58, no. 6, pp. 1255–1277, 1990.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 491973, 11 pages
http://dx.doi.org/10.1155/2013/491973

*Research Article*

# Towards Efficient and Secure Geographic Routing Protocol for Hostile Wireless Sensor Networks

**Chen Lyu, Dawu Gu, Yuanyuan Zhang, Tingting Lin, and Xiaomei Zhang**

*Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*

Correspondence should be addressed to Dawu Gu; dwgu@sjtu.edu.cn

Received 23 November 2012; Revised 25 June 2013; Accepted 2 July 2013

Academic Editor: Adel Soudani

Copyright © 2013 Chen Lyu 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.

Traditional routing protocols are quite vulnerable under the attacks from both external and internal attackers. In this work, we examine the impact of a wide variety of attacks in malicious wireless sensor networks scenario. An Efficient and Secure Geographic Routing protocol ESGR is proposed to exploit the geographic location, cryptography mechanisms, and broadcast nature of the wireless channel. ESGR utilizes the geographic leashes and the TESLA scheme to provide resistance against the Sybil attack and wormhole attack. Meanwhile, it employs a distributed trust model and the packets opportunistic forwarding to prevent black hole and gray hole attacks. We demonstrate the results through analysis and simulations that ESGR effectively mitigates a specific variety of attacks and ensures high packet delivery rate in malicious sensor network environment.

## 1. Introduction

Wireless sensor networks (WSNs) have become an attractive solution for environment monitoring, target tracking, and battlefield surveillance applications nowadays. The wireless network routing protocols that majorly address the time-varying topology can be roughly divided into three categories: flat routing, hierarchical routing, and geographic routing. The flat routing includes table-driven routing and on-demand routing, such as DSDV [1], DSR [2], and AODV [3]. The hierarchical routing, for example ZRP [4], provides scalable management to multiple network levels. Geographic routing, such as GPSR [5], is designed for the networks provided with specific positioning information, such as GPS (Global Positioning System) or other location determination techniques [6]. Based on the information, each node in the GR only keeps the local one-hop connectivity and makes the forwarding decision on the fly. As it does not use control packets to establish a path, the geographic routing is more resilient compared with other types of routings.

In the adversarial environments, attacks come from both the external and internal attackers. The external attackers who do not possess the credentials to participate could be excluded by the authentication mechanisms. However, due to the constrained resource, such as the network bandwidth, the processing capability, and the battery capacity of the sensor node, traditional asymmetric cryptography mechanisms cannot be directly deployed over the WSNs. For example, some Public Key Infrastructure inappropriate for the length of the ciphertext has led to costly transmission or storage. Other than the external attacks, a typical internal attack would compromise a node to gain access to the secret keys stored on it. The asymmetric cryptography mechanisms are not the practicable countermeasures against such attacks [7, 8]. In this paper, we study a variety of attacks against geographic routing in wireless sensor networks and propose a novel attack detection and defense algorithm. It would leverage the geographic locations of nodes, the efficient cryptography mechanisms, and the broadcast nature of wireless channel.

The location in GR is the key information in the geographic routing. An attacker or the compromised node could falsify its location to get more chances to disrupt the routing service if there is not a proper method to verify its claimed location. Unfortunately, most of the proposed secure routings do not deal with it. Only a few of prior works [9, 10] have suggested to use the secure anchor nodes to verify the location. However, these anchors composing a basic infrastructure are required to be trustful and communicate securely among each other. Unlike previous solutions, we consider

the issue in the design of the protocol without the extra cost of setting up and maintaining such a secure infrastructure. Our work verifies a sensor node's position by geographic leashes [11] combined with TESLA scheme [12] to detect attacks associated with location, such as the Sybil attack and wormhole attack.

Based on the verification results of location, we present an Efficient and Secure Geographic Routing protocol (ESGR) for WSNs. To further resist other routing attacks such as black hole or gray hole attack, the basic idea of ESGR is an opportunistic routing [13–15]. In addition, it adopts a novel trust model leading to an effective metric against a range of attacks from the external to internal attacks in hostile sensor networks. Provided by the routing metric in ESGR, more receivers would be opportunistically selected as the next hop candidates to forward the packet, and the transmission would not be interrupted as long as there are still candidates to choose from. Therefore, ESGR's per-hop packet transmission is controlled and observed by the sender instantly. Our simulation results show that ESGR enables more than 95% packet delivery rate (PDR) when one fifth of nodes become black hole nodes. Even when one half of sensor nodes drop packets, its PDR can still maintain more than 85%.

We summarize the main advantages of ESGR as follows.

First, it is based on the geographic forwarding with small per-node local state, which means there are no routing tables to maintain, so the forwarding decisions are made on the fly. State locality with minimal overhead is crucial to ESGR's efficiency.

Second, the detection of the malicious behavior is based on the broadcast nature of the wireless channel. Nodes verify the locations of their neighbors by both the geographic leashes and TESLA mechanisms. To reduce the latency that is brought about by TESLA, ESGR predicts its energy consumption to enable instant authentication.

Third, the opportunistic approach can provide a certain degree of redundancy and randomness to enhance the resiliency in face of malicious nodes' misbehaviors. In order to prevent packet loss, more nodes are selected to be the forwarding candidates. If one of candidates drops the packet, ESGR can select another one from the candidates until the packet is successfully transmitted.

Fourth, the direct trusted information based on watching the next hop's forwarding is used as ESGR's primary trust metric. A sensor node that relays traffic data will earn higher trust from the neighbor nodes. With every node estimating and acting on trust metrics, it is able to route around unreliable neighbors.

Finally, ESGR is shown to be secure and efficient through theoretical analysis and extensive simulations. It achieves 1.5 times higher PDR than the other two protocols and is scalable with an acceptable overhead.

The rest of the paper is organized as follows. The related work is in Section 2. We describe the network and security model in Section 3. Section 4 presents our protocol ESGR. The security and efficiency analysis is performed in Section 5. Section 6 provides experimental results that demonstrate the effectiveness of ESGR in addressing the considered attacks. We conclude our paper in Section 7.

## 2. Related Work

There are several secure flat routing schemes in the prior works, secure table-driven routing schemes, and secure on-demand routing schemes (such as, [16–18]). SEAD [16] protects distance-vector calculations from distance decreasing and uses hash chains to authenticate routing updates, which tends to have high communication overhead. As a result, it would not be suitable in the large-scale sensor networks. In [17], the authors propose an on-demand routing protocol ODSBR, which uses an adaptive probing technique to detect the malicious link caused by individual or collusion nodes. In [18], they present an algorithm detecting the Byzantine attacks during route discovery and then propose an on-demand routing whose metric combines a node's trustworthiness and performance. However, as a predetermined route must be established before a packet transmission, it introduces control messages for both table-driven and on-demand routings. They would become the attackers' major targets leading to vulnerability of the sensor networks.

To secure geographic routing, the authors in [19] propose a secure forwarding mechanism providing the authentication and integrity by the TIK protocol. It also combines with a secure grid location service to verify the location information. However, the protocol requires an assumption that all nodes are tightly time synchronized. In our work, we only require loosely synchronized clocks with an upper bound on the sending time [20]. SBGR [21] is a beaconless geographic routing where the nodes compete in a distributed way to acquire the chance to forward the packets. SIGF [22] presents a configurable secure geographic routing family for the wireless sensor networks and makes the tradeoff between security and performance. Neither of them responds to the attacks associated with their locations. In [9], the authors combine lightweight localization techniques with intrusion identification techniques for accurate location information. Their routing protocol then incorporates a distributed trust model to prevent some routing attacks. Liu et al. [10] deploy a location verification algorithm to address the attacks falsifying the location information and then propose a trust-based multipath routing. These works [9, 10] have some similarities with ours, but their location verification schemes rely on a large number of secure anchors to cover all the sensor nodes. Our protocol does not require such an infrastructure and makes use of the geographic leashes and TESLA scheme to defend against these location relevant attacks.

Geographic routing or position-based routing, such as GPSR, has drawn much attention in the literature for its simplicity and efficiency [23–28]. However, this scheme alone does not guarantee delivery due to the existence of local minima (or dead ends) [5]. To tackle the issue, [25, 26] assign for each node a virtual coordinate in a new plane and perform greedy forwarding based on the virtual coordinates. The authors in [28] decompose a given network into a minimum number of greedily routable components, where greedy

routing is guaranteed to work. In this paper, we do not discuss the extension since the increasing complexity does not add much in our work.

## 3. Network and Security Model

*3.1. Network Model.* We consider a wireless sensor network which consists of sinks and large number of sensor nodes randomly deployed in the network. Sinks are considered to be always trusted, and sensor nodes may be compromised by the adversary. Nodes participate in the data forwarding for other nodes. We assume that the wireless channel is symmetric. All the nodes can communicate with each other within the transmission range $R$. Each node is stationary in its location and sometimes turns off the transceiver for reducing energy consumption.

The locations of sinks as important resources are known in advance in the system. For the reason of geographic routing, all the sensors first need to obtain their own positions. We assume that each node's position is securely decided once it is deployed in the network. Moreover, all the nodes in the network are loosely time synchronized. Nodes can overhear and then verify the transmissions in their one-hop neighbors over wireless channel, which enables detection of malicious behaviors.

ESGR requires that, for each pair of end nodes, a source $s$ and a sink $d$ that wish to communicate securely across the network share a preestablished symmetric key. Moreover, each node has a unique identity and a pair of public and private keys predistributed in the network.

*3.2. Security Model.* We define an attack as any action by an entity that results in disruption or degradation of the routing service. Attacks can divide into the following types.

*Sybil Attack.* In the Sybil attack [8], a malicious node can behave as if it were a larger number of nodes either by impersonating other nodes or simply by claiming false identities. In the geographic routing, a Sybil node could appear in more than one location at once. It can then tamper or forward the packets in a black hole or gray hole attack.

*Rushing Attack.* A malicious node in the rushing attack attempts to tamper packets and hurry them to the next hop node. In the traditional routing protocol, since only one packet is forwarded by the node, it would discard the legal one if it was forwarded by the adversary firstly.

*Wormhole Attack.* Two compromised nodes can communicate with each other by a private channel in the wormhole attack [11]. The adversary can tunnel the packet to another location and replay it there. A sensor node that hears a packet transmission directly from one wormhole node will consider itself to be a neighbor of that node.

*Black Hole or Gray Hole Attack.* In this type of attack, a malicious node drops all or part of the received traffic. The black hole attack is a type of attack in which a node supposed to relay packets discards all instead. The malicious node can also accomplish this attack selectively, which is rather called the gray hole attack.

*Other Attacks.* There are some other types of attacks, such as the blacklist attack, the replay attack, and the node selfish attack. However, since no information about adversaries is changed between nodes in ESGR, the blacklist attack can be avoided. The time stamp with the authentication mechanism can be used to prevent the replay attack. We treat the node selfish attack as the gray hole attack as the adversary also selectively drops packets for its own benefit.

We consider the denial-of-service attack, such as an attacker sending a high number of messages with meaningless packets, will be quickly identified by the authentication scheme. Moreover, attacks against lower layers such as the link or the physical layer are not addressed.

## 4. ESGR: Efficient and Secure Geographic Routing

ESGR is under the category of the general geographic routing, so it contains the following two parts.

*4.1. Location Beacons.* In this part, our scheme implements geographic leashes and provides an efficient broadcast authentication in the wireless sensor networks. We deploy efficient TESLA scheme [12] as it is based on the symmetric cryptographic primitives. TESLA requires loose time synchronization among all the communicating parties. In the network, we assume that all nodes share a synchronized clock with a maximum clock synchronization error of $\Delta$. For simplicity, we assume the clock drift is negligible. ESGR location beaconing is composed of three phases: setup, sending location beacons, and verifying location beacons.

*4.1.1. Setup.* For the TESLA scheme, each node splits the time $t$ into a series of intervals $I_1, I_2, \ldots, I_N$, where $t$ is the amount of time between rekeying. The duration of each time interval is $T_{\text{int}}$ in common. We denote the starting time of the interval $I_i$ by $T_i$, and $T_i = T_0 + i * T_{\text{int}}$, where $T_0$ is the starting time of the entire hash chain. In each interval, the sensor node may send zero or multiple packets.

Consider the chain of length $N$ with the values $K_0$, $K_1, \ldots, K_{N-1}$ for time intervals $I_0, I_1, \ldots, I_{N-1}$. The sender picks the last key $K_{N-1}$ of the key chain randomly and calculates the entire key chain using a pseudorandom function $F$ to derive the previous keys: $K_i = F(K_{i+1})_{\forall i \in \{0,\ldots,N-2\}}$. The value $K_0$ serves as a commitment to the entire chain, which allows anybody to authenticate the following values of the chain. Moreover, TESLA uses a second pseudorandom function $F'$ to derive the key $K_i'$: $K_i' = F'(K_i)$, which can be used to compute the Message Authentication Code (MAC) of the message for each time interval.

As TESLA requires an initially packet to authenticate the neighbor nodes, we achieve this authentication with a digital signature scheme, where the message is signed by

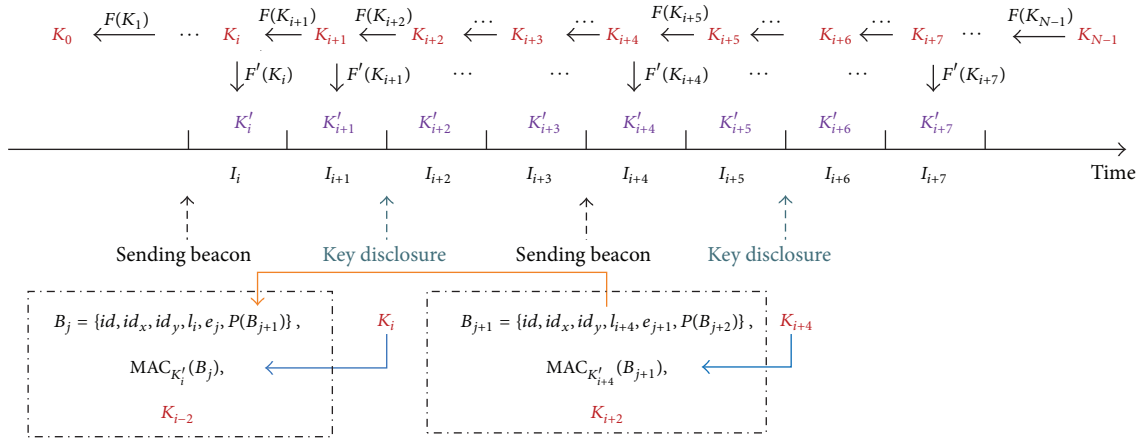



Figure 1: The TESLA scheme for location beacons when the key disclosure delay $d = 2$.

the private key of the node. The initially authentication packet is transmitted by the sensor node with the identity *id* as

$$m = \langle id, T_{\text{int}}, K_0, T_0, N, I_0, d \rangle, \text{Sign}(m), cert, \quad (1)$$

where $d$ is the key disclosure delay, Sign($m$) is the signature, and *cert* is issued by the trusted Certificate Authority and prestored into the node. On receiving the initially authenticated packet, the neighboring nodes will verify the sender and record the associated parameters with the sensor's *id* if valid.

*4.1.2. Sending Location Beacons.* The neighbors' coordinates are updated periodically through one-hop beacons. Each node periodically transmits a beacon $B_j$, containing its own identity *id* and location $l_{id} = (id_x, id_y)$. Our mechanism extends the beacon to include the time stamp $I_i$ in the current time interval. Then the sender computes the MAC value over the beacon with the private key $K_i$ and broadcasts the packet to all the neighbors at the same time. For example, as shown in Figure 1, the construction of the packet in interval $I_i$ is

$$B_j = \left\{id, id_x, id_y, I_i, e_j, P\left(B_{j+1}\right)\right\}, \text{MAC}_{K'_i}\left(B_j\right), K_{i-d}, \quad (2)$$

where $P(B_{j+1})$ is prediction outcome of the content of next beacon, $e_j$ is the sleep warning or work state with remaining energy, and the | stands for the message concatenation.

As the TESLA scheme inevitably introduces the additional delay by the key disclosure compared to other signature schemes, our beacon message appends the prediction outcome of next beacon $B_{j+1}$ to overcome the drawback and provide faster authentication for the beacon packet. Then, we explain how to build the prediction outcome, as illustrated in Figure 1, $P(B_{j+1}) = H(id, id_x, id_y, I_{i+4}, e_{j+1})$, where $H$ is a hash function. As the sensor node with the unique identity in our model is immobile, the identity and location information will be always the same. The changed information is the energy remainder of sensor node, which can be predicted by the sender in the current beacon according to the previous consumption model. Therefore, the location beacons are sent periodically and then predicted. However, how to build an accurate model to predict the energy consumption is orthogonal to our work.

The key remains secret for $d - 1$ time intervals. Packets sent at time interval $I_i$ can hence disclose the key $K_{i-d}$ if needed. Once the neighbors receive that key, they can verify the authenticity of the previous packets sent at interval $I_{i-d}$. Note that if there is no beacon or packet sent in $I_{i+d}$, the sender then transmits the secret key individually to the receivers for the purpose of verifying the beacon $B_j$.

*4.1.3. Verifying Location Beacons.* When a neighbor node receives a beacon, it would verify each packet that the key used to compute the MAC of that packet is not yet disclosed by the sender. Considering the beacon $B_j$ arrives at the receiver at its local time $t_r$, the security condition of TESLA [20] is that

$$\left\lfloor \frac{t_r + \Delta - T_0}{T_{\text{int}}} \right\rfloor < I_i + d. \quad (3)$$

If this security condition is not satisfied, the receiver would drop the packet. Otherwise, it stores the packet and verifies the authenticity till it knows $K_i$ later. Then it recovers the commitment $K_0$ by iteratively invoking the hash function and applies the valid key to check the stored MAC. Furthermore, when it has received an authenticated key $K_u$ in the reconstructed key chain, the receiver only needs to verify that $K_u = F^{i-u}(K_i)$. If the authentication is valid, the receiver can update the key chain and store the latest authenticated key to improve the efficiency of verification. Furthermore, to achieve instant authentication and verify the following $B_{j+1}$ after receiving $B_j$ in our example, the receiver first verifies $B_j$, gets the prediction outcome $P(B_{j+1})$ from the beacon $B_j$, and then checks whether the recomputed value matches $P(B_{j+1})$ given relevant values in the $B_{j+1}$.

After verifying the packet, the node could obtain the coordinates of the sender's location $l_S = (S_x, S_y)$ in the beacon and cache it with the identity in the neighbor list. Then the receiver measures the correctness of the location information based on the geographic leashes, which is based on the radio propagation model. Assuming that the transceiver

gain is set to be one and the channel bandwidth exactly matches, we define $\alpha$ as the path loss factor, which depends on the environment. The relationship between the received signal strength $P^d$ and the distance $d_{SQ}$ from the sender $S$ to the receiver $Q$ can be expressed as (4), where $C$ is the speed of the radio and $f$ is the frequency of the radio [9, 29]:

$$P^d = \frac{P_t C^2}{(4\pi)^2 d_{SQ}^{\alpha} f^2} + \varepsilon. \tag{4}$$

Here, $P_t$ means the setting transmission power and $\varepsilon$ is represented as a zero-mean Gaussian random variable. The maximum relative error of the distance is considered to be $\delta$. We denote the coordinates of the receiver's location by $l_Q$. The value of $d_{SQ}$ is constrained by the following inequality: $d_{SQ} \leq \|l_S - l_Q\| + \delta$. Due to the inequality, the receiver weights the location trusted information $\mathrm{LT}_S$ of the sender to be zero or one, which is used to construct the trust model in the next part.

### 4.2. Data Packets

*4.2.1. Distributed Trust Model.* For detecting routing attacks in a WSN, we design a distributed trust model which enables each node to define the trustworthiness of its neighbors combining the location trusted information and direct trusted information.

We now explain the direct trusted information. To further detect neighbor nodes dropping or selectively forwarding packets, the sender overhears the wireless channel to check whether the packet is actually forwarded by its selected next hop node. Meanwhile, it enforces efficient authentication according to the TESLA scheme. We assume that the average number of packets sent to node $F$ for forwarding is marked as $Q_F$ and the average number of valid data transmissions of node $F$ is marked as $V_F$. The metric of direct trust can be given by $\mathrm{DT}_F = V_F/Q_F$. Initially, $Q_F = 1$, $V_F = 0$. Finally, the total trust value for node $F$ is produced by combining the location trusted information and direct trusted information:

$$T_F = \eta \cdot \mathrm{DT}_F + (1 - \eta) \cdot \mathrm{LT}_F, \tag{5}$$

where $\eta$ is the factor with $0 < \eta < 1$. As the number of interactions increases, the direct trusted information becomes more significant than the location trusted information; thus $\eta$ can be modified to a larger value.

*4.2.2. Opportunistic Forwarding.* When the source prepares to send a message $M$, it will first encrypt the message $M$ with a symmetric key $K$ sharing with one of the sinks. As the intermediate node in the routing path needs some information such as the unique identifier of the packet, the source node will then put the necessary and constant information in the header field marked as $HI$, which is also contained in the message $M$. Then, the source (or intermediate) node $R$ begins to calculate the metrics of all the neighbors and establish its own forwarder list. Supposing that the distance from the sender to the sink is $D_R$ and the distance

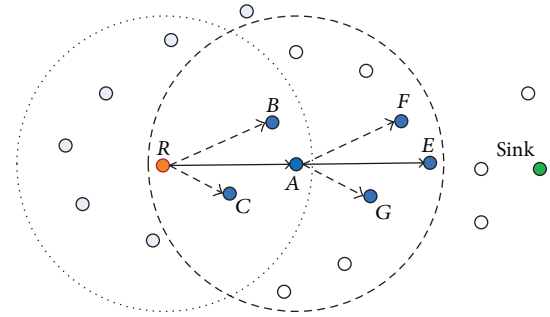



Figure 2: An illustration of the opportunistic forwarding scheme. For node $R$, the forwarder list includes nodes $A$, $B$, and $C$. The packet is then transmitted to node $A$, and node $A$ establishes its forwarder list including nodes $E$, $F$, and $G$.

from its neighbor $F$ to the sink is $d_F$, the routing metric can be given by

$$\mathrm{metric}_F = \beta_1 \cdot \left(1 - \frac{d_F}{D_R}\right) + \beta_2 \cdot T_F + \beta_3 \cdot e_F, \tag{6}$$

where $e_F$ is the remaining energy of node $F$ which is perceived by node $R$ through frequent location beacons and $\beta_i$ is the coefficient factor ($0 < \beta_i < 1$ and $\beta_1 + \beta_2 + \beta_3 = 1$). The routing metric is an integrated set of trust, energy, and progress to the sink. Based on (6), each valid neighbor is assigned a metric value, and the one with the largest metric will have the highest priority.

The forwarder list is established according to the routing metric, and we set the number of candidates for the next hop to be $N$: $C_1, C_2, \ldots, C_N$. For example, as shown in Figure 2, node $R$'s candidate nodes for the next hop are nodes $A$, $B$, and $C$ with $N = 3$. If $Num$ is less than $N$, it will be filled with pad to ensure the packet length unchanged. The payload of the forwarding data packet can be expressed as

$$HI, E_K(M), H(E_K(M)), Num, C_1, C_2, \ldots, C_N, I_d, \tag{7}$$

where $H$ is the hash function, $E_K(\cdot)$ is one of symmetric encryption algorithms with the secret key $K$, and $I_d$ is the current time interval. The format of $HI$ is given by:

$$id_{\text{source}} \| sequence \| id_{\text{sink}}. \tag{8}$$

The identity of the source $id_{\text{source}}$ plus the sequence number $sequence$ marks a unique data packet in the network. Therefore, $HI$ is unique, so is $H(E_K(M))$. The source (or intermediate) node will store the values of $HI$ and $H(E_K(M))$ for each new incoming packet in its ID list. The data packet with MAC attached is then broadcast to the neighbors. After the transmission, the sender would also monitor the next hop's transmission and adjust the direct trusted value by collecting the next hop's feedback information.

The receiver within the source (or intermediate) node's transmission range could receive and verify the packet. The validation process can be divided into two steps and operate as follows. First, the neighbor node will authenticate the sender of the incoming packet with the TESLA scheme.

Then, it will look for its ID list and compare the values of $HI$ plus $H(E_K(M))$ of the packet with the cached ones. If the packet has been received, these values are exactly equal and the receiver would drop the packet.

After successful validation, if a node finds itself to be the first candidate in the forwarder list (in our example, node $A$ or node $E$), it would establish its own forwarder list by calculating each neighbor's metrics and immediately send the data packet. However, if it finds itself not the first candidate in the forwarder list (e.g., node $B$, node $C$, node $F$, or node $G$), the packet will be cached for a while according to the priority determined by the forwarder list. For example, if there are $k$ nodes ahead, it will wait for $k$ time slots before forwarding that packet. The cached packet is attached a counter which will minus one every time. During the waiting, if it receives an incoming packet that has the same $HI$ plus $H(E_K(M))$ recorded in the ID list, the corresponding cached packet will be discarded as it must be relayed by another candidate with higher forwarding priority (e.g., node $A$ or node $E$). When the counter returns to zero, the cached packet will be forwarded. Subsequent nodes follow the same process until the packet reaches the destination.

## 5. Security and Efficiency Analysis

*5.1. Security Analysis.* In this section, we specify how ESGR addresses the major attacks described in Section 3.

*5.1.1. Sybil Attack.* In the Sybil attack, the authentication and geographic leashes mechanism in ESGR can be used to detect it. Each node in the network corresponds to a location point and a unique pair of public and private keys. For efficiency, when a beacon or a data packet is sent by a node, it requires to be signed with a key of TESLA scheme. Moreover, as the key can be authenticated by a commitment and the commitment is signed by the node's private key, the adversary cannot easily impersonate other node and appear in more than one location at the same time. If the adversary claims false location in the beacon, the neighbor nodes can use the geographic leashes for verification to defend against the attack. Even if the adversary is lucky to avoid the detection, our ESGR can deal with the misbehavior like black hole or gray hole attack as soon as these Sybil nodes become packet droppers.

*5.1.2. Rushing Attack.* In the rushing attack, the adversary tries to block the communication by hurrying its modified packets to the next hop node of routing path. As one of the opportunistic routings, ESGR has more candidates to relay the data packets. For each valid packet determined by both the fields of $HI$ and $H(E_K(M))$, it can avoid the tempering attack incurred by rushing attack. The nodes in ESGR forward every copy of the packets even if only one of the fields is identical. As a result, the correct packet will reach the destination though it also receives some redundant packets.

*5.1.3. Wormhole Attack.* ESGR verifies the locations of the neighbors by the geographic leashes and then obtains the results of validations. Followed by a low location trusted value, ESGR can alleviate the consequences of the wormhole attack. Furthermore, after the adversaries have taken control of a fraction of the traffic through creating the wormholes, they would maliciously drop or corrupt packets, which would be alleviated by the distributed routing metric in our scheme.

*5.1.4. Black Hole or Gray Hole Attack.* Under the black hole or gray hole attack, the number of the packets received by sink nodes in the network will decrease largely. ESGR can ensure the throughput when the malicious nodes drop or selectively drop data packets, as potential multiple paths are available in the opportunistic forwarding scheme. Even when there are many black hole attackers, our path can avoid selecting them for the next hop with their low trust values.

As illustrated in Figure 3, consider that node $A$ is a new black hole and node $F$ is an existing gray hole node. A packet is forwarded by node $R$ to the sink, and the forwarder list is node $A$, node $B$, and node $C$. In case node $A$ is selected as the first candidate, the packet would be dropped. Then, node $B$ as the second candidate will relay the packet to node $F$, $G$, or $H$ for the next hop instead of node $A$ and suppress the lower priority candidate's forwarding (node $C$). We assume that node $B$ suspects that node $F$ is an adversary due to its low trust value. With the largest routing metric, node $G$ would be selected as the first candidate with the highest priority by node $B$. The packet would be transmitted to node $G$ and then forwarded to the sink. Moreover, as node $R$ does not receive the packet sent by node $A$, our routing mechanism decreases the direct trusted value $\mathrm{DT}_A$ thus lowering the routing metric $\mathrm{metric}_A$ for node $A$. When node $B$ succeeds in delivering the packet, node $R$ increases and updates its routing metric $\mathrm{metric}_B$. Therefore, after a period of time, node $R$ may select node $B$ as the first candidate with the highest priority for the following packets.

*5.2. Efficiency Analysis .* We conduct the theoretical analysis of PDR and end-to-end delay performance.

*5.2.1. Packet Delivery Rate.* Assume there are maximum $N$ forwarding candidates. Suppose $p_x$ is the probability that the data packets are dropped by the candidate node. The analytical packet delivery ratio is the end-to-end probability that a packet is successfully delivered from a source to the sink:

$$P_{\mathrm{ESGR}} = (1 - p_x^N)^{L-1}, \tag{9}$$

where $L$ is the average hop for the packet. For simplicity, we assume the per-hop packet delivery is independent of each other.

For GPSR like unicast routing protocol, the packet delivery ratio is denoted as follows:

$$P_{\mathrm{unicast}} = (1 - p_x)^{L-1}. \tag{10}$$

We can see that the opportunistic routing significantly improves the packet delivery rate compared with the ordinary unicast routing.

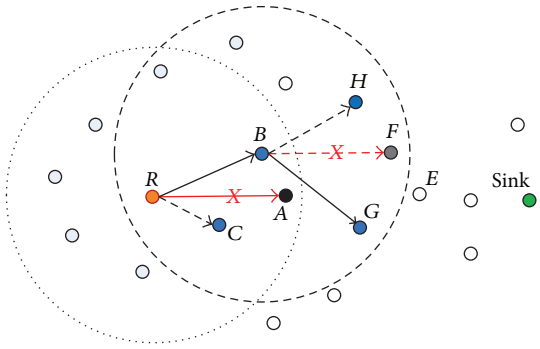



Figure 3: An illustration of black hole attack and gray hole attack. Node $A$ is a black hole node. Node $F$ is a gray hole node.

*5.2.2. End-to-End Delay.* In order to measure the end-to-end delay in theory, we study the per-hop packet forwarding delay of the $Q$th candidate. It is divided into two parts: one part is introduced by the sender and the other comes from the candidate coordination:

$$\Delta T_{\mathrm{ESGR}} = t_s + (Q-1)\cdot \Delta T, \tag{11}$$

where $\Delta T$ represents the given time slot and $t_s$ denotes the packet transmission delay in the corresponding normal situations. Assume the $Q$th forwarder relays the packet for one transmission on average; then

$$Q = \sum_{i=1}^{N} i\left(1-p_x\right)\cdot p_x^{i-1} + N\cdot p_x^N. \tag{12}$$

Since ESGR is one of the opportunistic routings, we make use of the broadcast nature that all nodes within the coverage of the sender would receive the signal. To decrease the packet loss, some alterations are encouraged on the RTS/CTS/ACK mechanism in 802.11b like [15]. The packet is sent in unicast form and takes the first candidate as the next hop. On the receiver side, if the message's next hop is not the receiver, it is also delivered to the upper layer and then processed by ESGR. The sender delay includes four parts: channel contention time for 802.11b ($T_c$), data transmission time ($T_{\mathrm{data}}$), propagation delay ($T_p$), and the key disclosure delay for authentication ($T_{\mathrm{key}}$). Then, $t_s = T_c + T_{\mathrm{data}} + T_p + T_{\mathrm{key}}$.

Therefore, the end-to-end packet transmission delay of ESGR can be expressed as follows:

$$\begin{aligned} T_{\mathrm{ESGR}} &= \Delta T_{\mathrm{ESGR}}\cdot L \\ &= \left(T_c + T_{\mathrm{data}} + T_p + T_{\mathrm{key}} \right. \\ &\quad \left. + \left(\sum_{i=1}^{N} i\left(1-p_x\right)\cdot p_x^{i-1} + N\cdot p_x^N - 1\right)\cdot \Delta T\right)\cdot L \end{aligned} \tag{13}$$

while the end-to-end delay of unicast routing with $N = 1$ is given by $T_{\mathrm{unicast}} = (T_c + T_{\mathrm{data}} + T_p + T_{\mathrm{key}})\cdot L$.

Table 1: Experimental setup.

| Parameter | Value |
|---|---|
| Plane size | 1 km by 1 km |
| Number of nodes | 50~300 |
| Medium access control | 802.11 b at 1 Mbps |
| Transmission range | 250 m |
| Traffic size | 512 Bytes |
| Number of candidates | $N = 3$ |
| Number of flows | 10 CBR, 2 packets/s |
| Parameter | $\eta = 0.7$, $\beta_1 = 0.4$, $\beta_2 = 0.4$, $\beta_3 = 0.2$ |

We can find that with a larger value of $N$, the packet delivery rate is higher but the end-to-end delay is also larger. In our simulation, we will show how the candidate number affects both the performances of PDR and delay.

## 6. Simulation

The performance of ESGR is evaluated by the OPNET network simulator, together with the geographic routing protocol GR (the version with perimeter forwarding switched off GPSR) and the zone routing protocol ZRP. We employ the propagation model of two-ray ground. Moreover, we use the AES algorithm for encryption with secret key of 128 bits. The length of MAC or the hash key is 80 bits. The time slot of opportunistic routing is set to be 10 ms. Some other important parameters are set in Table 1. The performance metrics consist of the packet delivery rate, the end-to-end delay, and the bandwidth of overhead.

*6.1. Node Density.* A number of nodes are placed randomly in a square area of 1000 m. We assume that one fifth of the nodes in the network are the black hole nodes. We measure the performance of all the protocols in different sized networks, with the increasing node density.

Figure 4 shows the performance of packet delivery rate. Due to one of the opportunistic routings, the PDR of ESGR is the highest in all the scenarios with different network size. In the network of 300 nodes, ESGR outperforms the GR and ZRP protocols and enables more than 95% packet delivery rate on a 1000 m $*$ 1000 m place with 60 black hole nodes. The differences of PDRs are not distinct for geographic routings under various node densities.

The overhead is measured by the extra bits per second for data transmissions. When the number of sensor nodes increases, the overheads of three protocols also increase as the density of neighbor nodes becomes larger. As shown in Figure 5, the node density impacts the effectiveness of ZRP by the decrease of the throughput. Although ESGR has a larger overhead than GR, it is still accessible for the network's load.

From the simulation results, ESGR has 1.5 times higher PDR than the other two protocols and keeps stable with an acceptable growing overhead when the network size grows large.

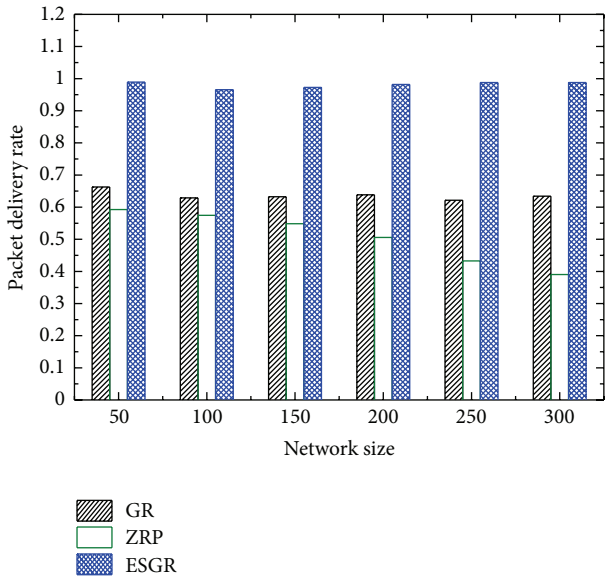



Figure 4: Packet delivery rate for different node densities.

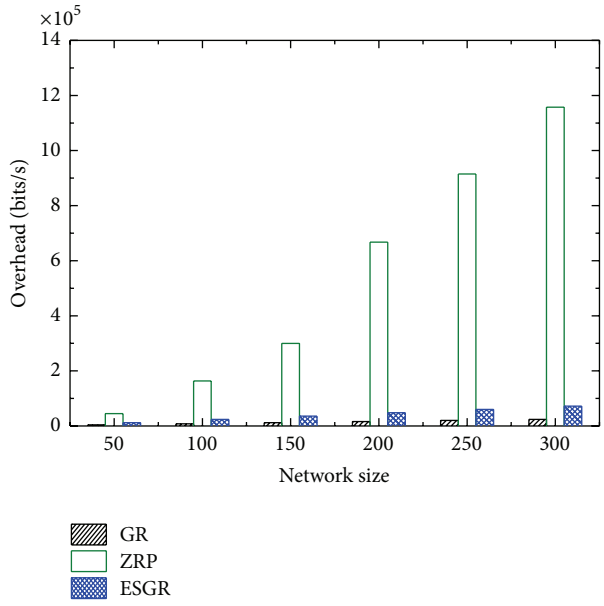
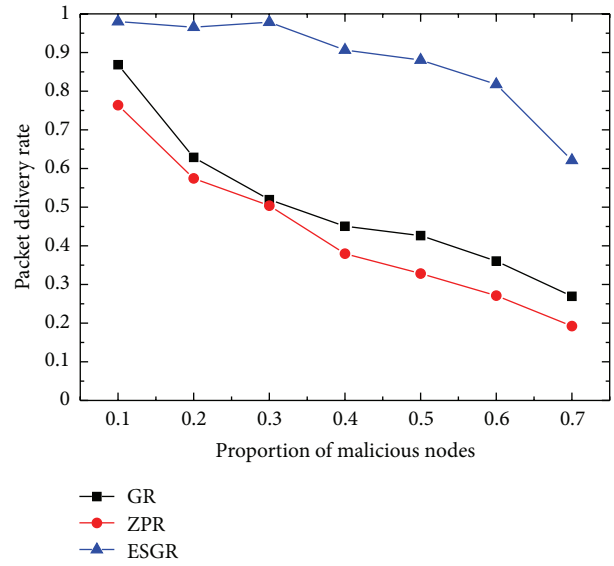
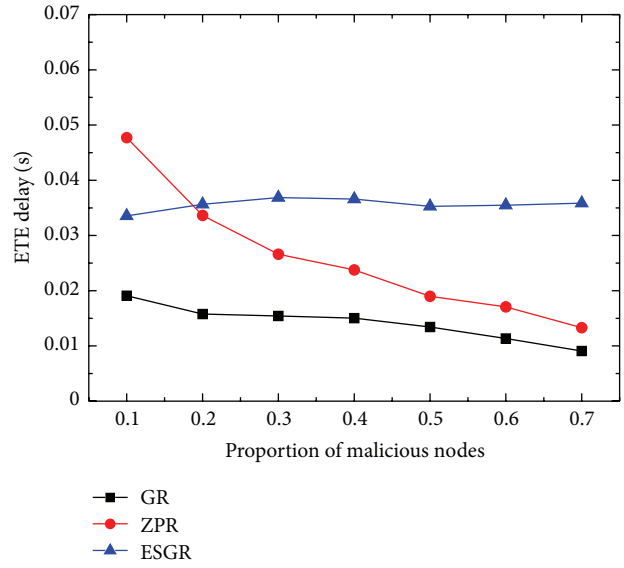



Figure 5: Overhead for different node densities.




(a) Packet delivery rate

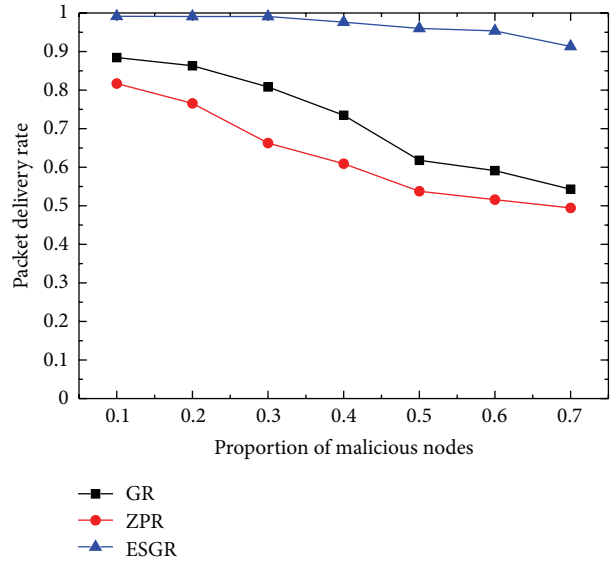



(b) End-to-end delay

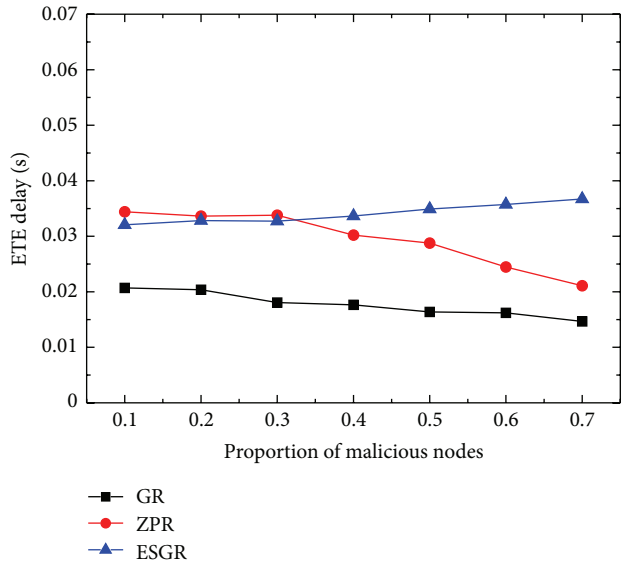
Figure 6: Packet delivery rate and end-to-end delay for different proportions of black hole nodes.

*6.2. Proportion of Malicious Nodes.* We assume that the network size is 100 nodes. The proportion of malicious nodes in the network varies from 0.1 to 0.7. The malicious nodes would drop or tamper the data packets that introduces packet loss. We consider these attacks leading to packet dropping to be black hole and gray hole attacks as they have the same consequences. Then, black hole nodes would drop all the packets that they have received. For gray hole nodes, the probability of packet loss is set to 0.5.

As shown in Figures 6 and 7, the insecure routings GR and ZRP are very sensitive to malicious nodes. The delivery rate falls rapidly even with a small percentage of gray hole nodes. The simulation results also indicate that malicious nodes do not deeply affect the delivery rate of our ESGR. The packet delivery rate falls from close to 100% to more than 85% when half of the sensor nodes are black hole nodes. They can be detected by our distributed trust model, and more candidates in the opportunistic routing are used for transmissions to improve the PDR. We can also find that the performances across all protocols against the gray hole attack are similar to those against the black hole attack. When half of sensor nodes become gray hole nodes and lose packets with a probability of 50%, ESGR can achieve the delivery rate of more than 95% while GR and ZRP only obtain the PDR of approximately 60% in the hostile sensor networks.




(a) Packet delivery rate




(b) End-to-end delay

Figure 7: Packet delivery rate and end-to-end delay for different proportions of gray hole nodes.

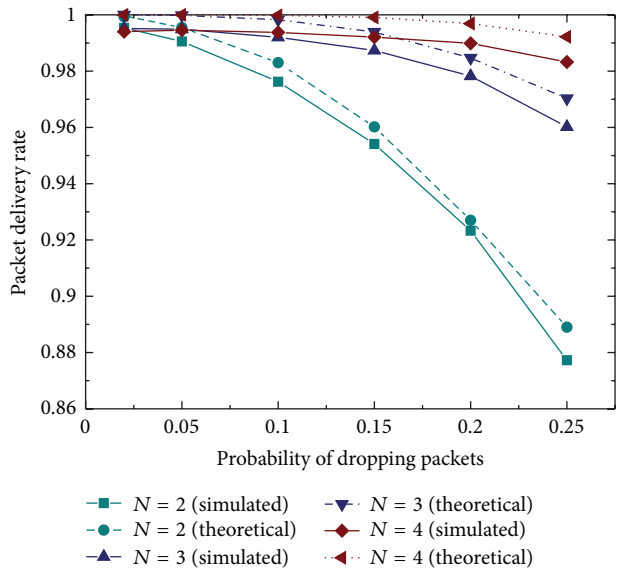



Figure 8: Packet delivery rate for different candidate numbers.

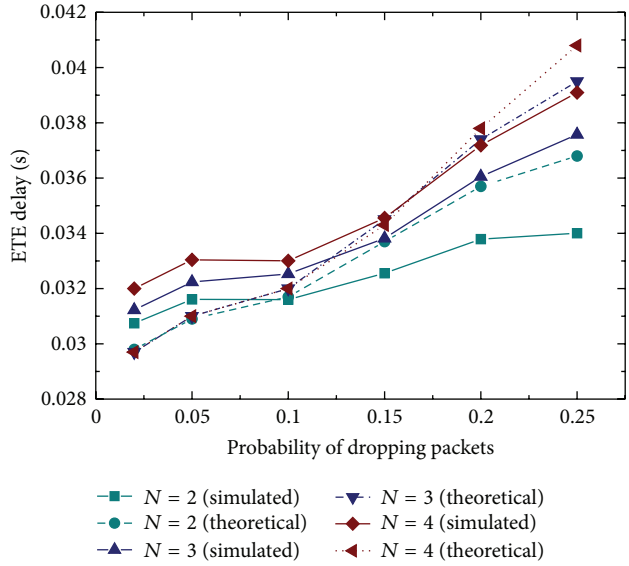



Figure 9: End-to-end delay for different candidate numbers.

As the proportion of malicious nodes is increasing, the average end-to-end delays of GR and ZRP decrease. However, the delay in ESGR grows in the hostile network, as it adopts more candidate nodes to ensure the throughput at the expense of some delay. The link breaks because the black hole or gray hole nodes would introduce suboptimal candidate to relay the data packets. The delay of such packets would raise the average delay. Moreover, the TESLA scheme also introduces the authentication delay, which is one part of the end-to-end delay performance.

Therefore, ESGR can maintain an uninterrupted and secure communication while the delay is growing within the scope of permission.

*6.3. Candidate Number.* The number of candidates could affect the performance of ESGR, as simulated in a network with 200 nodes. The more forwarding candidates are in the network, the higher delivery rate could be achieved to improve the robustness, as shown in Figure 8. Nevertheless, the improved gain becomes slower when $N$ continues increasing. The measured results are almost consistent with our theoretical analysis in Section 5.

The simulation result of the end-to-end delay performance is shown in Figure 9. When many normal nodes become black holes, they would often block the communication leading to a high probability of packet dropping. As a result, the increasing value of $Q$ aggravates the average delay. While many forwarding candidates have joined to

relay packets, the step of the PDR improvement becomes narrow and the end-to-end delay would increase due to the long one-hop delay during a successful delivery. In our comparisons, we ignore the correlation of per-hop in the theoretical analysis, which contributes to the difference between the simulated and the theoretical results.

## 7. Conclusion

In this paper, we design and present an efficient and secure geographical routing against a series of attacks. ESGR requires associative one-way hash function and TESLA mechanism for security. It also makes use of the broadcast nature of wireless channel and forwards packets based on the opportunistic approach. Furthermore, our routing metric is combined with a novel distributed trust model to defend against packet tempering and dropping incurred by the Sybil attack, wormhole attack, black hole attack, and so forth.

ESGR is evaluated under various network configurations such as node density, proportion of malicious nodes, and candidate number. We show that the protocol can tolerate the existence of the malicious nodes and still maintain a high throughput at the expense of some acceptable delay in hostile sensor networks. In the future work, the mobility of sensor nodes will be introduced in the network scenarios, and extended analysis against collusion attacks will be conducted.

## Acknowledgments

This paper is supported by the National Science and Technology Major Projects (Grant no. 2012ZX03002011-002), National Natural Science Foundation of China (Grant no. 61103040), and Doctoral Fund of Ministry of Education of China (Grant no. 20120073110094). Finally, the authors would like to thank the anonymous reviewers for their helpful comments.

## References

[1] C. E. Perkins and P. Bhagwat, "Highly dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers," in *Proceedings of the ACM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM '94)*, pp. 234–244, London, UK, October 1994.

[2] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*, vol. 353, pp. 153–181, Kluwer Academic Publishers, Norwell, Mass, USA, 1996.

[3] C. E. Perkins, E. Belding-Royer, and S. Das, "RFC3561: ad hoc ondemand distance vector (AODV) Routing," Internet RFCs, 2003.

[4] Z. J. Haas, "A new routing protocol for the reconfigurable wireless networks," in *Proceedings of the IEEE Conference on Universal Personal Communications (ICUPC '97)*, pp. 562–566, San Diego, Calif, USA, October 1997.

[5] B. Karp and H. T. Kung, "GPSR: greedy perimeter stateless routing for wireless networks," in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking (MOBICOM '00)*, pp. 243–254, Boston, Mass, USA, August 2000.

[6] J. Hightower and G. Borriello, "Location systems for ubiquitous computing," *Computer*, vol. 34, no. 8, pp. 57–66, 2001.

[7] H. Chan and A. Perrig, "Security and privacy in sensor networks," *Computer*, vol. 36, no. 10, pp. 103–105, 2003.

[8] J. Newsome, E. Shi, D. Song, and A. Perrig, "The Sybil attack in sensor networks: analysis & defenses," in *Proceedings of the 3rd International Symposium on Information Processing in Sensor Networks (IPSN '04)*, pp. 259–268, April 2004.

[9] M. García-Otero, T. Zahariadis, F. Álvarez et al., "Secure geographic routing in ad hoc and wireless sensor networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, Article ID 975607, 2010.

[10] K. Liu, N. Abu-Ghazaleh, and K.-D. Kang, "Location verification and trust management for resilient geographic routing," *Journal of Parallel and Distributed Computing*, vol. 67, no. 2, pp. 215–228, 2007.

[11] Y.-C. Hu, A. Perrig, and D. B. Johnson, "Packet leashes: a defense against wormhole attacks in wireless networks," in *Proceedings of the 22nd Annual Joint Conference on the IEEE Computer and Communications Societies (INFOCOM '03)*, pp. 1976–1986, April 2003.

[12] A. Perrig, R. Canetti, J. Tygar, and D. Song, "The TESLA broadcast authentication protocol," *RSA CryptoBytes*, vol. 5, no. 2, 2002.

[13] S. Biswas and R. Morris, "Exor: opportunistic multi-hop routing for wireless networks," in *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM '05)*, pp. 133–143, Philadelphia, Pa, USA, 2005.

[14] S. Chachulski, M. Jennings, S. Katti, and D. Katabi, "Trading structure for randomness in wireless opportunistic routing," in *Proceedings of the ACM Conference on Computer Communications (SIGCOMM '07)*, pp. 169–180, Kyoto, Japan, August 2007.

[15] S. Yang, F. Zhong, C. K. Yeo, B. S. Lee, and J. Boleng, "Position based opportunistic routing for robust data delivery in MANETs," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '09)*, pp. 1325–1330, December 2009.

[16] Y.-C. Hu, D. B. Johnson, and A. Perrig, "SEAD: secure efficient distance vector routing for mobile wireless ad hoc networks," *Ad Hoc Networks*, vol. 1, no. 1, pp. 175–192, 2003.

[17] B. Awerbuch, R. Curtmola, D. Holmer, C. Nita-Rotaru, and H. Rubens, "ODSBR: an on-demand secure Byzantine resilient routing protocol for wireless ad hoc networks," *ACM Transactions on Information and System Security*, vol. 10, no. 4, 2008.

[18] M. Yu, M. Zhou, and W. Su, "A secure routing protocol against byzantine attacks for MANETs in adversarial environments," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 1, pp. 449–460, 2009.

[19] J.-H. Song, V. W. S. Wong, and V. C. M. Leung, "Secure position-based routing protocol for mobile ad hoc networks," *Ad Hoc Networks*, vol. 5, no. 1, pp. 76–86, 2007.

[20] A. Perrig, R. Canetti, D. Song, and J. D. Tygar, "Efficient and secure source authentication for multicast," in *Proceedings of the Symposium on Network and Distributed System Security (NDSS '01)*, February 2001.

[21] R. Marin-Perez and P. M. Ruiz, "SBGR: a simple self-protected beaconless geographic routing for wireless sensor networks," in *Proceedings of the 8th IEEE International Conference on Mobile Ad-Hoc and Sensor Systems (MASS '11)*, pp. 610–619, October 2011.

[22] A. D. Wood, L. Fang, J. A. Stankovic, and T. He, "SIGF: a family of configurable, secure routing protocols for wireless sensor networks," in *Proceedings of the 4th ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN '06)*, pp. 35–48, Alexandria, Va, USA, October 2006.

[23] Y. Liu, S. Shi, and X. Zhang, "Balance-aware energy-efficient geographic routing for wireless sensor networks," in *Proceedings of the 8th Conference of Wireless Communications, Networking and Mobile Computing (WiCOM '12)*, Shanghai, China, 2012.

[24] J. A. Sanchez, R. Marin-Perez, and P. M. Ruiz, "Beacon-less geographic multicast routing in a real-world wireless sensor network testbed," *Wireless Networks*, vol. 18, no. 5, pp. 565–578, 2012.

[25] R. Kleinberg, "Geographic routing using hyperbolic space," in *Proceedings of the 26th IEEE International Conference on Computer Communications (INFOCOM '07)*, pp. 1902–1909, May 2007.

[26] R. Sarkar, X. Yin, J. Gao, F. Luo, and X. D. Gu, "Greedy routing with guaranteed delivery using Ricci flows," in *Proceedings of the International Conference on Information Processing in Sensor Networks (IPSN '09)*, pp. 121–132, April 2009.

[27] X. Xiang, X. Wang, and Z. Zhou, "Self-adaptive on-demand geographic routing for mobile ad hoc networks," *IEEE Transactions on Mobile Computing*, vol. 11, no. 9, pp. 1572–1586, 2012.

[28] G. Tan and A.-M. Kermarrec, "Greedy geographic routing in largescale sensor networks: a minimum network decomposition approach," *IEEE/ACM Transactions on Networking*, vol. 20, no. 3, pp. 864–877, 2012.

[29] T. S. Rappaport, *Wireless Communications: Principles and Practice*, Prentice-Hall, Englewood Cliffs, NJ, USA, 2002.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 620945, 8 pages
http://dx.doi.org/10.1155/2013/620945

*Research Article*

# Energy-Efficient Routing Algorithms Based on OVSF Code and Priority in Clustered Wireless Sensor Networks

**Xiaoling Wu,[1] Yangyang Wang,[1,2] Guangcong Liu,[3] Jianjun Li,[1] Lei Shu,[4] Xiaobo Zhang,[3] Hainan Chen,[1,3] and Sungyoung Lee[5]**

[1] *Guangzhou Institute of Advanced Technology, Chinese Academy of Sciences, Guangzhou 511458, China*
[2] *Jinan University, Guangzhou 510632, China*
[3] *Guangdong University of Technology, Guangzhou 510006, China*
[4] *Guangdong University of Petrochemical Technology, Maoming 525000, China*
[5] *Department of Computer Engineering, Kyung Hee University, Yongin 446701, Republic of Korea*

Correspondence should be addressed to Sungyoung Lee; sylee@oslab.khu.ac.kr

Received 25 December 2012; Revised 1 June 2013; Accepted 2 June 2013

Academic Editor: Adel Soudani

Copyright © 2013 Xiaoling Wu 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.

Energy awareness is a vital design issue in wireless sensor networks. Since the amount of sensing data may be large and sensor nodes are usually battery-powered, it is critical to design energy-efficient routing algorithms to prolong network lifetime. Given a certain sensor deployment, the routing strategy of sensor data would have profound effects on the communication cost. In this paper, based on low-energy adaptive clustering hierarchy (LEACH) series protocols which are low-energy consumption adaptive clustering routing protocols, we propose the OVSF mechanism based routing protocol and EERPP (Energy-Efficient Routing Protocol based on Priority). Simulation results of OVSF mechanism based protocol and EERPP demonstrate a significant improvement on the network metrics such as the lifetime and the end-to-end delay.

## 1. Introduction

Wireless sensor network (WSN) in which sensor nodes, sink, and management node are deployed, is a new kind of data-centric wireless network. It has been the focus of a lot of recent research and development. As widely used in military, biology, transportation, environmental science, health monitoring and space exploration, and so forth, it has been recognized as one of the most important technologies in the 21st century [1].

The primary purpose of a wireless sensor network is to sense the environment and collect or relay the collected information. Many routing, power management, and data dissemination protocols have been specifically designed for WSNs where energy awareness is an essential design issue [2, 3].

To prolong the lifetime of wireless sensor network, innovative techniques that improve energy efficiency are highly required. The above constraints combined with a typical deployment of a large number of sensor nodes pose many challenges to the design and management of WSNs and necessitate energy awareness at all layers of the networking protocol stack. At the network layer, it is highly desirable to find solutions for energy-efficient route discovery and relaying of data from the sensor nodes to a sink or base-station (BS) so that the lifetime of the network is maximized.

Routing in WSNs is very challenging due to its inherent characteristics that distinguish them from contemporary communication and wireless ad hoc networks. First, it is not possible to generate a global addressing scheme for the deployment of sensor nodes. Therefore, classical IP-based protocols cannot be directly applied to sensor networks. Second, in contrast to typical communication networks almost all applications of sensor networks require the transmission of sensed data from multiple sources to a particular sink. Third, the generated data traffic has significant redundancy since multiple sensors may generate same data within the vicinity of an event. Such redundancy needs to be well handled

by the routing protocols to improve energy and bandwidth utilization. Fourth, sensor nodes are constrained in terms of transmission power, energy, processing capacity, and storage. Thus, they require careful resource management.

Due to such differences, many new algorithms have been proposed for the routing problem in WSNs [4, 6–17]. One of the famous hierarchical network routing protocols is low-energy adaptive clustering hierarchy (LEACH), which has been widely utilized for its energy efficiency and simplicity [18]. In the clustering environment, sensing data gathered by the nodes is transmitted to BS through cluster heads (CHs). As the nodes will communicate data over shorter distances in such an environment, the energy spent in the network is likely to be substantially lower compared to when every sensor communicates directly to BS. To this end, various clustering algorithms have been proposed in different context. Most algorithms aim at generating the minimum number of clusters and minimum transmission distance. These algorithms also distinguish themselves by how the CHs are elected. The LEACH algorithm and its related extension [19] are based on continuous cycle of cluster reconstruction which can be described using "round" concept. Each cycle is divided into two stages: cluster building stage and the stability of the data transmission stage. The protocol is by means of selecting new CH to balance the energy of the nodes to improve the wireless sensor network lifetime. At the same time, to realize the irrelevant data transmission among nodes, it uses the TDMA (Time Division Multiple Address) mechanism [20]. To some extent, it improves the performance of wireless sensor network. However, the TDMA mechanism incises a transmission time frame into several slots which request the routing protocol to do time synchronization on the total system. It has higher delay and not very efficient if we put forward higher requirements on the energy utilization. Moreover, in TDMA based mechanism, the node always transmits the data in its time slot but ignores its own requirement, which may cause unnecessary energy consumption, while OVSF based mechanism solves this problem by checking the nodes' own requirement, for example, whether the nodes have sensed interesting information or not.

In this paper, we propose (1) an improved protocol based on LEACH series protocols which we select instead of the TDMA mechanism, and (2) a new transmission algorithm which is EERPP (Energy-Efficient Routing Protocol based on Priority) for each CH to transmit their own cluster members data or relay the data.

The remainder of the paper is organized as follows. Section 2 presents related work. Section 3 explains the improved mechanism based on OVSF code. Section 4 describes the proposed new transmission algorithm EERPP. In Section 5, we evaluate the performance of improved protocol by simulation. Section 6 concludes the paper.

## 2. Related Work

Many researchers have devoted themselves to the optimal routing protocol in WSNs. The proposed routing protocol can be divided into two categories: one is the plain routing protocol, such as direct diffusion (DD) [21] and security protocol for sensor networks, and the other is the clustering routing protocol, such as low-energy adaptive clustering hierarchy (LEACH). In this paper, we focus on the clustering routing category.

To address the issue of limited communication bandwidth and energy in WSN, Heinzelman et al. firstly propose a low-energy adaptive clustering hierarchy (LEACH) application-specific protocol [18]. It improves system lifetime compared with general-purpose multihop approaches. Recently, a variety routing protocols have been proposed which were improvement version of LEACH protocol. Younis et al. [22] propose REED (Robust Energy-Efficient-Distributed clustering) for clustering sensors deployed in hostile environments in an interleaved manner with low complexity. REED is a self-organized clustering method which constructs independent sets of CH overlays on the top of the physical network to achieve fault tolerance. Each sensor must reach at least one CH from each overlay. Attea and Khalil proposed a new evolutionary based routing protocol for clustered heterogeneous WSNs [23]. The aim was to alleviate the undesirable behavior of the Evolutionary Algorithms (EAs) when dealing with clustered routing problem in WSN by formulating a new fitness function that incorporates two clustering aspects, namely, cohesion and separation error. Cheng et al. proposed NHRPA, a novel hierarchical routing protocol algorithm for WSNs in [24]. The proposed routing protocol adopts routing technology for the nodes based on the distance of nodes to the base station, density of nodes distribution, and residual energy of nodes. The evaluation results show that the proposed routing protocol algorithm is more efficient for wireless sensor networks in terms of the energy usage, packet latency, and security in the presence of node compromise attacks. Jiang et al. present a QoS-guaranteed coverage precedence routing algorithm in [25] to accommodate both energy-balance and coverage-preservation for sensor nodes in WSNs. The energy consumption for radio transmissions and the residual energy over the network are taken into account when the proposed protocol determines an energy-efficient route for a packet. The proposed protocol is able to increase the duration of the on-duty network and provide extra service time with full sensing coverage compared with LEACH and the LEACH-Coverage-U protocols, respectively. Sun and Gu [26] propose an energy-efficient clustering scheme based on LEACH (low-energy adaptive clustering hierarchy), that is, LEACH-Energy Distance (LEACH-ED). In LEACH-ED, CHs are selected by a probability based on the ratio between residual energy of node and the total current energy of all of the sensor nodes in the network. LEACH-ED is a kind of self-organized protocol based on LEACH. Almost all these improvements based on LEACH are to research a better clustering method to improve the performance of WSNs. However, the TDMA (Time Division Multiple Access) mechanism used or assumed in these protocols may cause long end-to-end delay and consequently high energy consumption because nodes have to send data to the sink at its own distribution time slot. To address this issue, we propose an improved routing protocol based on OVSF code transmission instead of TDMA. Compared with TDMA

based protocol, the proposed OVSF based routing protocol greatly reduces time delay and improves the network energy utilization efficiency. Moreover, it also facilitates sensor nodes localization. As we know, LEACH is based on "round" and there are two stages in each "round." If we only consider the CH selection methods and ignore other aspects, it is not efficient on the performance improvement of WSNs. Thus we propose a new OVSF code based protocol to replace TDMA which greatly improves the utilization ratio of energy and the delay of WSNs. We do not focus on the clustering mechanism but the data transmission mechanism.

## 3. Improved Mechanism Based on OVSF

*3.1. The Principle and Mathematical Theory of OVSF.* OVSF (orthogonal variable spreading factor) which is firstly proposed in 1997 by Adachi is a variable spread spectrum factors based on orthogonal code [27–32]. Taking advantage of its orthogonal and inherence, it perfectly complete the data transmission in wireless communication.

The mathematical theory of OVSF is as follows.

Set $\{C_N(n)\}$ where $n = 1, \ldots, N$ is a collection of orthogonal spread spectrum code which has length $N$. If the $N$ orthogonal $C_N(n)$ is regarded as a row vector with length $N$, the row vector forms an $N * N$ matrix $C_N$ which can be recursively generated according to the following formula by $C_{N/2}$:

$$C_N = \begin{bmatrix} C_N(1) \\ C_N(2) \\ C_N(3) \\ C_N(4) \\ \vdots \\ C_N(N-1) \\ C_N(N) \end{bmatrix} = \begin{bmatrix} C_{N/2}(1)\, C_{N/2}(1) \\ C_{N/2}(1)\, C_{N/2}(1) \\ C_{N/2}(2)\, C_{N/2}(2) \\ C_{N/2}(2)\, C_{N/2}(2) \\ \vdots \\ C_{N/2}\left(\frac{N}{2}\right) C_{N/2}\left(\frac{N}{2}\right) \\ C_{N/2}\left(\frac{N}{2}\right) C_{N/2}\left(\frac{N}{2}\right) \end{bmatrix}. \quad (1)$$

The orthogonal code of variable length also can be formed with the structure of recursive tree. The spanning tree is shown in Figure 1.

The proposed mechanism based on OVSF code realizes the optimal transmission by selecting better code when the number of cluster members is smaller than the vector length. At the beginning of the data transmission, the mechanism selects an optimal $N$ as the layer of OVSF code spanning tree. The selection algorithm is shown in Figure 2.

*3.2. The Process of Data Transmission Based on OVSF.* The proposed OVSF based mechanism follows the process below:

(1) the CH sends the OVSF matrix to its cluster members;

(2) the nodes add a group of OVSF code in front of the data;

(3) at the CH, we use the same OVSF matrix multiplied by the received data. Transmission process and decoding principle are shown in Figure 3.

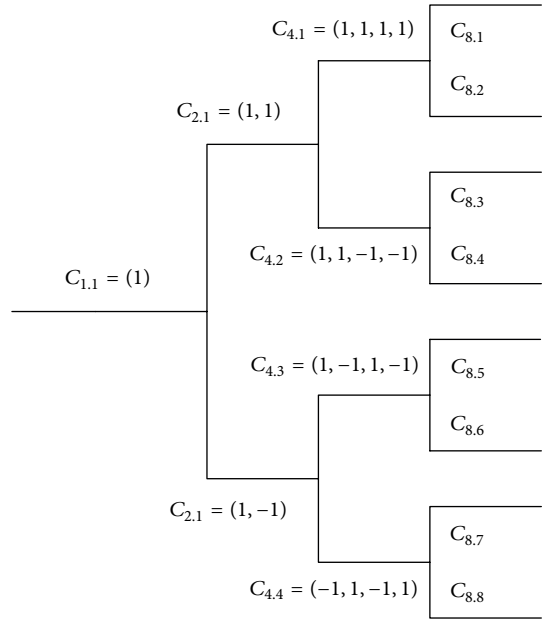



Figure 1: The spanning tree of OVSF.

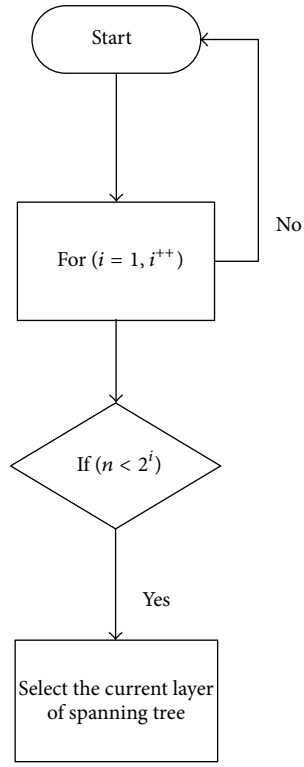



Figure 2: The selection algorithm of spanning tree layer.

For the OVSF code distribution mechanism, we select the DCA (Dynamic Code Allocation Scheme). As the name suggests, the allocated codes are dynamic in nature in the sense that a particular session may start transmitting with a particular OVSF code and end its session with a different code. The decision algorithm is shown in Figure 4.

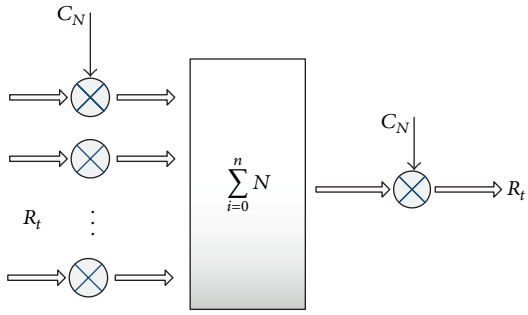



Figure 3: The transmission and decoding process.

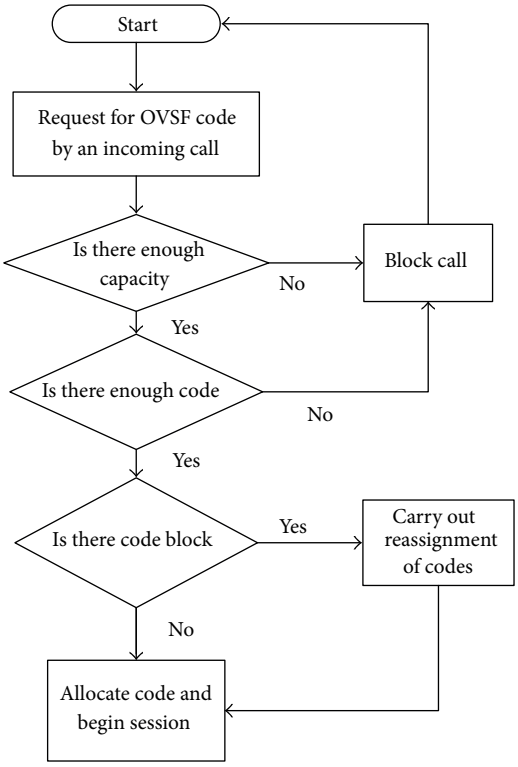



Figure 4: The distribution algorithm of DCA.

Through the DCA algorithm, while ensuring efficient use of the code resource the delay can be reduced and the throughput of WSNs can be improved. The proposed OVSF mechanism outperforms TDMA on the delay, energy consumption, and throughput.

As Figure 3 shows, the data flow $R(t)$ is transmitted by multiplying its distributed OVSF code and when the data flow arrives at the CH they are multiplied by the same OVSF matrix $C_N$. The CH identifies the data of its member through the OVSF code which is distributed at the stage of cluster building. As each OVSF code multiplies the matrix $C_N$ whose rows are orthogonal to each other, we can decode the signals using (2) which shows the decoding principle:

$$\sum_{i=1}^{N_i} C_i(N_1, n_1) * C_j(N_2, n_2) = \begin{cases} 0 & n_2 \neq n_1 + jn_2, \\ N_1 & n_2 = n_1 + jn_2. \end{cases} \quad (2)$$

The data transmission process can be represented by

$$R(t) * C_i * C_j = \begin{cases} 0 & i \neq j, \\ iR(t) & i = j, \end{cases} \quad (3)$$

where $i$ is the serial number of the distribution of OVSF code and $j$ is the row number of the matrix $C_N$ which is multiplied by the data flows at the CH. The orthogonality and incoherence of each node has been utilized not only for indiscriminate data transmission, but also for reducing the delay and improving the energy consumption rate compared with the TDMA.

## 4. The Energy-Efficient Routing Protocol Based on Priority (EERPP)

*4.1. The Model of EERPP.* As mentioned above, during the stability of the data transmission stage, the CHs collect the sensing data of cluster members and transmit or relay messages to the sink. In this mechanism, the duty of CHs is to transmit the data of their own cluster members or act as a relay. When the CH acts as a relay, it transmits the data all the time until the sink receives the sensing data. Considering the possible high energy consumption due to the repetitive transmissions, it is obvious that this mechanism is not efficient enough and there is some space to improve.

To solve the above inefficient mechanism, we propose the EERPP for CH to transmit or relay the data. With the EERPP, each CH is given a transmission priority according to the distance to the sink. In other words, the CH which is closer to the sink would have higher priority to send the data to sink and the other CHs would go to sleep stage to avoid the repetitive data transmission. The proposed EERPP is a transmission mechanism based on priority for better energy balance.

Firstly, we define a region to assign the transmission priority. According to the distance to the sink, we define the topology of WSNs to $k$ regions through (4). The illustration of the region division and transmission priority is shown in Figure 5,

$$k = \frac{[d]}{d_{\text{sensor}}}, \quad (4)$$

where the value of $k$ stands for the priority.

As shown above, the topology of the wireless sensor network is divided into $k$ ($k \geq 1$) areas. Then we assign the priority for every CH to send its own data or relay others' data. Under the EERPP, $R_1$ has the highest priority to send data to sink than all the other CHs. To sum up, the CH $R_j$ has higher transmission priority when $j$ is lower than other CHs' priority. The value of each CH's index $j$ is computed in

$$j = \begin{cases} 1 & d_{\text{sensor}} \geq d, \\ \left\{1 + d - \left[\frac{(d_{\text{sensor}})}{d}\right] * (k-1)\right\} * \beta & d_{\text{sensor}} < d. \end{cases} \quad (5)$$

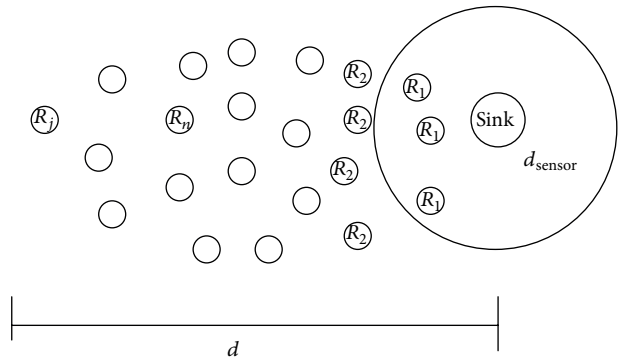



Figure 5: The illustration of transmission priority.

*4.2. The Algorithm of EERPP.* The algorithm is divided into two stages. Firstly, we compute the value of region $k$ where $d$ is the distance from CH to sink and $d_{\text{sensor}}$ is the sensor's detection distance. We calculate the number of region according to the sensor's sensing distance ($d_{\text{sensor}}$). To reduce the data replication, we select the $d_{\text{sensor}}$ as the distance of one hop. As the result, the algorithm could guarantee the successful data transmission and reduce unnecessary energy cost.

Then we calculate the index of CH in (5) where $\beta$ is the weight value which affects the transmission priority so as to affect the performance of WSN. The reason of introducing $\beta$ is that the multipath fading is considered when the WSN is built in certain severe environment. At different working environment, we can adjust the value of $\beta$ to get the optimal energy consumption rate.

The procedure of the proposed algorithm for efficient data transmission from the CH to the sink is as follows.

Step 1: calculate the value of $k$ using (3).

Step 2: calculate the value of $j$ using (4).

Step 3: receive the members' data or other clusters' data.

Step 4: go to the transmission mechanism

While the data is from its own members, transmit to next CH whose $R_{jt}$ is $R_{j+1}$.

While the data is not from its own members, go to relay mechanism.

If ($j == 1$) transmits data to sink, else the $R_{j++}$ transmits data to next CH.

When the CHs are not at the stage of transmission, they would be turned into sleep to save energy. Through this priority based transmission mechanism it would prolong the lifetime of WSNs.

## 5. Simulation

In this paper, we select NS-2.34 based on Ubuntu as the simulation platform to evaluate the OVSF and priority based transmission mechanisms that we proposed. Details of the simulation environment are given in Table 1.

Table 1: Simulation scenario parameters.

| Network parameters | Value |
|---|---|
| Simulation area (m ∗ m) | 100 ∗ 100 |
| Number of node | 101, 201 |
| The station of Sink | (200, 200) |
| Initial energy (J) | 5 |
| Excvr (J/bit) | $50E - 9$ |
| Packet-size/bit | 512 |
| Simulation time (s) | 600 |
| Distance of sensor sensing ($d_{\text{sensor}}$) | 10 |
| Shadowing model | Lognormal |
| Fading model | Ricean |
| Deployment | Random |

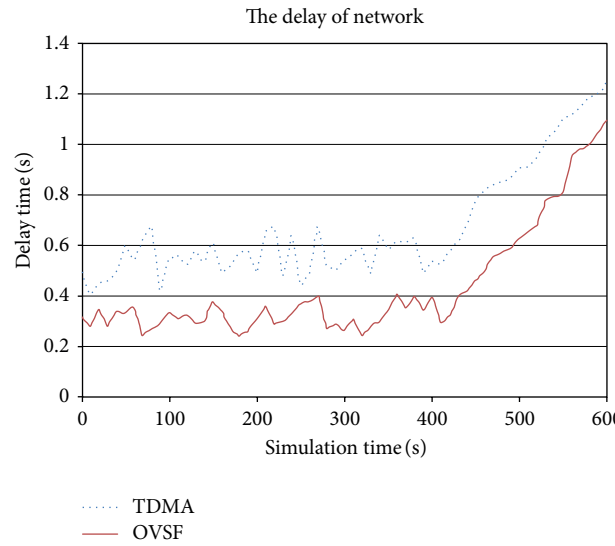



Figure 6: The delay of network [4, 5].

Firstly, we compare the network delay and energy consumption between the TDMA and OVSF mechanism. In the simulation, we use the number of alive-nodes to represent the energy consumption. We stipulate that the protocol in which there are more alive-nodes has lower energy consumption in the network. This paper is an extension of our previously published papers [4, 5] in which the relational results are shown in Figures 6 and 7 with initial random deployment of nodes in every time execution. It is clear that the OVSF mechanism outperforms TDMA mechanism which is applied in LEACH and its related extensions in terms of network lifetime and end-to-end delay.

Secondly, we apply the EERPP algorithm and compare the performance with the OVSF mechanism when we set $\beta = 1$. The simulation result is shown in Figure 8. As we explained before, the protocol which has more remaining nodes has lower energy consumption, so EERPP outperforms OVSF based mechanism on energy consumption; furthermore, it outperforms LEACH series protocols.

Finally, we simulate the performance of EERPP while the value of $\beta$ is different from 0 to 1. The simulation result is

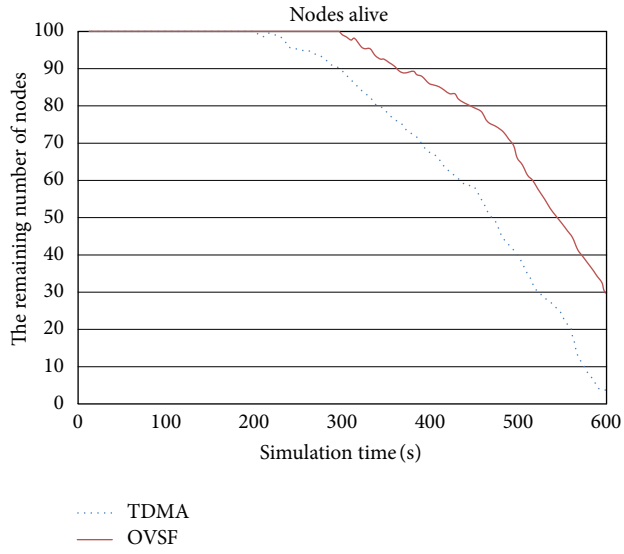



Figure 7: The remaining number of nodes [4, 5].

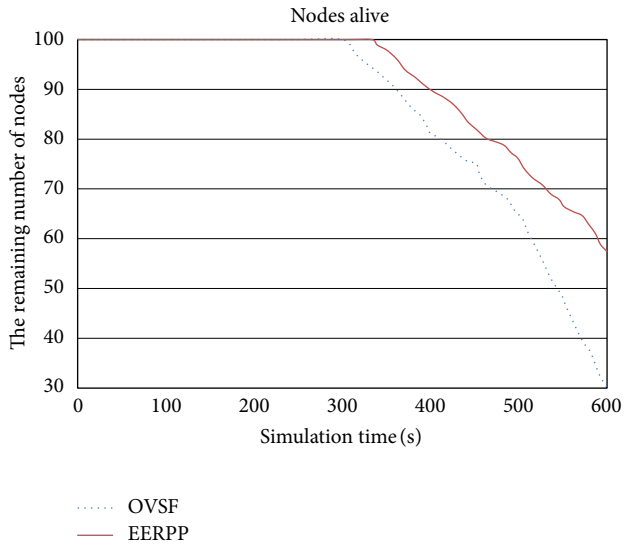



Figure 8: The comparison between EERPP and OVSF.

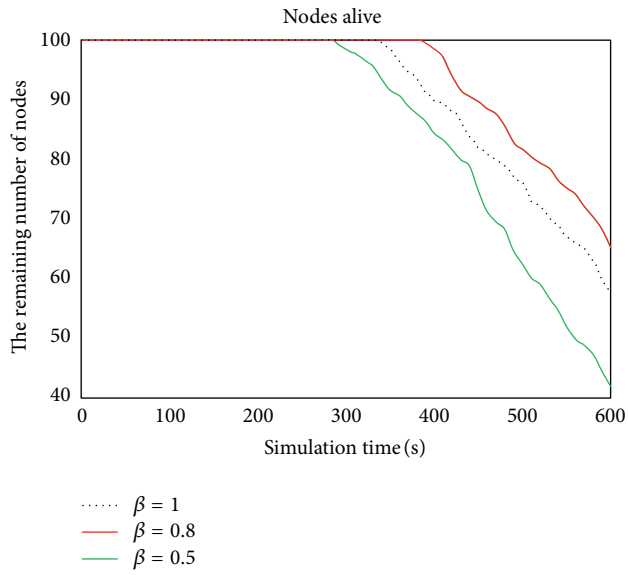



Figure 9: EERPP with different values of $\beta$.

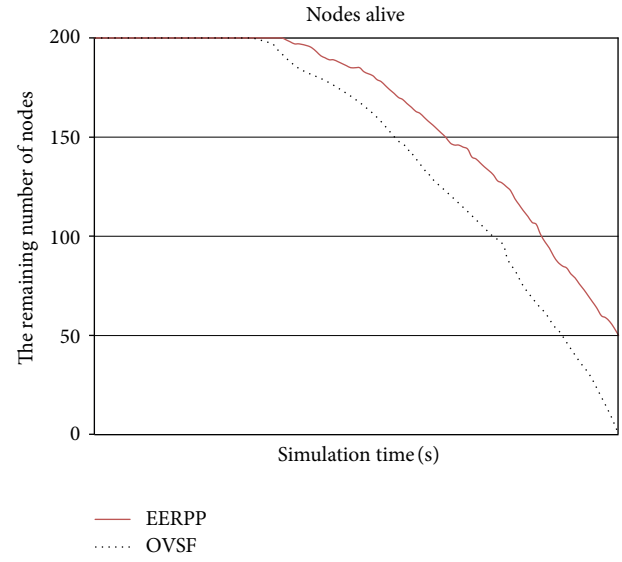



Figure 10: The comparison between EERPP and OVSF at 200 ∗ 200 ($m^2$).

shown in Figure 9. The performance shows the best when $\beta$ is around 0.8.

To prove that the improved routing protocol has scalability, we reset the simulation parameters and simulate the remaining number of nodes. We set the simulation area as 200 ∗ 200 ($m^2$) and the number of nodes is configured to 200. At the same time, we adjust the simulation time to 1000 (s). The value of $\beta$ is set to 0.8. The simulation result is shown in Figure 10 and shows similar characteristics with 100 ∗ 100 ($m^2$) area which proves on some aspect the scalability of the proposed algorithms.

As shown above, the OVSF mechanism is better than the LEACH and its extensions which use TDMA mechanism in terms of the delay and energy consumption. The proposed new CH transmission algorithm which is EERPP, the remaining number of nodes is higher than the OVSF mechanism while it is certainly higher than the TDMA mechanism. At the same time, we simulate the proposed mechanisms in the deployment area with different size to prove that the improved routing protocol based on OVSF and EERPP has scalability when evaluating the energy consumption rate. We also simulate with different values of $\beta$ to see how the weight values affect the energy consumption of nodes. Based on the above results, we can conclude that the improved protocol based on OVSF code and EERPP has significant advantages on the delay and energy consumption over LEACH series routing protocols.

## 6. Conclusion

In this paper, we propose an improved protocol based on OVSF code and an Energy-Efficient Routing Protocol based on Priority (EERPP) in a cluster based sensor networks. From the simulation results, we can conclude that the OVSF code mechanism has advantages compared with the LEACH series protocols utilizing TDMA mechanism in terms of both network delay and lifetime, while the proposed EERPP is better than LEACH and OVSF based mechanism in terms of energy consumption. At the same time, we can also decide an optimal value of $\beta$ which is near 0.8 to improve the performance to adapt to different environment where WSNs are deployed.

## Acknowledgments

This research was supported by the MKE (The Ministry of Knowledge Economy), Republic of Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency)" (NIPA-2011-(C1090-1121-0003)), and 2013 Guangdong Province Petrochemical Equipment Fault Diagnosis Key Laboratory Open Fund.

## References

[1] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Networks*, vol. 52, no. 12, pp. 2292–2330, 2008.

[2] G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: a survey," *Ad Hoc Networks*, vol. 7, no. 3, pp. 537–568, 2009.

[3] X. Wu, J. Cho, B. J. d'Auriol, and S. Lee, "Mobility-assisted relocation for self-deployment in wireless sensor networks," *IEICE Transactions on Communications*, vol. E90-B, no. 8, pp. 2056–2069, 2007.

[4] Y. Wang, X. Wu, J. Wang, W. Liu, and W. Zheng, "An OVSF code based routing protocol for clustered wireless sensor networks," *International Journal of Future Generation Communication and Networking*, vol. 5, no. 3, 2012.

[5] Y. Wang, X. Wu, J. Wang, W. Liu, and W. Zheng, "An improved routing protocol based on OVSF code transmission in wireless sensor networks," in *Proceedings of the AST/EEC/MMHS/AIA International Conference*, pp. 109–114, 2012.

[6] Z. A. Eu, H.-P. Tan, and W. K. G. Seah, "Opportunistic routing in wireless sensor networks powered by ambient energy harvesting," *Computer Networks*, vol. 54, no. 17, pp. 2943–2966, 2010.

[7] J. Wu, M. Lu, and F. Li, "Utility-based opportunistic routing in multi-hop wireless networks," in *Proceedings of the 28th International Conference on Distributed Computing Systems (ICDCS '08)*, pp. 470–477, July 2008.

[8] J. Shi, Z. Shan, and x. Liu, "A two-level routing scheme for wireless sensor network," *International Journal of Distributed Sensor Networks*, vol. 2012, Article ID 601389, 6 pages, 2012.

[9] P. Zhu and F. Jia, "A novel hybrid self-organizing clustering routing algorithm," *International Journal of Distributed Sensor Networks*, vol. 2012, Article ID 146516, 8 pages, 2012.

[10] W. Liu, S. Zhang, and J. Fan, "A diagnosis-based clustering and multipath routing protocol for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2012, Article ID 504205, 11 pages, 2012.

[11] W. Xiaoling, S. Lei, M. Min, C. Jinsung, and L. Sungyoung, "Coverage-driven self-deployment for cluster based mobile sensor networks," in *Proeedings of the 6th IEEE International Conference on Computer and Information Technology (CIT '06)*, September 2006.

[12] W. Fang, F. Liu, F. Yang, L. Shu, and S. Nishio, "Energy-efficient cooperative communication for data transmission in wireless sensor networks," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 4, pp. 2185–2192, 2010.

[13] J.-H. Ho, H.-C. Shih, B.-Y. Liao, and S.-C. Chu, "A ladder diffusion algorithm using ant colony optimization for wireless sensor networks," *Information Sciences*, vol. 192, pp. 204–212, 2012.

[14] L. Shu, M. Hauswirth, L. Cheng, J. Ma, V. Reynolds, and L. Zhang, "Sharing worldwide sensor network," in *Proceedings of the International Symposium on Applications and the Internet (SAINT '08)*, pp. 189–192, August 2008.

[15] L. Almazaydeh, E. Abdelfattah, M. Al-Bzoor, and A. Al-Rahayfeh, "Performance evaluation of routing protocols in wireless sensor networks," *International Journal of Computer Science and Information Technology*, pp. 64–73, 2010.

[16] X. Wu, B. J. d'Auriol, J. Cho, and S. Lee, "Optimal routing in sensor networks for in-home health monitoring with multi-factor considerations," in *Proceedings of the 6th Annual IEEE International Conference on Pervasive Computing and Communications (PerCom '08)*, pp. 720–725, Hong Kong, China, March 2008.

[17] X. Wu, J. Cho, B. J. D'Auriol, S. Lee, and Y.-K. Lee, "An integrated sleep-scheduling and routing algorithm in ubiquitous sensor networks based on AHP," *IEICE Transactions on Communications*, vol. E90-B, no. 12, pp. 3392–3401, 2007.

[18] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660–670, 2002.

[19] V. A. Geetha, P. V. Kallapurb, and S. Tellajeerac, "Clustering in wireless sensor networks: performance comparison of LEACH and LEACH-C protocols using NS2," *Procedia Technology*, vol. 4, pp. 163–170, 2012.

[20] S. Spinsante, S. Andrenacci, and E. Gambi, "De Bruijn sequences for spread spectrum applications: Analysis and results," in *Proceedings of the 18th International Conference on Software, Telecommunications and Computer Networks (SoftCOM '10)*, pp. 365–369, September 2010.

[21] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, "Directed diffusion for wireless sensor networking," *IEEE/ACM Transactions on Networking*, vol. 11, no. 1, pp. 2–16, 2003.

[22] O. Younis, S. Fahmy, and P. Santi, "An architecture for robust sensor network communications," *International Journal of Distributed Sensor Networks*, vol. 1, no. 3-4, pp. 305–327, 2005.

[23] B. A. Attea and E. A. Khalil, "A new evolutionary based routing protocol for clustered heterogeneous wireless sensor networks," *Applied Soft Computing Journal*, vol. 12, pp. 1950–1957, 2012.

[24] H.-B. Cheng, G. Yang, and S.-J. Hu, "NHRPA: a novel hierarchical routing protocol algorithm for wireless sensor networks," *Journal of China Universities of Posts and Telecommunications*, vol. 15, no. 3, pp. 75–81, 2008.

[25] J. -A. Jiang, T. -S. Lin, C. -L. Chuang et al., "A QoS-guaranteed coverage precedence routing algorithm for wireless sensor networks," *Sensors*, vol. 11, no. 4, pp. 3418–3438, 2011.

[26] Y. J. Sun and X. P. Gu, "Clustering routing based maximizing lifetime for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 5, no. 1, article 88, 2009.

[27] T. Erlebach, R. Jacob, M. Mihalak, M. Nunkesser, G. Szabo, and P. Widmayer, "An algorithmic view on OVSF code assignment," *Algorithmica*, vol. 47, no. 3, pp. 269–298, 2007.

[28] F. Y. L. Chin, Y. Zhang, and H. Zhu, "Online OVSF code assignment in cellular networks," in *Proceedings of the 3rd International Conference on Algorithmic Aspects in Information and Management*, vol. 4508 of *Lecture Notes in Computer Science*, pp. 191–200, Springer, Berlin, Germany, 2007.

[29] Z. Wang, X. Ma, and G. B. Giannakis, "OFDM or single-carrier block transmissions?" *IEEE Transactions on Communications*, vol. 52, no. 3, pp. 380–394, 2004.

[30] M. Karakoc, A. Soke, and A. Kavak, "Using diploidy genetic algorithm for dynamic OVSF code CT allocation in WCDMA networks," in *Proceedings of the IEEE Radio and Wireless Symposium (RWS '07)*, pp. 15–18, January 2007.

[31] D. S. Saini and S. V. Bhooshan, "Performance improvement in 3G and beyond CDMA systems using priority based code assignment scheme," in *Proceedings of the International Conference on Signal Processing, Communications and Networking (ICSCN '07)*, pp. 98–101, February 2007.

[32] F. Adachi, M. Sawahashi, and K. Okawa, "Tree-structured generation of orthogonal spreading codes with different lengths for forward link of DS-CDMA mobile radio," *Electronics Letters*, vol. 33, no. 1, pp. 27–28, 1997.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 564386, 14 pages
http://dx.doi.org/10.1155/2013/564386

*Research Article*

# A Nonuniform Sensor Distribution Strategy for Avoiding Energy Holes in Wireless Sensor Networks

**Guoxi Ma and Zhengsu Tao**

*Department of Electronic, Information and Electrical Engineering, Shanghai Jiaotong University, No. 800, Dongchuan Road, Shanghai 200240, China*

Correspondence should be addressed to Zhengsu Tao; zstao@sjtu.edu.cn

Received 6 June 2013; Accepted 12 June 2013

Academic Editor: J. Barbancho

Copyright © 2013 G. Ma and Z. Tao. 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 energy hole problem exerts great impact on the energy efficiency and lifetime of wireless sensor networks (WSNs) based on many-to-one communication model. Unequal cluster emerged in recent years is a good way to alleviate the energy hole problem by dispersing cluster heads' burden. However, it fails to address this problem fundamentally due to its inherent characteristics. The single non-uniform nodes distribution strategy can alleviate the energy hole problem well by setting more nodes in networks to achieve energy balance, yet it may result in low energy efficiency and high cost of the networks. In this paper, by analyzing and minimizing intra- and inter-cluster energy consumption, we construct a suboptimal unequal cluster for WSNs. We propose a non-uniform sensor distribution strategy based on the previous unequal cluster in accordance with the energy balance principle. Simulation results show that our proposed non-uniform sensor nodes distribution strategy can not only achieve good energy efficiency as the unequal cluster method, but also ensure the network energy consumption balance and resolve the energy hole problem completely as the non-uniform sensor distribution approach. Furthermore, our algorithm needs fewer sensors to be settled than single non-uniform node distribution.

## 1. Introduction

Advances in MEMS-based sensor technology and wireless communications in recent years have contributed to the development of low-cost, low-power, multifunctional sensor nodes that are small in size with short communication distance and weak computational ability. These small sensor nodes are capable of sensing the environment, storing and processing the collected sensor data, and interacting and collaborating with each other within the network [1]. Due to the characteristics of strict energy constraint and non-rechargeable energy provision, the energy resource of sensor networks should be wisely managed to extend the lifetime of sensors. Since energy consumption is of vital importance for wireless sensor networks, much attention has been paid to low-power hardware design, collaborative signal processing techniques, and energy-efficient algorithms at various WSNs [2, 3].

The main goal of WSNs is to collect useful information as much as possible in the monitoring area, which implies that energy efficiency and lifetime of networks are very important. In order to achieve high energy efficiency and increase network lifetime, sensors are often hierarchically organized into clusters. Within a cluster, each node has its own cluster head (CH) and transmits data to its CH over relatively short distance, which in turn forwards the data (or it is further aggregated) to the sink via a single-hop or a multihop path through other CHs. Previous researches have shown that multi-hop communication between a data source and a data sink is usually more energy efficient than direct transmission due to the characteristics of wireless channel. This method can ensure high energy efficiency. However, the energy hole problem arises when the multi-hop forwarding model is used in intercluster communication. As the cluster heads close to the data sink are burdened with heavy relay traffic, they will die much faster than the other cluster heads, reducing sensing coverage and causing network partitioning [2, 4–7]. Although the strategy of rotating the cluster head role ensures that sensors consume energy in a more consistent manner, the energy hole problem described previously still cannot be

eliminated. Experimental results in [4] show that when the network lifetime is over, up to 90% of the total initial energy of the nodes is left unused in uniform sensor distribution WSNs. The unbalanced distribution of communication loads caused by this problem has become a key factor that affects the lifetime of the networks.

In order to solve this problem, many strategies have been proposed. Some proposed clustering algorithms to alleviate this problem by wisely arranging the cluster size. The result is that clusters closer to the base station are expected to have smaller cluster size. By this way, the CHs will consume less energy during the intracluster data processing and preserve more energy for the inter-cluster relay traffic to prevent their premature death. This kind of unequal cluster strategy can alleviate the energy hole problem and achieve better energy efficiency than equal cluster structure in WSNs. But the energy hole problem described previously still cannot be eliminated completely due to its inherent characteristics of many-to-one communication. In nonhierarchical flat structure WSNs, Lian et al. [4] explicitly proposed a nonuniform node distribution strategy to solve the energy hole problem and enhance the data capacity. In [5], the authors also argue that energy depletion balance may be achieved by using the non-uniform node distribution method. Wu et al. [8] proposed a non-uniform node distribution method to address the energy hole problem. These single nonuniform node distribution strategies can solve the energy hole problem but need more nodes to be deployed for sharing the task of data relay without any data aggregation, thus resulting in lower energy efficiency and high cost.

In this paper, by combining the advantages of both unequal cluster and non-uniform node distribution approach, we develop a non-uniform sensor distribution strategy based on unequal cluster to eliminate the energy hole problem. We adopt the corona model of WSNs and divide energy consumption in each cluster into two parts including intra- and intercluster energy consumption. By analyzing the energy consumption relationship between intra-cluster data processing and inter-cluster data forwarding, we derive the expression of the per node energy consuming rate (ECR) based on the distance from clusters to the sink. Through ECR model, we find that if the cluster size (the same as corona width) is reasonably arranged, ECR will be minimized, which means minimal average energy consumption of nodes and longer network lifetime. With the minimal ECR in each corona, we deduce the non-uniform sensor distribution method.

In our strategy, nodes with different density are placed in different coronas according to the previous non-uniform sensor distribution results. If we let $\rho_{\min}$ denote the lowest node density that can meet the requirement of networks coverage and connectivity, the set of sensors with density higher than $\rho_{\min}$ in some coronas will be selected to be put into sleep mode, which means no participation in data sensing and forwarding tasks. All active nodes are organized into unequal clusters. When some nodes die prematurely due to the energy hole problem, the redundant sleeping nodes will wake up in time to ensure the normal operation of networks [9]. Our goal is to ensure that there is nearly no residual energy in networks when networks lose information collection and transmission abilities, rather than that all settled sensors use up their energy at the same time. Theoretical analysis and simulation results show that our non-uniform sensor distribution strategy based on unequal cluster can obtain both the energy efficiency of unequal hierarchy cluster and achieve the energy consumption balance of non-uniform sensor distribution for WSNs. The energy hole problem can be resolved completely using the least sensor nodes. Moreover, the lifetime of network can be longer compared with single unequal cluster protocol or non-uniform node distribution strategy.

The remainder of the paper is organized as follows. Section 2 covers related researches in this area; Section 3 introduces the system assumptions used throughout this paper; Section 4 establishes and analyzes the ECR model of sensor node with intra- and inter-cluster communication and introduces the calculation algorithm for non-uniform sensor distribution strategy; Section 5 elaborates on our simulation efforts and the analysis of the results obtained; Section 6 offers concluding remarks and points out research directions in the future.

## 2. Literature Review

In recent years, a large number of papers have been published on how to extend the network lifetime and solve energy hole problem for WSNs. Up till now, many algorithms about hierarchical cluster and nodes distribution have been proposed for addressing them. In the following part, we will give a brief review to the most important researches and findings related to our approach.

Li and Mohapatra [10] initiated the energy hole problem study in a large many-to-one sensor network. They described the energy hole in a corona model and defined the per node traffic load and the per node energy consuming rate (ECR), both of which are used in our paper. Based on the observation for ECR in each corona, they proved that nodes in inner coronas consume energy much faster and have shorter lifetime. They developed a mathematical model to analyze the energy hole problem and proved that hierarchical deployment and data compression have a positive effect in a uniformly distributed sensor network. Olariu and Stojmenović [11] are the first researchers to study the issue of whether energy hole can be avoided from a theoretical perspective. Assuming a wireless sensor network with uniform node distribution and uniform data reporting functions, they proposed an energy model to analyze the relationship between the network lifetime and the width of each corona in concentric corona model. By further assuming that the transmission range of sensor is adjustable, they demonstrated that when all the coronas have the same width, the energy consumed by routing can be minimized. Moreover, they points out conditions for avoiding the unbalanced energy depletion problem.

The major purpose for eliminating the energy hole problem is to enhance the network energy efficiency and prolong the lifetime. In flat WSNs, many authors [9, 12, 13] adopted the

strategy of adjusting the transmission power of nodes to avoid the energy hole problem. By assigning different transmission radii according to the distance from sensors to sink, energy hole problem can be alleviated. Considering the energy consumption distribution for single-hop and multihop communication to the sink, Perillo et al. [14] proposed an alternate mode between multihop and singlehop to achieve energy consumption balance. They calculated the optimization of network lifetime in a linear programming problem. But the authors only consider that nodes use up their energy simultaneously without thinking of using energy efficiently to collect useful data. For hierarchical network algorithms based on cluster architecture, Heinzelman et al. [6] proposed the LEACH algorithm. Due to the communication with the base station by single hop, lots of energy is consumed in the long-distance communication. Many other cluster-based hierarchy algorithms are improved from LEACH [15–20]. Benefiting from data aggregation for redundant sensing data and the decrease of communication distance among nodes, hierarchical network algorithms can achieve better energy efficiency and energy consumption balance than flat algorithms. The energy hole problem is alleviated to some extent but is far from being solved.

To avoid the energy hole problem in multi-hop communication WSNs, unequal cluster concept which is different from the general equal hierarchical cluster algorithms has been adopted to extend the network lifetime in these years [21–24]. Soro and Heinzelman [21] investigated firstly an unequal clustering size model (UCS) to balance the energy consumption of cluster heads in multi-hop heterogeneous WSNs. Through both theoretical and experimental analyses, they proved that unequal cluster could be useful, especially for heavy traffic applications. EEUC [24, 25] has been proved as an efficient algorithm to address energy hole problem. In EEUC, clusters closer to the base station are smaller in size than those farther away from sink; thus, cluster heads closer to the base station can save some energy for forwarding inter-cluster data. This algorithm works well in balancing the energy consumption among cluster heads and slowing down their premature death. However, due to the inherent characteristics of a many-to-one communication, energy hole problem cannot be eliminated completely.

Another strategy to avoid the energy hole is the node density control, which can balance the energy consumption effectively. Lian et al. [4] proposed the non-uniform node distribution strategy in wireless sensor networks. The energy hole is caused by massive energy consumption near the sink. The energy hole problem can be solved completely by deploying more nodes near the sink to relay the distant data. Wu et al. [26] proposed a non-uniform node distribution strategy to achieve the subbalanced energy depletion. The authors point out that if the number of nodes in every corona increases in geometric progression with a predetermined ratio, the network can achieve balanced energy depletion. Olariu and Stojmenović [11] discussed the non-uniform node distribution strategy in wireless sensor networks. Assuming an energy consumption model in which only energy consumption for data transmission is considered, they proved that balanced energy depletion can be achieved when the node density of each corona is arranged proportionally. In their scheme, nodes near the sink have to send data at a lower rate. Compared with the hierarchical cluster algorithms, single node density control strategy can guarantee simultaneous energy depletion of all sensors but will result in lower energy efficiency and high cost because of the massive redundant data transmission and redundant node distribution.

## 3. Preliminaries

Before elaborating on our algorithm, we will introduce the characteristics of the network model used in our study. We consider a WSN consisting of sensors and sink and make the following assumptions.

*3.1. Assumptions on Node and Energy of the Network.* We consider a sensor network with $N$ nodes in a circular area within a radius of $R$ to continuously monitor the sensing area. We denote $i$th sensor by $v_i$ and the corresponding sensor node set only $V = \{v_1, v_2, \ldots, v_k\}$ where $|V| = N$. The sink is located at the center as shown in Figure 1. Our assumptions about the sensor nodes and the network model are as follows.

(1) Each sensor has the maximum transmission range denoted by $r_{\max}$, and the sink node and all sensors are stationary after disposition.

(2) Nodes can estimate the approximate distance to another node based on the received signal strength.

(3) Nodes can use power control to adjust the transmission power according to the distance to the receiver.

(4) The links between nodes are symmetric, and time division multiple access (TDMA) scheduled data transmission from normal nodes to its cluster head.

(5) The network runs a periodic data gathering application. The sensor generates traffic at an average rate of $\lambda$ bits/second and sends it to its CH, which in turn delivers it to the sink using multi-hop communication.

A typical sensor node includes three basic units: sensing unit, processing unit, and transceivers. For our energy model of multi-hop forwarding scheme, we assume a free space propagation channel model [23]. We ignore the power consumption of node for sensing because it is constant at any time and cannot be reduced with whatever means. Thus, the energy model of a sensor involves the power for data aggregating, data receiving, and data transmission according to this radio hardware energy dissipation in both the free space ($d^2$ Power loss) and the multipath fading ($d^4$ Power loss) channel models. If the distance is shorter than the threshold $d_0$, the free space (FS) model is used; otherwise, the multi-path (MP) model is used. If only $E_{\text{Tx}}$ denotes the energy consumption for transmitting data, $E_{\text{Rx}}$ the energy consumption for receiving data, and $E_{\text{Ag}}$ the energy consumption of aggregating data, the energy for transmitting,

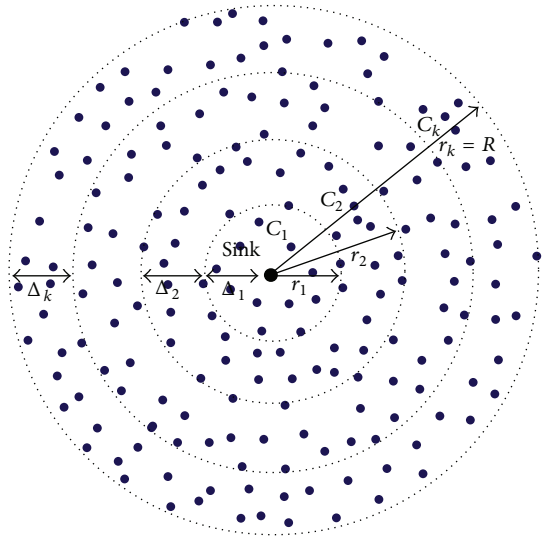



Figure 1: Sensor network area consisting of coronas.

receiving, and aggregating $l$ bits data over distance $d$ can be calculated as follows:

$$E_{\mathrm{Tx}} = \begin{cases} E_{\mathrm{Tx}} = \left(e_1 + e_2^* d^2\right) l, & d < d_0, \\ E_{\mathrm{Tx}} = \left(e_1 + e_4^* d^4\right) l, & d \geq d_0, \end{cases}$$
$$E_{\mathrm{Rx}} = e_1 l, \tag{1}$$
$$E_{\mathrm{Ag}} = e_3 l.$$

Here, the electronics energy $e_1$ depends on factors such as the digital coding and modulation. The amplifier energies $e_2$ and $e_4$ are the energies required for power amplification in the two radio models, respectively. $e_3$ is the energy consumption for dealing with the aggregation of unit sensor data. Some typical values for the parameters above in current sensor technologies are as follows: $e_1 = 50\,\mathrm{nJ/bit}$, $e_2 = 10\,\mathrm{pJ/bit/m^2}$, $e_3 = 5\,\mathrm{nJ/bit/signal}$, and $e_4 = 0.0013\,\mathrm{pJ/bit/m^4}$.

*3.2. The Sensor Network Model Based on Corona.* We adopt the corona model in this paper. The area around sink is divided into coronas of a dynamic width as illustrated in Figure 1. We assume that the sink has a steady energy supply and a powerful radio that can cover the whole monitoring area. $K$ concentric circles of radius $0 < r_1 < r_2 < \cdots < r_k = R$ are centered at the sink with the corresponding node distribution density $\rho_1 > \rho_2 > \cdots > \rho_k = \rho_{\min}$. The width of corona $C_i$ is $\Delta_i$ which is delimited by the circles of radius $r_{i-1}$ and $r_i$ (equivalent to $r_i - r_{i-1}$). Sensors can adjust their transmission ranges to save energy. All the sensors are deployed in such a way that reliable communication between sensors in adjacent coronas can be guaranteed, and the width of each corona does not exceed the sensor's maximum transmission range $r_{\max}$.

In many other unequal cluster algorithms like EEUC, the cluster's size far from the sink is always bigger than the cluster's size near the sink, so it is assumed that a sensor in $C_i$ corona uses a transmission radius of $\Delta_i$ to reach a sensor in $C_{i-1}$ corona. This assumption can ensure effective communication between adjacent cluster heads with the cluster sizes being strictly reduced. In this paper, we do not define such a prerequisite for transmission radius of cluster head. Furthermore, we find that the cluster size will not always decrease as the distance from cluster to the sink is shortened under the premise that ECR of each corona is minimized. To ensure effective data transmission in hop-by-hop communication, for instance, in $C_i$ corona in our model, we select the maximum width between corona $C_i$ and corona $C_{i-1}$ as the sensor transmission radius to ensure effective communication with each sensor in corona $C_{i-1}$. Cosider

$$t_{ix} = \max\left\{\Delta_i, \Delta_{i-1}\right\}. \tag{2}$$

We have the following assumptions about the clustering approach in each corona: sensors whose distance to the sink is in $(r_{i-1}, r_i]$ are organized into clusters to cover the corona $C_i$. A sensor located in the corona $C_i$ is assigned to the nearest CH in the same corona. All sensors are organized into clusters, and their data is relayed by the closest CH in the adjacent corona to the sink in multi-hop communication. Such a corona-based model enables us to analyze theoretically the relationship among ECR, the traffic volume relayed by CH, and the distance from cluster to the sink.

## 4. Nonuniform Sensors Distribution Strategy

As mentioned previously, when WSNs are sensing and collecting data, the redundant distribution sensor nodes are in the sleep mode. They turn off most of their components and only run a timer circuitry to listen to the channel. Hence, energy consumption of sensors in the sleep mode can be ignored. The active nodes are distributed uniformly with the minimal node density $\rho_{\min}$ in the network. Based on this, the optimal ECR of each corona is calculated in uniform node distribution network. In the following part of this chapter, we will analyze the energy consumption for communication in cluster and construct the optimal unequal cluster network. Then, we will create a mathematical model for ECR and calculate its minimal value. Finally, we will derive the non-uniform node distribution strategy based on unequal cluster according to the optimal ECR in each corona.

*4.1. ECR Model Based on Intra- and Intercluster Energy Consumption.* In this section, we will analyze the energy consumption in cluster and deduce the mathematical model among the ECR, the cluster size, and the distance from node to sink. We formulate the suboptimal cluster size calculation approach in each corona based on the ECR. Instead of considering the energy consumption balance in every corona directly [3], we optimize the ECR in each corona. Let $N_i$ denote the number of nodes in corona $C_i$ and $E_i$, the energy consumed per unit time by all nodes in it. We can divide the energy consumption in corona $C_i$ into two parts: the intra- and inter-cluster energy consumption. Let $E_{\mathrm{intra}\_i}$ denote the intra-cluster energy consumption for local sensing, information processing, and communicating,

and $E_{\text{inter\_}i}$ the inter-cluster energy consumption used for relaying the traffic data from the outside coronas. According to the assumptions and previous network model, $E_i$ can be calculated by

$$E_i = E_{\text{intra\_}i} + E_{\text{inter\_}i}. \tag{3}$$

Let $p_i$ denote the probability that the nodes become CHs in corona $C_i$. Like many distance-based cluster formation algorithms, we assume that each CH is located at the center of its cluster. Then, we can calculate $E_{\text{intra\_}i}$ as below:

$$E_{\text{intra\_}i} = N_i p_i E_{\text{Cluster\_}i}. \tag{4}$$

In this equation, $E_{\text{Cluster\_}i}$ is the energy consumption per second for one cluster in corona $C_i$ and can be calculated as follows [27]:

$$E_{\text{Cluster\_}i} = E_{\text{CH\_}i} + E_{\text{nonCH\_}i}. \tag{5}$$

In the equation, $E_{\text{CH\_}i}$ is the energy used by CH to receive data from the member nodes, aggregate the data, and transmit the aggregate data to CHs in the next corona. $E_{\text{nonCH\_}i}$ is the energy used by each noncluster head node to transmit its data to the cluster head during a unit time. $E_{\text{CH\_}i}$ and $E_{\text{nonCH\_}i}$ can be given by

$$\begin{aligned} E_{\text{CH\_}i} &= \left(\frac{1}{p_i} - 1\right) e_1 \lambda + \frac{e_3 \lambda}{p_i} + \frac{\alpha\lambda\left(\Delta_i^2 e_2 + e_1\right)}{p_i}, \\ E_{\text{nonCH\_}i} &= \left(e_1 + \Delta_i^2\ e_2\right)\lambda, \end{aligned} \tag{6}$$

where $\alpha$ denotes the aggregation coefficient for all CHs. Substituting (5), (6) into (4), the expected power consumption for intracluster of all CHs in the corona $C_i$ can be given by

$$E_{\text{intra\_}i} = N_i \lambda \left\{\left[2\left(1 - p_i\right) + \alpha\right] e_1 + \left(1 - p_i + \alpha\right)\Delta_i^2\ e_2 + e_3\right\}. \tag{7}$$

As the sensing data relay from the outside coronas is done hop by hop, the total traffic load carried by the CHs in the corona $C_i$ is equal to the total traffic volume originating from all clusters in corona $i+1$ to $K$. So $E_{\text{inter\_}i}$ for all nodes in the corona $C_i$ can be given approximately by:

$$E_{\text{inter\_}i} = \rho_{\min}\pi\left(R^2 - r_i^2\right)\left(2e_1 + e_2\Delta_1^2\right)\alpha\lambda. \tag{8}$$

Let $E_{\text{node\_}i}$ denote the ECR in the corona $C_i$; it can be given by

$$E_{\text{node\_}i} = \frac{E_i}{N_i}. \tag{9}$$

Accordingly, the number of active nodes in corona $C_i$ can be expressed by $\rho_{\min}\pi(2r_i\Delta_i - \Delta_i^2)$; substituting (7), (8), and (3) in (9), $E_{\text{node\_}i}$ is more specifically given by

$$\begin{aligned} E_{\text{node\_}i} &= \lambda\left\{\left[2\left(1 - p_i\right) + \alpha\right] e_1 + \left(1 - p_i + \alpha\right)\Delta_i^2\ e_2 + e_3\right\} \\ &\quad + \alpha\lambda\ \frac{\left(R^2 - r_i^2\right)\left(2e_1 + e_2\Delta_i^2\right)}{2r_i\ \Delta_i - \Delta_i^2}. \end{aligned} \tag{10}$$

In [27], the author has proved that approximate $2\pi r_i/\Delta_i$ CHs are needed to cover corona $C_i$ at least in corona model. Then we can get the CH election probability $p_i$ as follows:

$$p_i = \frac{2r_i}{\rho_{\min}\left(2r_i\Delta_i^2 - \Delta_i^3\right)}. \tag{11}$$

For simplicity, we set $\lambda = 1$ bit/second which will not affect the results of our analysis. Substituting (11) into (10), we can simplify $E_{\text{node\_}i}$ as follows:

$$\begin{aligned} E_{\text{node\_}i} &= \Bigg[\left(2e_1 + \alpha e_1 + e_3\right) + \left(e_2 + \alpha e_2\right)\Delta_i^2 \\ &\quad - \frac{4e_1 r_i}{\rho_{\min}\left(2r_i\Delta_i^2 - \Delta_i^3\right)} - \frac{2e_2\ r_i}{\rho_{\min}\left(2r_i - \Delta_i\right)}\Bigg] \\ &\quad + \alpha\ \frac{\left(R^2 - r_i^2\right)\left(2e_1 + e_2\Delta_i^2\right)}{2r_i\Delta_i - \Delta_i^2}. \end{aligned} \tag{12}$$

In order to further study the relationship of ECR between intra- and inter-cluster communications, we will split $E_{\text{node\_}i}$ into two parts. Let $E_{\text{ave\_intra\_}i}$ denote the node's average energy consumption rate in intra-cluster communication and $E_{\text{ave\_inter\_}i}$ the node's average energy consumption rate in inter-cluster communication. Without loss of generality, we assign a reasonable value to $\alpha$, for example $\alpha = 0.1$. Obviously, we have $\Delta_i \le r_i$ ($\Delta_i = r_i$ if and only if $i = 1$). According to (5), (10), and (11), $E_{\text{ave\_intra\_}i}$ and $E_{\text{ave\_inter\_}i}$ can be calculated as folows:

$$\begin{aligned} E_{\text{ave\_intra\_}i} &= 2e_1 + \alpha e_1 + e_3 + \left(e_2 + \alpha e_2\right)\Delta_i^2 \\ &\quad - \frac{4e_1 r_i}{\rho_{\min}\left(2r_i\Delta_i^2 - \Delta_i^3\right)} - \frac{2e_2 r_i}{\rho_{\min}\left(2r_i - \Delta_i\right)} \end{aligned} \tag{13}$$

$$E_{\text{ave\_inter\_}i} = \alpha\frac{\left(R^2 - r_i^2\right)\left(2e_1 + e_2\Delta_i^2\right)}{2r_i\Delta_i - \Delta_i^2}. \tag{14}$$

*4.2. The ECR Model Analysis.* It can be seen from (10) that the data aggregation coefficient $\alpha$ and the cluster head probability $p_i$ always affect the node's average energy consumption rate. When $\alpha$ varies from 1 to 0 and cluster heads compress data more efficiently, the node's average energy consumption gets smaller and the energy hole of network can be alleviated as proved by many researchers. Equivalently, when the probability of being a cluster head $p_i$ increases, the energy used for intra-cluster communication will decrease due to shorter communication distance between nodes and their cluster heads. The data aggregation and unequal hierarchical cluster protocol can effectively reduce the ECR in WSNs, which can lead to higher energy efficiency and longer network lifetime. This is why non-uniform sensor distribution strategy based on unequal cluster is better than general non-uniform sensor distribution method.

Next, we will analyze the relationship among $E_{\text{node\_}i}$, $\Delta_i$, and $r_i$. To enhance the energy efficiency and prolong the network lifetime, we need to define the optimal $\Delta_i$ to minimize ECR in corona $C_i$. We assume that a WSN covers

a circular sensing region with $R = 200$ meters and $P_{\min} = 0.00318$, which implies that 400 sensors are uniformly distributed in this monitoring area. By studying the nodes' energy consumption rate with different cluster radius $\Delta_i$ based on fixed $r_i$ through simulations, we find that ECR can be minimized for each constant distance from cluster to the sink when an appropriate value is assigned to cluster size $\Delta_i$. Our analytical findings can be found in the following simulation tests.

Figures 2, 3, 4, and 5 demonstrate the changes for ECR, average energy consumption rate for intra-cluster data processing and inter-cluster communication versus cluster size $\Delta_i$ when $r_i$ is, respectively, defined as 35 meters, 60 meters, 90 meters, and 150 meters. As shown in Figures 2 and 3, when the upper bound of corona $C_i$ is closer to sink ($r_i$ is small), $E_{\text{node\_}i}$ will be dominated by $E_{\text{ave\_inter\_}i}$ as $\Delta_i$ increases. The reason is that the nodes' energy is mainly used for relaying the traffic data coming from the outside coronas, and inter-cluster communication is the key factor to cause the energy hole in this scenario. On the contrary, $E_{\text{node\_}i}$ expands rapidly as $\Delta_i$ decreases. This result proves that it is not always the best option to reduce the sizes of clusters that are closer to the sink, even though it has been adopted in many existing unequal cluster protocols by many researchers. The cluster size cannot be defined as too small when the cluster is close to sink.

As shown in Figures 4 and 5, when the upper bound of corona $C_i$ is away from sink ($r_i$ is big), we can see that $E_{\text{ave\_intra\_}i}$ will determine the value of $E_{\text{node\_}i}$ as $\Delta_i$ increases. This phenomenon can be explained as follows: when corona $C_i$ is away from sink, the data traffic that needs to be relayed will decrease, which means that the energy consumption for inter-cluster communication is reduced. The energy used for intra-cluster communication becomes the major factor to determine the value of $E_{\text{node\_}i}$. But it is worth mentioning that the cluster size increase in these coronas will not always reduce the value of ECR. This result shows that it is not always the best way to save energy by simply increasing the size of clusters away from the sink. It is of no help to alleviate the energy consumption rate in inner coronas but increases their own energy consumption. The author has proved [28] that the network lifetime is determined by nodes in the innermost corona in multi-hop communication no matter how the clusters in outer corona in a uniform network are adjusted and organized. Furthermore, we can see that there is an optimal cluster size $\Delta_i$ for each $r_i$ to minimize ECR as indicated in Figures 2–5.

Figure 6 illustrates the change of $E_{\text{node\_}i}$ versus the distance $r_i$ from sink to coronas and unequal cluster size $\Delta_i$ in the whole sensing area. In practical applications, we have the limited condition $\Delta_i \le r_i$. From the simulation result, we can see that the curved surface of $E_{\text{node\_}i}$ is given with an irregular concave surface. There is an extreme small point for ECR with a group of corresponding $r_i$ and $\Delta_i$. At the same time, with $r_i$ increasing from 0 to 200 meters, the unequal cluster's size $\Delta_i$ will increase firstly and then decrease when $E_{\text{node\_}i}$ is minimized. It proves that the cluster will not always become smaller in size as its distance to the sink decreases, under the precondition that ECR in unequal clusters is minimized.

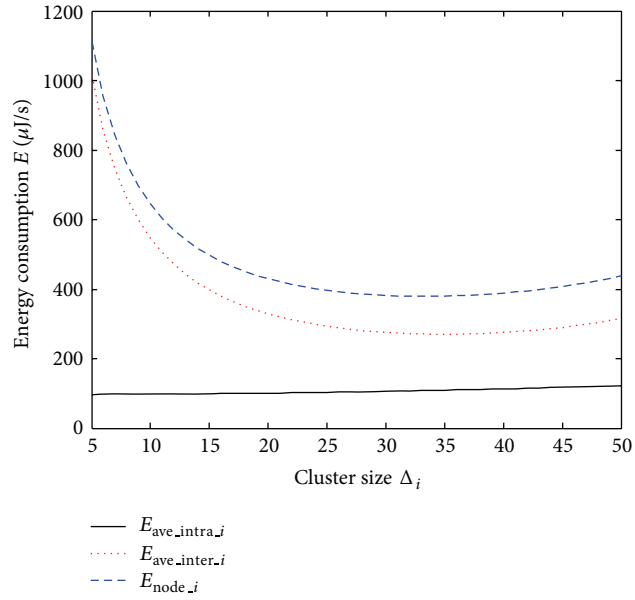



Figure 2: Node energy consumption rate versus cluster radius $\Delta_i$ ($r_i = 35$ meters).

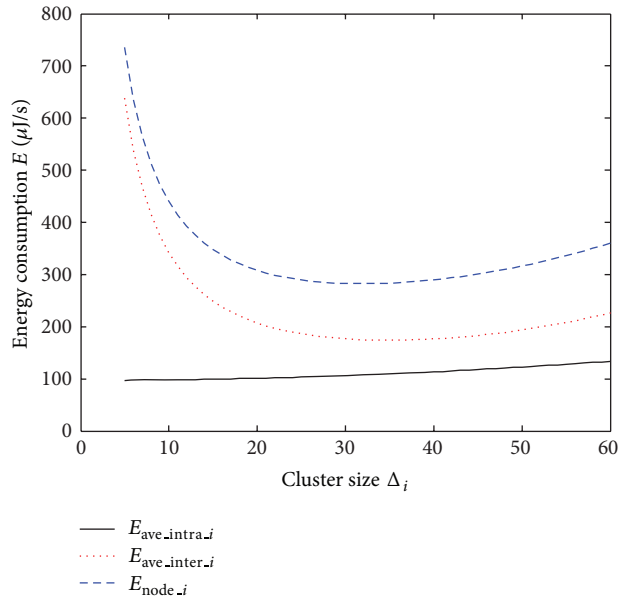



Figure 3: Node energy consumption rate versus cluster radius $\Delta_i$ ($r_i = 60$ meters).

Based on such results, we can calculate a group of optimal $\Delta_i$ and divide the monitoring area into several coronas in which ECR is minimized, under the premise that the network parameters are fixed. In other words, the unequal cluster can also been constructed. With the minimal ECR in each corona, the network can obtain longer lifetime and high energy efficiency, with the energy hole problem being further alleviated.

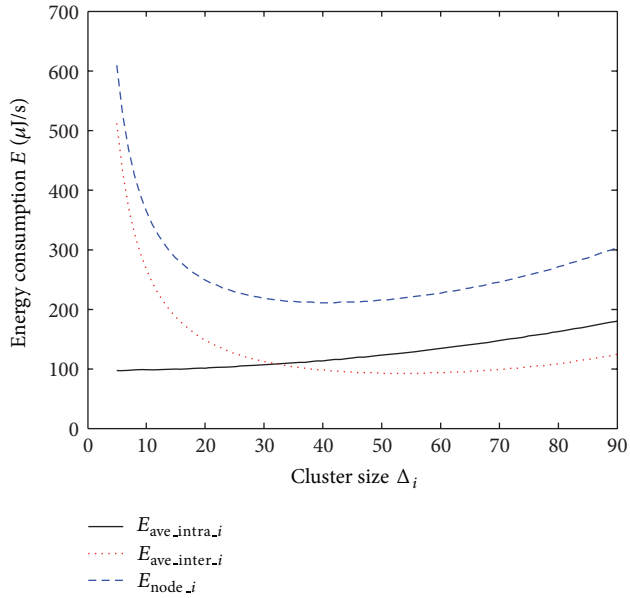



Figure 4: Node energy consumption rate versus cluster radius $\Delta_i$ ($r_i$ = 90 meters).

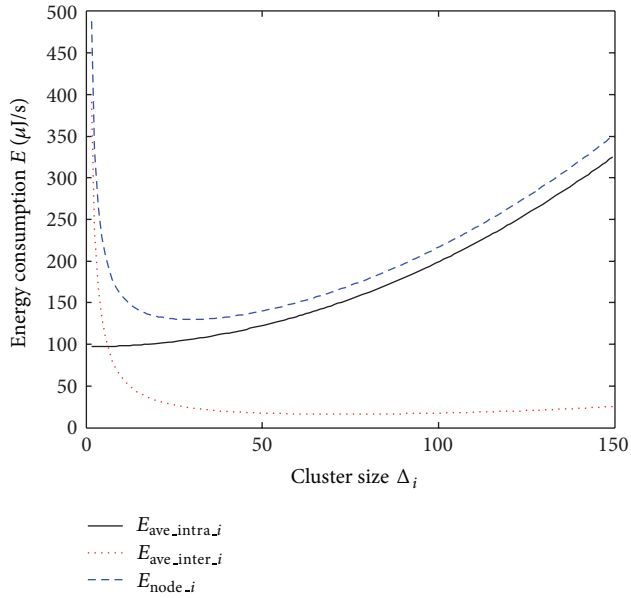



Figure 5: Node energy consumption rate versus cluster radius $\Delta_i$ ($r_i$ = 150 meters).

*4.3. The Computation for Sub-optimal Unequal Cluster and Minimal ECR.* From the previous analysis, we can calculate a group of optimal $\Delta_i$ to achieve minimal ECR. In the following section, we will give the optimal algorithm to determine the minimal ECR and the optimal $\Delta_i$ of each corona. Within each corona, we can construct the unequal hierarchical cluster. Then, we can calculate the number of sensors that need to be distributed in every corona, by following the principle that all nodes will use up their energy when network is no longer able to collect and transmit information.

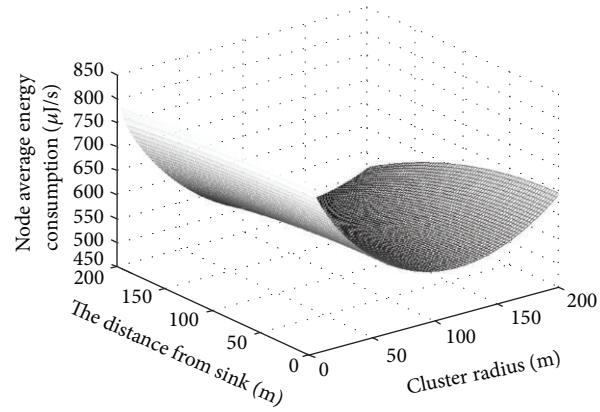



Figure 6: Node energy consumption rate versus the distance form sink $r_i$ and cluster size $\Delta_i$.

We notice that the cluster size is equal to the corona width in corona $C_1$; hence $\Delta_1 = r_1$. Then, according to (11), the ECR in corona $C_1$ can be calculated with the following formula:

$$E_{\text{node\_1}} = 2e_1 + \alpha e_1 + e_3 - \frac{2e_2}{\rho_{\min}} + \left(e_2 + \alpha e_2\right)\Delta_1^2 - \frac{4e_1}{\rho_{\min}\Delta_1^2} + \alpha\frac{\left(R^2 - \Delta_1^2\right)\left(2e_1 + e_2\Delta_1^2\right)}{\Delta_1^2}. \tag{15}$$

In (15), $\Delta_1$ is the only unknown parameter for the function $E_{\text{node\_1}}$ if the monitoring area radius $R$ is predetermined. Figure 7 illustrates the change of $E_{\text{node\_1}}$ with cluster size $\Delta_1$ when $R$ = 200 m. As can be seen from this curve, there is an optimal cluster size to minimize the per node energy consumption rate. If $R$ = 200 meters, we can find that ECR in corona $C_1$ gets the minimal value when $\Delta_{\text{opt\_1}}$ = 32.16 m, according to the numeric computation in (15). Through further analysis, we can find that once the radius of monitoring area is defined, the corresponding optimal cluster size $\Delta_{\text{opt\_1}}$ can be obtained. With this optimal $\Delta_{\text{opt\_1}}$, the ECR in innermost corona is minimized, and the network lifetime can be maximized.

After $\Delta_1$ is obtained, we begin to calculate other corona width (cluster size) through the following iterative algorithm. We have the relation

$$r_i = r_{i-1} + \Delta_i. \tag{16}$$

Submitting (16) into (12), we have

$$E_{\text{node\_}i} = \left(2e_1 + \alpha e_1 + e_3\right) + \left(e_2 + \alpha e_2\right)\Delta_i^2 - \frac{4e_1\left(r_{i-1} + \Delta_i\right)}{\rho_{\min}\left(2r_{i-1}\Delta_i^2 + \Delta_i^3\right)} - \frac{2e_2\left(r_{i-1} + \Delta_i\right)}{\rho_{\min}\left(2r_{i-1} + \Delta_i\right)} + \alpha\frac{\left[R^2 - \left(r_{i-1} + \Delta_i\right)^2\right]\left(2e_1 + e_2\Delta_i^2\right)}{2r_{i-1}\Delta_i + \Delta_i^2}. \tag{17}$$

Differentiating (17) for $\Delta_i$, $E_{\text{node\_}i}$ is minimized by the value of $\Delta_i$ that is a solution of

$$A\Delta_i^6 + B\Delta_i^5 + C\Delta_i^4 + D\Delta_i^3 + E\Delta_i^2 + F\Delta_i + G = 0. \tag{18}$$

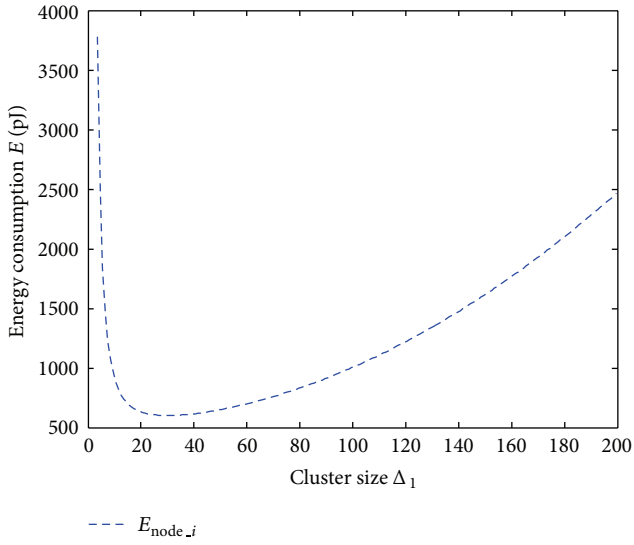



FIGURE 7: Node energy consumption rate in corona $C_1$ versus cluster size $\Delta_i$.

In (18)

$$\begin{aligned} A &= -6\alpha\rho_{\min}e_2, \\ B &= \alpha\rho_{\min}\left(-10e_2r_{i-1} - 10e_3\right), \\ D &= \alpha\rho_{\min}\left(4R^2e_2 - 4e_2r_{i-1}^2 - 16e_3r_{i-1} - 8e_1\right) - 4e_2, \\ C &= \alpha\rho_{\min}\left(6R^2e_2r_{i-1} - 6e_2r_{i-1}^3 - 12e_1r_{i-1} - 12e_1\right) \\ &\quad - 6e_2r_{i-1}, \\ E &= \alpha\rho_{\min}\left(2R^2 - 2r_{i-1}^2 - 16e_1r_{i-1}\right) - 16e_1, \\ F &= \alpha\rho_{\min}\left(2R^2r_{i-1} - 2r_{i-1}^3\right) - 36e_1r_{i-1}, \\ G &= -16e_1r_{i-1}^2. \end{aligned} \tag{19}$$

Obviously, if $r_{i-1}$ is a known number, $\Delta_i$ will be the only unknown parameter in (19). According to the Galois theory [28], the root of (18) cannot be obtained by elementary algebra. However, we can use numeric solutions to calculate the roots of general polynomial equation. Define

$$f(x) = Ax^6 + Bx^5 + Cx^4 + Dx^3 + Ex^2 + Fx + G. \tag{20}$$

Then,

$$\begin{aligned} f'(x) &= 6Ax^5 + 5Bx^4 + 4Cx^3 + 3Dx^2 + Ex + F, \\ f''(x) &= 30Ax^4 + 20Bx^3 + 12Cx^2 + 6Dx + E. \end{aligned} \tag{21}$$

It is obvious that $f(x)$, $f'(x)$, and $f''(x)$ are continuous functions in the whole real domain. Assume that there is a number $q \in [0,1]$, where $f(q) = 0$. According to Newton-Raphson theorem, if $f'(q) \neq 0$, there is a $\delta > 0$ and for any initial approximation $q_0$, $q_0 \in [q-\delta, q+\delta]$, the sequence $\{q_k\}_{k=0}^{\infty}$ defined by the iteration

$$q_{k+1} = q_k - \frac{f(q_k)}{f'(q_k)} \quad (k = 0, 1, 2, \ldots) \tag{22}$$

will converge to $q$. Thus, given an initial value for $x_0$, for example, $x_0 = 0.01$, we start the iteration based on $x_n = x_{n-1} - f(x_{n-1})/f'(x_{n-1})$ to update $x_n$ until the difference between $x_n$ and $x_{n-1}$ is smaller than the certain reservation threshold. Consequently, we can get the approximate optimal cluster size $\Delta_{\text{opt\_}i} = x_n$. Because all sensors have the maximum transmission range $r_{\max}$ and the restriction of $\Delta_i \leq r_{\max}$, we have $\Delta_i = r_{\max}$ when $\Delta_{\text{opt\_}i} \geq r_{\max}$. Substituting $\Delta_{\text{opt\_}i}$ and $r_i$ into (11), we can get the minimal value of $E_{\text{node\_}i}$ in corona $C_i$.

The values of $\Delta_{\text{opt\_}i}$ and $r_1$ in corona $C_1$ that we have obtained are the initial conditions for our iterative algorithm. For the relation $r_2 = r_1 + \Delta_{\text{opt\_}2}$, if we put it into (17), we can obtain $\Delta_{\text{opt\_}2}$ through the previous iterative algorithm. Repeat the same procedure and submit $r_3 = r_2 + \Delta_{\text{opt\_}2}$ into (17), we can obtain $\Delta_{\text{opt\_}3}$. Repeating this iterative process (16)–(22) constantly, we can calculate all the optimal cluster size in every corona $\Delta_{\text{opt\_}1}, \Delta_{\text{opt\_}2}, \ldots, \Delta_{\text{opt\_}k}$ and $r_1, r_2, \ldots, r_k$.

From (13), in the outermost corona $C_k$, $E_{\text{ave\_inter\_}k} = 0$ and $E_{\text{ave\_intra\_}k}$ is an increasing function of $\Delta_i$ when $\Delta_i < r_i$. When ECR is minimized, cluster size will be close to zero, which means that cluster is degraded into one node. In order to guarantee the energy efficiency of the cluster, we require that iterative process be terminated when $R - r_{i-1} \leq r_1$ or $R - r_{i-1} \leq R\sqrt[4]{(R^2e_2 - 2e_1)/Ne_2R^2}$ and define the cluster size of outermost corona $C_k$ as $R - r_{i-1}$. $R\sqrt[4]{(R^2e_2 - 2e_1)/Ne_2R^2}$ is the optimized cluster size to ensure the coverage and connectivity for circle model of WSNs in hierarchical cluster structure [29]. Since the optimal size of the outermost cluster cannot be determined, our algorithm is only a suboptimal solution. Fortunately, the average energy consumption rate in the outermost corona is always the smallest in the entire network, thus exerting impact on our algorithm for eradicating the energy hole problem.

*4.4. The Nonuniform Nodes Distribution Strategy.* After we have obtained $\Delta_{\text{opt\_}1}, \Delta_{\text{opt\_}2}, \ldots, \Delta_{\text{opt\_}k}$ and $r_1, r_2, \ldots, r_k$, the per node energy consuming rate (ECR) in each corona $E_{\text{node\_}1}, E_{\text{node\_}2}, \ldots, E_{\text{node\_}k}$ can be obtained as well. Since the sink node is not limited to energy, the iterative algorithm can run on it. In the following part, we will firstly introduce node deployment redundancy rate coefficient $\psi_i$, the ratio of the nodes density in corona $C_i$, and the nodes density in the outmost corona $C_k$. We assume that there is no sleeping node in corona $C_k$ and the node density is $\rho_{\min}$ Consider

$$\psi_i = \frac{\rho_i}{\rho_k} = \frac{\rho_i}{\rho_{\min}}. \tag{23}$$

To put it simple, we assume that each sensor has the power of $\tau$ joule and use $S_i$ to denote the area of corona $C_i$. Then, there are $\rho_{\min}S_i$ active nodes to participate in the network operation simultaneously in this corona. When all

the distributed nodes use up all their energy, their survival time can be calculated as follows:

$$T_i = \frac{\rho_i S_i \tau}{\rho_{\min} S_i E_{\text{node\_}i}} = \frac{\psi_i \tau}{E_{\text{node\_}i}}. \quad (24)$$

Ideally, if there is no residual energy for all nodes when network loses its function, the network lifetime and the energy efficiency are maximized. That is,

$$\frac{\psi_1 \tau}{E_{\text{node\_1}}} = \frac{\psi_2 \tau}{E_{\text{node\_2}}} = \cdots = \frac{\tau}{E_{\text{node}_k}}. \quad (25)$$

By (25), we can obtain

$$\psi_i = \frac{E_{\text{node}_i}}{E_{\text{node}_k}} \quad (1 \le i < k). \quad (26)$$

According to (23) and (23), the optimal sensor distribution density and the number of sensors in $C_i$ corona can be given by

$$\rho_i = \frac{E_{\text{node\_}i}}{E_{\text{node\_}k}} \rho_{\min},$$
$$N_i = \rho_{\min} \pi \Delta_i (2r_i - \Delta_i) \frac{E_{\text{node\_}i}}{E_{\text{node\_}k}}. \quad (27)$$

As the minimal values of $E_{\text{node\_1}}, E_{\text{node\_2}}, \ldots, E_{\text{node}_i}, \ldots, E_{\text{node\_}k}$ have been determined as above, we can obtain the sensor distribution density $\rho_1, \rho_2, \ldots, \rho_{k-1}$ and the sensors distribution number $N_1, N_2, \ldots, N_k$ in each corona. $\rho_{\min}$ is the preliminary parameter of WSNs. As a result, energy hole problem of the network is resolved by using this non-uniform sensor distribution strategy based on unequal cluster.

*4.5. The Data Transmission Mechanism.* In order to get higher data transmission efficiency, we design an energy-balancing layered data transmission mechanism based on the previous non-uniform sensor distribution algorithm. Firstly, the sink constructs the unequal coronas, and sensor node can be deployed according to the results derived in the previous section. Sensor nodes select CHs and join the adjacent CH based on its distance to sink. Upon completion of the suboptimal unequal clusters, CHs are able to deliver their data to the sink. Each CH firstly aggregates the data received from its cluster members and then transfers them to the sink node via a multi-hop path through other intermediate CHs. The organization of intra-cluster data transmission is similar to LEACH, so we will not elaborate it again in this section. The pseudocode of the inter-cluster data relay algorithm is presented in Algorithm 1. In order to achieve balanced energy depletion among the CHs, they firstly exchange their energy and position information *CHMessage* to maintain a real-time table *CoronaList* about the neighbor cluster head. By this algorithm, CH can select one relay CH with maximum energy resource. At the same time, it has to exchange the residual energy message with all candidate relay CHs in lines 4–12 of Algorithm 1. After selecting the relay CH with the maximum residual energy, the CH can forward its own data and the data coming from its upper corona. The process of selecting relay CH and forwarding data will be repeated until the data arrive at sink node in lines 13–17 of Algorithm 1. If there is more than one candidate with the same maximum residual energy, choose one of them randomly. When SystemMessage is received, for instance, instructions for completing data transfer or reelecting cluster head, and another interrupt instruction is trigged by sink node, the network will terminate the data transmission process.

On the other hand, as indicated in Algorithm 1, since the non-uniform sensor distribution strategy based on unequal cluster can be calculated by the sink node, the complexity of network cluster protocol is determined by the clustering algorithm and inter-cluster data transmission mechanism. The non-uniform sensor distribution strategy based on unequal cluster does not increase the complexity of network.

## 5. Simulation Results

In this section, both the numerical results and the performance results of our strategy will be presented via simulation by MATLAB. At the beginning, we will calculate the optimal corona width and build the unequal hierarchical cluster network based on network parameters. Then we calculate the node distribution density in each corona. As the node distribution algorithm is based on unequal cluster, the performance of unequal clustering will directly influence the energy efficiency of network. Therefore, we firstly examine the performance of our unequal cluster algorithm, by comparing with two typical cluster protocols EEUC and LEACH. Then, we verify the effectiveness of our algorithm in eliminating the energy hole problem. At the same time, we prove that our distribution strategy has better energy efficiency and network lifetime with less sense nodes compared with the single node distribution strategy [26].

In order to conduct the experiments, proper parameters for both the sensor nodes and the network should be defined. We assign $\rho_{\min}$ with the initial value 0.00318, which is the minimum node density to ensure the effective coverage and data collection in the monitoring area. Once sensor nodes have been settled, the active nodes in each corona are organized into clusters and the other redundant nodes are kept in sleep-listening model. Then, each ordinary node forwards certain bits of data to its cluster head, which in turn aggregates and forwards the received data to sink by multi-hop communication. When some nodes die prematurely, nodes in sleep model will wake up and join the nearest cluster to fill in the vacancy. We adopt the STEM [30] sleep-listening mechanism in this paper. Every sensor node will keep a table of neighboring nodes in its competing range. Because the cluster head node always dies earlier than general nodes, it can select the nearest node to wake up according to STEM mechanism. The node's energy consumption for wakening, listening, and detection in sleep model is much lower than that in active model, so for simplicity, we do not consider it. The simulation parameters for our proposed mechanism are defined in Table 1.

```
(1) CalculateCHMessage(ID, Layer_ID, ResEnergy)
(2) Exchange(CHMessage) in adjacent corona
(3) Rec_Message(CHMessage) from Neighbor_CH
(4) if (Neighbor_CH.Layer_ID = CHMessage.Layer_ID + 1)
(5)     add node Neighbor_CH to CH.Up_CoronaList
(6)     Return (CHMessage(ID, Layer_ID, ResEnergy)) to Neighbor_CH
(7) end if
(8) if (Neighbor_CH.Layer_ID = CHMessage.Layer_ID – 1)
(9)     add node Neighbor_CH to CH.Down_CoronaList
(10)    Request (CHMessage(ID, Layer_ID, ResEnergy)) from Neighbor_CH
(11)    Update (CH.Down_CoronaList)
(12) end if
(13) while (Data_arrival = TRUE)
(14)    { if (Rec_Message(Neighbor_CH.Data) from CH.Up_CoronaList)
(15)        {CH.Down = MaxResEnergy(CH.Down_CoronaList)
(16)        Transfer_Data(CH.Down, Neighbor_CH.Data)}
(17)    else{ if (SystemMessage = TRUE)
(18)        EXIT}
(19)    }
```

Algorithm 1: The data transmission mechanism.

Table 1: Parameters and characteristics of the network.

| Parameter | Value |
|---|---|
| Network size (circle) | $R = 200$ meters |
| Sink location | (0, 0) |
| Data packet size | 256 bytes |
| Initial energy | 0.5 J |
| Initial $\rho_{\min}$ | 0.00318 |
| Aggregation ratio | 0.1 |

*5.1. Calculation for Nonuniform Sensor Distribution Based on Unequal Cluster.* According to our non-uniform sensor distribution strategy, the circle sensing area is partitioned into unequal clusters firstly based on the parameters in Table 1. The size of the innermost corona will be determined firstly, followed by size of outermost corona until the conditions of terminating iteration is triggered. Then, the corresponding non-uniform distribution sensors density with minimal ECR in each corona can be achieved as indicated in Table 2.

As mentioned in the previous section, the energy hole problem may arise in multi-hop wireless sensor networks if the cluster heads close to the data sink are burdened with heavy relay traffic. Different from the general opinions on unequal cluster, for instance, EEUC which argues that clusters closer to sink should be smaller in size, we find that cluster will not always be downsized as the distance from clusters to sink decreases so long as ECR in each corona obtains the minimal value. Downsizing the cluster will result in cluster increase in inner corona. Though it can alleviate relay traffic burden on the CH nodes, ECR in the whole cluster will increase dramatically, which will exacerbate the energy hole problem because of the decrease of cluster members. This result serves as an important guideline for the design of unequal cluster. In the following simulation test, we will prove that our unequal cluster can achieve better performance than EEUC.

*5.2. Performance of Nonuniform Sensor Distribution Strategy Based on Unequal Cluster.* Since the sensor distribution strategy is based on unequal clusters, the performance of the unequal cluster based on our optimal algorithm will exert great impact on network. We will firstly review the characteristics of the unequal cluster by comparing with cluster algorithms such as LEACH and EEUC. Simulation results show that our unequal cluster can achieve better energy consumption balance between cluster heads and minimal ECR in the whole network. Secondly, we will look into our non-uniform sensor distribution strategy in terms of the residual energy of nodes, the number of nodes that need to be distributed, and the network lifetime.

*5.2.1. The Performance of Unequal Cluster.* In our scenarios, we use the same parameters for EEUC mechanism [24] with the node density being $\rho_{\min}$. We also conduct lots of experiments to determine the optimal number of clusters for LEACH. Figure 8 shows the total energy consumed by all cluster heads in three algorithms after thirty rounds of simulations. The energy consumed by cluster heads per round in our unequal cluster and EEUC is much lower than that of LEACH. Due to the need for sending their packets to sink via single hop, the energy consumption of cluster heads is much higher in LEACH. Moreover, the CHs' energy consumption in our unequal cluster network is slightly higher than the CHs' energy consumption in EEUC because there are more cluster heads sharing the tasks of data forwarding in EEUC than in our unequal cluster network. If evaluated only from the perspective of energy consumption on cluster head, EEUC can balance the energy consumption of cluster head better than our approach. However, if the energy consumption of all

Table 2: Optimal nonuniform sensor distribution density and unequal clusters.

| Corona | Cluster size (m) | Per node energy consuming rate (ECR) ($\mu$J/s) | Redundancy rate coefficient $\psi_i$ | Nonuniform sensor distribution density $\rho_i$ | The number of sensors in each corona |
|---|---|---|---|---|---|
| $C_1$ | 32.16 | 518.73 | 4.761 | 0.01514 | 55 |
| $C_2$ | 51.22 | 204.72 | 1.879 | 0.00598 | 115 |
| $C_3$ | 47.95 | 156.81 | 1.439 | 0.00458 | 150 |
| $C_4$ | 38.35 | 131.74 | 1.209 | 0.00385 | 139 |
| $C_5$ | 29.32 | 108.95 | 1 | 0.00318 | 107 |

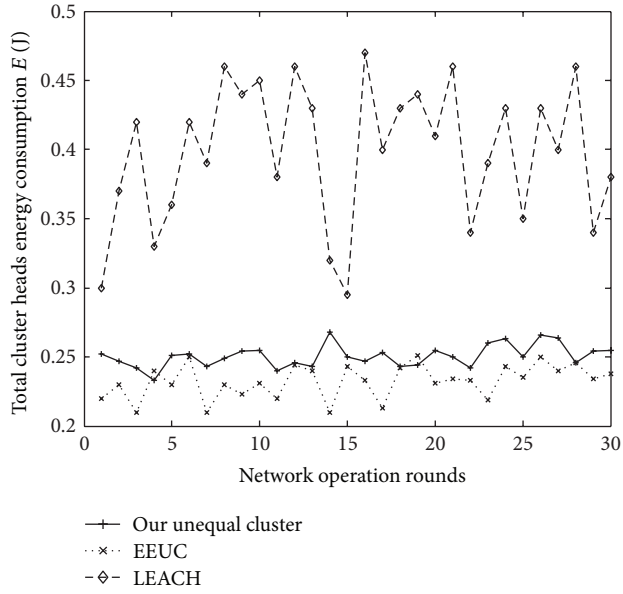



Figure 8: Total energy consumption of cluster heads.

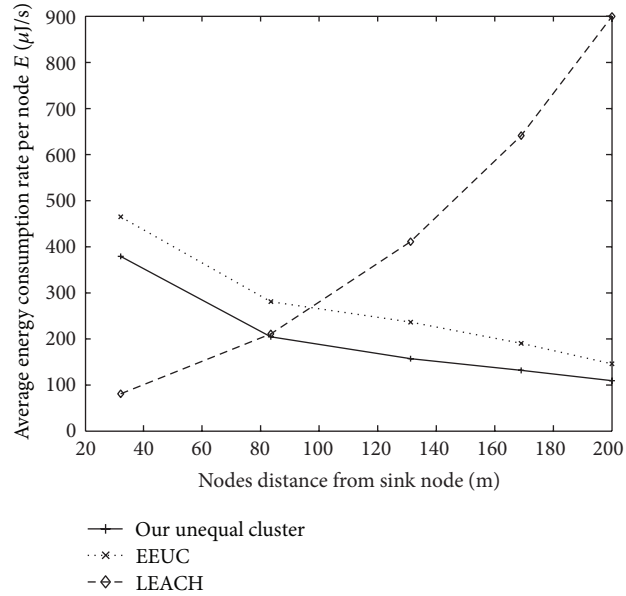



Figure 9: The energy consumption per node.

nodes in the cluster is taken into account, our unequal cluster approach consumes less energy than EEUC.

Generally speaking, in accordance with the basic principle to avoid energy hole problem using unequal cluster algorithm, more nodes need to participate in traffic data relay, in inner corona in particular. However, to ensure longer network lifetime, we should not only consider the energy balance among the cluster head nodes but also consider the energy balance among all cluster members. ECR is an important parameter to measure energy consumption of all cluster members. In EEUC, the balance of energy consumption for cluster heads is the only factor that is taken into consideration, with ECR which has direct bearings on network lifetime being ignored. Figure 9 shows the change of ECR based on the distance from nodes to sink for three algorithms. It can be seen that our algorithm is better than EEUC in terms of nodes' average energy consumption rate. For LEACH, the average energy consumption of network nodes will increase sharply when nodes are away from sink because of the long distance in single-hop communications.

*5.2.2. Energy Efficiency and Lifetime.* In this part, we will analyze the energy efficiency and lifetime of our strategy. Firstly, we will study the residual energy of each node when the network lifetime ends. Then, we compare the sensor density of our strategy with that of other non-uniform sensor distribution methods. Finally, we verify that our non-uniform sensor distribution strategy based on unequal cluster can obtain longer network lifetime and better energy efficiency.

Firstly, we examine residual energy of nodes when the network lifetime ends. We define the network lifetime as the time period until WSNs cannot guarantee the effective coverage and data sensing and collection for the monitoring area. And sensor death in the network means the sensor loses the ability to sense data or send data to its cluster head. Figure 10 shows the cumulate residual energy of nodes in each corona when the network lifetime ends. The fitting fragments of line indicating the total energy of the nodes belong to the five coronas $C_1$–$C_5$ from right to left. It can be seen that when the network lifetime ends, there is nearly no residual energy in the network since only few sensors have residual energy. Through further analysis, we can find that the little residual energy is the result of energy consumption unbalance in the cluster topology maintenance and node sleep-listening mechanism in different coronas. For data relay, unbalanced energy consumption is prevented, or the energy consumption balance of the entire network is achieved indirectly, and the energy hole problem is solved completely. On the other hand,

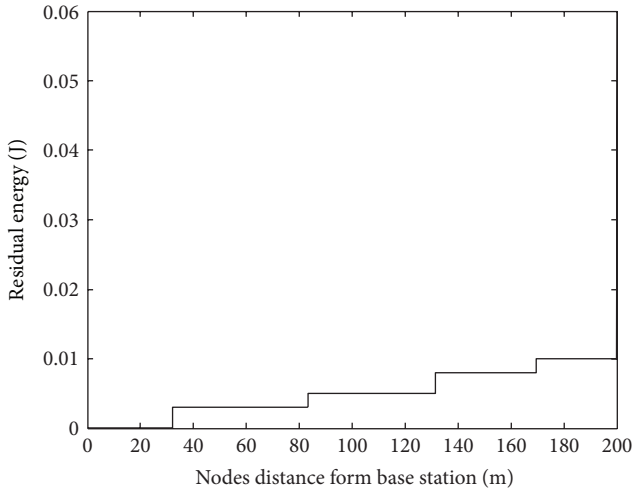



Figure 10: The nodes' residual energy when WSNs lose function.

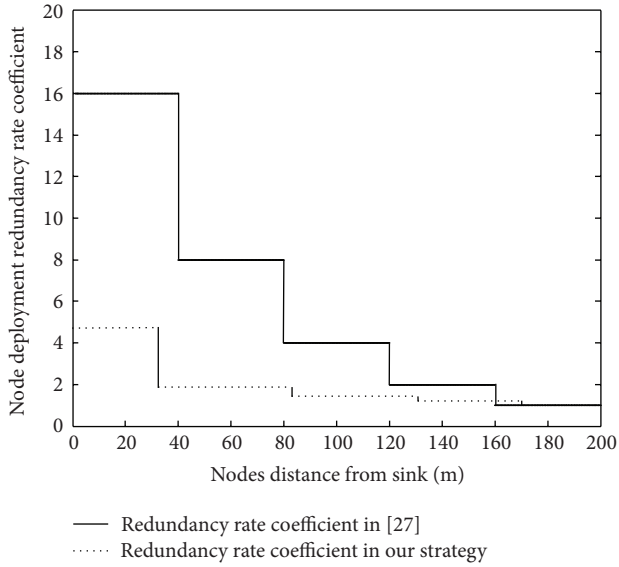



Figure 11: Node distribution redundancy rate coefficient.

it also proves that the nonoptimal ECR in the outermost corona in our strategy exerts no adverse impact on the elimination of energy hole problem.

Secondly, we compare the deployed sensors' density of our strategy with that of other non-uniform sensor distribution methods. In [26], we propose a non-uniform node distribution strategy to achieve the balance of energy depletion. We also point out that if the density of nodes in coronas increases in geometric progression with common ratio, energy depletion balance in the network can be achieved and the energy hole problem could be solved.

In this distribution strategy, sensor nodes are not organized into cluster hierarchy structure, and the monitoring area is divided into coronas with the same width. Each node in the network has several candidate relay nodes in the next inner corona. All experiments are done with the same minimized node density $\rho_{\min}$ in the outmost corona to ensure the effective network coverage and data collection in monitoring area.

In Figure 11, we can see that node distribution density is much lower in our non-uniform sensor distribution strategy. Figure 12 shows the specific number of distribution nodes in each corona for the two non-uniform sensor distribution strategies. It is obvious that few sensors are needed to effectively monitor the sensing area and eliminate the energy hole problem if our strategy is adopted. This is mainly due to the inherent advantages of the unequal cluster, including data aggregating, efficient data routing mechanism, and lower radio communication conflicts. Data aggregating can significantly reduce the frequency of data forwarding, and cluster hierarchy can be used to select efficient routing path more easily.

Finally, we verify the network lifetime and energy efficiency for our non-uniform sensor distribution strategy based on the unequal cluster. In order to estimate the lifetime of the WSNs, the metrics of First Node Dies (FND), Percent of the Nodes Alive (PNA), and Last Node Dies (LND) are always used [11]. FND is useful in sparsely deployed WSNs.

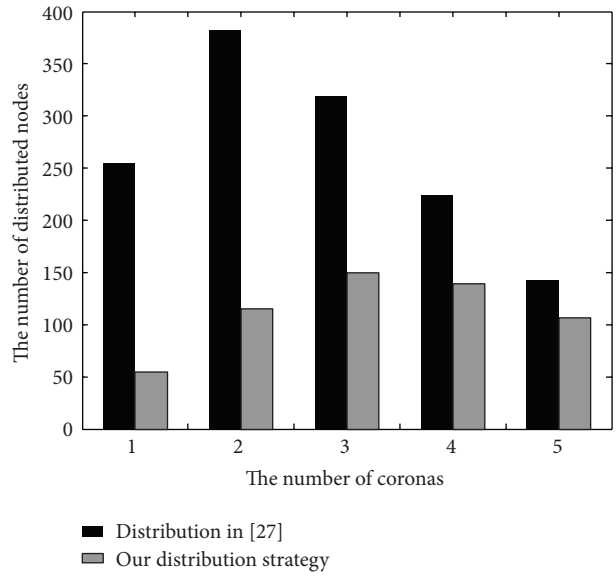



Figure 12: The number of distributed nodes in each corona.

However, PNA is more suitable to measure the network lifetime in densely deployed WSNs. LND does not have much practical value. FND and LND are not suitable for our node distribution strategy. When WSNs lose its ability to collect and transmit data, the network is regarded as dead.

Figure 13 shows the number of sensor nodes that are still alive during simulation test, and Figure 14 illustrates the rounds of data transmission for the four strategies during their lifetime. For EEUC, as more cluster heads participate in the traffic relay in inner corona, it can balance the energy consumption between cluster heads. Moreover, it can achieve better energy efficiency and longer lifetime than LEACH with the FND metric. But the energy hole problem which is inevitable in uniform sensor distribution WSNs, even in

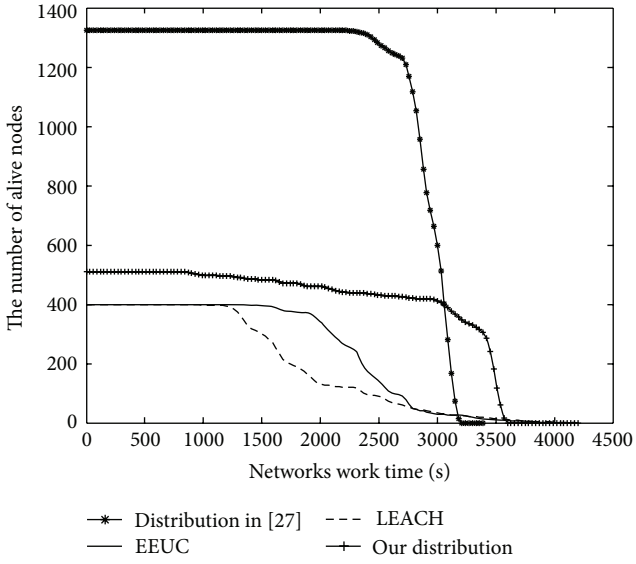



Figure 13: The network lifetime.

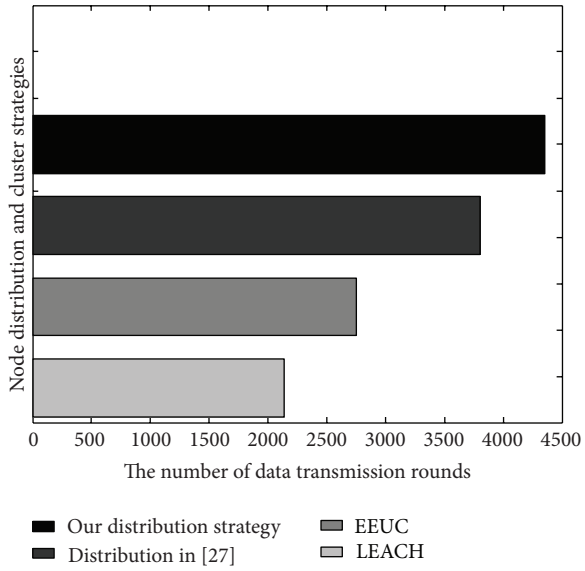



Figure 14: The rounds of data transmission.

cluster hierarchy structure, will cause much residual energy when the network lifetime ends. For single non-uniform node distribution strategy, energy consumption balance of sensors can be ensured, and all the sensor nodes will die nearly at the same time. Therefore, the energy hole problem has been resolved completely. But consider the nodes that need to be settled, we will find that this strategy may result in very low energy efficiency and high cost for WSNs.

In our non-uniform sensor distribution strategy based on unequal cluster, we adjust the node distribution density according to the minimal value of ECR in each corona while taking into account the energy consumption of cluster heads and cluster members. From Figures 13 and 14, we can see that the FND occurs early in our strategy, but the network still works well in network coverage and data collection through sleep-listening mechanism of redundant nodes. When the number of survival nodes in the network approaches 400, all the remaining nodes will use up their residual energy almost at the same time, and then the network lifetime ends. In this way, the energy consumption balance for the entire network is achieved, and the energy hole problem is eliminated completely. From the simulation results, it can be seen that our strategy can ensure longer network lifetime and better energy efficiency compared with the other three strategies.

## 6. Conclusion

In this paper, we have proposed a nonuniform sensor distribution strategy based on unequal cluster for WSNs. The major objective of our algorithm is to resolve the energy hole problem and improve the performance of WSNs based on multi-hop communication.

By focusing on intra- and intercluster energy consumption and using the ECR model, we calculate the minimal value of ECR according to the distance to the base station and deduce the non-uniform node distribution strategy. Our non-uniform sensor distribution mechanism works well in eliminating the energy hole problem by balancing average energy consumption speed of cluster nodes. Through sleep-listening mechanism of redundant nodes and suboptimal unequal cluster protocol, our non-uniform sensor distribution strategy can ensure better energy efficiency and longer network lifetime. Moreover, we have proved that our strategy ensures better performance than LEACH, EEUC, and other single non-uniform sensor distribution algorithms, as evidenced by the simulation results.

We also find that minor unbalance in energy consumption will be caused by cluster topology maintenance and node sleep-listening mechanism in different coronas if our strategy is adopted. In this paper, we only use a simple sleep-listening schedule approach. How to get the efficient cluster topology maintenance approach and node sleep-listening mechanism for our algorithm are the major issues we need to solve in the future.

## References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, 2002.

[2] G. Chen, C. Li, M. Ye, and J. Wu, "An unequal cluster-based routing protocol in wireless sensor networks," *Wireless Networks*, vol. 15, no. 2, pp. 193–207, 2009.

[3] D. Culler and W. Hong, "Wireless sensor networks," *Communications of the ACM*, vol. 47, no. 6, pp. 30–33, 2004.

[4] 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.

[5] F. Ingelrest, D. Simplot-Ryl, and I. Stojmenović, "Target transmission radius over LMST for energy-efficient broadcast protocol in ad hoc networks," in *Proceedings of the IEEE International*

*Conference on Communications*, pp. 4044–4049, Paris, France, June 2004.

[6] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660–670, 2002.

[7] V. Mhatre and C. Rosenberg, "Design guidelines for wireless sensor networks: communication, clustering and aggregation," *Ad Hoc Networks*, vol. 2, no. 1, pp. 45–63, 2004.

[8] X. Wu, G. Chen, and S. K. Das, "On the energy hole problem of nonuniform node distribution in wireless sensor networks," in *Proceedings of the IEEE International Conference on Mobile Ad Hoc and Sensor Sysetems (MASS '06)*, pp. 180–187, Vancouver, Canada, October 2006.

[9] S. Chachra and M. Marefat, "Distributed algorithms for sleep scheduling in wireless sensor networks," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '06)*, pp. 3101–3107, Orlando, Fla, USA, May 2006.

[10] 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.

[11] 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)*, pp. 1–12, Barcelona, Spain, April 2006.

[12] J. Hill, R. Szewczyk, A. Woo, S. Hollar, D. Culler, and K. Pister, "System architecture directions for networked sensors," *SIGPLAN Notices*, vol. 35, no. 11, pp. 93–104, 2000.

[13] C. Song, M. Liu, J. Cao, Y. Zheng, H. Gong, and G. Chen, "Maximizing network lifetime based on transmission range adjustment in wireless sensor networks," *Computer Communications*, vol. 32, no. 11, pp. 1316–1325, 2009.

[14] M. Perillo, Z. Cheng, and W. Heinzelman, "On the problem of unbalanced load distribution in wireless sensor networks," in *Proceedings of the IEEE Global Telecommunications Conference Workshops (GLOBECOM '04)*, pp. 74–79, Dallas, Tex, USA, December 2004.

[15] S. Bandyopadhyay and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," in *Proceedings of the 22nd Annual Joint Conference on the IEEE Computer and Communications Societies*, pp. 1713–1723, April 2003.

[16] S. Bandyopadhyay and E. J. Coyle, "Minimizing communication costs in hierarchically-clustered networks of wireless sensors," *Computer Networks*, vol. 44, no. 1, pp. 1–16, 2004.

[17] T. Moscibroda and R. Wattenhofer, "Maximizing the lifetime of dominating sets," in *Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS '05)*, p. 242, April 2005.

[18] X. Fan and Y. Song, "Improvement on LEACH protocol of wireless sensor network," in *Proceedings of the International Conference on Sensor Technologies and Applications (SENSORCOMM '07)*, pp. 260–264, Valencia, Spain, October 2007.

[19] 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.

[20] O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 366–379, 2004.

[21] S. Soro and W. B. Heinzelman, "Prolonging the lifetime of wireless sensor networks via unequal clustering," in *Proceedings of the 19th IEEE International Parallel and Distributed Processing Symposium (IPDPS '05)*, April 2005.

[22] M. Ye, C. Li, G. Chen, and J. Wu, "EECS: an energy efficient clustering scheme in wireless sensor networks 10a.2," in *Proceedings of the 24th IEEE International Performance, Computing, and Communications Conference (IPCCC '05)*, pp. 535–540, April 2005.

[23] J. Lian, L. Chen, K. Naik, T. Otzu, and G. Agnew, "Modeling and enhancing the data capacity of wireless sensor networks," in *IEEE Monograph on Sensor Network Operations*, pp. 91–183, IEEE Press, 2004.

[24] C. Li, M. Ye, G. Chen, and J. Wu, "An energy-efficient unequal clustering mechanism for wireless sensor networks," in *Proceedings of the 2nd IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS '05)*, pp. 597–604, Washington, DC, USA, November 2005.

[25] J. Yu, Y. Qi, G. Wang, Q. Guo, and X. Gu, "An energy-aware distributed unequal clustering protocol for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2011, Article ID 202145, 8 pages, 2011.

[26] 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.

[27] T. Shu and M. Krunz, "Coverage-time optimization for clustered wireless sensor networks: a power-balancing approach," *IEEE/ACM Transactions on Networking*, vol. 18, no. 1, pp. 202–215, 2010.

[28] H. Edwards, *Galois Theory*, Springer, New York, NY, USA, 1984.

[29] N. Amini, A. Vahdatpour, W. Xu, M. Gerla, and M. Sarrafzadeh, "Cluster size optimization in sensor networks with decentralized cluster-based protocols," *Computer Communications*, vol. 35, no. 2, pp. 207–220, 2012.

[30] C. Schurgers, V. Tsiatsis, and M. B. Srivastava, "STEM: topology management for energy efficient sensor networks," in *Proceedings of the IEEE Aerospace Conference*, 2002.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 102508, 6 pages
http://dx.doi.org/10.1155/2013/102508

*Research Article*

# On the Security of Certificateless Signature Schemes

**Gaurav Sharma, Suman Bala, and Anil K. Verma**

*Computer Science and Engineering Department, Thapar University, Patiala 147004, India*

Correspondence should be addressed to Gaurav Sharma; gaurav.sharma@thapar.edu

Received 21 December 2012; Revised 19 May 2013; Accepted 20 May 2013

Academic Editor: J. Barbancho

Copyright © 2013 Gaurav Sharma 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.

Wireless Sensor Network (WSN) has proved its presence in various real time applications and hence the security of such embedded devices is a vital issue. Certificateless cryptography is one of the recent paradigms to provide security. Certificateless public key cryptography (CL-PKC) deals effectively with the twin issues of certificate management in traditional public key cryptography and key escrow problem in identity-based cryptography. CL-PKC has attracted special attention in the field of information security as it has opened new avenues for improvement in the present security architecture. Recently, Tsai et al. proposed an improved certificateless signature scheme without pairing and claimed that their new construction is secure against different kinds of attacks. In this paper, we present a security analysis of their scheme and our results show that scheme does not have resistance against malicious-KGC attack. In addition, we have found some security flaws in the certificateless signature scheme of Fan et al. and proved the scheme vulnerable to Strong Type I attack.

## 1. Introduction

The validation of public keys by a trusted third party, also known as Certificate Authority (CA), makes traditional Public Key Infrastructure (PKI) uneconomical. The user selects a public key and then CA provides a digital certificate to associate the public key with the user's identity. The management of these certificates is a complex issue and increases the computation and storage cost manifold. To resolve the issues of PKC a revolutionary ID-based infrastructure was introduced by Shamir [1] in 1984. This seminal concept of Identity Based Cryptography (IBC) allows the user to choose a public key of its own choice such as email ID, phone number, and name. In IBC, users do not generate their own private keys as in traditional PKC. Private keys are generated by Key Generation Centre (KGC), maintains the private keys of all the users, but there is always a possibility of the misuse of these private keys as they can be used to decrypt any ciphertext and forge the signature of user on any message for signature generation. Eventually, this new paradigm solved the problem of certificate management but gave birth to inherent problem of key escrow.

In 2003, Al-riyami and Paterson [2] proposed a novel approach to eliminate the inherent key escrow problem of IBC as well as the use of certificates in traditional PKC. This approach is known as CL-PKC, where KGC generates a partial-private key for the user while user's secret key and partial-private key are used to generate the public key of the user. In other words, CL-PKC differs from IBC in terms of arbitrary public key, and when a signature is transmitted, user's public key is attached with it but not certified by any of the trusted authority. Moreover, KGC is not aware of the secret key of the user.

However, Al-riyami and Paterson's [2] scheme has been proved insecure against Type I adversary by Huang et al. [3] and proposed an improved scheme. A generic construction has been proposed by Yum and Lee [4] in 2004 which is based on identity based signature. Later, Hu et al. [5] found it insecure against key replacement attack and proposed an improved version. Meanwhile Libert and Quisquater [6] proposed another generic construction without precomputations, which is based on Al-riyami and Paterson's work. In 2005, Gorantla and Saxena [7] proposed an efficient CLS scheme but it was found to be insecure against the key replacement attack by Cao et al. [8]. Li et al. [9] and Zhang et al. [10] proposed CLS schemes based on elliptic curve but verification algorithms in their schemes require four pairing computations. To improve the performance, Yap et al. [11] proposed an efficient CLS scheme which required only two bilinear pairings. However, Park and Kang [12] found that

the scheme [11] is insecure against a key replacement attack. Recently, Au et al. [13] suggested a new kind of malicious-but-passive-KGC attack where adversary may get access to the secret/public key of KGC and then modified Hu et al.'s model [5] for capturing the attack. In 2007, Huang et al. [14] proposed two new short CLS schemes and claimed their first scheme is provably secure against a Normal Type I adversary as well as Super Type II adversary and the second scheme is secure against Super Type I and Type II adversaries. Unfortunately, Shim [15] claimed that the first scheme in [14] is universally forgeable by the Type I adversary. Later, Tso et al. [16–18] presented efficient short CLS schemes. Recently two CLS schemes were proposed by Xu et al. in [19, 20] for mobile wireless cyber-physical systems, and emergency mobile wireless cyber-physical systems respectively. They were claimed to provide high efficiency and provable security. However, Zhang et al. [21] has shown that these two schemes are universally forgeable against public key replacement attack. Wang et al. [22] proposed a scheme which need not compute the pairing $e(P, P) = g$ at the sign stage, rather it precomputes and publishes the system parameters.

Recently, Du and Wen [23] presented a short CLS scheme and claimed that it is secure against Strong adversaries. However, Fan et al. [24] and Choi et al. [25] independently showed it to be insecure against Strong Type I adversary. Further, Fan et al. [24] proposed a CLS scheme from bilinear pairing with additional property of nonrepudiation but later it was found in [26] that the scheme does not acheive Girault's level 3 security. Later, Tian et al. [27] claimed that the scheme [25] didnot withstand against Strong Type II adversary.

In certificateless infrastructure, the majority of the schemes lacks in some common security issue. To attack a CLS scheme broadly two types of adversaries have been defined: Type I and Type II. A Type I adversary can replace a user's public key but is not able to obtain KGC's master secret key and a Type II adversary is a malicious KGC who knows the master secret key but cannot replace user's public key. Although Huang et al. [28] divide the potential adversaries according to their attack power and enrich the CL-PKC with three more categories. A clear definition of all the three categories of adversaries, Normal, Strong, and Super, has been provided together with the security models. On association with the existing categorization of Type I and Type II adversaries, six types of adversaries can be obtained. These are Normal Type I, Strong Type I, Super Type I, Normal Type II, Strong Type II, and Super Type II. In fact, if a scheme is secure against a Super Type I (II) adversary, it will guarantee the security against Normal and Strong Type I (II) adversaries but the reverse may not be true.

In any certificateless scheme, it is always a good idea to avoid pairing operation as it leads to the increase in computation cost manifold as compared to any other operation. An interesting attempt has been made by He et al. [29] in 2011. He et al. developed an efficient short CLS scheme without pairing. The advantage of the scheme is that it does not use any pairing operation and the length of signature is short. However, in 2012, Tian and Huang [30] proved that the scheme cannot resist against Strong Type II adversary having an access to the master secret key of the KGC. Later Tsai et al. [31] discovered that the short CLS scheme [29] cannot withstand against Type II adversary and proposed an improved scheme to overcome the weaknesses of He et al.'s [29] scheme. In this paper, we provide a cryptanalysis on the Tsai et al. [31] scheme by using two Type II attacks.

As all the schemes based on ID-based cryptography have been implemented on sensor network, so these schemes are similarly applicable to Wireless Sensor Network [32]. Mica2, Micaz, Tmote sky, and TelosB are the commonly available motes and can be used for implementation. Evaluation of these schemes can be on the basis of various factors like energy consumption, computation time, and security provided. The schemes discussed here in this papers are very much of interest because they are free from pairing, so easily applicable to WSN. But with less resource consumption scheme should not compromise with security. These schemes are found to be vulnerable and few flaws have been reported. In this paper few attacks have been given which will help to improve the scheme.

The rest of the paper is organized as follows. Section 2 presents some preliminaries and complexity assumptions. Section 3 reviews the Tsai et al.'s scheme [31]. In Section 4, we discuss the security analysis of Tsai et al.'s scheme and prove that the scheme is insecure against Strong Type II attack. Section 5 reviews the Fan et al.'s scheme [24]. In Section 6, we discuss the security analysis of Fan et al.'s scheme and proved in insecure against Strong Type I attack followed by the concluding remarks on the presented work.

## 2. Preliminaries

This section revisits the fundamentals used in the CLS scheme.

*2.1. Overview of Elliptic Curve Cryptography.* An elliptic curve [33, 34] is a set of points over a finite field $GF(p)$, a Galois Field of order $p$, which satisfies the Weierstra $\mathscr{B}$ equation [35]

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \tag{1}$$

but for simplification of computations, cryptographic applications prefer the simple form of Weierstra $\mathscr{B}$ equation as

$$y^2 = x^3 + ax + b, \tag{2}$$

where $a, b \in GF(p)$.

*2.2. Complexity Assumptions.* The security of elliptic curve based cryptosystem is based on the assumption that the Elliptic Curve Discrete Logarithm Problem (ECDLP) is hard, which can be defined as follows.

Let $E$ be an elliptic curve over a finite field $F_p$. Suppose, there are points $P$, $Q$ on the curve $E(F_p)$ for given generator $P$. Determine $k$ such that $Q = [k]P$.

## 3. Review of Tsai et al.'s Short CLS Scheme

In this section, we briefly review the short certificateless signature scheme based on ECDLP [31]. The scheme works as follows.

*Setup*. Let $G$ be a cyclic additive group, let $E/F_p$ be an elliptic curve $E$ over a prime finite field $F_p$ defined by an equation $y^2 = x^3 + ax + b$, and let $p$ be $k$-bit prime number, where $p \in G$. Initially, the KGC computes its master public key $P_{\text{pub}} = xP$ and chooses two secure one-way hash functions: $H_1 : \{0,1\}^* \times G \times G \rightarrow Z_n^*$ and $H_2 : \{0,1\}^* \times G \times G \times G \rightarrow Z_n^*$, where $x \in Z_n^*$ is the master key chosen by KGC. The KGC then publishes public parameters $\{F_p, E/F_p, G, P, P_{\text{pub}}, H_1, H_2\}$ and keeps master key $x$ secret.

*Set-Secret Value*. A signer chooses his/her identity ID and his/her secret value $x_{\text{ID}}$. The signer then computes $P_{\text{ID}} = x_{\text{ID}}P$ and keeps master key $x$ secret $x_{\text{ID}}$.

*Partial-Private-Key Extract*. The KGC computes $R_{\text{ID}} = r_{\text{ID}}P$ and $h_{\text{ID}} = H_1(\text{ID}, R_{\text{ID}}, P_{\text{ID}})$ for each signer with his/her identity $\text{ID} \in \{0,1\}^*$, where $r_{\text{ID}} \in Z_n^*$ is a random number. The KGC then computes $s_{\text{ID}} = r_{\text{ID}} + h_{\text{ID}}x \bmod n$ and sends $(s_{\text{ID}}, R_{\text{ID}})$ to the user via a secure channel. Notably, the tuple $(s_{\text{ID}}, R_{\text{ID}})$ is the partial-private key of the user and the user can confirm its validity by checking the following equation: $s_{\text{ID}}P = R_{\text{ID}} + h_{\text{ID}} \cdot P_{\text{pub}}$. If the equation holds, the partial-private key $(s_{\text{ID}}, R_{\text{ID}})$ is valid; otherwise, the signer rejects the partial-private key $(s_{\text{ID}}, R_{\text{ID}})$.

*Set-Private Key*. The signer uses $sk_{\text{ID}} = (x_{\text{ID}}, s_{\text{ID}})$ as his/her private key.

*Set-Public Key*. The signer adopts $pk_{\text{ID}} = (P_{\text{ID}}, R_{\text{ID}})$ as his/her public key.

*Sign.* Assume a signer wants to sign a message $m$, he/she performs the following steps to generate signature $(R, s)$ on chosen message $m$.

(i) The signer computes $R = l \cdot P$, $h_1 = H_2(m, R, P_{\text{ID}}, R_{\text{ID}})$, $h_2 = H_2(m, R, P_{\text{ID}}, R_{\text{ID}}, P_{\text{pub}})$, where $r_{\text{ID}}$ is a random number.

(ii) The signer checks whether $gcd(l + h_1, n)$ equals 1. If it does not hold, the signer returns to step (i).

(iii) The signer computes $s = (l + h_1)^{-1}(h_2 \cdot x_{\text{ID}} + s_{\text{ID}}) \bmod n$ and then sends $(R, s)$ to the verifier.

*Verify.* Upon receiving the signature $(R, s)$ on message $m$ from the signer, the verifier can confirm the validity of signature $(R, s)$ using the following equation:

$$s \cdot (R + h_1 \cdot P) = h_2 \cdot P_{\text{ID}} + R_{\text{ID}} + h_{\text{ID}} \cdot P_{\text{pub}}, \tag{3}$$

where $h_1 = H_2(m, R, P_{\text{ID}}, R_{\text{ID}})$, $h_2 = H_2(m, R, P_{\text{ID}}, R_{\text{ID}}, P_{\text{pub}})$, and $h_{\text{ID}} = H_1(\text{ID}, R_{\text{ID}}, P_{\text{ID}})$.

If the above equation holds, signature $(R, s)$ is valid; otherwise, the verifier rejects the signature.

## 4. Cryptanalysis of Tsai et al.'s Short CLS Scheme

In this section, we prove that the He et al. [29] CLS scheme is forgeable by the Strong Type II adversary; that is, the adversary can forge users certificateless signatures by using malicious-KGC attack. Tsai et al. proposed an improvement in the He et al.'s [29] scheme and claimed that the scheme is secure under discrete logarithm assumption in random oracle model. Unfortunately, the scheme was found to be insecure against the malicious-KGC attack.

*4.1. Attack 1.* The adversary $\mathcal{A}_{\text{II}}$ will perform the following steps.

(i) The adversary $\mathcal{A}_{\text{II}}$ choose random numbers $t, l' \in Z_n^*$ and a message $m'$ and computes

$$R' = l'P. \tag{4}$$

The adversary $\mathcal{A}_{\text{II}}$ replaces the KGC's master public key $P_{\text{pub}}$ with

$$P'_{\text{pub}} = \frac{t - R_{\text{ID}}}{h'_{\text{ID}}}, \tag{5}$$

where, $h'_{\text{ID}} = H_1(\text{ID}, P_{\text{ID}}, R_{\text{ID}})$.

And, the adversary generates the signature as

$$s' = \frac{t + h'_2 P_{\text{ID}}}{(l' + h'_1)P} \bmod n, \tag{6}$$

where $h'_1 = H_2(m', R', P_{\text{ID}}, R_{\text{ID}})$, $h'_2 = H_2(m', R', P_{\text{ID}}, R_{\text{ID}}, P'_{\text{pub}})$. Clearly, $(R', s')$ is the forged signature on the message $m'$.

(ii) To check the validity of the signature, the verifier can perform the following verification by using the following equation:

$$\begin{aligned} s' \cdot (R' + h'_1 \cdot P) &= \frac{t + h'_2 P_{\text{ID}}}{(l' + h'_1)P} \cdot (l'P + h'_1 P) \\ &= t + h'_2 P_{\text{ID}} \\ &= h'_2 \cdot P_{\text{ID}} + \left[\frac{t - R_{\text{ID}}}{h'_{\text{ID}}} \cdot h'_{\text{ID}} + R_{\text{ID}}\right] \\ &= h'_2 \cdot P_{\text{ID}} + R_{\text{ID}} + h'_{\text{ID}} \cdot P'_{\text{pub}}. \end{aligned} \tag{7}$$

*4.2. Attack 2.* The adversary $\mathcal{A}_{\text{II}}$ will perform the following steps to forge a signature.

(i) The adversary $\mathcal{A}_{\text{II}}$ selects a random number $t' \in Z_n^*$ and computes $R' = t' \cdot P$.

(ii) $\mathcal{A}_{\text{II}}$ chooses a random number $r'_{\text{ID}} \in Z_n^*$ and computes $R'_{\text{ID}} = r'_{\text{ID}} \cdot P$.

(iii) The adversary obtains the hash values $h_1' = H_2(m', R', P_{ID}, R_{ID}')$, $h_2' = H_2(m', R', P_{ID}, R_{ID}', P_{pub})$, and $h_{ID}' = H_1(ID, P_{ID}, R_{ID}')$.

(iv) $\mathcal{A}_{II}$ assesses whether $gcd(l + h_1, n)$ equals 1. If it does not hold, the signer returns to step (i).

(v) As the the adversary is of Type II, the value of $x$ is known. Then, $\mathcal{A}_{II}$ computes

$$s' = \left(t' + h_1'\right)^{-1} \left(r_{ID}' + h_{ID}' \cdot x + \frac{h_2' \cdot P_{ID}}{P}\right) \bmod n. \quad (8)$$

The signature is $(R', s')$ on message $m'$.

(vi) To check the validity of the signature, the verifier can perform the following verification as follows:

$$s' \cdot \left(R' + h_1' \cdot P\right) = h_2' \cdot P_{ID} + R_{ID} + h_{ID}' \cdot P_{pub}, \quad (9)$$

where $h_1' = H_2(m', R', P_{ID}, R_{ID}')$, $h_2' = H_2(m', R', P_{ID}, R_{ID}', P_{pub})$, and $h_{ID}' = H_1(ID, P_{ID}, R_{ID}')$

$$\begin{aligned} & s' \cdot \left(R' + h_1' \cdot P\right) \\ &= \left(t' + h_1'\right)^{-1} \left(r_{ID}' + h_{ID}' \cdot x + \frac{h_2' \cdot P_{ID}}{P}\right) \\ &\quad \times \left(t' \cdot P + h_1' \cdot P\right) \\ &= \left(t' + h_1'\right)^{-1} \left(r_{ID}' + h_{ID}' \cdot x + \frac{h_2' \cdot P_{ID}}{P}\right) \left(t' + h_1'\right) \cdot P \\ &= \left(r_{ID}' + h_{ID}' \cdot x + \frac{h_2' \cdot P_{ID}}{P}\right) \cdot P \\ &= \left(r_{ID}' \cdot P + h_{ID}' \cdot x \cdot P + h_2' \cdot P_{ID}\right) \\ &= R_{ID}' + h_2' P_{ID} + h_{ID}' \cdot P_{pub}. \end{aligned} \quad (10)$$

## 5. Review of Fan et al.'s Short CLS Scheme

In this section, we briefly review the short certificateless signature scheme based on ECDLP [24]. The scheme works as follows.

*Setup.* Let $G_1$, $G_2$, and $G_T$ be three cyclic additive groups of prime order $q \leq 2^k$ where $k$ is a security parameter, and let $e$ be an efficiently computable bilinear pairing $e : G_1 \times G_2 \to G_T$, which satisfies the properties of bilinearity and nondegeneracy. Suppose that a message $m$ which will be signed is an element in $Z_q^*$. KGC chooses two random generators $P_1 \in G_1$ and $P_2 \in G_2$ and a random integer $s \in Z_q^*$. It then computes $P_{pub} = sP_2 \in G_2$ and $g = e(P_1, P_2) \in G_T$. It then selects two distinct cryptographic hash functions $H_1 : \{0,1\}^* \to Z_q^*$ and $H_2 : \{0,1\}^* \times G_2 \to Z_q^*$. KGC publishes the system parameters, params $= \{k, G_1, G_2, e, q, P, g, P_{pub}, H_1, H_2\}$, and keeps its master key $s$ secret.

*User-Key Gen.* A user with identity ID randomly chooses $r \in Z_q^*$ and then computes $pk_{ID} = rP_2$ and $pk_{ID}' = r(P_{pub} + Q_{ID}P_2)$ where $Q_{ID} = H_1(ID)$. The user keeps $r$ secretly and sets $(pk_{ID}, pk_{ID}')$ as its public key.

*Partial-Private-Key Gen.* KGC takes params, the user's partial public information $(Q_{ID}, pk_{ID})$ as inputs, and then generates the user's partial-private key $d_{ID} = 1/(s + Q_{ID} + H_1(ID \parallel pk_{ID}))P_1$. Then KGC returns $d_{ID}$ to the user via a secure manner. After receiving $d_{ID}$, the user checks the correctness of $d_{ID}$ by examining if $e(d_{ID}, P_{pub} + Q_{ID}P_2 + H_1(ID \parallel pk_{ID})P_2) = g$. The private key of the user is $(d_{ID}, r)$.

*CL Sign.* To produce the signature on message $m \in \{0,1\}^*$, the user with identity ID performs the following steps:

(i) set $h = H_2(m, pk_{ID})$,

(ii) compute $S = (1/(r + h))d_{ID}$, where $S$ is the signature on message $m$ of the user.

*CL Verify.* Given params, message $m$, $pk_{ID}$, $pk_{ID}'$, and the signature $S$ on message $m$ of the user with identity ID, the signature can be verified as follows:

(i) let $h = H_2(m, pk_{ID})$;

(ii) if the following formula holds, the signature $S$ is valid:

$$\begin{aligned} & e\left(S, pk_{ID}' + H_1\left(ID \parallel pk_{ID}\right) pk_{ID}\right. \\ &\quad \left. + h\left(P_{pub} + Q_{ID}P_2 + H_1\left(ID \parallel pk_{ID}\right) P_2\right)\right) = g. \end{aligned} \quad (11)$$

## 6. Cryptanalysis of Fan et al.'s Short CLS Scheme

In this section, we demonstrate that the Fan et al. [24] CLS scheme is forgeable by the Strong Type I adversary; that is, adversary can replace a user's public key but is not able to obtain KGCs master secret key. $\mathcal{A}_I$ is able to retrieve the partial-private key of the user.

*6.1. Attack.* The $\mathcal{A}_I$ will perform the following steps.

(i) The adversary $\mathcal{A}_I$ chooses a random number $r' \in Z_n^*$ and replaces a user's public key $PK_{ID}$ with $PK_{ID}^* = r'P_2$ and $PK_{ID}'$ with $PK_{ID}'^* = r'(P_{pub} + Q_{ID}P_2)$.

(ii) $\mathcal{A}_I$ makes a strong sign query with ID, $m$, and $r'$ as input and then the challenger returns a valid signature $S' = (1/(r' + h'))d_{ID}$ where $h' = H_2(m, PK_{ID}^*)$.

(iii) $\mathcal{A}_I$ obtains the hash value $h'$ on $m$, $PK_{ID}^*$ by making a hash query.

(iv) $\mathcal{A}_I$ can then compute the user's partial-private key $d_{ID} = (r' + h')S'$ as he knows the value of $r'$ and $h'$.

## 7. Conclusion

The schemes discussed here are of much interest because they are free from pairing and hence can easily be applicable to WSN. But less resource consumption is not enough reason to compromise security. In this paper, security attacks have been applied on two different schemes. Tsai et al. proposed the CLS scheme without pairing which is claimed to be more efficient than the existing schemes (since pairing is always an expensive operation). An exhaustive cryptanalysis has been shown in Section 4 and the results indicate that the improved scheme by Tsai et al. does not resist against the Strong Type II attacks and hence is forgeable. Moreover, we have found that Fan et al's. CLS scheme is forgeable by the Strong Type I adversary. Therefore, to construct a secure certificateless signature scheme without bilinear pairing needs more attention.

## References


[1] A. Shamir, "Identity-based cryptosystems and signature schemes," in *Advances in Cryptology*, vol. 196 of *Lecture Notes in Computer Science*, pp. 47–53, Springer, Berlin, Germany, 1984.

[2] S. S. Al-riyami and K. G. Paterson, "Certificateless public key cryptography," in *Advances in Cryptology-ASIACRYPT 2003*, vol. 2894 of *Lecture Notes in Computer Science*, pp. 452–473, Springer, Berlin, Germany, 2003.

[3] X. Huang, W. Susilo, Y. Mu, and F. Zhang, "On the security of certificateless signature schemes from asiacrypt 2003," in *Cryptology and Network Security*, vol. 3810 of *Lecture Notes in Computer Science*, pp. 13–25, Springer, Berlin, Germany, 2005.

[4] D. H. Yum and P. J. Lee, "Generic constructin of certificateless signature," in *Information Security and Privacy*, vol. 3108 of *Lecture Notes in Computer Science*, pp. 200–211, Springer, Berlin, Germany, 2004.

[5] B. C. Hu, D. S. Wong, Z. Zhang, and X. Deng, "Key replacement attack against a generic construction of certificateless signature," in *Information Security and Privacy*, vol. 4058 of *Lecture Notes in Computer Science*, pp. 235–246, Springer, Berlin, Germany, 2006.

[6] B. Libert and J. J. Quisquater, "On constructing certificateless cryptosystems from identity based encryption," in *Proceedings of the 9th International Conference on Theory and Practice of Public-Key Cryptography (PKC '06)*, vol. 3958 of *Lecture Notes in Computer Science*, pp. 474–490, Springer, Berlin, Germany, 2006.

[7] M. Gorantla and A. Saxena, "An efficient certificateless signature scheme," in *Computational Intelligence and Security*, vol. 3802 of *Lecture Notes in Computer Science*, pp. 110–116, Springer, Berlin, Germany, 2005.

[8] X. Cao, K. G. Paterson, and W. Kou, "An attack on a certificateless signature scheme," Cryptology EPrint Archive 2006/367, 2006, http://eprint.iacr.org/.

[9] X. Li, K. Chen, and L. Sun, "Certificateless signature and proxy signature schemes from bilinear pairings," *Lithuanian Mathematical Journal*, vol. 45, no. 1, pp. 76–83, 2005.

[10] Z. Zhang, D. S. Wong, J. Xu, and D. Feng, "Certificateless public-key signature: security model and efficient construction," in *Applied Cryptography and Network Security*, vol. 3989 of *Lecture Notes in Computer Science*, pp. 293–308, Springer, Berlin, Germany, 2006.

[11] W. S. Yap, S. H. Heng, and B. M. Goi, "An efficient certificateless signature scheme," in *Emerging Directions in Embedded and Ubiquitous Computing*, vol. 4097 of *Lecture Notes in Computer Science*, pp. 322–331, Springer, Berlin, Germany, 2006.

[12] J. Park and B. Kang, "Security analysis of the certificateless signature scheme proposed at Sec Ubiq 2006," in *Emerging Directions in Embedded and Ubiquitous Computing*, vol. 4809 of *Lecture Notes in Computer Science*, pp. 686–691, Springer, Berlin, Germany, 2007.

[13] M. H. Au, J. Chen, J. K. Liu, Y. Mu, D. S. Wong, and G. Yang, "Malicious KGC attacks in certificateless cryptography," in *Proceedings of the 12th Australasian Conference on Information Security and Privacy (ACISP '07)*, vol. 4586 of *Lecture Notes in Computer Science*, pp. 308–322, Springer, 2007.

[14] X. Huang, Y. Mu, W. Susilo, D. S. Wong, and W. Wu, "Certificateless signature revisited," in *Information Security and Privacy*, vol. 4586 of *Lecture Notes in Computer Science*, pp. 308–322, Springer, Berlin, Germany, 2007.

[15] K. Shim, "Breaking the short certificateless signature scheme," *Information Sciences*, vol. 179, no. 3, pp. 303–306, 2009.

[16] R. Tso, X. Yi, and X. Huang, "Efficient and short certificateless signature," in *Cryptology and Network Security*, vol. 5339 of *Lecture Notes in Computer Science*, pp. 64–79, Springer, Berlin, Germany, 2008.

[17] R. Tso, X. Yi, and X. Huang, "Efficient and short certificateless signatures secure against realistic adversaries," *Journal of Supercomputing*, vol. 55, no. 2, pp. 173–191, 2011.

[18] R. Tso, X. Huang, and W. Susilo, "Strongly secure certificateless short signatures," *Journal of Systems and Software*, vol. 85, no. 6, pp. 1409–1417, 2012.

[19] Z. Xu, X. Liu, G. Zhang, W. He, G. Dai, and W. Shu, "A certificateless signature scheme for mobile wireless cyber-physical systems," in *28th International Conference on Distributed Computing Systems Workshops, ICDCS Workshops 2008*, pp. 489–494, chn, June 2008.

[20] Z. Xu, X. Liu, G. Zhang, and W. He, "McCLS: certificateless signature scheme for emergency mobile wireless cyber-physical systems," *International Journal of Computers, Communications and Control*, vol. 3, no. 4, pp. 395–411, 2008.

[21] F. Zhang, S. Miao, S. Li, Y. Mu, W. Susilo, and X. Huang, "Cryptanalysis on two certificateless signature schemes," *International Journal of Computers, Communications and Control*, vol. 5, no. 4, pp. 586–591, 2010.

[22] C. Wang, D. Long, and Y. Tang, "An efficient certificateless signature from pairings," *Journal of Information Science and Engineering*, vol. 8, no. 1, pp. 96–100, 2009.

[23] H. Du and Q. Wen, "Efficient and provably-secure certificateless short signature scheme from bilinear pairings," *Computer Standards and Interfaces*, vol. 31, no. 2, pp. 390–394, 2009.

[24] C. Fan, R. Hsu, and P. Ho, "Truly non-repudiation certificateless short signature scheme from bilinear pairings," *Journal of Information Science and Engineering*, vol. 27, no. 3, pp. 969–982, 2011.

[25] K. Y. Choi, J. H. Park, and D. H. Lee, "A new provably secure certificateless short signature scheme," *Computers and Mathematics with Applications*, vol. 61, no. 7, pp. 1760–1768, 2011.

[26] Y. C. Chen and G. Horng, "On the security models for certificateless signature schemes achieving level 3 security," IACR Cryptology EPrint Archive 554, 2011.

[27] M. Tian, L. Huang, and W. Yang, "On the security of a certificateless short signature scheme," Cryptology EPrint Archive, 2011, http://eprint.iacr.org/2011/419.

[28] X. Huang, Y. Mu, W. Susilo, D. S. Wong, and W. Wu, "Certificateless signatures: new schemes and security models," *Computer Journal*, vol. 55, no. 4, pp. 457–474, 2012.

[29] D. He, J. Chen, and R. Zhang, "An efficient and provably-secure certificateless signature scheme without bilinear pairings," *International Journal of Communication Systems*, vol. 25, no. 11, pp. 1432–1442, 2011.

[30] M. Tian and L. Huang, "Cryptanalysis of a certificateless signature scheme without pairings," *International Journal of Communication Systems*, 2012.

[31] J. Tsai, N. Lo, and T. Wu, "Weaknesses and improvements of an efficient certificateless signature scheme without using bilinear pairings," *International Journal of Communications Systems*, vol. 25, no. 11, pp. 1432–1442, 2012, Wiley-Blackwell.

[32] 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.

[33] "2000. *Standards for efficient cryptography SEC 1: Elliptic curve cryptography*," Certicom Research, http://www.secg.org/collateral/sec1_final.pdf.

[34] "2000. *Standards for efficient cryptography SEC 2: Recommended Elliptic Curve Domain Parameters. Standards for Efficient Cryptography*," Version 1.0. Certicom Research, http://www.secg.org/collateral/sec2_final.pdf.

[35] D. Hankerson, A. J. Menezes, and S. Vanstone, *Guide to Elliptic Curve Cryptography*, Springer, New York, NY, USA, 2004.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 328747, 14 pages
http://dx.doi.org/10.1155/2013/328747

*Research Article*

# Implementing a Distributed WSN Based on IPv6 for Ambient Monitoring

**D. F. Larios, J. M. Mora-Merchan, E. Personal, J. Barbancho, and C. León**

*Department of Electronic Technology, University of Seville, C. Virgen de Africa 7, 41011 Seville, Spain*

Correspondence should be addressed to D. F. Larios; dflarios@dte.us.es

Received 21 December 2012; Accepted 19 April 2013

Academic Editor: Adel Soudani

Copyright © 2013 D. F. Larios 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.

Traditionally, Wireless Sensor Networks (WSNs) are used for monitoring an extensive area. In these networks, a centralized server is usually used to collect and store the sensor information. However, new distributed protocols allow connections directly to the WSN nodes without the need of a centralized server. Moreover, these systems are able to establish communications among heterogeneous networks. The new protocols strategy is focused on considering several WSNs as a unique distributed one. This way, a user of the system is able to analyze a process under study as a whole instead of considering it as a set of different subsystems. This is the case in the evaluation of migratory waterbirds' environment. In this case, it is usual to deploy several WSNs in different breeding areas. They are all interconnected and they measure different environmental parameters. However, this improvement in the data access flexibility may result in a loss of network performance and an increase in network power consumption. Focused on this problem, this paper evaluates different communication protocols: distributed and centralized, in order to determine the best trade-off for environmental monitoring in different migratory areas of waterbirds.

## 1. Introduction

Nowadays, the number of electronic devices present in our environment is increasing continuously. The technological improvements and cost reduction of these smart devices provide a significant increase in their capabilities, among which are their connectivity. The *Internet of Things* (IoT) [1] is a paradigm which is gaining importance in the current scenario of modern telecommunication. This concept represents a novel scenario in which we are completely surrounded by smart devices. These devices can interact among them, providing different information or adding capabilities to the network [2]. An important group of these devices is formed by the Wireless Sensor Networks (WSNs) [3].

A WSN consists of many small devices deployed in a physical environment [4]. Each device, called a node, has special capabilities such as communication with its neighbors, sensing, and data storage and processing [5]. All nodes make a mesh network of devices that can collaborate among them allowing the implementation of distributed solutions to solve complex problems. Due to this, WSNs have many applications [6], among which environmental monitoring as an area where the potential impact is huge [7]. It allows monitoring an area at a low cost and little need of human presence. However, they have some technical requirements, such as the following.

(i) *Autonomy*. Batteries must be able to power the nodes during the whole network lifetime. Despite typical WSNs using low power radio devices (such as IEEE 802.15.4 [8] radio transceivers), the radio transceiver spends most of the energy consumption in a node in typical monitoring applications. Due to it, the network has to reduce data traffic as much as possible.

(ii) *Robustness*. In this kind of application, human maintenance is usually difficult because of the hardness of the terrain. Therefore, it is important to design robust networks that are adaptable to any incident.

(iii) *Flexibility*. The network must to be able to add, move, or remove nodes to meet the application requirements. The network must automatically detect the changes, organizing the communications in consequence.

(iv) *Low Price*. To save energy, the transmission range is limited too. This decrease in communication coverage is solved by increasing the density of the network. A high cost of the nodes would make unfeasible the use of a WSN versus other technologies.

WSNs are normally used in environmental monitoring applications to collect information through the sensors incorporated into each node. There is a special device called "Base Station" whose mission is to request and store the network information [9]. The Base Station generally acts also as a gateway, allowing the user to access the collected data through an infrastructured network, such as Internet.

To monitor an environment, the nodes of a WSN construct mesh networks that allow communications between devices. Due to the special characteristics of WSNs, such as their reduced energy available or low bandwidth, WSNs require the use of adapted protocols to establish communications. Nowadays, routing protocols with low power consumption continue being a main issue for WSNs. Several routing solutions have been proposed in the literature [10], mainly focused on energy consumption [11, 12], security issues [13, 14], or fault-tolerant capacity [15].

However, these proposed communication algorithms are classically designed to solve communication problems in an ad hoc manner. They solve particular scenarios, but they do not provide a general framework that allows intercommunication between heterogeneous networks with different communication requisites.

As an alternative to this classical WSN scheme and directly related to the IoT philosophy, some authors are currently proposing the use of IPv6 implementation over WSNs [16, 17]. This implementation are focused on reducing as much as possible the requirements of the Base Station, that is, using the Base Station only as a gateway between the two networks [18] and without intelligence. The use of IPv6 provides a common framework where additional protocols can be added, maintaining a basic interoperability between networks. IPv6 framework for low power consumption systems (also known as 6LoWPAN) is currently a main research area. Some authors are proposing new uses of 6LoWPAN technology [19, 20]. Other authors proposed additional schemes, for example, to compact addressing between devices [21] or to increase security [22, 23]. Other authors, as classical architectures, focused on provided efficient routing protocols [24, 25].

As can be seen, 6LoWPAN adds the IPv6 advantages of robustness and flexibility to WSNs, increasing the connectivity of the nodes. This implementation not only solves the retrieving information problem from the Base Station but it allows the access to every node in the WSN from anywhere in Internet [26]. Thanks to 6LoWPAN each node can be uniquely identified [27]. However, because IPv6 was not initially designed to operate over WSNs, it also has some constraints [28]. It could not make it interesting for all applications.

In this paper, we propose a comparison between IPv6 protocol implementation versus classical WSN communication protocols. This evaluation is focused on its use for monitoring application.

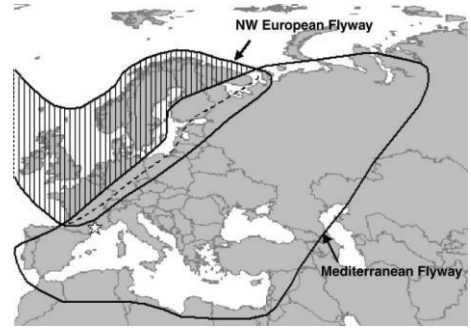



Figure 1: New potential delineations of Teal populations in western Eurasia [30].

The rest of the paper is organized as follows. Section 2 describes the application of our interest. A description and some common WSN protocols are described in Section 3. These protocols are evaluated in Section 4. Finally, Section 5 introduces a comparison between the evaluated protocols, before describing the final conclusions and remarks in Section 6.

## 2. Application Description

The evaluation of flooding areas, especially the estimation of flood level, is very important in monitoring waterbird colonies. Some studies relate the status of waterbird habits to the population of these kinds of birds [29]. Moreover, a change in the behavior and pathways of migratory waterbirds has been observed by biologists in the last decades [30], as can be seen in Figure 1. Due to this, biologists are very interested in this kind of environmental monitoring.

To capture that information, we already have a WSN deployed into the Doñana Biological Station [31]. Doñana is a wildlife reserve protected by the Spanish Government (located in the south of Spain). It covers a huge area of about 542 $km^2$ with little human interference through its entire history. This network is called ICARO [32] ("Inteligencia Computacional Aplicada a Redes de Observación," a Spanish acronym which means computational intelligence applied to monitoring networks). ICARO has been built based on TelosB [33] nodes and TinyOS [34] operating system. Each node (Figure 2) has a meteorological station with the ability to measure temperature, rainfall, wind speed and direction, and humidity. These nodes use a TelosB platform for processing and transmissions. Finally each node gets energy from a solar power system. These nodes use available information to predict the flood level of the marsh areas of Doñana [35].

ICARO network architecture is a centralized one, where the flood level estimation of each node is sent to a Base Station and stored later in a server. Node information is only accessible offline, and data is stored in a external database.

However, an effective monitoring of the habitat of migratory species requires deploying several networks in each

Figure 2: A node of ICARO WSN.

breeding area. Therefore, we are designing new WSNs (i.e., ICARO2, ICARO3, etc.) to cover new areas (the new design of the improved nodes is shown in Figure 3). One of our goals is to provide a unified interface for all the ICARO networks. This set of networks is called eSapiens, a Spanish acronym that means intelligent acquisition and processing system integrated in natural environments.

Although we can keep the ICARO architecture (for reusing the hardware and the deployments), this scheme suffers several problems: ICARO depends on several hard-defined bottle necks. All information goes through a sole Base Station. Inside nodes are not accessible from outside nodes. Therefore, it is not possible to access data directly from sources, but it is in a data server.

A new approach is chosen for eSapiens: flatten all hierarchical structures, allow direct communication to every node (Figure 4) in all networks, and keep historical data in each node. To embrace this approach, a change to a more complete communication protocol is necessary.

One advantage obtained from a more complete communication protocol is that it makes easy the communication between nodes of different networks, redirecting the information through a heterogeneous networks. Therefore, a node in an ICARO network can exchange information with another node of another ICARO network. It is depicted in Figure 5. Additionally, the information stored in a node can be retrieved by a PC using a standard web-page browser, where a user can access individually and transparently each device of the different ICARO networks into eSapiens infrastructure.

## 3. WSN Communications in Environmental Monitoring

Typically, WSN protocols allow mesh typologies of nodes with multihop structures. Thanks to this, it is possible to monitor huge areas with low cost and low power devices.

Figure 3: Node example for the new ICARO networks.

Moreover, it requires the use of low power radio transceivers. Typically, these radio transceivers are based on the IEEE 802.15.4 Standard [8]. This standard defines the physical and the MAC layers as is detailed in Figure 6. Based on this standard, several protocols have been defined. These protocols are necessary to enable the communication between nodes with multiple hops. Each node is required to know its neighbors and what routes it can use to send a message to another node, without sight line. Due to this, WSN protocols used for environmental monitoring require two classes of messages.

(i) *Network Messages*. These messages are used to maintain the network. It is used to discover node neighbors and obtain the routing table.

(ii) *Information Messages*. These messages are used to transmit useful information for application purposes. In environmental monitoring, these messages typically contain sensor measurements.

Both types of messages are needed; however the use of network messages increases the power consumption and increases the occupation index of the channel with information not directly related to the proposed application. Moreover, obtaining a high reliability requires the use of big headers. Therefore, to reduce the power consumption it is necessary to obtain a trade-off between the required reliability of the network and the amount of additional information sent between nodes.

Currently, there are different solutions for routing protocols in WSNs. Communication protocols designed specifically for WSNs, such as Collection Tree Protocol or ZigBee, are less flexible but their implementation over WSNs has less power consumption. On the other hand, IPv6 implementation, such as BLIP, takes advantage of the information messages to update the routing table.

In this section, we are going to describe some common WSN protocols, highlighting their advantages and disadvantages to use on the proposed application.

*3.1. ZigBee.* The ZigBee standard [36] is a project supported by the ZigBee Alliance. ZigBee defines the network layer and the application layer, both on top of IEEE 802.15.4 link layer. This standard defines several net structures in which three

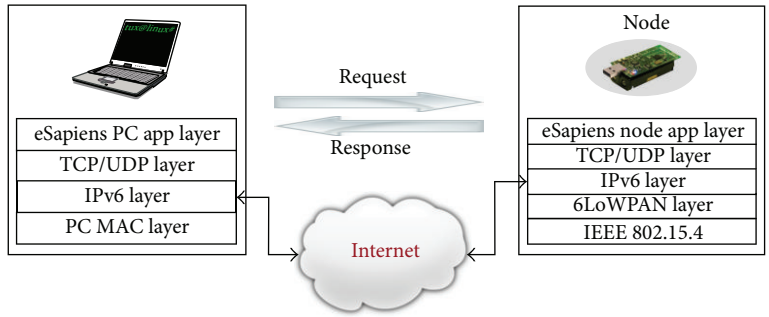



Figure 4: Direct communication to the node through Internet.

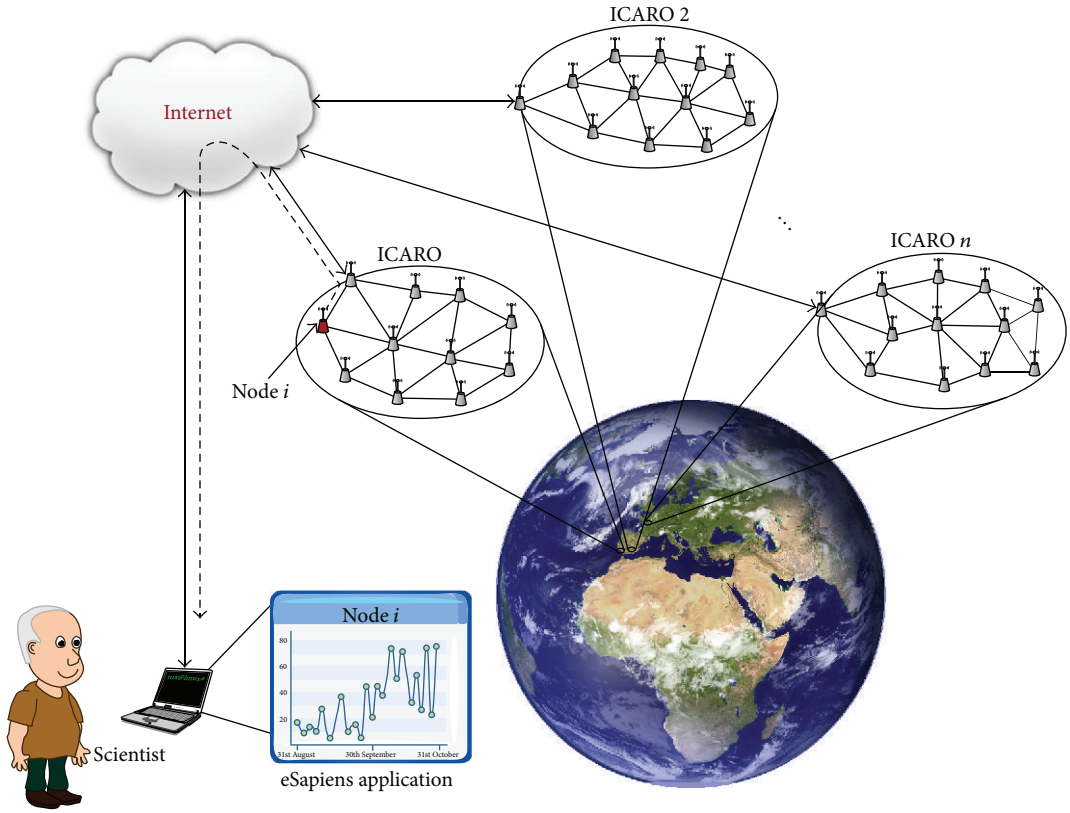



Figure 5: WSNs mutipoint access through Internet.

types of nodes are identified: *a coordinator*, *routers*, and *end devices*. The application layer ZigBee defines an application support sublayer (APS) on which the denominated end points (similar to TCP/IP ports) are defined. An application or object can be defined for each end point. However, there is a reserved end point for the ZigBee device object (ZDO) which is responsible for communicating the *binding tables* to the network. These tables keep information about the network nodes' presence and their services. In the ZigBee network the services define the *profiles* whereby the devices are grouped in *clusters* (devices within the same profile).

*3.2. Collection Tree Protocol.* The Collection Tree Protocol (CTP) [37] is a tree-based collection protocol. In this protocol, a single or small group is shown as sink nodes (tree roots) which are the main branches of the tree (connectivity between nodes). Based on this tree, when each node sends a message to a sink node, it looks for the most suitable neighbor node, using a routing gradient. With this technique, the protocol is responsible for network management discovering the neighbors and estimating the best transmission path (minimizing the number of hops and the global consumption). This is possible because each node has a table with a list of the best neighbors to reach a root node. The CTP is common in data collection applications where nodes periodically send information to a root node (typically called Base Station). A specific implementation of this protocol is CTP Noe [38].

*3.3. 6LoWPAN.* The IPv6 in Low Power Wireless Personal Area Networks (6LoWPAN) [39, 40] is a project supported

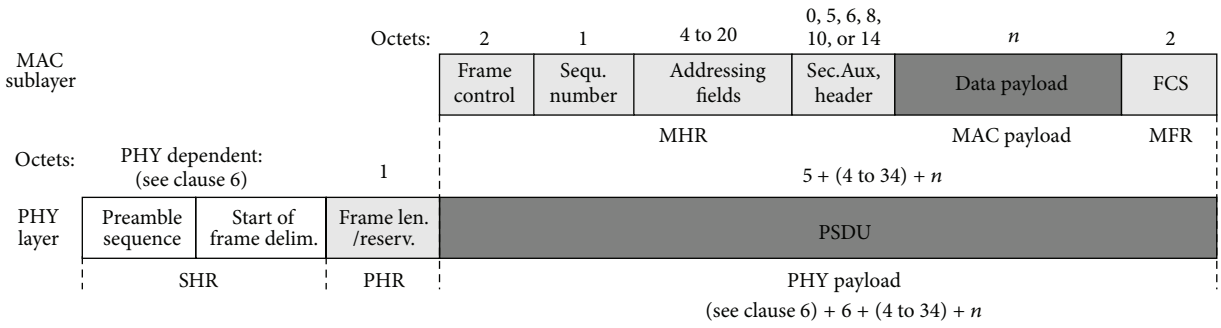



Figure 6: IEEE 802.15.4 datagram.

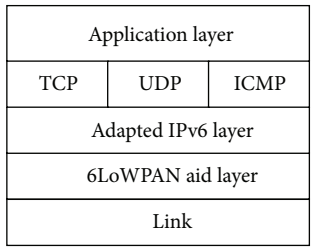



Figure 7: 6LoWPAN layer description.

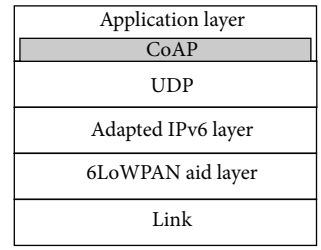



Figure 8: CoAP layer description.

by the IETF (Internet Engineering Task Force). It was created to adapt IPv6 datagrams over IEEE 802.15.4 links. This adaptation presents the problem of great differences in MTU sizes between IPv6 and IEEE 802.15.4 (much larger in typical IPv6 implementation). To resolve this problem, 6LoWPAN implements a packet fragmentation mechanism [41] and a header compression algorithm [42]. On the other hand, the use of IPv6 simplifies tasks of network management and expands connectivity with Internet protocols (e.g., TCP, UDP, and ICMP) and services [26] (e.g., HTTP, SMTP, FTP, and SOAP), as can be seen in Figure 7.

*3.3.1. BLIP.* The Berkeley Low power IP stack (BLIP) [43] is implementation of IPv6 for WSNs over TinyOS. Currently, it is not full IETF standard compliant. However, it supports 6LoWPAN/HC-01, header compression, IPv6 neighbor discovery, default route selection, point-to-point routing, and network programming. If there are communications in the network, the device that receives a message analyzes the sender information and the ACK to refresh its neighbor and the routes tables. Only in the case that no communication occurs during a long time, a node sends a message to refresh its routing table. Additionally, some of the implementation allow the use of compressed headers. That reduces the overhead of IPv6 communications. With all this, BLIP provides significant interoperability with other IP networks.

*3.3.2. Constrained Application Protocol.* The Constrained Application Protocol (CoAP) [44] is a specialized web transfer protocol for machine-to-machine (M2 M) applications. It is proposed by CoRE (Constrained RESTful Environments), a work group at the IETF. The main goal of CoAP is to provide a mechanism to easily translate a protocol like HTTP to a less complex one. It allows integration between constrained networks (such as WSNs) and standard Internet networks. CoAP is a single protocol over UDP (as is shown in Figure 8) and is subdivided into two sublayers. The first one, the CoAP message sublayer, is responsible for dealing with UDP (CoAP operates over UDP) and defines four types of messages: *Confirmable*, *Nonconfirmable*, *Acknowledgement*, and *Reset*. The second one is the request/response sublayer. It contains methods and response codes. An implementation of this protocol is `libcoap` which provides the same methods as the ones used by HTTP: GET, PUT, POST, and DELETE. However, there is specific implementation of these libraries for WSNs [45] with a more reduced set (only GET and PUT) which is typically sufficient for most applications.

## 4. Experimental Evaluations

In this paper, we are going to compare different protocols which allow us to use IPv6 in a WSN (a network of TelosB nodes with TinyOS Operating System). Additionally, some other well-known protocols are included in the comparison.

A developed test application sends the same application data using each studied protocol. A sniffer captured the whole wireless traffic between nodes. To acquire data traffic, our testbench comes with a USB dongle 802.15.4 sniffer (IA OEM DAUB1 2400 by Adaptive Modules Ltd.) [46] and "Perytons Protocol Analyzers" [47] (Figure 9) software to analyze the packets detected.

The network used to test the protocols is made of 3 nodes with the architecture described in Figure 10.

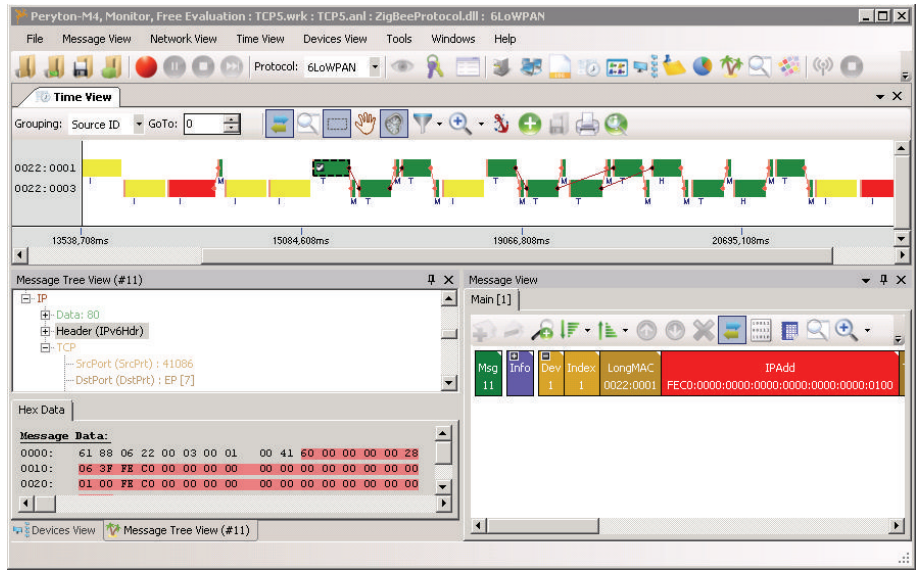
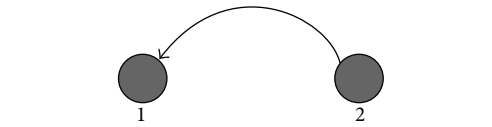
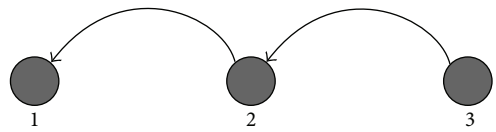



Figure 9: Perytons Protocol Analyzers screenshot.




(a) Topology without hop




(b) Topology with hop

Figure 10: Studied topologies.

A limitation of output power in the transceiver of nodes (down to −12 dB) allows the experimentation in a close controlled scenario. It causes each node not to have full connectivity with the rest of the nodes. Due to this, in some cases nodes have to route messages through other nodes. Same coverage is used for all the experiments.

All the IPv6 dependent protocols use the Berkeley implementation for WSNs (6LoWPAN).

Our test application is developed to imitate the behavior of the proposed flooding estimation application.

(i) Every node is cyclically acquiring environmental information.

(ii) Once per day, a WSN node executes the data aggregation algorithm, to predict flood level.

(iii) Only if the flood level is different from the last estimation, it sends a message to the Base Station with the new estimation.

Based on this simple architecture, a count of sent bytes per data has been done. In all measurements, it only showed size in bytes of any layer above MAC layer. To get real size of messages, it is necessary to add the MAC size. Usually WSNs based on IEE802.15.4 radio transceiver use short addresses (9 bytes) to reduce the overhead. When some protocols depend on responses in MAC layer (i.e., wait for MAC ACK), those messages are shown too.

Moreover, all experiments done in both topologies show an $n$-hop communication which is an equivalent to $n$ times 1-hop. The overhead of frames for the evaluated protocols does not increase with multiple hops. So, the results only show the last transmission (1-hop).

All the expressed results are obtained without considering fragmentation. To maintain a low power consumption, it is necessary to reduce as much fragmented messages as possible because it increases the required transmissions with their associated ACK and therefore it increases the overhead.

*4.1. Static Ad Hoc Implementation.* This implementation is used only for comparison purposes. An Ad hoc implementation is specific to application and network. The efficiency of an ad hoc implementation is maximum because its topology is fixed in programming time and routing tables are static. Due to this, these networks do not require network messages to obtain the topology neither to provide information to help the routing of the packets.

Against that, it needs to know beforehand the location of the network nodes. A change in its topology requires reprogramming all nodes. Neither of these problems have been considered here.

This kind of implementation is not very common due to the difficulty in the evaluation of topologies in networks.

Despite their lack of flexibility, these implementation have very few overheads and do not require additional messages to build a routing table. So, it is considered as an ideal case and it is used as reference to compare other routing protocols.

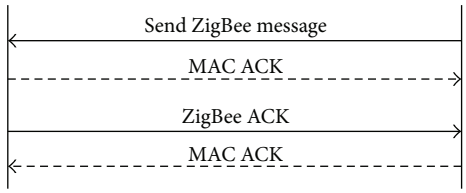



Figure 11: ZigBee message transmission.

With the proposed testbench, we talk about two different protocols: a one *similar* to UDP without ACK (1 byte to destination, 1 byte to source and payload) and a one *similar* to TCP with ACK (2 bytes more with source and target to send ACK) (see Table 1). If both protocols do not coexist, it is not necessary to have an extra byte to identify the protocol.

This implementation does not allow communication between devices. Moreover, they cannot acquire information from a remote WSN. Therefore, all the acquired information must be stored in a central server.

*4.2. ZigBee.* The structure of ZigBee transmissions is depicted in Figure 11. In this case, a remote node is sending a non-requested message. ZigBee requires the use of ACK messages at link and MAC layers. Figure 12 shows a ZigBee frame. As can be seen, it has a reduced overhead, similar to CTP one.

Table 2 sums up the size of messages transmitted with ZigBee. These messages do not increase in size with multihop transmissions.

ZigBee allows direct communication between devices of a network, but it does not allow to redirect messages with a remote network. ZigBee has not been designed to allow fragmentation.

*4.3. Collection Tree Protocol Implementation.* Collection protocols are useful to gather information from different sensors to a central device in a network, generally to the Base Station. This is the typical application in environmental monitoring where all the information gathered in the Base Station is stored in a database.

Collection nodes only store a reduced number of routes. They try to send information through an optimum path to a Base Station. If the nodes in the path are busy, they search for other paths, always trying to minimize the cost (estimated as number of hops).

This implementation does not allow communication between devices. The network only maintains routes to the Base Station. Moreover, it cannot acquire information from other remote WSNs. Therefore, all the acquired information must be stored in a central server.

Figure 13 shows a generic frame of a collection message. Its overhead is reduced.

Typical CTP schema is depicted in Figure 14. This protocol only sends ACK at MAC layer, maintaining the number of exchanged messages low.

Table 3 sums up the size of messages transmitted with the evaluated CTP protocol (from TinyOs Stack). CTP messages do not increase their size in multihop nodes.

Table 1: Ad hoc protocol.

| Msg type | Bytes sent |
|---|---|
| Data message | 2 + payload (up to 114) |
| MAC ACK | 5 |

Table 2: ZigBee protocol.

| Msg type | Bytes sent |
|---|---|
| ZigBee message | 18 + payload (up to 100) |
| ZigBee ACK | 18 |
| MAC ACK | 5 |

Table 3: Collection protocol.

| Msg type | Bytes sent |
|---|---|
| Data message | 12 + payload (up to 106) |
| MAC ACK | 5 |

*4.4. TCP BLIP Implementation.* BLIP implementation supports over the 6LoWPAN aid layer. The scheme of a TCP message using 6LoWPAN is depicted in Figure 15.

TCP transmission require a complex message negotiation. This negotiation ensures the integrity of the information, but drastically increases the overhead. Messages vary in function of the application. For example, the scheme to transmit a webpage over 6LoWPAN is depicted in Figure 16.

As can be seen, a webpage structure increases the number of required messages even more. Due to this, in transmission with constrained communications, such as that used in WSNs, it is better to use transmissions without a protocol in application layer, that is, sending the information after establishing the socket connection.

TCP allows communication between devices, whether they are in the same network or not. Moreover, the accessibility of each device in a TCP network allows the user to request information stored locally. So, transmission is only by demand.

TCP BLIP permits message fragmentation. If a message is higher than the maximum payload, BLIP automatically fragments it. To do this, it adds fragmentation header to the first fragment (4 bytes). This header is added between the MAC header and the 6LoWPAN header.

The rest of the fragments are sent only with MAC header and a 5-byte fragmentation header that identifies the full message. The rest of the headers in these fragments are avoided to reduce the overhead as much as possible.

*4.4.1. TCP BLIP without Header Compression.* Figure 17 depicts a standard payload message, using TCP frames without header compression. TCP frames require the use of big headers. It considerably reduces the size of the payload or fragments the message.

Table 4 sums up the size of messages with a TCP connection without header compression. As can be seen, the number of transmitted messages and the number of bytes sent in communications are high.

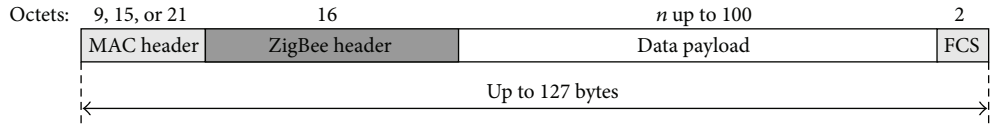



Figure 12: ZigBee frame.

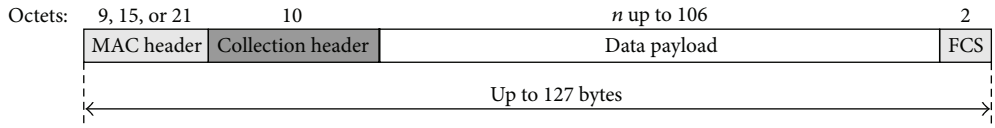



Figure 13: CTP frame.

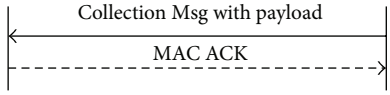



Figure 14: CTP message transmission.

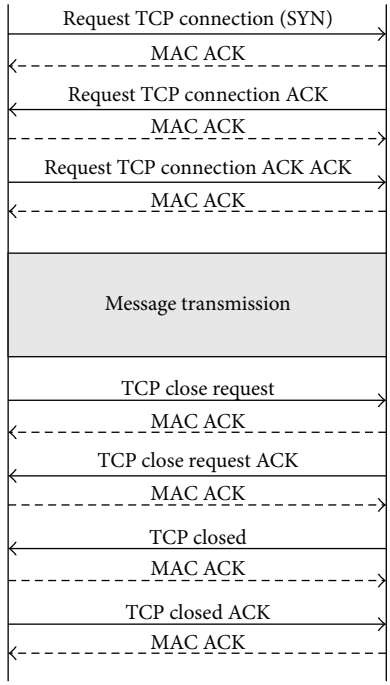



Figure 15: TCP negotiation.

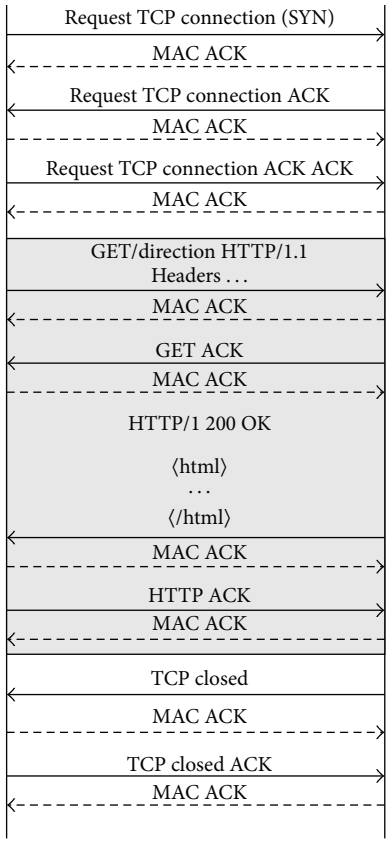



Figure 16: Webpage transmission.

TCP requires a high number of transmissions. Firstly, it requires to start explicit connection with an SYN message. After that, we need to send the message with the payload. Finally, it is necessary to close the connection explicitly. Moreover, TCP requires the use of the ACK messages link layer, in addition to the ACK in MAC layer. Due to this, the TCP protocol is in general too costly to be used in devices with limited bandwidth, such as WSN devices.

*4.4.2. TCP BLIP with Header Compression.* Figure 18 depicted a TCP payload frame. The size of messages with header compression is lower than that of messages without it. However, their headers are still too high for a protocol with constrained maximum message size, such as in 802.15.4.

Table 5 sums up the size of transmission messages in a TCP connection with header compression. As can be seen, TCP header compression reduces the overhead, but it does not reduce the number of transmissions.

As conclusion, the TCP over 802.15.4 radio transceiver requires a large number of transmissions and it has a big overhead, even with header compression. TCP is only useful with not so frequent communication with reduced payload,

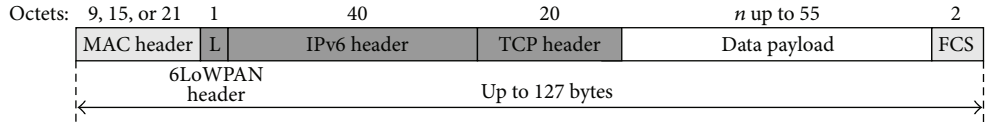



Figure 17: TCP frame without compression.

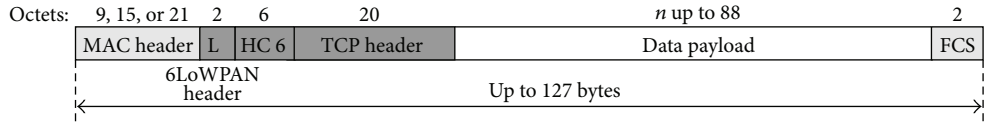



Figure 18: TCP frame with compression.

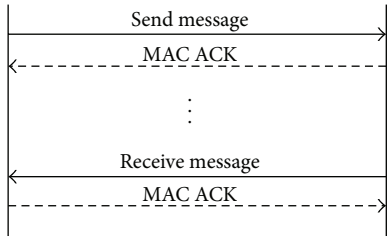



Figure 19: UDP transmission.

Table 4: TCP without compression.

| Msg type | Bytes sent |
|---|---|
| Request TCP connection (SYN) | 83 |
| Request TCP connection ACK | 71 |
| Request TCP connection ACK ACK | 63 |
| Data message | 63 + payload (up to 55) |
| ACK | 71 |
| MAC ACK | 5 |
| Close request | 73 |
| Close ACK | 63 |
| Closed | 63 |
| Closed ACK | 63 |

where the system reliability is more important than its power consumption.

*4.5. UDP BLIP Implementation.* In contrast to TCP connection, UDP does not have negotiation to ensure correct transmission. Due to it, UDP needs to send less messages. Therefore, it is not reliable (i.e., there is no guarantee that sent UDP messages or packets would reach their destinations at all). Figure 19 shows a scheme of an UDP connection.

Despite its less overhead, the lack of reliability of UDP messages is a drawback, especially in wireless communication with a nonnegligible packet error rate, such as 802.15.4 communications.

Due to this, to increase the reliability in WSNs using UDP, it is necessary to add a communication protocol at the application layer, such as CoAP or an adaptation [48] of IEEE1451 Standard [49].

Table 5: TCP with compression.

| Msg type | Bytes sent |
|---|---|
| Request TCP connection (SYN) | 50 |
| Request TCP connection ACK | 38 |
| Request TCP connection ACK ACK | 30 |
| Data message | 30 + payload (up to 88) |
| ACK | 38 |
| MAC ACK | 5 |
| Close request | 40 |
| Close ACK | 30 |
| Closed | 30 |
| Closed ACK | 30 |

Table 6: UDP without compression.

| Msg Type | Bytes sent |
|---|---|
| Data message | 51 + payload (up to 67) |
| MAC ACK | 5 |

Like in the TCP BLIP implementation, if messages are longer than the available payload, they are fragmented. This fragmentation has the same structure of TCP communications: first fragment adds a header fragmentation (4 bytes) to the original header. The rest of the fragments only have fragmentation header of 5 bytes without any other header.

As TCP protocol, UDP allows communication between devices, whether they are in the same network or not, and it also allows to store information locally, transmitting it only on demand to a user.

*4.5.1. UDP BLIP without Header Compression.* Figure 20 summarizes a UDP frame between two nodes, without multihop, without header compressions or security.

As can be seen, it has less overhead than TCP, but it does not ensure the receiving of the message.

Table 6 summarizes the number of bytes required to send a message between nodes with UDP without compression.

UDP without compression significantly reduces the overhead in comparison with TCP connection.

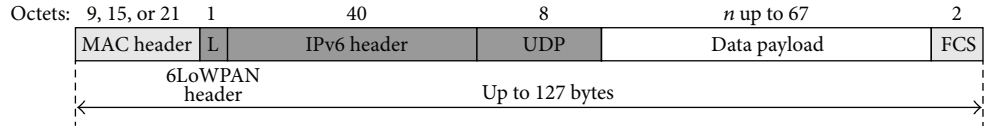



Figure 20: UDP frame without compression.

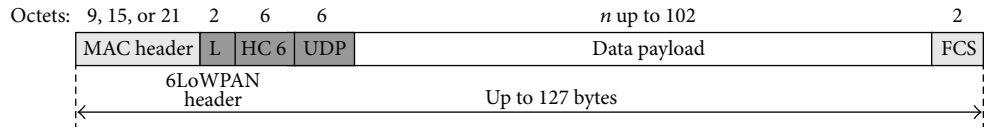



Figure 21: UDP frame with compression.

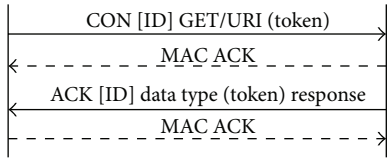



Figure 22: CoAP transmission.

*4.5.2. UDP BLIP with Header Compression.* Figure 21 summarizes a UDP frame between two nodes, without multihop, with header compression and security.

As can be seen, it has very low overhead, especially if short MAC addresses are used.

Table 7 summarizes the number of bytes needed to send a UDP message between nodes with compression.

The overhead of UDP communications is slightly higher than the case of CTP networks, but it allows more flexibility. As it was mentioned before, its main drawback is the lack of reliability.

*4.6. CoAP Implementation.* CoAP implementation add a minimal negotiation to UDP messages with the purpose of increasing the reliability, but maintaining the overhead low. Figure 22 shows the structure of a CoAP negotiation.

The use of Piggy-backed messages prevents the use of a link level ACK, reducing the traffic. Moreover, it uses an implicit socket connection. Due to this, it does not require additional messages to establish or close connections. Like TCP or UDP, CoAP is able to communicate between devices which allow transmittng data on demand.

*4.6.1. CoAP without Header Compression.* The overhead of most common CoAP messages are depicted in Table 8 and Figure 23.

This protocol is a good trade-off between reliability and overhead. It does not increase too much the overhead, but establishes messages to ensure a correct transmission.

*4.6.2. CoAP with Header Compression.* Figure 24 depicted a CoAP frame with header compression. As can be seen, this protocol has a reduced overhead.

Table 7: UDP with compression.

| Msg type | Bytes sent |
|---|---|
| Data message | 16 + payload (up to 102) |
| MAC ACK | 5 |

Table 8: CoAP w/o compression and w/o fragmentation.

| Msg type | Bytes sent |
|---|---|
| GET | 56 + token (0 to 4) + payload (up to 62) |
| POST | 56 + token (0 to 4) + payload (up to 62) |
| MAC ACK | 5 |

Table 9: CoAP w/ compression and w/o fragmentation.

| Msg type | Bytes sent |
|---|---|
| GET | 21 + token (0 to 4) + payload (up to 97) |
| POST | 21 + token (0 to 4) + payload (up to 97) |
| MAC ACK | 5 |

Table 9 sums up the results obtained with the most common CoAP messages used in 802.15.4 WSNs. It is important to consider that CoAP implementation for *TinyOS* is still a work in progress, and not all the methods are currently available.

The reliability and extra cost are similar to CTP protocols or ZigBee, but CoAP provides more flexibility.

In conclusion, CoAP is presented as an interesting compromise between the reliability of TCP and the reduced overhead of UDP. Its overhead is reduced, but nonetheless, it is higher in the case of CTP messages.

## 5. Comparison between Protocols

This section describes a comparison between the different tested protocols.

*5.1. Routing Overheads and Evaluation of Power Consumption.* Routed messages between networks require the use of headers. But headers increase the number of total sent bytes in

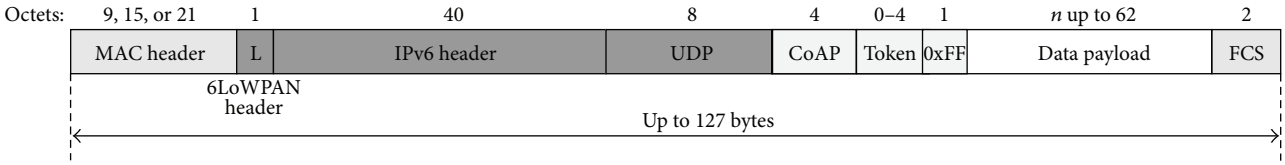



Figure 23: CoAP frame without compression.

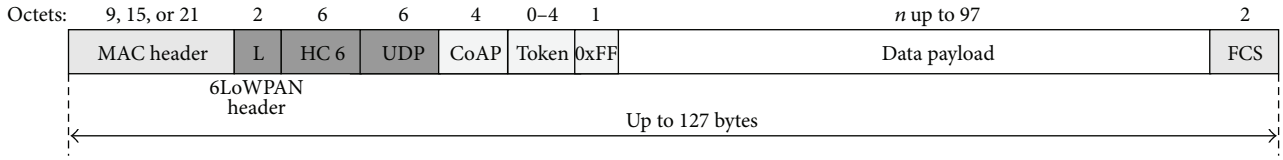



Figure 24: CoAP frame with compression.

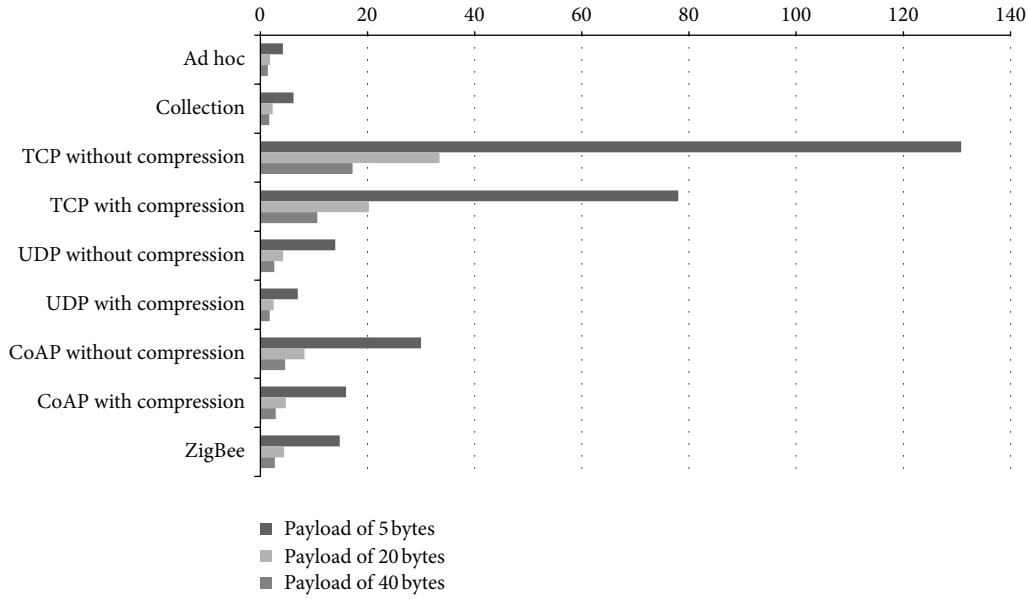



Figure 25: Overhead factor versus payload.

the network. Therefore, it is necessary to search for a trade-off between size of header and reliability. Table 10 shows the number of messages interchanged in order to send a packet of data (5 bytes) and total number of sent bytes (including our payload of 5 bytes).

The ratio between total number of bytes versus payload is the overhead factor. Overhead factor (also shown in Table 10) is related directly to energy consumption. The time of transmission or, in other words, the number of sent bytes are the main influence on energy consumption. Bigger overhead implies bigger energy consumption for the same payload, as energy consumption is the main issue in Wireless Sensor Networks [11] because of limited autonomy in nodes.

To evaluate the energy consumption of sending a packet, it can be estimated as described in [35]:

$$E_{Tx,i} = T_{Tx} \cdot P_{Tx}, \tag{1}$$

where $T_{Tx}$ is time of transmission and $P_{Tx}$ is power of emission. $P_{Tx}$ is estimated as 38 mW [33].

At the same time, there is some energy consumption in the receiving node. It can be estimated as

$$E_{Rx,j} = T_{\text{on}} \cdot P_{Rx}, \tag{2}$$

where $T_{\text{on}}$ is the time that its transceiver is active and $P_{Rx}$ is the power consumed in reception (41 mW [33]). $T_{\text{on}}$ depends on energy saving policies on each protocol.

So, if we only consider energy losses in transmission time, that energy can be modeled as $E_L$:

$$E_L = E_{Tx} + E_{Rx} = \frac{l_s}{r_t} \cdot \left(P_{Tx} + P_{Rx}\right), \tag{3}$$

where $r_t$ is the transmission rate of the platform (i.e., $r_t$ = 250 kb/s) and $l_s$ is the total amount of bits sent in communication process. Based on data of Table 10, Table 11 shows estimation of energy losses in transmission calculated on each protocol.

Table 10: Protocol comparison (payload 5 bytes without fragmentation).

| Protocol | Msg interchanged | Bytes sent | Overhead factor |
|---|---|---|---|
| Ad-hoc with ACK | 2 | 21 | 4.2 |
| Collection | 2 | 31 | 6.2 |
| TCP (w/o compression) | 16 | 654 | 130.8 |
| TCP (w/ compression) | 16 | 390 | 78.0 |
| UDP (w/o compression) | 2 | 70 | 14.0 |
| UDP (w/ compression) | 2 | 35 | 7.0 |
| CoAP (w/o compression) | 4 | 75 | 30.0 |
| CoAP (w/ compression) | 4 | 40 | 16.0 |
| ZigBee | 4 | 74 | 14.8 |

The overhead factor depends on the length of headers and payload. As payload increases, the overhead will decrease. Figure 25 shows the evolution of the overhead factor versus payload size for each tested protocol.

As can be seen, the use of an ad hoc solution has reduced headers, but this system provides almost no services. The use of complete TCP/IP headers only justified sending large messages. UDP and CoAP offer good ratios of overhead with a limited set of services.

To estimate total global consumption, it is necessary to add all the messages of network construction. This amount depends on protocol and topology and they are arranged based on the design time of the application.

*5.2. Latency.* Table 12 depicts the latency between the different evaluated protocols. Latency is obtained measuring the time between starting a request of new information and the time when the node receives this information.

TCP latency is several times higher than the rest of the evaluated protocols. UDP and CoAP have similar latency, with CTP having the lowest latency (approximately, 5 times lower than UDP). Despite of the extra information added by CoAP, it presents a similar latency to UDP.

*5.3. Evaluation of the Comparison.* According to the above results, the main advantages and disadvantages of the evaluated protocols are depicted in Table 13.

For a distributed application, such as the proposed flood level monitoring for waterbirds, 6LoWPAN based protocols are the best trade-off between flexibility and power consumption.

Among the evaluated protocols based on 6LoWPAN, CoAP is the best option for constrained networks. This protocol has advantages such as reliability, but it maintains a low overhead. Moreover, its last draft [50] provides techniques to reduce power consumption like using local proxies or sleeping radio transceiver.

Table 11: Energy consumption (estimated for transmission of 5-byte message without fragmentation).

| Protocol | Energy consumption/(Ws) |
|---|---|
| Ad hoc with ACK | $53 \cdot 10^{-6}$ |
| Collection | $78 \cdot 10^{-6}$ |
| TCP (w/o compression) | $1653 \cdot 10^{-3}$ |
| TCP (w/ compression) | $986 \cdot 10^{-6}$ |
| UDP (w/o compression) | $177 \cdot 10^{-6}$ |
| UDP (w/ compression) | $88 \cdot 10^{-6}$ |
| CoAP (w/o compression) | $190 \cdot 10^{-6}$ |
| CoAP (w/ compression) | $101 \cdot 10^{-6}$ |
| ZigBee | $187 \cdot 10^{-6}$ |

Table 12: Latency comparison between protocols.

| Protocol | Compressed/ms | Not compressed/ms |
|---|---|---|
| TCP | 497 | 395 |
| UDP | 18 | 25 |
| CoAP | 21 | 29 |
| ZigBee | — | 13 |
| CTP | — | 4 |

The use of local storage and direct communication between devices reduces interchanged messages to Base Station and they avoid the maintenance of a central server. The maximum storage of information depends on the memory of platform and the number of nodes. That is, TelosB nodes have an external Flash of 1 Mb, and for flood level estimation, we need 5 bytes in each change (4 bytes with a timestamp + a byte with the estimated flood level). In a worst case scenery with a daily flood level modification, every node can retain up to 7 months of information.

## 6. Conclusions

This paper proposes eSapiens, a distributed WSN for monitoring flood level in several breeding areas of migratory waterbirds. The used data aggregation algorithm allows the use of local storage. Thus it avoids the use of a central server, simplifies the architecture, and reduces the cost. Moreover, the proposed infrastructure requires communication between remote devices. This architecture offers a trade-off between power consumption and reliability.

Focusing on these issues some algorithms have been evaluated. These algorithms can be divided into two families: classic centralized algorithms and fully distributed algorithms.

According to our conclusions, current fully distributed algorithms, such as IPv6 over WSNs, provide flexibility without too much extra cost. For all this, eSapiens has chosen CoAP as best option for its IEEE 802.15.4 WSN devices.

Currently, the authors are developing additional WSNs to spread in other flooded areas where waterbirds live.

Table 13: Comparison of evaluated protocols.

| Protocol | Ad-hoc | ZigBee | CTP | 6LoWPAN (TCP) | 6LoWPAN (UDP) | 6LoWPAN (CoAP) |
|---|---|---|---|---|---|---|
| Mesh network (allows direct communication between nodes) | No | Yes | No | Yes | Yes | Yes |
| Redirection (allows communication between nodes of different networks) | No | No | No | Yes | Yes | Yes |
| Central node required to store gathered information | Yes | No | Yes | No | No | No |
| Communication node → Internet | No | No | No | Yes | Yes | Yes |
| Communication node ← Internet | No | No | No | Yes | Yes | Yes |
| Overhead | Very low | Low | Low | Very high | Low | Medium |
| Latency | Very low | Low | Very low | Very high | Low | Low |
| Energy consumption | Very low | Low | Very low | Very high | Medium | Medium |

## Acknowledgments

This research has been supported by the Consejería de Innovación, Ciencia y Empresa, Junta de Andalucía, Spain, through the projects of excellency ARTICA (Reference no. P07-TIC-02476) and eSapiens (Reference no. TIC-5705) and by the "Cátedra de Telefónica, Inteligencia en la Red," Seville, Spain, through the project TORTUGA.

## References

[1] L. Atzori, A. Iera, and G. Morabito, "The internet of things: a survey," *Computer Networks*, vol. 54, no. 15, pp. 2787–2805, 2010.

[2] A. Jara, M. Zamora, and A. Skarmeta, "Glowbal IP: an adaptive and transparent IPv6 integration in the internet of things," *Mobile Information Systems*, vol. 8, no. 3, pp. 177–197, 2012.

[3] S. Tozlu, M. Senel, W. Mao, and A. Keshavarzian, "Wi-Fi enabled sensors for internet of things: a practical approach," *IEEE Communications Magazine*, vol. 50, pp. 134–1143, 2012.

[4] M. Islam, M. Hassan, G.-W. Lee, and E. -N. Huh, "A survey on virtualization of wireless sensor networks," *Sensors*, vol. 12, no. 2, pp. 2175–2207, 2012.

[5] C.-Y. Chong and S. P. Kumar, "Sensor networks: evolution, opportunities, and challenges," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1247–1256, 2003.

[6] I. Akyildiz, Y. Su, W. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, pp. 393–422, 2002.

[7] L. Ruiz-Garcia, L. Lunadei, P. Barreiro, and I. Robla, "A Review of wireless sensor technologies and applications in agriculture and food industry: state of the art and current trends," *Sensors*, vol. 9, no. 6, pp. 4728–4750, 2009.

[8] *IEEE Std 802.15.4-2003: IEEE Standard for Information Technology—Telecommunications and Information Exchange between Systems—Local and Metropolitan Area Networks Specific Requirements Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LRWPANs)*, 2003.

[9] R. Machado, N. Ansari, G. Wang, and S. Tekinay, "Adaptive density control in heterogeneous wireless sensor networks with and without power management," *IET Communications*, vol. 4, no. 7, pp. 758–767, 2010.

[10] M. Haneef and D. Zhongliang, "Design challenges and comparative analysis of cluster based routing protocols used in wireless sensor networks for improving network life time," *Advances in Information Sciences and Service Sciences*, vol. 4, no. 1, pp. 450–459, 2012.

[11] L. Liu and C. Wu, "A novel power intensity routing in WSN," *International Journal of Digital Content Technology and Its Applications*, vol. 6, no. 1, pp. 178–184, 2012.

[12] D. Ding, L. Fangai, L. Qianqian, and Y. Guangxu, "An improved clustering algorithm based on backup path," *Advances in Information Sciences and Service Sciences*, vol. 4, no. 8, pp. 207–216, 2012.

[13] S. Zihao and L. Shufen, "Security threats and security policy in wireless sensor networks," *Advances in Information Sciences and Service Sciences*, vol. 4, no. 10, pp. 166–173, 2012.

[14] C. Fenhua, "A distributed dynamic pairwise key establishment scheme for wireless sensor networks," *International Journal of Advancements in Computing Technology*, vol. 4, no. 4, pp. 261–267, 2012.

[15] H. Li, L. Pang, and Y. Wang, "A domain-based secure communication scheme with fault-tolerant capacity," *Advances in Information Sciences and Service Sciences*, vol. 4, no. 5, pp. 44–52, 2012.

[16] J. Ko, A. Terzis, S. Dawson-Haggerty, D. Culler, J. Hui, and P. Levis, "Connecting low-power and lossy networks to the internet," *IEEE Communications Magazine*, vol. 49, no. 4, pp. 96–101, 2011.

[17] S. Hong, D. Kim, M. Ha et al., "SNAIL: an IP-based wireless sensor network approach to the Internet of things," *IEEE Wireless Communications*, vol. 17, no. 6, pp. 34–42, 2010.

[18] M. Ha, S. Kim, H. Kim, K. Kwon, N. Giang, and D. Kim, "SNAIL gateway: dual mode wireless access points for WiFi and IP based wireless sensor networks in the internet of things," in *Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC '12)*, pp. 169–173, 2012.

[19] N.-T. Dinh and Y. Kim, "Restful architecture of wireless sensor network for building management system," *KSII Transactions on Internet and Information Systems*, vol. 6, no. 1, pp. 46–63, 2012.

[20] A. Jara, D. Fernndez, P. Lpez, M. Zamora, L. Marin, and A. Skarmeta, "YOAPY: a data aggregation and pre-processing module for enabling continuous healthcare monitoring in

the internet of things," in *Proceedings of the 4th International Workshop on Ambient Assisted Living (IWAAL '12)*, vol. 7657, pp. 248–255, Lecture Notes in Computer Science, 2012.

[21] Y.-J. Wang, Z.-H. Qian, X. Wang, and D.-Y. Sun, "Addressing scheme for internet of things based on ipv6 over low-power wireless personal area network (6LoWPAN)," *Journal of Electronics and Information Technology*, vol. 34, no. 4, pp. 763–769, 2012.

[22] H. Yu and J. He, "Trust-based mutual authentication for bootstrapping in 6LoWPAN," *Journal of Communications*, vol. 7, no. SPL.ISS. 8, pp. 634–642, 2012.

[23] L. Oliveira, J. Rodrigues, A. De Sousa, and J. Lloret, "Denial of service mitigation approach for ipv6-enabled smart object networks," *Concurrency Computation Practice and Experience*, vol. 25, no. 1, pp. 129–142, 2013.

[24] R. Lu, X. Li, J. Wang, and F. Sun, "Research on route protocol and architecture for wsn based on ipv6," *Advanced Materials Research*, vol. 616–618, pp. 2233–2238, 2013.

[25] A. Castellani, M. Rossi, and M. Zorzi, "Back pressure congestion control for CoAP/6LoWPAN networks," *Ad Hoc Networks*, 2013.

[26] Z. Shelby, "Embedded web services," *IEEE Wireless Communications*, vol. 17, no. 6, pp. 52–57.

[27] P. Rajasekaran, R. Janardhan, and R. Chander, "A smarter toll gate based on web of things," in *Proceedings of the IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT '13)*, Bangalore, India, 2013.

[28] C. Sammarco and A. Iera, "Improving service management in the internet of things," *Sensors*, vol. 12, no. 9, pp. 11888–11909, 2012.

[29] M. A. Rendón, A. J. Green, E. Aguilera, and P. Almaraz, "Status, distribution and long-term changes in the waterbird community wintering in Doñana, south-west Spain," *Biological Conservation*, vol. 141, no. 5, pp. 1371–1388, 2008.

[30] M. Guillemain, N. Sadoul, and G. Simon, "European flyway permeability and abmigration in Teal Anas crecca, an analysis based on ringing recoveries," *Ibis*, vol. 147, no. 4, pp. 688–696, 2005.

[31] "Dõnana Biological Station," http://www.ebd.csic.es/website1/Principal.aspx.

[32] J. Mora-Merchan, F. Molina, D. Larios, G. Rodriguez, J. Barbancho, and C. León, "Architecture for environmental data access in WSN," in *Proceedings of the International Conference on Data Communication Networking and the International Conference on Optical Communication Systems (DCNET/OPTICS '11)*, pp. 102–106, Seville, Spain, 2011.

[33] J. Polastre, R. Szewczyk, and D. Culler, "Telos: enabling ultra-low power wireless research," in *Proceedings of the 4th International Symposium on Information Processing in Sensor Networks (IPSN '05)*, vol. 48, pp. 364–369, April 2005.

[34] P. Levis, S. Madden, J. Polastre et al., *TinyOS: An Operating System for Sensor Networks*, Springer, 2005.

[35] D. Larios, J. Barbancho, G. Rodríguez, J. Sevillano, F. Molina, and C. León, "Energy efficient wireless sensor network communications based on computational intelligent data fusion for environmental monitoring," *IET Communications*, vol. 6, no. 14, pp. 2189–2197, 2012.

[36] Z. Alliance, *Zigbee Specification*, 2007.

[37] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis, "Collection tree protocol," in *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems (SenSys '09)*, pp. 1–14, ACM, New York, NY, USA, November 2009.

[38] R. Fonseca, O. Gnawali, K. Jamieson, S. Kim, P. Levis, and A. Woo, "The collection tree protocol," TEP 123, 2006.

[39] J. W. Hui and D. E. Culler, "Extending IP to low-power, wireless personal area networks," *IEEE Internet Computing*, vol. 12, pp. 37–45, 2008.

[40] J. Hui and D. Culler, "IPv6 in low-power wireless networks," *Proceedings of the IEEE*, vol. 98, pp. 1865–11878, 2010.

[41] G. Montenegro, N. Kushalnagar, J. Hui, and D. Culler, *RFC 4944: Transmission of IPv6 Packets over IEEE 802.15.4 Networks*, 2007.

[42] J. Hui and P. Thubert, *RFC 6282: Compression Format for IPv6 Datagrams over IEEE 802.15.4-Based Networks*, 2011.

[43] "Berkeley Low-power IP stack project," http://smote.cs.berkeley.edu:8000/tracenv/wiki/blip.

[44] Z. Shelby, K. Hartke, C. Bormann, and B. Frank, *Draft-Ietf-Core-Coap-13: Constrained Application Protocol (CoAP)*, 2012.

[45] K. Kuladinithi, O. Bergmann, T. Ptsch, M. Becker, and C. Grg, *Implementation of CoAP and Its Application in Transport Logistics*, 2011.

[46] "USB Dongle—IA OEM-DAUB1 2400 by Adaptive Modules Ltd.," http://www.adaptivem2m.com/zigbee-technology/zigbee-usb-dongle.htm.

[47] "Perytons Protocol Analyzers," http://www.perytons.com.

[48] J. E. Higuera and J. Polo, "IEEE 1451 standard in 6LoWPAN sensor networks using a compact physical-layer transducer electronic datasheet," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 8, pp. 2751–2758, 2011.

[49] *IEEE Std 1451.5-2007: IEEE Standard for a Smart Transducer Interface for Sensors and Actuators Wireless Communication Protocols and Transducer Electronic Data Sheet (TEDS) Formats*, 2007.

[50] Z. Shelby, K. Hartke, C. Bormann, and B. Frank, "Constrained application protocol (CoAP)," Tech. Rep., IETF Secretariat, Fremont, Calif, USA, 2013.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 652495, 11 pages
http://dx.doi.org/10.1155/2013/652495

*Research Article*

# Improved Reliable Trust-Based and Energy-Efficient Data Aggregation for Wireless Sensor Networks

**Chen-xu Liu, Yun Liu, and Zhen-jiang Zhang**

*Key Laboratory of Communication & Information Systems, Beijing Municipal Commission of Education, Beijing Jiaotong University, Beijing 100044, China*

Correspondence should be addressed to Chen-xu Liu; hbtslcx@gmail.com

Received 16 December 2012; Revised 2 May 2013; Accepted 9 May 2013

Academic Editor: J. Barbancho

Copyright © 2013 Chen-xu 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.

In wireless sensor networks, secure data aggregation is very important for reducing the quantity of data transmitted and prolonging the lifetime of wireless sensor networks. When wireless sensor networks are deployed in untrusted and hostile environments, their nodes are often compromised, which reduces the security and reliability of the transmitted data. Compromised nodes can inject erroneous data, selectively forward data to an adversary, impersonate legal nodes to join routing paths, and disrupt data transmission during the data-aggregation operation. Previous researchers have relied on reputation system to find compromised nodes and prevent attacks during the data-aggregation operation. In this paper, we propose an improved reliable, trust-based, and energy-efficient data-aggregation protocol for wireless sensor networks. We call the protocol the iRTEDA protocol, and it combines the reputation system, residual energy, link availability, and a recovery mechanism to improve secure data aggregation and ensure that the network is secure, reliable, and energy-efficient. Simulations have shown that the iRTEDA protocol exceeds the performances of other protocols from the perspectives of the accuracy of the data, the reliability of the routing path, the consumption of energy, and the lifetime of secure data aggregation.

## 1. Introduction

A wireless sensor network (WSN) is a highly distributed network composed of sensor nodes with special capabilities, and they are deployed in large-scale area to monitor the environment and collect related information [1–4]. Recently, the use of WSNs has become more popular and promising in various research areas, such as environmental monitoring [4–7], military target tracking [8, 9], natural disaster relief [10], and health monitoring [11]. WSNs have become one of the key modern information technologies, and the technology is changing people's lives and the way people interact with the physical world.

Generally, sensor nodes are deployed so densely that the sensing scales of neighboring nodes often have serious overlaps, resulting in redundant sensing of data and unnecessary expense in correlating the same data. The amount of data received by the base station is much greater than necessary. Data-aggregation technology [12, 13] is used to process the raw data, eliminate redundant or superfluous data, and save energy by ensuring that the network works efficiently. Obviously, the security of WSNs must be taken into consideration when they are deployed in insecure and hostile environments, so secure data-aggregation gradually is becoming a key technology. There are several approaches to keep the data-aggregation process secure [14], such as cryptography, key management, authentication mechanisms, privacy-preserving technologies, and reputation-based mechanisms. To the best of our knowledge, cryptographic-based technologies generally are utilized to keep the network free from the attacks during the data-aggregation process. However, the security of data-aggregation cannot be guaranteed by pure cryptography, because cryptographic-based technologies cannot provide adequate defense against node capture attack that results in the occurrence of compromised nodes. Compromised nodes allow easy access to the cryptographic keys that are used for cryptographic-based data-aggregation, and, when such access occurs, it cannot be detected.

To overcome the shortcoming of cryptographic-based secure data-aggregation, reputation and trust systems are

being developed to complement existing technologies for monitoring network activities and events. Reputation and trust systems are utilized to detect, collect, process, and disseminate feedback concerning the sensors' recent behaviors and to assess their trustworthiness for specific applications. The goals of using such systems are to defend against node capture attack that results in the occurrence of compromised nodes, identified nodes that have been compromised, and exclude them from further participation in data-aggregation. The trustworthiness of sensors is evaluated based on their various activities, including data collection, data transmission, aggregator selection, and routing path selection. The reputation of each node refers to the expectation of neighboring nodes concerning a node's behavior based on their observations of its past actions. Thus, trust and reputation in a WSN often are mentioned together. Therefore, a node's expectation will affect its choices and activities. Trust of a node generally is defined as the expected value of that node's reputation.

In this paper, we propose an improved reliable, trust-based, and energy-efficient data-aggregation protocol for WSNs. We call the protocol the iRTEDA protocol, and simulations have shown that the iRTEDA protocol exceeds the performances of other protocols from the perspectives of the accuracy of the data, the reliability of the routing path, the consumption of energy, and the lifetime of secure data-aggregation. It improves Suat Ozdemir's RDAT protocol [15, 16] and provides information concerning residual energy and link availability to evaluate the trustworthiness and reliability of sensor nodes based on the observations of neighboring nodes. The iRTEDA protocol uses a reputation function that is based on a Beta distribution to assess the reputation and trustworthiness of nodes in performing their tasks. Information concerning the residual energy and link availability of the nodes was introduced to help the network reselect the aggregators and improve the robustness of the selected routing path, when the aggregators are judged as compromised nodes. In addition, a recovery mechanism is proposed in the iRTEDA protocol to keep nodes from becoming isolated, improve the structure of the clusters, and reduce the energy consumed for data transmission after the compromised nodes have been excluded.

The organization of the paper is listed in the following. Section 2 introduces previous works about trust-based systems in wireless sensor networks. Descriptions of system model, Beta reputation system, and problem statement are provided in Section 3. Section 4 proposes improved reliable, trust-based, and energy-efficient data-aggregation. Performance evaluation and analysis are described in Section 5. Section 6 is the conclusions.

## 2. Related Work

In this section, previous work is presented to introduce the development of trust-based system for WSNs. In accordance with the characteristics of WSNs, centralized trust-based systems are not feasible because there is a centralized trusted center that controls the systems in the network. As the domain grows, one centralized trusted center can reduce the scalability and expandability of the network. Thus, decentralized trust-based systems are being developed and used in WSNs. Recently, trust-based systems that are used in WSNs have been divided into five types according to their different applications, that is, generic, routing, access, location, and aggregation. In [17, 18], Boukerche and Ren proposed a trust computation and management system (TOMS) that develops a generic trust model to evaluate all the actions of the nodes, including making credential assignments, managing the trust values of the nodes, updating the keys, and judging the actions of the nodes to decide their access rights. Similar to Boukerche et al.'s scheme, RFSN [19, 20] first combines first-hand and second-hand information to compute the reputation and trust values of the nodes and then develops a general, trust-based model. In [21], Shaikh et al. proposed a hybrid trust management architecture for clustered WSNs, which they called GTMS. Their scheme divided the evaluation of the reputation values into three levels, that is, sensor node, cluster head, and base station. The innovation of the GTMS scheme was that it eliminated the use of one single reputation value to evaluate the actions of the nodes by dividing the evaluations of the reputation and trust values into three levels according to the architecture of the network. Michiardi et al. [22] and Srinivasan et al. [23] developed special trust-based models, referred to as CORE and DRBTS, to prevent nodes from exhibiting selfish behavior and to exclude compromised nodes just for routing and location, respectively. Compared with the models above, RDAT protocol [15, 16] introduced multiple functions to compute the reputation and trust values based on three specific aspects, that is, sensing, aggregating, and routing. The protocol combines the evaluation of these three aspects to evaluate the trustworthiness of nodes. Our proposed scheme is based on the RDAT protocol, but it has significant and beneficial advantages over that protocol because it considers energy efficiency, link availability, reselection of aggregators, and recovery mechanisms.

## 3. Preliminary

*3.1. System Model.* In this paper, the hierarchal cluster architecture was used to construct WSNs composed of sensor nodes that were densely deployed in clusters. It was assumed that the operation of each cluster was relatively independent and that very few, if any, nonoverlapping areas would be sensed between the clusters. Hence, the reputation and trust of the sensor nodes are evaluated only by the nodes in their own cluster. A watchdog mechanism is used to monitor and detect the actions of target nodes, and those actions are characterized as cooperative or noncooperative; judgments are made concerning whether the action of the nodes are right or not. Thus, the reputation and trust system is responsible for maintaining the reputation and trust of a node, and this duty includes many tasks. The system updates reputation information based on new observations made by the watchdog mechanism and creates new evaluations of the trustworthiness of the nodes.

Each cluster has a cluster head called an aggregator, which is in charge of a certain number of sensor nodes

and has the capability of performing the data-aggregation operation. Aggregators are utilized to process received data from children nodes and transmit the aggregated results to base station. Sensor nodes, with the exception of the aggregator, sense the data, monitor the activities of other nodes, exchange observations with neighboring nodes, evaluate the trustworthiness of the nodes, and transmit data and observations to the aggregator. When the sensor nodes report their readings to the aggregators, the messages are encrypted and decrypted by pairwise keys that are generated and that are possessed only by the two communicating parties through their negotiation. In addition, the aggregators, namely, the cluster heads that are in the hierarchal cluster architecture, are not maintained for a long term and dynamically changing, because, when the system is running, adversaries can follow the aggregators more and more closely, so security problems associated with those aggregators will increase. In addition, the energy consumption of those aggregators increases rapidly and significantly when they remain unchanged for extended periods. Thus, to solve the security and energy problem, sensor nodes in the cluster must be reselected dynamically as aggregators at intervals.

*3.2. Beta Reputation System.* Reputation and trust systems are used extensively in various domains for WSNs, such as gathering, sharing, and modeling information, routing, decision making, and dissemination, which help the network identify malicious and compromised nodes and eliminate their adverse effects. Sensor nodes use the watchdog mechanism to monitor the behaviors of neighboring nodes and to decide whether those behaviors are acceptable or not. Then, the information obtained about the activities of the sensor nodes is used to evaluate the trustworthiness of the nodes and to decide whether possible compromised nodes exist.

In [19, 20], the researchers proposed that the use of a binary rating for the behaviors of sensor nodes was adequate for the data-aggregation operation, because their activities were evaluated as only good or bad. Therefore, a Bayesian formulation, called the Beta reputation system, was introduced to represent and update the trust of sensor nodes. This system can be expressed as

$$P(\varphi \mid \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\cdot\Gamma(\beta)}\varphi^{\alpha-1}(1-\varphi)^{\beta-1}, \quad \text{where} \quad 0 \le \varphi \le 1, \ \alpha > 0, \ \beta > 0. \tag{1}$$

The parameter $\varphi$ is the probability that binary events will occur, and $P(\varphi \mid \alpha, \beta)$ represents the distributions determined by the parameters $\alpha$, $\beta$, and the gamma function, $\Gamma$. The probability expectation of this distribution $P(\varphi \mid \alpha, \beta)$ is given in [24] as

$$E(\varphi) = \frac{\alpha}{\alpha+\beta}. \tag{2}$$

The function is based on the theory of statistics, and it scales the values in the interval with any condition, even for any peaking time.

Now, we present a description of how the Beta reputation system works in WSNs. Consider that there are two nodes, $i$ and $j$, in the network. Each node detects and monitors the action of the other node, and they use their observations to update their reputation and trust. For simplicity, it is assumed that node $j$ is the target node and that node $i$ is the detecting node. Since the Beta reputation system is used as the binary rating function for the behaviors of the sensor nodes, detecting node $i$'s observation of the action of target node $j$ is divided into two categories, that is, correct and false. Let $m$ be the number of correct behaviors of target node $j$, and let $n$ be the number of false behaviors. Then, the parameters $\alpha$, $\beta$ in the Beta function are set as follows:

$$\alpha = m + 1, \quad \beta = n + 1, \quad m, n \ge 0. \tag{3}$$

Therefore, the variable $\varphi$ is redefined as correct observation the target node, while $P(\varphi \mid \alpha, \beta)$ is the probability that $\varphi$ has a special value. The probability expectation $E(\varphi)$ may be expressed as the most likely value of the variable $\varphi$. Then, we discuss the values of the probability expectation $E(\varphi)$, including three possible conditions as follows:

$$\begin{aligned} &(1)\ \alpha = \beta, \quad E(\varphi) = \frac{\alpha}{\alpha+\beta} = 0.5, \\ &(2)\ \alpha < \beta, \quad E(\varphi) = \frac{\alpha}{\alpha+\beta} = \frac{1}{1+\beta/\alpha} < 0.5, \\ &(3)\ \alpha > \beta, \quad E(\varphi) = \frac{\alpha}{\alpha+\beta} = \frac{1}{1+\beta/\alpha} > 0.5. \end{aligned} \tag{4}$$

(1) When $\alpha = \beta$, $E(\varphi) = 0.5$. The number of correct behaviors of the target node is equal to the number of false behaviors. The result shows that the probability that target node is legal is the same as the probability that it is not; that is, it has been compromised.

(2) When $\alpha < \beta$, $E(\varphi) < 0.5$. The number of correct behaviors of the target node is less than the number of false behaviors. This result indicates that the probability that target node has been compromised is greater than the probability that it is a legal node.

(3) When $\alpha > \beta$, $E(\varphi) > 0.5$. The result indicates the opposite of the condition in (2), where the condition was $\alpha < \beta$.

In addition to those described above, normal sensor nodes and aggregators will store their observations of neighboring nodes in a table and exchange the table with neighboring nodes. In this way, the nodes will combine first-hand and second-hand information to evaluate reputation and trust. Then, the network can use the trust to judge whether the aggregators and normal sensor nodes are compromised.

*3.3. Problem Statement.* The RDAT protocol first was proposed in [15, 16] to achieve the operation of reliable data-aggregation and transmission based on the reputation system. It combines the reputation functions of sensing, aggregating, and routing to achieve the security of data-aggregation

operation. However, security issues in RDAT are focused only on reputation and trust. During data transmission, the aggregators select the routing path and forward the aggregated results to base station along that path. However, security issues can occur when the same secure routing paths are used repeatedly. Some nodes in those paths will consume excessive energy, and others will consume little energy, resulting in the disruption of the balance of energy consumption and shortening of the lifetime of the network. In addition, when aggregators or normal sensors are judged as compromised nodes by the reputation and trust system, a recovery operation must be used for those nodes. Therefore, the focal points of our work are to incorporate in the reputation and trust system with nodes' residual energy, link availability between nodes, and recovery mechanism to achieve a more secure, reliable, and energy-efficient data-aggregation operation.

## 4. Improved Reliable Trust-Based and Energy-Efficient Secure Data-Aggregation

Our proposed iRTEDA protocol is developing from RDAT and improving the problem that occurs in protocol RDAT. The basic idea of iRTEDA protocol is to combine reputation and trust system, nodes' energy consumption, robustness of selecting routing path, and a recovery mechanism to decrease energy consumption and keep data-aggregation more reliable, when sensor nodes are compromised. Let us introduce iRTEDA protocol in detail as follows.

*4.1. Reputation and Trust Computation.* In the proposed iRTEDA protocol, the reputation and trust system defines the tasks of the nodes as sensing, aggregating and routing, and evaluating the trustworthiness of each task based on first-hand and second-hand information. It is assumed that there is a cluster that is composed of a group of sensor nodes, $N_k$ $(k = 1, 2, \ldots, n)$. Node $N_i$ is assigned to evaluate the reputation and trust for the tasks of its neighboring node $N_i$ in the cluster. It monitors the actions of node $N_j$, decides whether the actions are good or bad, and records the judgments in an observation table. The information recorded in the table by node $N_j$ is the first-hand information for node $N_j$. First-hand information, as the name suggests, is the information in the reputation table about neighboring node $N_j$ that is observed and recorded by node $N_i$. At the same time, the other nodes in the cluster also have recorded their judgments of node $N_j$'s tasks in their own observation tables.

Then, the information regarding node $N_j$ in the observation tables is exchanged between nodes. Node $N_i$ receives the observation tables from other nodes in the cluster and obtains neighboring nodes' observations of node $N_j$. The information about node $N_j$ in the other nodes' observation tables is second-hand information for node $N_i$. The observation tables are exchanged in two ways, that is, on demand and broadcasting. The on demand procedure takes place when node $N_i$ requests that neighboring nodes exchange observation tables with it and uses the observations to evaluate the reputation of $N_j$. Broadcasting refers to the case in which the nodes broadcast their observation tables for a certain time period. For the sake of simplicity, in this paper, we used broadcasting to exchange the reputation tables.

Calculating trust for the actions of one node based on first-hand and second-hand information is introduced in the following. The task of sensing is taken as an example for the explanation of the reputation and trust evaluation. When the node $N_i$ is detecting and monitoring the sensing task of node $N_j$, the numbers of behaviors and misbehaviors of node $N_j$, as judged by node $N_i$, are recorded as $\alpha_{i,j}^{\text{new}}$ and $\beta_{i,j}^{\text{new}}$, respectively. Reputation value and trust for the sensing task of node $N_j$ are represented by $R_{i,j}^{\text{sensing}}$ and $T_{i,j}^{\text{sensing}}$, respectively. The formula of reputation and trust calculation is

$$\begin{aligned} R_{i,j}^{\text{sensing}} &= \text{Beta}\left(\alpha_{i,j}^{\text{new}} + 1, \beta_{i,j}^{\text{new}} + 1\right) \\ &= \frac{\Gamma\left(\alpha_{i,j}^{\text{new}} + 1 + \beta_{i,j}^{\text{new}} + 1\right)}{\Gamma\left(\alpha_{i,j}^{\text{new}} + 1\right) \cdot \Gamma\left(\beta_{i,j}^{\text{new}} + 1\right)} \varphi^{(\alpha_{i,j}^{\text{new}}+1)-1} (1-\varphi)^{(\beta_{i,j}^{\text{new}}+1)-1}, \end{aligned} \tag{5}$$

$$T_{i,j}^{\text{sensing}} = E\left(R_{i,j}^{\text{sensing}}\right) = \frac{\alpha_{i,j}^{\text{new}} + 1}{\alpha_{i,j}^{\text{new}} + \beta_{i,j}^{\text{new}} + 2}. \tag{6}$$

The parameters $\alpha_{i,j}^{\text{new}}$ and $\beta_{i,j}^{\text{new}}$ are the new numbers of correct and false actions of node $N_j$ calculated by first-hand and second-hand information. The process of integrating first-hand and second-hand information into an overall reputation was proposed in [16, 19, 20], and it is shown below:

$$\begin{aligned} \alpha_{i,j}^{\text{new}} &= p * \alpha_{i,j}^{\text{now}} + m_{i,j} + \sum_{k \in N} R\left(m_{k,j}\right), \\ \beta_{i,j}^{\text{new}} &= p * \beta_{i,j}^{\text{now}} + n_{i,j} + \sum_{k \in N} R\left(n_{k,j}\right), \end{aligned} \tag{7}$$

where parameters $\alpha_{i,j}^{\text{now}}$ and $\beta_{i,j}^{\text{now}}$ are the last observations about correct and bad actions of $N_j$ in the observation table, respectively, and $m_{i,j}$ and $n_{i,j}$ represent the number of recent observations of correct and bad sensing actions, respectively. Old feedback cannot always work effectively for the new reputation and trust rating during the operation, so the old observations are less important than most recent observations, and they will be eliminated gradually. Therefore, the elimination parameter $p < 1$ is introduced to achieve the characteristic of last observations described above. In addition, second-hand information is exchanged between node $N_i$ and $N_k$ $(k = 1, 2, \ldots, n)$, and the observed numbers of correct and bad behaviors are expressed as $R(m_{k,j})$ and $R(n_{k,j})$ $(k = 1, 2, \ldots, n)$, respectively. Second-hand information for correct and bad actions is defined in [20] and is shown below:

$$\begin{aligned} R\left(m_{k,j}\right) &= \frac{2 * \alpha_{i,k}^{\text{now}} * m_{k,j}}{\left(\beta_{i,k}^{\text{now}} + 2\right) * \left(m_{k,j} + n_{k,j} + 2\right) * \left(2 * \alpha_{i,k}^{\text{now}}\right)}, \\ R\left(n_{k,j}\right) &= \frac{2 * \beta_{i,k}^{\text{now}} * n_{k,j}}{\left(\beta_{i,k}^{\text{now}} + 2\right) * \left(m_{k,j} + n_{k,j} + 2\right) * \left(2 * \alpha_{i,k}^{\text{now}}\right)}. \end{aligned} \tag{8}$$

Then, we can compute node $N_j$'s trust $T_{i,j}^{\text{sensing}}$ about the sensing task using formula (6). According to the value of node $N_j$'s trust, node $N_i$ can use formula (4) to make the judgment about whether node $N_j$ is compromised. Node $N_i$ will transmit the judgment to the aggregator in the cluster. The aggregator receives the judgment about Node $N_j$ from other nodes in the cluster and decides whether node $N_j$ is compromised. If node $N_j$ is a compromised node, the aggregator will exclude node $N_j$ from the network and use the measures to eliminate compromised node $N_j$'s negative effects.

*4.2. Residual Energy and Link Availability.* The rating system above is based on reputation and trust, and it only pays attention to the reputation of sensing, routing, and aggregating actions; it cannot guarantee that the energy consumed by the sensor nodes is reasonable. Nodes with high reputations, which are selected for routing paths, will be utilized repeatedly, and they will take much more energy than those with low reputations. Obviously, selecting nodes for routing paths based only on their reputation and trust system ignores the nodes' energy usage. That will result in the repeated selection of nodes with high reputation, and these nodes will be overused for forwarding the data, disrupting the equilibrium of nodes' energy consumption. The nodes in the paths use up their energy rapidly, and others consume less energy, and this leads to significant differences between the energy consumptions of the nodes and decreases the lifetime of the network. Therefore, the residual energy and link availability between nodes must be taken into consideration and combined with the reputation and trust system to keep the network secure, reliable, and energy-efficient.

Sensor nodes in the routing path, in addition to their sensing task, must relay data towards the aggregators and base station. Sensor nodes and aggregators record their residual energy and exchange this information with neighboring nodes. Then, the nodes can use the energy information to determine the link availability between them. Thus, the aggregators also can obtain the status of the energy levels of the nodes in their cluster and identify the best nodes for forwarding data to the base station. Energy tables and reputation tables are exchanged simultaneously, and, when they are received, there are two advantages in the data-aggregation operation; that is, (1) in each cluster, nodes can select the best aggregator of the cluster in a certain time period according to reputation and energy information. The selection requires the consensus of all the nodes' points and is determined based on the equilibrium of reputation and energy, and (2) after the aggregation is completed in each cluster, then results of the aggregation are transmitted to the base station along the routing path. Combining reputation and energy information could confirm link availability between nodes and identify a better path from each aggregator to the base station. It is also useful to select the best path from the sensor nodes to their clusters' aggregators.

The parameter $N_{ET}$ is defined below to represent a node's combined information of reputation and energy. It takes both reputation and energy into consideration and is used to judge whether a node's trust and energy are sufficient for it to be an eligible aggregator or a routing node. Calculating the parameter $N_{ET}$ helps the system identify the best aggregator and routing nodes in each cluster:

$$N_{ET} = \frac{E \times T}{\text{Init} - E \times \text{Init} - T} \quad E > \theta_{Eg}^{Ag},\ T > \theta_{T}^{Ag}, \tag{9}$$

where $N_{ET}$ is the parameter that combines reputation and energy for each node.

Then, we define the link availability $L_{AB}$ between node $A$ and $B$ in the following:

$$L_{AB} = \frac{\text{Init} - T_{AB} \cdot \text{Init} - E_B}{T_{AB} \cdot E_B} \quad E_B > \theta_{Eg}^{\text{relay-node}},\ T_{AB} > \theta_{T}^{\text{link}}, \tag{10}$$

where $T_{AB}$ represents the reputation of node $B$ evaluated by node $A$, $E_B$ is the residual energy of node $B$, $\theta_{Eg}^{\text{relay-node}}$ is the minimum value of the residual energy of node $B$ to transmit the data, and $\theta_{T}^{\text{link}}$ is the minimum acceptable value for trust of link, which is node $B$'s reputation as evaluated by node $A$.

Assume that there are $p$ two-hop nodes with link availability in the routing path between $i$ and $j$, each with the middle nodes $s_l$ $(1 < l < p)$, and denote $L(i, j)$ as the link availability of nodes between $i$ and $j$. The whole link availability between $i$ and $j$ is denoted as Link$(i, j)$:

$$\text{Link}(i, j) = L(i, j) + \sum_{l=1}^{p} \min\left(L(i, s_l), L(s_l, j)\right). \tag{11}$$

In addition, each node $i$ assigns a weight to all its linked neighbors that are closer to the sink than it is. We denote $U(i)$ as the set of node $i$'s neighbors' link availability. And recall Link$(i, j)$ is the link availability between node $i$ and $j$. We assign link availability to each node $j$ in set $U(i)$ as

$$w_j = \frac{\text{Link}(i, j)^{\alpha}}{\sum_{m \in U(i)} \text{Link}(i, m)^{\alpha}}, \tag{12}$$

where $w_j$ is the link availability that $i$ chooses $j$ as the forwarder. When $\alpha = 0$, all nodes in $U(i)$ are given equal priority regardless of link resilience. When $\alpha$ is positive, the more resilient links are given higher priority. When $\alpha$ approaches infinity, only the more resilient links are chosen for routing. An intermediate value can be used to achieve a good balance between security and lifetime. Therefore, to choose a neighbor on the secure and reliable path to the base station, we use $U(i)$ as the set of link availability of node $i$ to decide whether the nodes are linked reliably to be on the routing path. If there is a need for balance between security and lifetime, a weight $w_j$ can be set to each of its link neighbors $j$ in the set $U(i)$. In order to simplify the analysis of our protocol in the following, we will set the parameter $\alpha = 0$ and the weight $w_j = 1$.

The network uses residual energy and link availability to reselect the aggregators and routing paths every $T_x$ time period. It is possible that some links between nodes will fail during the time period. So this operation is done each $T_x$ time period, and the new information of new aggregators and new upstream nodes in the cluster is shared and recorded.

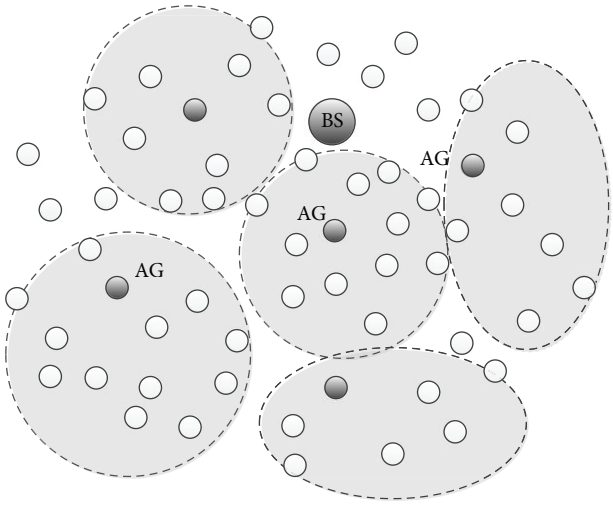



Figure 1: Cluster structure for a WSN.

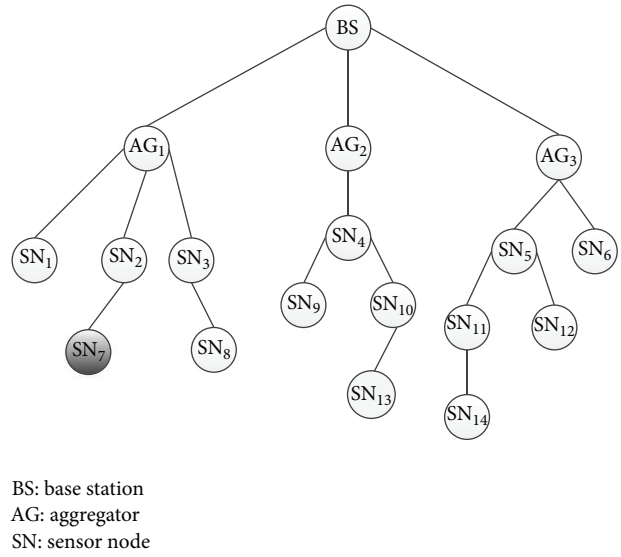


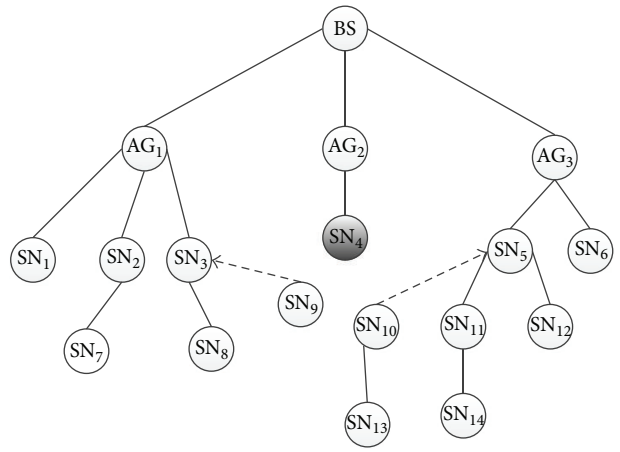

Figure 2: Compromised node is the leaf node.

*4.3. Recovery Mechanism.* After the computation of reputation and trust, the aggregators could determine and judge whether neighboring nodes are compromised or not. Once the nodes are judged as compromised nodes, those nodes will be excluded from the network. Their father nodes no longer receive data from the compromised nodes, and their children nodes do not forward data to them. After compromised nodes are marked and excluded, the structure of the network will be adjusted to guarantee that compromised nodes' children nodes find the right father nodes to keep these children nodes from becoming isolated nodes. Thus, a recovery mechanism is required to keep compromised nodes' children nodes from becoming isolated. The structure of a cluster in a WSN is shown in Figure 1. The entire network consists of the base station, aggregators, and the sensor nodes. Each aggregator, namely, the cluster head, manages all the sensor nodes in the cluster. In the following, the recovery mechanism is shown for different situations.

*(1) Leaf Node Is Compromised.* When the aggregator makes an evaluation that a leaf node is a compromised node, the aggregator will send messages to the sensor nodes in the cluster instructing them to ignore the compromised node's messages and to exclude it from the cluster. After that, the aggregator also will send warning messages to the aggregators in the neighboring clusters. In Figure 2, aggregator $AG_1$ makes a judgment that leaf node $SN_7$ has been compromised and excluded from the cluster. Then, aggregator $AG_1$ sends warning messages to the other aggregators $AG_2$ and $AG_3$. After those aggregators receive the warning messages, they will take precautions against compromised node $SN_7$ and ignore its messages.

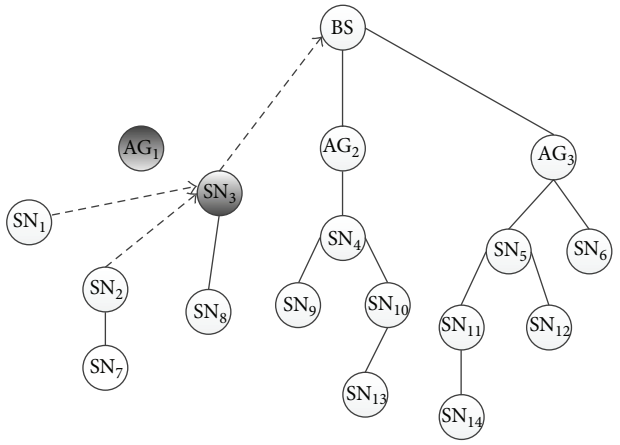
*(2) Intermediate Node Is Compromised.* When an intermediate node is judged to be a compromised node, the aggregator excludes the compromised node from this cluster and sends warning messages to this node's children nodes. Its children nodes will become isolated and tend to find candidate father nodes. If a child node cannot find its candidate father node, it will check its ID list and look for uncle nodes that were stored in the invitation message received from the cluster heads when the cluster structure was built. The node will send a request message to its uncle nodes asking to join their clusters and wait for their responses. When children nodes are authenticated legally by uncle nodes, a response will be transmitted to the children nodes admitting them to the cluster. This operation makes children nodes reselect a father node, reduces the hops to the base station, and reduces the energy required for communication during the transmission of data. The operation is shown in Figure 3. Aggregator $AG_2$ makes a judgment that intermediate node $SN_4$ has been compromised. Then node $SN_4$'s children nodes $SN_9$ and $SN_{10}$ send request messages to their uncle nodes $SN_3$ and $SN_5$. When the legality of $SN_9$ and $SN_{10}$ has been guaranteed by $SN_3$ and $SN_5$, they will join the cluster managed by $SN_3$ and $SN_5$.

*(3) Aggregator Is Compromised.* When the network determines that the aggregator is compromised, the aggregator will be reselected by nodes in the cluster. Reselecting the aggregator must involve evaluating the nodes' reputations and residual energies simultaneously and balancing the two aspects. Figure 4 shows that aggregator $AG_1$ is judged as a compromised node and excluded from the cluster. Neighboring clusters will be informed of the decision about compromised aggregator $AG_1$. Then, all the nodes in this cluster will exchange reputation and residual energy tables and combine reputation and residual energy to decide which node is most appropriate to be the aggregator in next time period. In Figure 4, sensor node $SN_3$ is selected as the new aggregator, and it collects the messages from other




BS: base station
AG: aggregator
SN: sensor node

Figure 3: Compromised node is intermediate node.

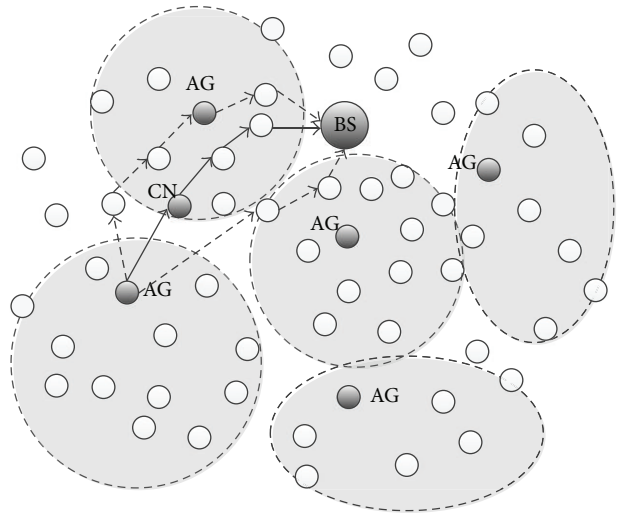



BS: base station
AG: aggregator
SN: sensor node

Figure 4: Compromised node is aggregator.

sensor nodes in the cluster and executes the data-aggregation operation.

*(4) Node in the Routing Path Is Compromised*. When node CN in the routing path is compromised, the subsequent operation is shown in Figure 5. The solid line is the original routing path, and the broken lines are the lines that may possibly be reselected to join the routing paths. In order to prevent forwarding and misdirecting attacks launched by a compromised node, the recovery mechanism makes a reliable multipath for data transmission after the compromised nodes are identified. It ensures that the nodes reselect reliable nodes, from the perspectives of trust and energy path, to transmit and route the data to the base station. The reselection of the routing path is based on trust, residual energy, and link




BS: base station
AG: aggregator
CN: cheating node

Figure 5: Compromised node is the node in the routing path.

availability, which keeps the quantity of reselected paths and excludes compromised nodes from the routing paths. To perform the recovery mechanism, each data aggregator must determine the possible paths to the base station and decide which path will be reselected to transmit the aggregated data.

## 5. Performance Evaluation

To evaluate the performance of the proposed iRTEDA protocol, we compared it with the RDAT protocol for different aspects, including average reputation values, data accuracy, routing path reliability, energy consumption, and lifetime. Both protocols were developed by using the Tiny OS 2.0 simulator (TOSSIM) and its variant, PowerTOSSIM. Energy is a crucial constraint in wireless sensor networks. It is an important issue in performance evaluation. So we use TinyOS simulator PowerTOSSIM, a power modeling extension to TOSSIM. PowerTOSSIM accurately models power consumed by TinyOS applications. One hundred sensor nodes were deployed in the network that had an area of 300 m × 300 m. The sensor nodes were distributed in different areas and organized based on the structure of the cluster. The base station was located in the central area, and each cluster had an aggregating node for the cluster header. A fixed number of nodes were assumed to be compromised. They transmitted false data to the aggregators, which aggregated the false data into the transmitted data and selectively forwarded it to base station. In addition, link failure and packet loss were set as fixed values to better represent practical situations in a WSN.

*5.1. Comparison of Reputation Value*. In the section, we used average reputation values to assess reputation and trust system based on the statistics associated with the running time period of the network. There were 30% of compromised

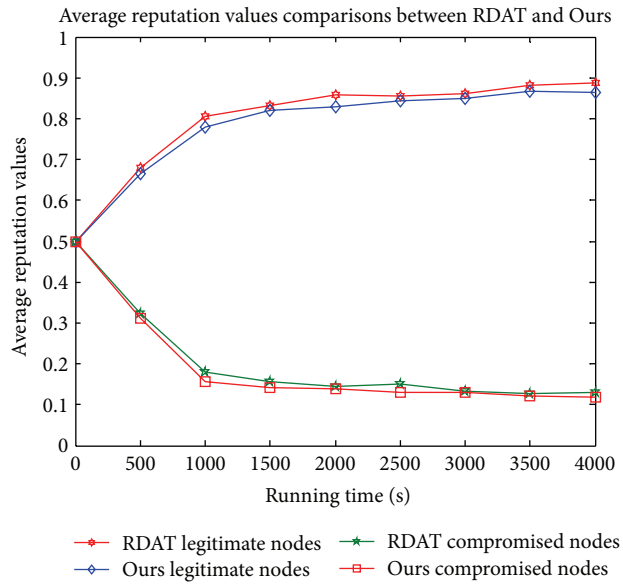



FIGURE 6: Comparison of average reputation values of $R_{i,j}^{\text{sensing}}$ between RDAT and iRTEDA protocols.

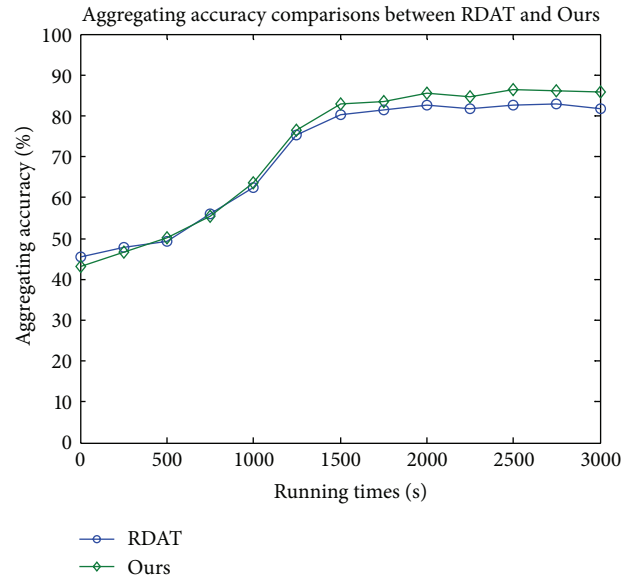



FIGURE 7: Comparison of aggregating accuracy of the RDAT and iRTEDA protocols when 30% of the nodes were compromised.

nodes in the network to provide sensing misbehaviors. So nodes should monitor and detect the misbehaviors of their neighboring nodes and compute those nodes' average reputation values. Figure 6 shows the average reputation values $R_{i,j}^{\text{sensing}}$ of the compromised and the legitimate nodes' sensing actions. We used those values to evaluate the nodes' trustworthiness. From Figure 6, it was concluded that the variation tendency of average reputation value for legitimate and compromised nodes in both protocols was almost the same because average reputation values $R_{i,j}^{\text{sensing}}$ were assessed by combining first-hand and second-hand information in both protocols. The compromised nodes cannot hide their sensing misbehaviors to get higher average reputation values. So, the compromised nodes' average reputation value was reduced in stages until the ultimate value was reached. However, as Figure 6 shows, the average reputation values of the legitimate nodes in iRTEDA protocol always were lower than those of the RDAT protocol. This occurred because iRTEDA protocol evaluated the nodes' average reputation values based on the nodes' behaviors and misbehaviors as well as taking residual energy and link availability into consideration. The evaluation standard will result in decreasing the security and average reputation values of nodes in the routing path, but it leads to better energy efficiency, stability, and reliability.

*5.2. Comparison of Data-Aggregation.* The accuracy of data-aggregation always has been regarded as a crucial criterion for aggregation performance in the network. It is denoted in this section the ratio of the sum of the data from legal nodes collected by the base station to the sum of all of the data collected by the base station.

When compromised nodes are normal nodes, the data collected by those nodes are illegal, and they are not included in the sum of the data collected from legal nodes by the base station. In the cluster, the data sensed by normal nodes are transmitted to the aggregators, and they are aggregated. When the aggregators are compromised, the results they producte become unbelievable. Figure 7 shows the increasing trend of aggregating accuracy for the RDAT protocol and our proposed protocol. Initially, the aggregating accuracies of both protocols increase at slow rates because the systems were just beginning to work, and the compromised nodes' misbehaviors were not detected yet, meaning they were not excluded from the network. When the reputation and trust system is running, the aggregating accuracy increases sharply when the running time is in the range of 500 to 1500 s. The aggregating accuracy of the iRTEDA protocol increased to almost 83%, whereas the RDAT protocol increased to about 80%. This is because the reputation system began to work and excluded the compromised nodes from the network. The aggregating results received by the base station belonged to the legal nodes to a greater extent as time passed. When the running time of about 1500 s was reached, the rate of growth of the aggregating accuracy slowed significantly. However, the final aggregating accuracy for the iRTEDA protocol was greater than that of the RDAT protocol. This occurred because the RDAT protocol does not pay enough attention to the energy and link availability issues, and this results in a greater death rate of nodes, meaning that the base station received less data from legal nodes. In addition, the recovery mechanism ensures that isolated legal nodes can rejoin the network and find a high-reputation parent node. This mechanism will increase the number of legal nodes that finally are received by the base station in the iRTEDA protocol.

*5.3. Comparison of Routing Path Reliability.* The routing-path reliability metric is defined as the ratio of the sums

of the amounts of sensing data and the aggregating results received by the aggregators and the base station to the total amount of transmitted data. The aggregators eliminate a lot of redundant data. Thus, the reliability of the routing-path should be compared in two ways, that is, normal nodes to aggregators and aggregators to the base station. The issues that may influence the performance of routing-path reliability are including link availability and residual energy. As seen in Figure 8, with the network running, the percentage of failed links increased. The number of failed links in the RDAT protocol increased much more quickly than that it in the iRTEDA protocol. The reason for this was that the RDAT protocol does not consider residual energy, and it does not have a recovery mechanism to keep the network running efficiently. The RDAT protocol only considers reputation, which leads to excessive use of nodes that have been assessed as having higher reputation. Such nodes consume more energy, and their death rate helps to balance the energy consumption of the nodes. Nodes with higher reputation will consume energy more quickly than others with low reputation, so they die much faster, which leads to higher failed links during data transmission. Figure 8 shows that the failed links in the RDAT protocol were similar to those in the iRTEDA protocol from the beginning to about 600 s. Because there is enough energy for nodes and aggregators in the beginning, few nodes are dead for using up the energy. During the time period, the fail links are occurring for the reason that some nodes have been compromised and excluded from the network. After the network has been running for 600 s, the number of failed links in the RDAT protocol is increasing much more rapidly that were in the iRTEDA protocol. Thus, in the RDAT protocol, the transmission of data among nodes, aggregators, and the base station became more and more unreliable with time. At the same time, the reliabilities of both protocols were significantly different when the network was running. Figure 9 shows that, when the failed links reached 30%, the reliability of RDAT and iRTEDA decreased to about 79.2% and 89.5%, respectively. The larger number of failed links is responsible for the lower reliability. With high unreliability of data transmission, there will be large amount of data loss, and data accuracy will decrease quickly. Thus, the iRTEDA protocol is more reliable for maintaining data transmission and improving the performance of data-aggregation, when the reputation system improves the security of the network.

*5.4. Comparison of Energy Consumption.* The concept of lifetime describes the death rate for the nodes in the network, which is defined as the number of dead nodes overtime. Energy consumption shows us the percentage of energy consumption of all the nodes in the network. Combining lifetime and energy consumption, we can obtain the detailed performance for the total energy consumption and energy balancing of nodes. Figure 10 shows that the death rates of nodes in the two protocols were significantly different. The death rate of nodes in RDAT was much faster than it was in our protocol. There is a large number of dead nodes in the beginning. The reason is that, in RDAT, there are excessive uses for those nodes with higher reputation in routing and transmission, which results in quick death for those nodes.

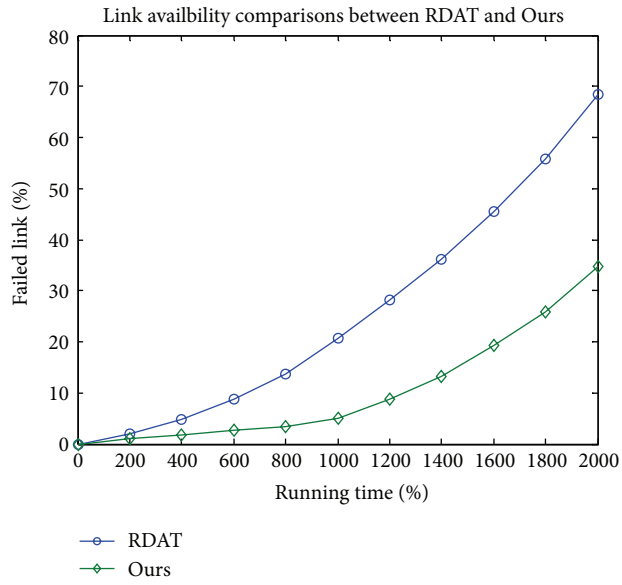



Figure 8: Comparison of the link availability of the RDAT and iRTEDA protocols when the percentage of compromised nodes was 30%.

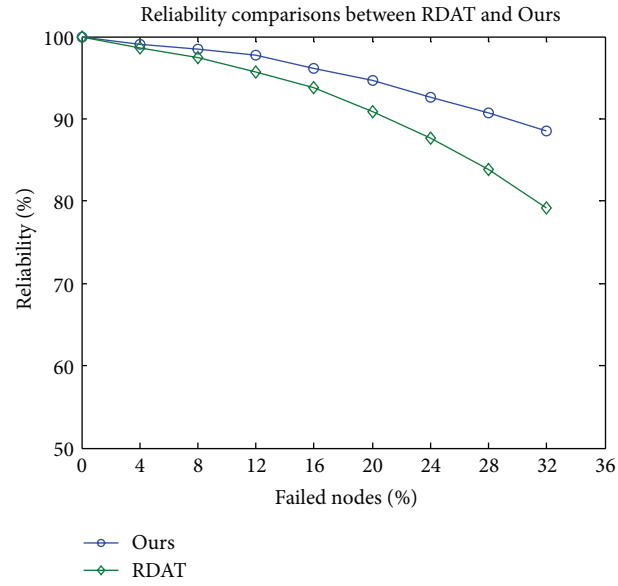



Figure 9: Comparison of the reliabilities of the RDAT and iRTEDA protocols when the percentage of compromised nodes was 30%.

With the network running for more than 2000 seconds, the nodes with lower reputation, isolated nodes, and the death rate are decreased to very low rates. The selection of the nodes in routing and transmission in our protocol was based on reputation, but it also emphasized energy consumption and link availability. Thus, unlike the RDAT protocol, it will not lead to excessive use of the same nodes, and it provides better energy balancing. All of the energy that is consumed will be distributed over large number of nodes in iRTEDA protocol, instead of that all of the energy consumption will focused on few nodes with high reputations and those nodes consume

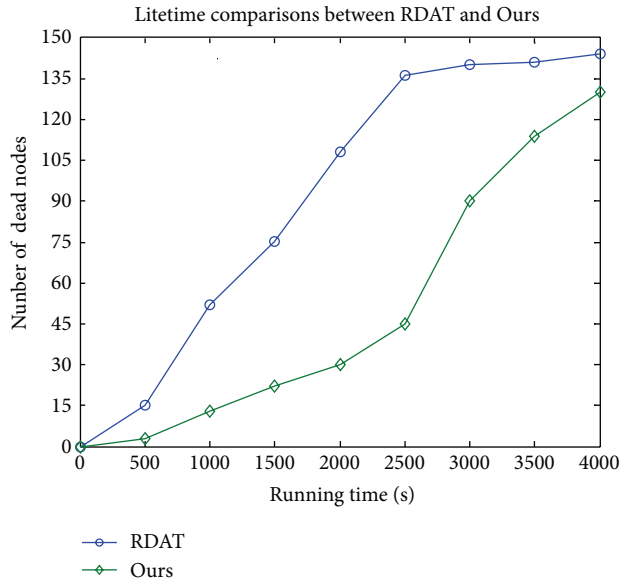



Figure 10: Lifetime comparison between the RDAT and iRTEDA protocols when the percentage of compromised nodes was 30%.

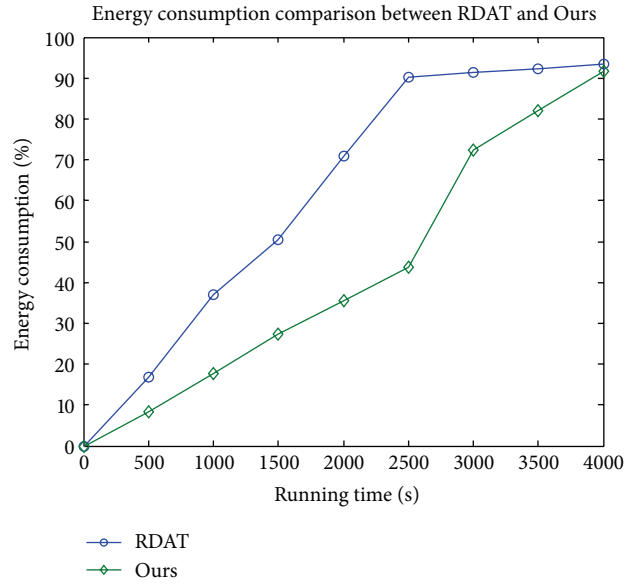



Figure 11: Comparison of the energy consumption for the RADT and iRTEDA protocols when the percentage of compromised nodes was 30%.

energy more rapidly. Figure 10 shows that the death rate of nodes in our protocol was much lower than that in the RDAT protocol in the beginning, and the rate increased rapidly at a running time of 2500 s. Because more energy is consumed for each node after that time period, the death of nodes will increase as long as the network keeps running. When the network had operated for 4000 s, only six nodes were left in the RDAT protocol, whereas 20 nodes were left in our protocol.

Figure 11 compares the entire energy consumptions of the two protocols. Before about 2500 s, the rate of increase in energy consumption was much lower in iRTEDA than in RDAT, with iRTEDA reaching about 43.6% and RDAT reaching about 90.1%. The difference in the energy consumption of the two protocols was caused by the proposed recovery mechanism, because the nodes with low reputation are judged as compromised nodes and excluded from the network. Those compromised nodes' children nodes will become isolated. Those isolated nodes transmit data directly to the base station and consume much more energy than before. Thus, the recovery mechanism will help the dispersing or isolated nodes in the network finish the reselection of father nodes and re-access to the clusters. This will lead to decreases in the average hops from those nodes to the base station which will save communication overhead and reduce the total energy consumption. At a running time of 4000 seconds, the percentage of energy consumption was about 93.4% for the RDAT protocol and about 83.8% for the iRTEDA protocol. Figure 11 shows that the total energy consumption was much lower for the iRTEDA protocol than for the RDAT protocol.

## 6. Conclusion

An improved data aggregation method for WSNs is presented in this paper. The method is reliable, trust-based, energy-efficient, and secure. The iRTEDA protocol combines residual energy and links availability to improve trust-based data-aggregation. Introducing residual energy and link availability facilitates the reputation system's ability to keep aggregators and nodes in the routing path from being used excessively and guarantees that the routing path selected by the reputation system will be much more reliable. In addition, the recovery mechanism prevents compromised nodes' children nodes from being isolated and helps those nodes to reselect new parent nodes. Simulation results showed that the proposed protocol outperformed the RDAT protocol with respect to its improved performance of data accuracy, routing path reliability, and the lifetime of data-aggregation, while reducing energy consumption. Thus, the proposed iRTEDA protocol achieved its goal of keeping secure data-aggregation in WSNs more reliable and energy-efficient.

## Acknowledgments

This research is supported by National Natural Science Foundation of China Under Grant 61071076, Beijing Natural Science Foundation Under Grant 4132057, National High-tech Research And Development Plans (863 Program) Under Grant 2011AA010104-2, and The Academic Discipline and Postgraduate Education Project of Beijing Municipal Commission of Education.

## References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–105, 2002.

[2] K. Römer, "Programming paradigms and middleware for sensor networks," in *Proceedings of the GI/ITG Workshop on Sensor Networks*, pp. 49–54, Karlsruhe, Germany, 2004.

[3] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Networks*, vol. 52, no. 12, pp. 2292–2330, 2008.

[4] D. Culler, D. Estrin, and M. Srivastava, "Overview of sensor networks," *Computer*, vol. 37, no. 8, pp. 41–49, 2004.

[5] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proceedings of the 1st ACM International Workshop on Wireless Sensor Networks and Applications (WSNA '02)*, pp. 88–97, Atlanta, Ga, USA, September 2002.

[6] N. Xu, S. Rangwala, K. K. Chintalapudi et al., "A wireless sensor network for structural monitoring," in *Proceedings of the Second International Conference on Embedded Networked Sensor Systems (SenSys '04)*, pp. 13–24, ACM Press, New York, NY, USA, November 2004.

[7] J. Polastre, R. Szewczyk, A. Mainwaring, D. Culler, and J. Anderson, "Analysis of wireless sensor networks for habitat monitoring," *Wireless Sensor Networks*, vol. 4, pp. 399–423, 2004.

[8] T. He, P. Vicaire, T. Yant et al., "Achieving real-time target tracking using wireless sensor networks," in *12th IEEE Real-Time and Embedded Technology and Applications Symposium (RTS '06)*, pp. 37–48, San Jose, Calif, USA, April 2006.

[9] G. Simon, G. Balogh, G. Pap et al., "Sensor network-based countersniper system," in *Proceedings of the Second International Conference on Embedded Networked Sensor Systems (SenSys '04)*, pp. 1–12, Baltimore, Md, USA, November 2004.

[10] E. Cayirci and T. Coplu, "SENDROM: sensor networks for disaster relief operations management," *Wireless Networks*, vol. 13, no. 3, pp. 409–423, 2007.

[11] A. Milenković, C. Otto, and E. Jovanov, "Wireless sensor networks for personal health monitoring: issues and an implementation," *Computer Communications*, vol. 29, no. 13-14, pp. 2521–2533, 2006.

[12] K. Akkaya, M. Demirbas, and R. S. Aygun, "The impact of data aggregation on the performance of wireless sensor networks," *Wireless Communications and Mobile Computing*, vol. 8, no. 2, pp. 171–193, 2008.

[13] R. Rajagopalan and P. K. Varshney, "Data aggregation techniques in sensor networks: a survey," *IEEE Communications Surveys and Tutorials*, vol. 8, no. 4, pp. 48–63, 2006.

[14] S. Ozdemir and Y. Xiao, "Secure data aggregation in wireless sensor networks: a comprehensive overview," *Computer Networks*, vol. 53, no. 12, pp. 2022–2037, 2009.

[15] S. Ozdemir, "Functional reputation based data aggregation for Wireless sensor networks," in *4th IEEE International Conference on Wireless and Mobile Computing, Networking and Communication (WiMob '08)*, pp. 592–597, October 2008.

[16] S. Ozdemir, "Functional reputation based reliable data aggregation and transmission for wireless sensor networks," *Computer Communications*, vol. 31, no. 17, pp. 3941–3953, 2008.

[17] A. Boukerche and Y. Ren, "A trust-based security system for ubiquitous and pervasive computing environments," *Computer Communications*, vol. 31, no. 18, pp. 4343–4351, 2008.

[18] Y. Ren and A. Boukerche, "Modeling and managing the trust for wireless and mobile ad hoc networks," in *IEEE International Conference on Communications (ICC '08)*, pp. 2129–2133, Beijing, China, May 2008.

[19] S. Ganeriwal and M. B. Srivastava, "Reputation-based framework for high integrity sensor networks," in *Proceedings of the ACM Workshop on Security of Ad Hoc and Sensor Networks (SASN '04)*, pp. 66–77, October 2004.

[20] S. Ganeriwal, L. K. Balzano, and M. B. Srivastava, "Reputation-based framework for high integrity sensor networks," *ACM Transactions on Sensor Networks*, vol. 4, no. 3, article 15, 2008.

[21] R. A. Shaikh, H. Jameel, S. Lee, S. Rajput, and Y. J. Song, "Trust management problem in distributed wireless sensor networks," in *12th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA '06)*, pp. 411–414, Sydney, Australia, August 2006.

[22] P. Michiardi and R. Molva, "CORE: a collaborative reputation mechanism to enforce node cooperation in mobile ad hoc network," in *Proceedings of the IFIP Conference on Communications and Multimedia Security*, vol. 228, pp. 107–121, Portoroz, Slovenia, 2002.

[23] A. Srinivasan, J. Teitelbaum, and W. Jie, "DRBTS: distributed reputation-based beacon trust system," in *2nd IEEE International Symposium on Dependable, Autonomic and Secure Computing (DASC '06)*, pp. 277–283, Ind, USA, October 2006.

[24] A. Jøsang and R. Ismail, "The beta reputation system," in *Proceedings of the 15th Bled Conference on Electronic Commerce*, p. 41, 2002.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 307038, 9 pages
http://dx.doi.org/10.1155/2013/307038

*Research Article*

# Energy-Aware Routing in Wireless Sensor Networks Using Local Betweenness Centrality

**Xiao-Hui Li[1,2] and Zhi-Hong Guan[1]**

[1] *College of Automation, Huazhong University of Science and Technology, Wuhan 430074, China*
[2] *College of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China*

Correspondence should be addressed to Zhi-Hong Guan; zhguan@mail.hust.edu.cn

Received 11 December 2012; Revised 19 April 2013; Accepted 22 April 2013

Academic Editor: J. Barbancho

Copyright © 2013 X.-H. Li and Z.-H. Guan. 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 propose an energy-aware dynamic routing strategy in order to provide balanced energy consumption in wireless sensor networks, hence, prolonging the lifetime of the network. The proposed routing algorithm uses *local betweenness centrality* to estimate the energy consumption of the neighboring nodes around a given local sensor node, without requiring global information about the network topology or energy consumption, and to divert traffic from nodes that are more heavily used. Because nodes with large local betweenness centrality consume energy more quickly, the network lifetime can be prolonged by redistributing energy consumption to nodes with smaller local betweenness centrality. Simulation results showed that the proposed routing strategy has advantages over shortest path routing with respect to extending network lifetime and balancing energy consumption in wireless sensor networks, yet does not introduce significant additional transmission overhead or a longer average path length.

## 1. Introduction

Wireless sensor networks (WSNs) have many desirable characteristics, including easy deployment and self-organization, and are becoming increasingly important in modern society. The applications of WSNs range from important societal issues such as environmental surveillance, intelligent transportation, disaster relief, and health care to military issues including battlefield biological and nuclear monitoring as well as target tracking [1]. WSNs generally consist of a large number of small, embedded, low-power sensor nodes with sensing, data processing, and wireless communication capabilities [1]. They are deployed in a wide distribution area and collaborate to form an *ad hoc* network. The sensor nodes are battery-operated, most of which are not rechargeable, and cease to function once the battery expires. Owing to logistical issues such as remoteness or inaccessibility of distribution areas, it is not straightforward to replace sensor nodes with expired batteries. Therefore, balancing the energy consumption of WSNs to prolong the network lifetime is of considerable importance. In WSNs, the energy is consumed in different ways in different nodes; however, the primary energy cost is in data transmission once the networks have been organized.

Experimental measurements have shown that data transmission is generally very expensive in terms of energy consumption, while data processing consumes significantly less energy [2]. Data transmission depends on the routing strategy; therefore, designing a reasonable routing strategy to balance energy consumption has significant potential to prolong the lifetime of WSNs [3].

Along with the discovery of scale-free [4] properties, complex networks have been applied to a wide range of natural and social systems, such as the Internet, social networks, scientific collaboration networks, metabolic networks, and WSNs [5, 6]. Recently, routing in complex networks has attracted considerable interest across many fields of science. Research on routing strategies has overwhelmingly focused on improving transport capacity and controlling congestion, which are crucial problems on many large-scale

networks such as the Internet, phone networks, and airport networks [7–14]. WSNs have a low data rate and low energy consumption [6, 15], and so the critical problem in WSNs is the longevity of the network, and the major constraint is energy consumption rather than congestion control.

Much work has been done developing WSN routing (e.g., see [16–27] for extensive reviews). Most existing energy-aware routing algorithms assume that the communication load is evenly distributed; however, this assumption is not consistent with the data usage requirements of individual nodes within many WSNs. Almost no studies have considered energy-aware routing in WSNs from the point of view of the structure and dynamics of large complex networks.

In this paper, we propose an energy-aware routing in WSNs using local betweenness centrality (EAR-LBC). The proposed routing algorithm has two features that differ from those of most current research

(1) It utilizes the basic principles of complex network theory, in particular the concept of betweenness centrality (BC) to estimate the remaining energy of the sensor nodes. Routing in WSNs takes place according to the criterion of the shortest available path from a given source to its destination. The nodes with the largest BC, which are usually called the central nodes of networks and are located in the shortest path, are susceptible to more frequent data transmission due to heavier traffic load [3]. Energy consumption at these central nodes occurs at a greater rate than that of other sensor nodes, leading to unbalanced energy consumption in the network. Once the central node runs out of energy, the WSN will cease to function. Consequently, we propose the EAR-LBC algorithm, which uses local BC (defined in Section 3) to consider the energy cost of forwarding data, and takes both the shortest path and the remaining energy of the sensor node into account, rather than simply employing traditional metrics such as the shortest path.

(2) A systematic approach is taken to verify the validity of the EAR-LBC algorithm. Because most network topological structures (including WSN) exhibit scale-free behavior, we designed a simulation model based on Barabasi-Albert (BA) scale-free networks (discussed in Section 4.1). This simulation model is consistent with most WSN topological structures. For definiteness and without loss of generality, we investigated the performance of the EAR-LBC algorithm using three simulation scenarios generated by the model.

The remainder of this paper is organized as follows. Section 2 presents a review of relevant prior work. The proposed routing strategy is described in Section 3. Section 4 discusses the simulation results. Finally, the paper is concluded in Section 5.

## 2. Related Works

Applying complex network theory to the design of energy-aware routing in WSNs is an interdisciplinary field, which combines WSNs with complex networks. Accordingly, in this sections we describe related research progress in two ways: energy-aware routing in WSNs and routing in scale-free networks.

*2.1. Energy-Aware Routing in WSNs.* Many energy-aware routing algorithms for WSNs have been presented in recent years. Those routing algorithms can be categorized into two types. One type is the clustering routing algorithms that divide sensor nodes into clusters and balance the energy consumption by cluster head selection to extend network lifetime [19, 23, 24, 26, 27]. Cluster-based routing is an efficient way to reduce energy consumption and extend network lifetime within a cluster. The number of messages transmitted to the base station is reduced by data aggregation and fusion, which reduces the overall energy consumption. Cluster-based routing is mainly implemented as a two-layer strategy: one layer is used to select cluster heads, and the other layer is used for routing. High-energy nodes can be used to process and send information, whereas low-energy nodes can be used to perform sensing in close proximity to the target. The clustering algorithm is based on cluster selecting, which incurs an additional energy cost. The other type is centralized routing, which uses probabilistic forwarding [18] or an optimization strategy, such as ant colony optimization, linear programming, or heuristic approaches, to find an energy-balanced route based on the global information on the network topology and energy consumption [16, 17, 20–22, 25].

However, most existing energy-aware routing algorithms assume that energy consumption in WSNs is evenly distributed or that a WSN is deployed as a specific scenario when analyzing the validity of its routing algorithms, which is not consistent with most WSN topological structures. Almost all studies have failed to consider energy-aware routing from the point of view of the structure and dynamics of large complex networks.

*2.2. Routing in Scale-Free Networks.* Because of the importance of large complex communication networks in modern society (such as the Internet, which has scale-free properties), the dynamics of the underlying structure (such as traffic congestion) have drawn much attention from both physics and engineering standpoints. Making full use of complex network theory, the routing strategies in scale-free networks overwhelmingly focus on improving transport capacity and congestion control. To avoid congestion and improve the transmission capacity of networks, many routing strategies have been proposed in scale-free networks, including random walk, efficient routing, local routing, optimal routing, and hub avoidance strategies [8–14].

WSNs typically have scale-free properties, and it is necessary to study the routing process according to the particular requirements of these types of networks. Energy awareness is a central design issue for WSNs; to extend network lifetimes, energy-aware routing should be considered in these scale-free networks.

## 3. Routing Strategy

In this section, we present an overview of the EAR-LBC routing strategy then provide the pseudocode to implement it. Finally, we use a simple routing example to illustrate the proposed algorithm.

*3.1. Greedy Forwarding.* The proposed routing strategy is a local routing search algorithm based on greedy forwarding. Greedy forwarding aims to bring messages closer to the destination using only local information in each stage of the journey. Thus, each node forwards the message to the neighbor that is most suitable from a local point of view, which can be the one that minimizes the energy cost to the destination in each step.

Because we apply greedy forwarding to find a route from the source to the destination, we must choose a selection function that describes which of the candidates is the most promising. That is, we must identify the optimum next stage of the journey at each sensor node during the process of route finding. Obviously, it is desirable for each sensor node to forward the data packet to a neighboring node that is both close to the destination and has sufficient energy to forward the data packet. This greedy forwarding criterion can be described as a selection function that determines which candidate is nearest to the destination [16]. Suppose that node $X$ has $M$ direct (one-hop) neighbors, $n_1, n_2, \ldots, n_M$, and the destination node is $D$. The selection function is

$$f = \min\left(\text{cost}_1, \text{cost}_2, \ldots, \text{cos t}_M\right), \tag{1}$$

where $\text{cost}_i$ is the cost between $n_i$, the $i$th neighbor of $X$, and the destination $D$, for $i \in [1, M]$. We define cost as follows:

$$\text{cost}_i = \alpha d_i + (1 - \alpha)\, e_i, \tag{2}$$

where $\alpha$ is an adjustable parameter [9], $d_i$ is the distance from the $i$th neighbor to the destination, and $e_i$ is the energy consumed at the $i$th neighbor. The parameter $\alpha$ determines the weightings of $d_i$ and $e_i$ in the cost calculation.

*3.2. Local Betweenness Centrality.* The availability of small, low-power global positioning system (GPS) receivers for calculating relative coordinates makes it possible to obtain the distance from the $i$th neighbor to the destination. Therefore, in (2), we can readily obtain the value of $d_i$. However, for large WSN applications, it can be very difficult for a given sensor node to determine the energy $e_i$ consumed at a neighboring node because of the additional overhead. In this section, we introduce the definition of local BC to estimate $e_i$.

Recent studies [7, 10] have reported that BC plays an important role in the traffic on networks. For a given network, the BC of a node is defined as

$$B_i = \sum_{s \neq d} \frac{\sigma_{sd}(i)}{\sigma_{sd}}, \tag{3}$$

where $\sigma_{sd}$ is the number of shortest paths going from $s$ to $d$, and $\sigma_{sd}(i)$ is the number of shortest paths going from $s$ to $d$ and passing through $i$. BC quantifies the number of times a node appears in the shortest paths between two other nodes. The BC is a useful measure of the load placed on a given node in a network, as well as the importance of the node in the network. It has become a popular measurement to characterize complex networks. Based on complex network theory, if the number of nodes on the networks is denoted by $N$, and there are $R$ data packets that need to be transmitted at every time step, then the average number of packets passing through a given node in $t$ time steps can be obtained as follows [10]:

$$\frac{RtB_i}{N(N-1)}. \tag{4}$$

Assuming that the sensor node forwarding a data packet consumes energy $E_f$, combining (4) and (2) we obtain

$$\text{cost}_i = \alpha d_i + (1 - \alpha)\, \frac{RtB_i}{N(N-1)} E_f, \tag{5}$$

where $B_i$ is BC of the $i$th neighbor.

However, there remains a problem that must be solved. The BC is calculated based on global topology information, which is generally not available in large-scale wireless *ad hoc* sensor networks. To deal with this problem, we propose an energy-aware routing system based on local BC. We extend the concept of BC from the global topology to a local routing table, which consists of the destination and the next hop information only. Similarly, the local BC of a node for a local routing table is defined as

$$b_i = \sum_{s,d,i \in \text{Local}, s \neq d} \frac{\sigma_{sd}(i)}{\sigma_{sd}}, \tag{6}$$

where $\sigma_{sd}$ is the number of paths going from $s$ to $d$ in the local routing table of the local sensor node, $X$, and $\sigma_{sd}(i)$ is the number of paths going from $s$ to $d$ and passing through the $i$th neighbor in the local routing table of $X$. The local BC gives an estimate of the traffic handled by the neighbors around $X$. Therefore, the neighbor with the greatest local BC delivers more data packets, thus; consumes more energy. If we combine (5) and (6), we arrive at

$$\text{cost}_i = \alpha d_i + (1 - \alpha)\, \frac{Rtb_i}{N(N-1)} E_f. \tag{7}$$

*3.3. Routing Algorithm.* The proposed routing algorithm is a distributed routing algorithm based on local BC. For each sensor node $X$, which receives a data packet $P$, the next hop is determined as follows.

(1) For each neighbor $n_i$, the cost, $\text{cost}_i$, is calculated using (7) with the current routing table.

(2) Among the neighbors of $X$, we choose the $n_i$ with the minimum $\text{cost}_i$ as the next hop and forward the data packet $P$.

(3) If there is no routing information about the current data packet $P$ in the routing table, routing paths for $P$ are added. If routing information already exists, we update the next hop information to that determined in step (2).

**Input:** the received packet ($P$)
**Output:** the next hop ($nexthop$)
/* for each neighbor of X, calculate its cost */
**foreach** *neighbor* $n_i$ of $X$ **do**
  **if** *routing table is empty* **then**
    $b_i = 0$
  **else**
    according to (6), computes $b_i$ based on routing table of $X$;
  **end**
  computes $d_i$;
  $cost_i = \alpha d_i + (1 - \alpha)\dfrac{Rtb_i}{N(N-1)}E_f$;
**end**
/* among the neighbors of X, chose the minimum cost as the next hop */
$cost_{min} = cost_1$
**foreach** *neighbor's* $cost_i$ *of* $X$ **do**
  **if** $cost_{min} > cost_i$ **then**
    $cost_{min} = cost_i$;
    $nexthop = n_i$;
  **end**
**end**
/* update the routing table of $X$ */
flagexist = 0;
**while** $flagexist$==0 *and not end of routing table of* $X$ **do**
  /* $r_i$ is the $i$th routing entry in the routing table of $X$ */
  **if** $r_i.destination == P.destination$ **then**
    flagexist == 1;
    $r_i.nexthop = nexthop$
  **end**
**end**
**if** $flagexist$==0 **then**
  /* there is no routing entry for this packet P, need to append new rout information */
  Add a new routing entry $r_{new}$;
  $r_{new}.destination = P.destination$;
  $r_{new}.nexthop = nexthop$;
**end**
**if** $P.destination$==$nexthop$ **then**
  directly send to the destinaiton
**end**

Algorithm 1: Greedy route finding.

The pseudocode to implement the algorithm is shown in Algorithm 1.

Initially, the routing table of $X$ will be empty. From (6), the value of $b_i$ is 0, and so the cost at a neighboring node is determined from $d_i$ in (7). That is, the cost is determined by the distance from the $i$th neighbor to the destination. The packet is forwarded to the neighbor that is closest to the destination. The next hop information of the sensor node $X$ is recorded in the routing table. As the amount of information in the routing table increases, the value of $b_i$ plays a larger role in the cost calculation. A neighbor with greater $b_i$ will have already consumed more energy due to local packet forwarding. If there are several neighbors with the same distance to the destination, the data packet should be forwarded to the neighbor with the lowest $b_i$, as this will have more energy remaining than the other neighbors.

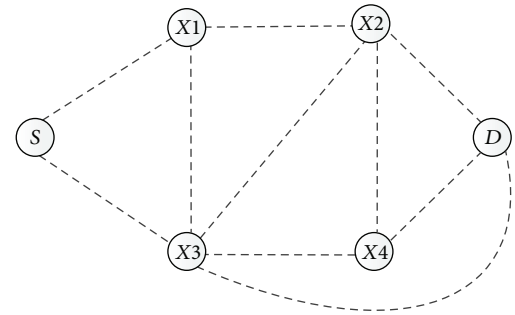



Figure 1: An example of WSN topology.

The key to the algorithm is dynamic calculation of cost informed by updating the local routing table. By choosing the optimal next hop by minimizing cost, the data packet can be routed by sensor nodes with more energy. Energy consumption becomes more uniform, and the network lifetime can be prolonged. Obviously, the proposed algorithm is equal to the traditional shortest path routing when $\alpha = 1$.

*3.4. A Routing Example.* Here, we illustrate the algorithm using a simple routing example. Consider a wireless network topology shown in Figure 1. To demonstrate the function of local BC, we consider the distance from one node to another to be the number of hops between them and focus on the change in the local BC. In Figure 1, suppose that the source node $S$ sends several data packets, denoted by $P1$, $P2$,..., to the destination node $D$ and that $\alpha = 0.125$.

When $S$ sends $P1$ to $D$, the routing table of $S$ is empty at the beginning. $S$ has two one-hop neighbors, $X1$ and $X3$. According to (6), the values of the local BC for these two nodes are 0. So the cost of neighbors around node $S$ is only determined by the distance between them, and the cost of node $X1$, $\text{cost}_{X1}$, is calculated as follows:

$$\text{cost}_{X1} = \alpha d_{X1\to D} + (1-\alpha)\, b_{X1} = \frac{1}{8}\times 2 + \left(1 - \frac{1}{8}\right)\times 0 = \frac{2}{8}. \tag{8}$$

The term $d_{X1\to D}$ is the distance between node $X1$ and the destination node $D$, and $b_{X1}$ is the local BC of node $X1$. In a similar manner, we can obtain the cost of node $X3$ as follows:

$$\text{cost}_{X3} = \alpha d_{X3\to D} + (1-\alpha)\, b_{X3} = \frac{1}{8}\times 1 + \left(1 - \frac{1}{8}\right)\times 0 = \frac{1}{8}, \tag{9}$$

where $\text{cost}_{X3}$ is less than $\text{cost}_{X1}$, so $X3$ is chosen as the next hop, and node $S$ updates its routing table as shown in Table 1.

Node $X3$ receives data packet $P1$ and finds the destination node $D$ is in its neighbor set, then forwards the data packet to the destination node $D$. Node $S$ sends $P1$ to $D$ using the route $S \rightarrow X3 \rightarrow D$.

When $S$ sends $P2$ to $D$, according to (6), the local BC of $X1$, denoted by $b_{X1}$, is 0 because there is no route passing through $X1$ in the routing table of $S$. The local BC of $X3$, denoted by $b_{X3}$, is 1 because the number of paths going from $S$ to $D$ in the local routing table of $S$ is 1, and the number of paths going from $S$ to $D$ and passing through $X3$ in the local routing table of $S$ is 1. $S$ calculates the cost of its neighbors, $X1$ and $X2$, again using (7); thus, we obtain

$$\begin{aligned} \text{cost}_{X1} &= \alpha d_{X1} + (1-\alpha)\, b_{X1} = \frac{1}{8} \times 2 + \left(1 - \frac{1}{8}\right) \times 0 = \frac{2}{8}, \\ \text{cost}_{X3} &= \alpha d_{X3} + (1-\alpha)\, b_{X3} = \frac{1}{8} \times 1 + \left(1 - \frac{1}{8}\right) \times 1 = 1, \end{aligned} \tag{10}$$

where $\text{cost}_{X1}$ is less than $\text{cost}_{X3}$, $X1$ is chosen as the next hop, and node $S$ updates its routing table as shown in Table 2. Node $X1$ receives data packet $P2$ and determines that the next hop is $X2$ using EAR-LBC. Finally, $S$ sends $P2$ to $D$ using route $S \rightarrow X1 \rightarrow X2 \rightarrow D$.

In this example, data packets are routed from $S$ to $D$ via $S \rightarrow X3 \rightarrow D$ and route $S \rightarrow X1 \rightarrow X2 \rightarrow D$. If the shortest path algorithm was to be used, all network traffic from $S$ to $D$ would be routed via $S \rightarrow X3 \rightarrow D$, and energy would be consumed at a greater rate at $X3$ than at nodes $X1$ and $X2$ because of the unbalanced network traffic. Using the proposed routing strategy, network traffic from $S$ to $D$ is shared by nodes $X1$, $X2$, and $X3$, and the the energy consumption is more balanced. The local BC is calculated based on the routing information recorded in the routing table. If the routing table is changed, the local BC is also changed, which in turn feeds back to change the routing table. From the interaction between the routing table and the calculation of the local BC, network traffic can be allocate in a manner that provides more balanced energy consumption in the network.

## 4. Simulation

In this section, we describe simulations that were performed to evaluate the performance of the proposed routing strategy, developed using MATLAB. Our goal was to determine the advantages of the routing strategy in terms of network lifetime, average path length, and residual energy by comparing the performance to that of other routing algorithms. Most routing processes in WSNs take place according to the criterion of the fewest hops from a given source to the destination. This is equivalent to the shortest path routing from a given source to its destination in a graph with the same edge weight on all available wireless links. We compared the performance of the EAR-LBC to that of the shortest path routing (SP).

Table 1: Routing table of node $S$ when sending data packet $P1$.

| Destination | Next hop |
|---|---|
| $D$ | $X3$ |

Table 2: Routing table of node $S$ when sending data packet $P2$.

| Destination | Next hop |
|---|---|
| $D$ | $X1$ |

*4.1. Simulation Model.* The BA model is one of several proposed models that generates scale-free networks, and the networks studied in our simulation were generated using this model. BA scale-free network incorporates two important general concepts: growth and preferential attachment. Both growth and preferential attachment exist widely in real networks. Growth means that the number of nodes in the network increases over time. Preferential attachment means that the more connected a node is, the more likely it will be to receive new links. Nodes with a higher degree have a greater ability to grab links added to the network.

The generation of networks begins with an initial network of $m_0$ nodes, where $m_0 \geq 2$ and the degree of each node in the initial network should be at least 1; otherwise the node will always remain disconnected from the rest of the network. New nodes are added to the network one at a time. Each new node is connected to $m$ existing nodes with a probability that is proportional to the number of links that the existing nodes already have. Formally, the probability $p_i$ that the new node is connected to node $i$ is [1]

$$p_i = \frac{k_i}{\sum_j k_j}, \tag{11}$$

where $k_i$ is the degree of node $i$ and the summation is over all preexisting nodes $j$. Heavily linked nodes tend to quickly accumulate even more links, while nodes with only a few links are unlikely to be chosen as the destination for a new link. The new nodes have a "preference" to attach themselves to nodes that are already heavily linked. The average degree of a scale-free network using BA model, denoted by $\langle k \rangle$, is approximately equal to $2m$ [4].

In our model, the number of sensor nodes in the networks is denoted by $N$. All nodes are treated as both sensors and routers for generating and transporting data packets, and each link has the same packet-delivery capacity. Consistent with the low data rate in WSNs, we assume that each node has sufficient processing and buffering capacity to deliver and handle all of the data packets it receives in each time step. Transport on the network proceeds in discrete time steps and is driven by inserting $R$ new data packets, with randomly chosen sources and destinations. At each time step, every node delivers the packets toward the neighboring node with the optimum next hop as determined by the routing strategy. For the sake of comparison, the number of new data packets generated by the nodes per time step was fixed at 1 (i.e., $R = 1$); however, it is trivial to change this so as to meet the demands of various example networks. The initial energy

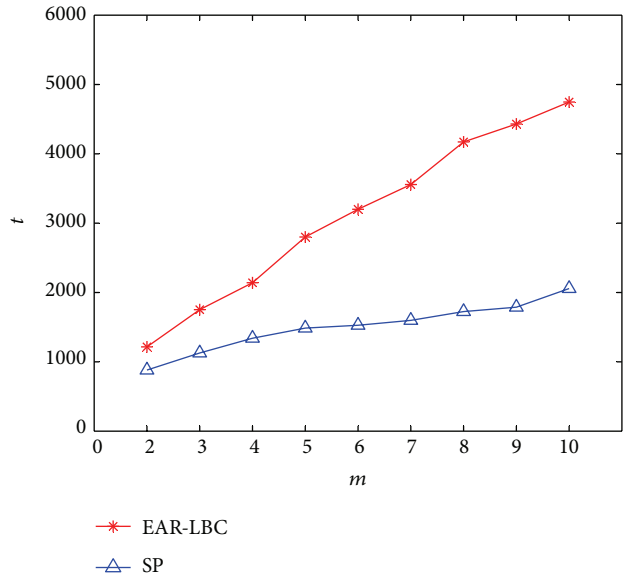



FIGURE 2: The network lifetime, $t$, as a function of parameter $m$ with $N = 600$ for both EAR-LBC and SP routing.

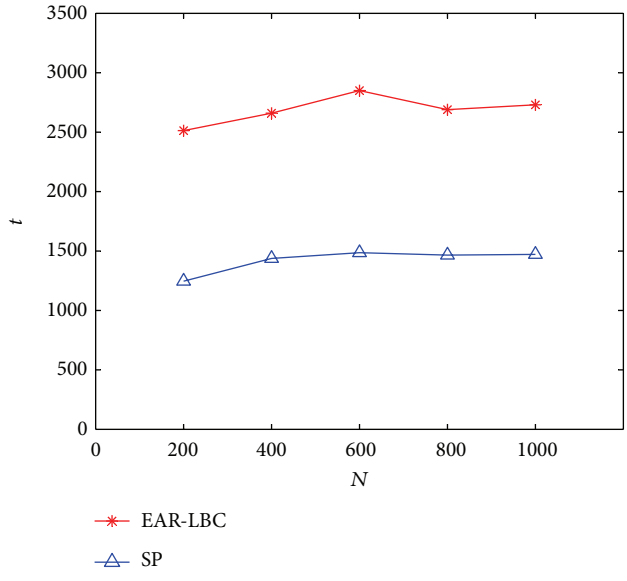



FIGURE 4: The network lifetime, $t$, as a function of the network scale, $N$, with $m = 5$, for both EAR-LBC and SP routing.

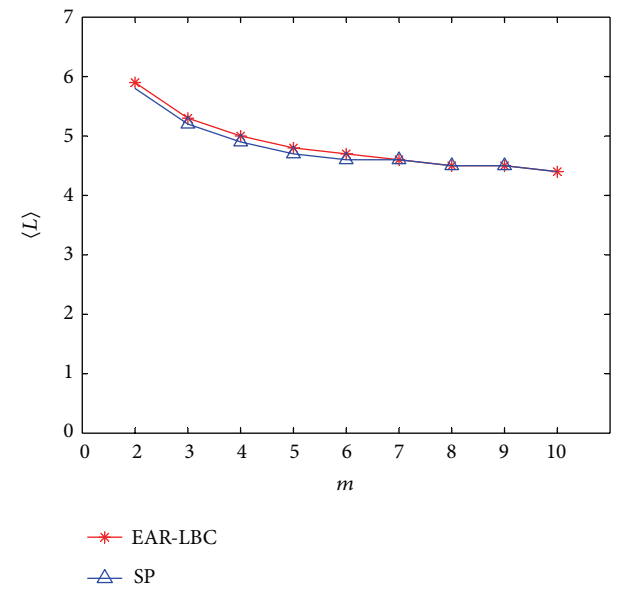



FIGURE 3: The average path length, $\langle L \rangle$, as a function of the parameter $m$ with $N = 600$ for both EAR-LBC and SP routing.

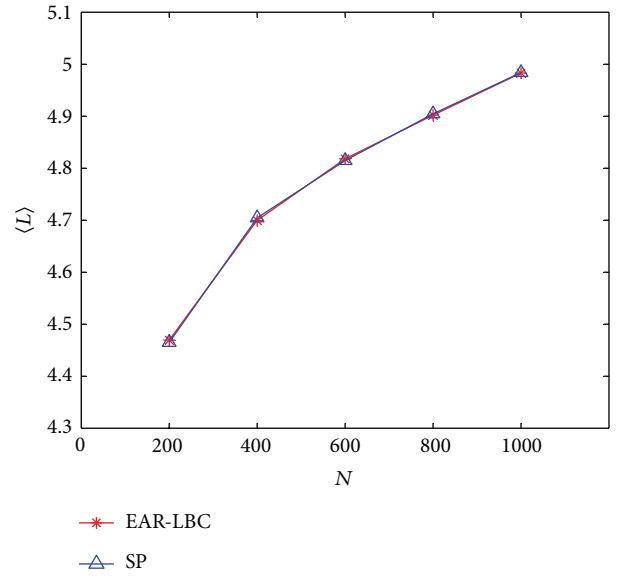



FIGURE 5: The average path length, $\langle L \rangle$, as a function of the network scale, $N$, with $m = 5$, for both EAR-LBC and SP routing.

of the every node was 5 joules, and $E_f$ is 0.02 joules. The adjustable parameter $\alpha$ was 0.75.

*4.2. Simulation Scenarios.* Three simulation scenarios were designed as follows.

(1) The number of nodes was fixed at $N = 600$, and the network lifetime and the average path length were investigated as a function of the average degree. In other words, we analyzed the performance of both EAR-LBC and SP routing when the average degree of the scale-free network increases, which corresponds to an increasing number of links between nodes.

(2) We fixed the average degree so that $m_0 = m = 5$ and $\langle k \rangle \approx 10$ and investigated the network lifetime and the average path length as a function of the number of nodes. In other words, we analyzed the performance of both EAR-LBC and SP routing when the number of nodes increases, but the average number of links between nodes does not change.

(3) We analyzed the distribution of the residual energy in a network using both EAR-LBC and SP routing. The

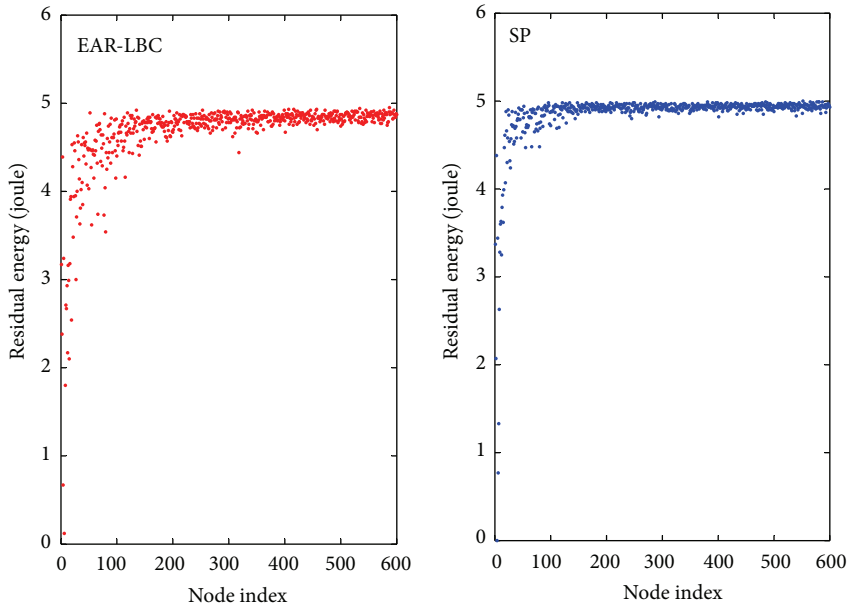



FIGURE 6: The distribution of residual energy plotted as a function of the node index, for a network where $N = 600$ and $m_0 = m = 6$, when the first failure node appears, for both EAR-LBC and SP routing.

average degree of this network was 12 ($m_0 = m = 6$, $\langle k \rangle \approx 12$), and the number of nodes was $N = 600$.

*4.3. Simulation Results.* Figure 2 shows the simulated network lifetime, and Figure 3 shows the simulated average path length with $N = 600$ as a function of the average degree of the network for the two routing strategies. The network lifetime is considerably longer for the EAR-LBC scheme. Furthermore, the larger the average degree of the network, the longer the network lifetime can be. This can be explained by considering that there are more paths among central nodes; in EAR-LBC, the network traffic can be distributed among a larger number of different paths, and, hence, nodes with excessive energy consumption can be avoided.

As shown in Figure 3, the average path length of two routing strategies decreased as the average degree increased, which can be explained by considering that there are more links between the sensor nodes, and so more direct routes are likely to be available. It can be seen that, although EAR-LBC enhances the network lifetime, it does not lead to a significant increase in the average path length. Both methods resulted in almost the same average path length. This is a particular useful result because an increase in the number of hops would result in an increased energy consumption for the network as a whole. In fact, EAR-LBC only incurs an additional computational cost in processing the expression in (7) in exchange for an increase in the network lifetime. This additional computational cost is slight because data processing consumes significantly less energy than data transmission in WSNs.

The average degree was fixed at $m = 5$ ($\langle k \rangle \approx 10$), and we investigated the network lifetime and the average path length as a function of the number of nodes. As shown in Figure 4, network lifetime was greatly enhanced using EAR-LBC; in both methods, the lifetime changes little a as the number of nodes is varied because if the average degree is invariant, the number of links among the sensor nodes does not increase. The network lifetime is primarily related to the traffic density rather than the number of nodes. Figure 5 shows the average path length as a function of the number of nodes; the results are similar to those shown in Figure 3. EAR-LBC enhanced network lifetime, but did not increase the average path length. EAR-LBC had almost the same average path length as SP routing.

Figure 6 shows the distribution of residual energy plotted against the node index for a network with $N = 600$ and $m_0 = m = 6$ when the first failure node appears (because of energy depletion) in EAR-LBC and SP routing. The network was generated using the BA model with $m_0 = m = 6$, which means that of the network initially had 6 nodes, and each new node was connected to 6 existing nodes. We view the order in which the nodes join the network as the index of the nodes, so that a smaller index corresponds to higher-degree nodes. Such nodes carry more networks traffic, so the energy consumption will be greater and faster at these nodes. This is why the residual energy of low-index nodes is less than that of high-index nodes. EAR-LBC resulted in a more uniform residual energy distribution than SP routing, even when the EAR-LBC network had a lifetime that was twice as long as that the SP routing. This is because EAR-LBC distributes the network traffic over several routes, decreasing the energy consumption of high-degree nodes.

Figure 7 shows a columnar comparison chart for EAR-LBC and SP routing, illustrating the distribution of nodes as a function of the residual energy. We rate the residual energy of each node on a scale of 1 to 4. A residual energy

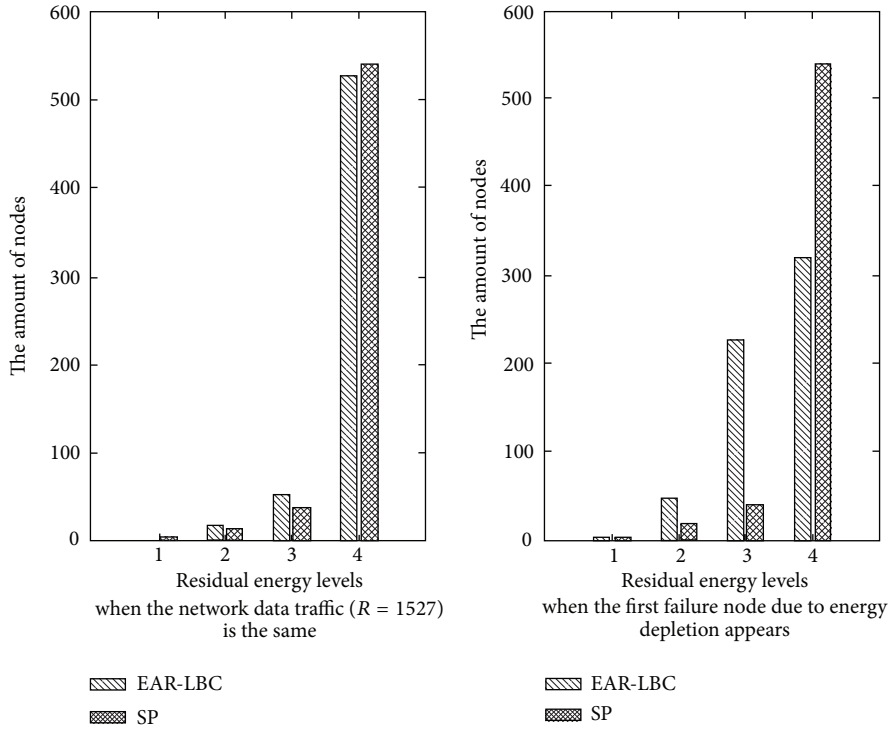



Figure 7: Columnar comparison chart for EAR-LBC and SP routing showing the number of nodes in different residual energy levels, for a network with $N = 600$ and $m_0 = m = 6$.

in range of 0 to 200 is level 1. A residual energy in range of 200 to 440 is levels 2. A residual energy in range of 440 to 480 is level 3. A residual energy in range of 480 to 500 is level 4. As shown in the left column, when EAR-LBC and SP routing had the same network data traffic with $R = 1527$, the first failure node appeared under SP, while the EAR-LBC network continued to function. In this case, the number of nodes in the EAR-LBC network with a residual energy in level 2 and 3 was larger than the number of nodes in that energy range for the network with SP routing. As shown in the right column of the figure, at failure, the residual energy of most nodes was in level 4 when using SP routing, but was evenly distributed in levels 3 and 4 when using EAR-LBC. That is, the residual energy distribution was more uniform in EAR-LBC compared to SP routing.

## 5. Conclusions and Future Works

We proposed an EAR-LBC for WSNs based on local BC and investigated the network lifetime and the average path length assuming that all nodes had the same initial energy. The proposed routing strategy considers realistic network structures and the dynamic energy consumption of large WSNs.

Our proposed algorithm improves upon the existing methods in two ways regarding the extension of network lifetime. First, our strategy uses greedy forwarding, which takes both the shortest path and the remaining energy at each node into account, rather than simply using traditional metrics such as the shortest path. A tunable weight, $\alpha$, is used as a parameter to determine the share of path length and remaining energy in route finding. This can be easily adjusted to optimize the routing strategy in line with the particular demands of a given network. Second, the routing strategy introduces the local BC to dynamically estimate the energy consumption of the neighboring nodes. Because of these features, even in the absence of global information on network topology and energy consumption, data packets can be routed to the sensor nodes with more residual energy, which provides a more balanced energy consumption in the network.

Because of these two improvements, EAR-LBC extends network lifetime without introducing additional transmission overhead or a longer average path length. Our results have applications to the design and optimization of routing for WSNs, including environmental monitoring systems, health care systems, and target-tracking systems.

Our future work will aim to improve and extend EAR-LBC algorithms, taking into account many additional characteristics of WSNs. There are a number of directions for future research. First, we will analyze and compare the performance of the EAR-LBC strategy described here and other energy-aware routing algorithms. Based on this comparative research, we will gain insight into which applications are best suited to EAR-LBC algorithms. Second, we will create a physical implementation to evaluate the performance of the EAR-LBC algorithms experimentally and study how to select design parameters according to the WSN application.

## Acknowledgment

This work was supported in part by the National Natural Science and Foundation of China under Grants 61105070, 61073025, 61170031, and 61272069.

## References

[1] D. F. Larios, J. Barbancho, G. Rodriguez, J. L. Sevillano, F. J. Molina, and C. Leon, "Energy efficient wireless sensor network communications based on computational intelligent data fusion for environmental monitoring," *IET Communications*, vol. 6, no. 14, pp. 2189–2197, 2012.

[2] G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: a survey," *Ad Hoc Networks*, vol. 7, no. 3, pp. 537–568, 2009.

[3] F. Ishmanov, A. S. Malik, and S. W. Kim, "Energy consumption balancing (ECB) issues and mechanisms in wireless sensor networks (WSNs): a comprehensive overview," *European Transactions on Telecommunications*, vol. 22, no. 4, pp. 151–167, 2011.

[4] A. L. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, p. 509, 1999.

[5] K. A. Hawick and H. A. James, "Small-world effects in wireless agent sensor networks," *International Journal of Wireless and Mobile Computing*, vol. 4, no. 3, pp. 155–164, 2010.

[6] H. Zhu, H. Luo, H. Peng, L. Li, and Q. Luo, "Complex networks-based energy-efficient evolution model for wireless sensor networks," *Chaos, Solitons and Fractals*, vol. 41, no. 4, pp. 1828–1835, 2009.

[7] G. Yan, T. Zhou, B. Hu, Z.-Q. Fu, and B.-H. Wang, "Efficient routing on complex networks," *Physical Review E*, vol. 73, no. 4, pp. 6108–6111, 2006.

[8] B. Danila, Y. Sun, and K. E. Bassler, "Collectively optimal routing for congested traffic limited by link capacity," *Physical Review E*, vol. 80, no. 6, Article ID 066116, pp. 6116–6122, 2009.

[9] M. Tang, Z. Liu, X. Liang, and P. M. Hui, "Self-adjusting routing schemes for time-varying traffic in scale-free networks," *Physical Review E*, vol. 80, no. 2, pp. 6114–6121, 2009.

[10] Z. H. Guan, L. Chen, and T. H. Qian, "Routing in scale-free networks based on expanding betweenness centrality," *Physica A*, vol. 390, no. 6, pp. 1131–1138, 2011.

[11] X. G. Tang, E. W. Wong, and Z. X. Wu, "Integrating networks structure and dynamic information for better routing strategy on scale-free networks," *Physica A*, vol. 388, no. 12, pp. 2547–2554, 2009.

[12] X. Ling, M. B. Hu, R. Jiang, R. Wang, X. B. Cao, and Q. S. Wu, "Pheromone routing protocol on a scale-free network," *Physical Review E*, vol. 80, no. 6, pp. 6110–6115, 2009.

[13] S. Meloni and J. Gomez-Gardenes, "Local empathy provides global minimization of congestion in communication networks," *Physical Review E*, vol. 82, no. 5, pp. 6105–6112, 2009.

[14] M. Tang and T. Zhou, "Efficient routing strategies in scale-free networks with limited bandwidth," *Physical Review E*, vol. 84, no. 2, pp. 6116–6120, 2011.

[15] E. Biagioni, "Topics in ad hoc and sensor networks," *IEEE Communications Magazine*, vol. 50, no. 7, p. 120, 2012.

[16] X. H. Li, S. H. Hong, and K. Fang, "A greedy and heuristic routing algorithm for wireless sensor networks in home automation," *IET Communications*, vol. 5, no. 13, pp. 1797–1805, 2011.

[17] D. Cheng, Y. Xun, T. Zhou, and W. Li, "An energy aware ant colony algorithm for the routing of wireless sensor networks," *Intelligent Computing and Information Science*, vol. 134, pp. 395–401, 2011.

[18] R. Shah and J. Rabaey, "Energy aware routing for low energy adhoc sensor networks," in *Proceedings of the IEEE Wireless Communications and Networking Conference (IEEE WCNC '02)*, pp. 350–355, Orlando, Fla, USA, March 2002.

[19] A. A. Abbasi and M. Younis, "A survey on clustering algorithms for wireless sensor networks," *Computer Communications*, vol. 30, no. 14-15, pp. 2826–2841, 2007.

[20] F. Ren, J. Zhang, T. He, C. Lin, and S. K. D. Ren, "EBRP: energybalanced routing protocol for data gathering in wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 12, pp. 2108–2125, 2011.

[21] M. Saleem, G. A. Di Caro, and M. Farooq, "Swarm intelligence based routing protocol for wireless sensor networks: survey and future directions," *Information Sciences*, vol. 181, no. 20, pp. 4597–4624, 2011.

[22] Y. H. Zhu, W. D. Wu, and V. C. M. Leung, "Energy-efficient tree-based message ferrying routing schemes for wireless sensor networks," *Mobile Networks and Applications*, vol. 16, no. 1, pp. 58–70, 2011.

[23] W. Liu, S. Zhang, and J. Fan, "A diagnosis-based clustering and multipath routing protocol for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2012, Article ID 504205, 11 pages, 2012.

[24] A. Wang, D. Yang, and D. Sun, "A clustering algorithm based on energy information and cluster heads expectation for wireless sensor networks," *Computer and Electrical Engineering*, vol. 38, no. 3, pp. 662–671, 2012.

[25] J. Pu, X. Tang, F. Wang, and Z. Xiong, "A multicast routing protocol with pruning and energy balancing for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2012, Article ID 632836, 11 pages, 2012.

[26] N. Aslam, W. Phillips, W. Robertson, and S. Sivakumar, "A multi-criterion optimization technique for energy efficient cluster formation in wireless sensor networks," *Information Fusion*, vol. 12, no. 3, pp. 202–212, 2011.

[27] L. Karim and N. Nasser, "Reliable location-aware routing protocol for mobile wireless sensor network," *IET Communications*, vol. 6, no. 14, pp. 2149–2158, 2012.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 463605, 8 pages
http://dx.doi.org/10.1155/2013/463605

*Research Article*

# A Design Approach for Controlled Self-Organization-Based Sensor Networks Focused on Control Timescale

**Daichi Kominami and Masayuki Murata**

*Graduate School of Information Science and Technology, Osaka University, 1-5 Yamadaoka, Suita, Osaka 565-0871, Japan*

Correspondence should be addressed to Daichi Kominami; d-kominami@ist.osaka-u.ac.jp

Received 12 December 2012; Accepted 26 April 2013

Academic Editor: J. Barbancho

Copyright © 2013 D. Kominami and M. Murata. 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.

Many researches on network control with a design principle of self-organization have been studied for large-scale networks. Since self-organized control is based on local interactions between system elements, it has high scalability, adaptability, and robustness; however, the management of the whole system is very difficult. In order to solve this problem, a controlled self-organization scheme has been proposed, which aims for desired system behavior by controlling a part of self-organized nodes. Although there are many practical proposals on the scheme, no design approach for it has ever been investigated. In this paper, we propose and evaluate a design approach for the network based on controlled self-organization, paying attention to the control timescale. Through computer simulations, we show the adaptability and stability of the proposed design approach.

## 1. Introduction

Future wireless sensor networks will have massive numbers of elements and should have highly scalable and adaptable properties. However, controlling such a large-scale network is very difficult challenges. For this purpose, self-organization has attracted an increasing attention due to its nature of scalability, adaptability, and robustness [1–3]. Each element in self-organization makes a decision on the basis of local interactions and local rules, which leads to the emergence of global behavior. However, this pure self-organization has some problems because of its bottom-up design [4], such as difficulty of managing the whole network and slow convergence speed after perturbations.

Practical realization of a self-organized network requires complicated emergent behavior to be manageable. However, decision-making based on local interactions in large systems results in emergent behavior, and precise management or control of such behavior is unrealistic. To solve these problems, [5] proposes *controlled self-organization*. The authors of that paper suggest the use of an *observer/controller architecture*, where an *observer* and a *controller* are responsible for correcting system-level behavior. In controlled self-organization, an external observer and controller are responsible for "external control," guaranteeing that system behavior remains within constraints set by the system manager. The main task of the observer is to monitor system behavior by sampling information from a subset of system elements. The controller evaluates the system behavior reported by the observer and performs control actions that influence the system toward a given objective function. This observation/control loop is performed periodically to satisfy system goals. The *observer/controller* architecture is responsible for ensuring the desired behavior of the system, for guaranteeing high system performance, and for encouraging convergence of the system state, thus making the self-organized system *manageable* by controlling some of the self-organization components. Various applications of controlled self-organization are found in [6–8].

Although controlled self-organization is important for the realization of large-scale wireless sensor networks, the potential for unexpected situations due to simultaneous external and self-organized control remains poorly understood. Robustness to network topology change is also important for wireless sensor networks, where changes due to wireless channel conditions, node positions, and the number of nodes are commonplace. If communications protocols are not sufficiently flexible regarding environmental perturbations, various types of performance degradation may occur, such as data collection failures, data delivery delays, and increased energy consumption.

These perturbations and controls in each layer in the wireless sensor network architecture operate on widely different timescales. MAC layer protocols support one-hop communication, where data transmission takes a few milliseconds in most IEEE 802.15.4 sensor networks [9]. Energy-efficient MAC protocols with sleep scheduling for prolonging network lifetime are often assumed in wireless sensor networks, which raises the lower limit of one-hop communication timescales due to the sleep cycles of tens of milliseconds to seconds [10–12]. Routing layer protocols have to deal with topological changes to realize source-to-destination communications. In [13–15], static sensor nodes manage the network topology by using periodic Hello messages every several tens of seconds. The timescale of the external control in controlled self-organization should be longer than that of the routing layer because global behavior of a self-organized network arises as a result of that routing process. Thus, because these control timescales substantially differ, it is insufficient to discuss robustness within only one layer.

In this paper, we propose a design approach for a scalable and robust network based on controlled self-organization, paying attention to the control timescale. We show that a design for robustness in only one layer cannot improve various types of perturbations that cause topological changes. As a solution to these problems, we propose a controlled self-organization-based routing protocol. We apply the controlled self-organization scheme to a potential-based routing and thereby propose *controlled potential-based routing (CPBR)*. Our study considers periodic environmental monitoring systems where sensor nodes deliver monitored data to multiple static sink nodes with CPBR. Then, we discuss how the timescale of control in the MAC, routing, and external control layers should be designed and investigate this through computer simulation.

The rest of this paper is organized as follows. In Section 2, we briefly present each layer's control, and in Section 3, we give descriptions of perturbation models. Then, in Section 4, we explain how to design them. We present the simulation results in Section 5. Finally, we conclude our paper in Section 6.

## 2. Overview of the Each Layer's Control

In this section, we give overviews of CPBR [17], and especially we discuss the control timescale in a MAC layer, routing layer, and external control.

*2.1. Sleep Control in MAC Layer.* One-hop communication is performed in the MAC layer, which takes several milliseconds in the most sensor network scenarios. Therefore, it is difficult to deal with perturbations that cause the topology changes with cycle of a few milliseconds or less. Moreover, in many MAC protocols in the sensor network, the sleep control is assumed, where power-saving operation is expected. For example, B-MAC [10], which is a widely known MAC protocol with the sleep control, allows nodes to sleep every tens of milliseconds to several seconds. Since each node can communicate with its neighbor nodes only when it is awake, the cycle of this sleep control means the minimum unit time of one-hop data transmissions. Hence, such frequent environmental perturbations (~1 ms) are dealt with retransmissions in the MAC layer. We use the intermittent receiver-driven data transmission protocol [16] as a MAC protocol. This protocol is one of the receiver-driven or receiver-initiated MAC protocols where nodes periodically sleep and transmit a beacon, containing an identifier (ID) of the nodes, to inform their neighbors that they are ready to receive data as shown in Figure 1. When a sender node receives a message containing an ID, it returns a send-request message (SREQ) so as to communicate with the sender of the ID.

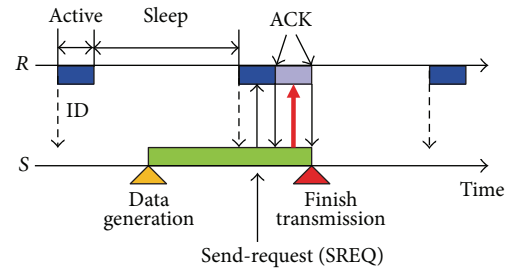



Figure 1: IRDT protocol [16].

*2.2. Route Management in Routing Layer.* CPBR is a kind of potential-based routing protocols, and it utilizes the proactive route management. In a potential-based routing, all nodes have a scalar value "potential." This potential of a node is lower as the hop count from the nearby sink node is smaller. Therefore, a node only forwards data to the neighbor with lower potential than its own for delivering data toward a sink node.

In CPBR, a potential of node $n$ at time $t$, denoted by $\phi(n,t)$, is given by (1). $Z(n)$ is a set of neighbors of node $n$ and $|Z(n)|$ is the size of it. For the calculation of potentials, each node has to manage its neighbors' potential. In order to do that, each node informs its potential to its neighbors periodically. When a node receives a neighbor's potential, it registers the potential of the neighbor, and when it cannot receive any potential from a neighbor during a certain time period, it clears the memory of the neighbor's potential received previously:

$$\phi(n,t+1) = \phi(n,t) + \frac{1}{|Z(n)|}\sum_{k\in Z(n)}\{\phi(k,t) - \phi(n,t)\}. \quad (1)$$

*2.3. External Control.* CPBR presumes a multisink sensor network and it performs global control of a potential field in order to balance the traffic loads of sink nodes, which is difficult to manage only by local interactions and rules. In CPBR, a control node, which is able to communicate with all sink nodes, is responsible for observing and controlling of potentials of all sink nodes. The control node controls potential of sink node $d$ at time $t$, denoted by $\Phi(d,t)$, via (2). $m$ is a metric for the control given by the network manager. Then, $m(d,t)$ is collected from sink node $d$ periodically (every $T_m$), and $\overline{m(t)}$ is the average of the metric at time $t$. Potentials

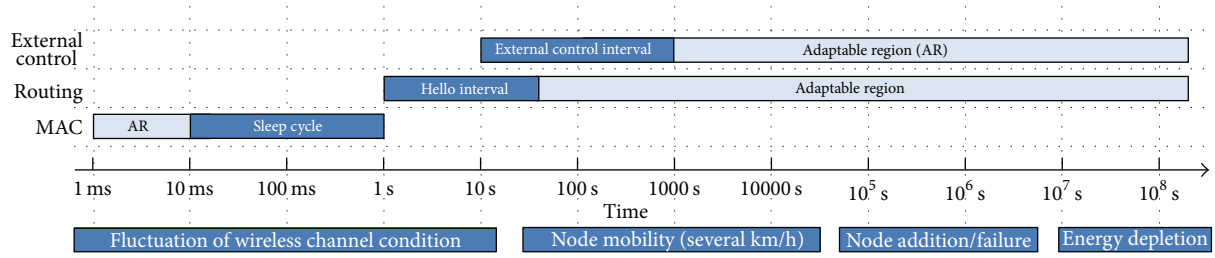



Figure 2: Intervals of environmental changes and each layer's control.

of all sink nodes are controlled by (2) and calculated simultaneously. The constant value of $\theta$ means the intensity of the control. In this paper, we employ the number of received data just before $T_m$ as a metric of the external control. This can result in the global load balancing of sink nodes:

$$\Phi(d, t+1) = \Phi(d, t) * \left(1 - \theta \frac{m(d,t) - \overline{m(t)}}{\overline{m(t)}}\right). \quad (2)$$

## 3. Perturbation Model

We assume four types of perturbations that cause topological changes in the network.

*3.1. Varying Wireless Channel Condition.* The burst packet errors occur in various timescale as mentioned in [18]. Then, we assume that burst packet errors happen due to varying wireless channel condition according to the Gilbert-Elliot model [19]. In this model, wireless channel is described with two-state Markov chain; that is, each link has two conditions "good" and "bad," respectively and alternates the conditions stochastically. In this paper, when a condition of a link is "good," no bit error occurs in the link and when "bad," bit error and packet loss always happen. The probabilistic transition of the channel condition occurs at fixed cycles $T_c$.

*3.2. Node Mobility.* The mobility of individual sensor node (except for sink nodes) is based on the random waypoint model [20]. A node determines a destination and moves there with constant speed. After arriving at the destination, it pauses for a definite period of time and moves for a new destination again. This destination and speed are randomly chosen. This mobility brings about quasicyclic changes to the network topology.

*3.3. Node Addition/Failure.* We assume a random addition and failure of a number of sensor nodes. Since such node addition is carried out in a planned manner, we refer to it as periodical perturbations. Mean time to failure of devices in the network also means that node failure is of a cyclical nature. This node addition occurs at the same time in the simulation, and the same is true for node failures.

## 4. Design Approaches of the Control Timescale

In this section, we present design approaches for a controlled self-organization-based network particularly focused on control timescales in a MAC layer, a routing layer, and an external control. An overview of their timescales is illustrated in Figure 2 along with those of perturbations.

*4.1. MAC Layer Design.* As to changes of wireless channel conditions that arise with the cycle of 1 ms to 1,000 ms, retransmission in a MAC layer is important. In a MAC layer, a node obtains more opportunities to detect a next-hop node when the cycle of sleep control is shorter, in case the node holds a data for a certain period of time (denoted by $T_d$) until it finishes forwarding the data to the next-hop node. However, as to the changes with a cycle shorter than this, a MAC layer cannot handle them fundamentally, and we need to choose a robust modulation method against severe changes of radio in the physical layer.

*4.2. Routing Layer Design.* When movement, additions, and failures of nodes occur, latest route information is necessary for data delivery. Therefore, correct selection of a next-hop node is attained as the updating cycle gets shorter. Similar to possibly supposed scenarios on wireless sensor networks, our research supports static and comparatively slow mobility of nodes, taking account of monitoring application of human health, animal behavior, and so forth. Then, every tens of seconds of periodical messages are used for neighbor detection, and message exchanges to maintain route information.

*4.3. External Control Design.* Comparatively, long-term perturbations such as movement, additions, and failures of sensor nodes may cause global topological changes, which cannot be dealt with by self-organized routing protocols based on only local information. Thus, since these perturbations degrade the performance of such routing protocols, the control and observation mechanism is required for normal operation.

Similar to the principle of routing layer design, the shorter control and observation cycle seem to be better. However, this cycle is closely bound together with the cycle of self-organized route construction in the routing layer, and therefore, the external control process and self-organized routing process can interfere mutually. In addition, convergence speed of self-organized methods is generally slow, and when the external control is conducted before routes do converge, a system does not satisfy the desired performance.

In order to examine the convergence speed of self-organized potential calculation, first we show analytical solution

of the 2-dimensional diffusion equation: $\partial\phi(x,y,t)/\partial t = D\nabla^2\phi(x,y,t)$. Here, we change the Cartesian coordinates $(x,y)$ to polar coordinates $(r,\theta)$ in order to reduce one of variables ($r_{\min} \le r \le r_{\max}$ and $-\pi \le \theta \le \pi$). Since we consider symmetric diffusion of potential from the origin, the solution of the equation is independent of angular coordinate $\theta$. Then, the diffusion equation is converted into polar coordinates as follows:

$$\frac{\partial}{\partial t}\phi(r,t) = D\left(\frac{\partial^2}{\partial r^2}\phi(r,t) + \frac{1}{r}\frac{\partial}{\partial r}\phi(r,t)\right). \quad (3)$$

Various boundary conditions can be found in natural world, and we assume two simple Dirichlet boundary conditions: $\phi(r_{\min},t) = \phi_{\min}$ and $\phi(r_{\max},t) = \phi_{\max}$ ($\phi_{\min} < \phi_{\max}$). The solution of (3) under the conditions is represented by (4), which is a sum of exponential functions:

$$\phi(r,t) = \sum_{n=0}^{\infty} A_n e^{-q_n^2 Dt} R(r,n) + C(r). \quad (4)$$

In the solution, $q_n$ and $A_n$ are functions of constant number $n$. Here, $q_n$ ($n = 0,1,2,\ldots$) is the real root of the following equation and satisfies the condition $q_k < q_{k+1}$ for any nonnegative integer number $k$:

$$J_0(\phi_{\min}q_n)\,Y_0(\phi_{\max}q_n) - Y_0(\phi_{\min}q_n)\,J_0(\phi_{\max}q_n) = 0, \quad (5)$$

where $J_0(x)$ and $Y_0(x)$ are the zero-order Bessel function of the first kind and the zero-order Bessel function of the second kind, respectively. $A_n$ depends on an initial condition and given an initial condition $\phi(r,0) = 0$, $A_n$ is calculated according to the following equation:

$$A_n = -\frac{\pi^2 q_n^2}{2}\frac{Y_0^2(q_n r_{\max})\,J_0^2(q_n r_{\min})}{J_0^2(q_n r_{\min}) - J_0^2(q_n r_{\max})} \cdot \int_{r_{\min}}^{r_{\max}} r\,(a\log(r) + b) \cdot \left(J_0(q_n r) - \frac{J_0(q_n r_{\max})}{Y_0(q_n r_{\max})}Y_0(q_n r)\right) dr. \quad (6)$$

$R(r,n)$ is a function only dependent on $n$ and radial coordinate $r$ as represented in the following equation:

$$R(r,n) = J_0(q_n r) - \frac{J_0(q_n r_{\min})}{Y_0(q_n r_{\min})}Y_0(q_n r). \quad (7)$$

$C(r)$ is represented by a basic logarithm function, $a\log(r)+b$, where $a$ and $b$ are constant number and calculated as follows: $a = (\phi_{\max}-\phi_{\min})/(\log(r_{\max})-\log(r_{\min}))$, $b = \phi_{\min} - ((\phi_{\max}-\phi_{\min})/(\log(r_{\max})-\log(r_{\min})))\log(r_{\min})$.

From (4), it can be found that the potential $\phi(r,t)$ exponentially converges without relying on the distance from the potential source, but relying on time. In [21], the authors point that the solution of the discrete diffusion equation also exponentially converges. From the above discussion, we could obtain an approximate solution of the diffusion equation. If the solution is represented by a basic exponential function, $f(x) = u\,e^{-x/\tau} + v$, convergence of the potential can be estimated using time constant $\tau$. It is worth noting that calculation of $\tau$ requires the value of $\phi$ after convergence. Therefore, in order to understand the convergence behavior of the system, computer simulation is one of the means.

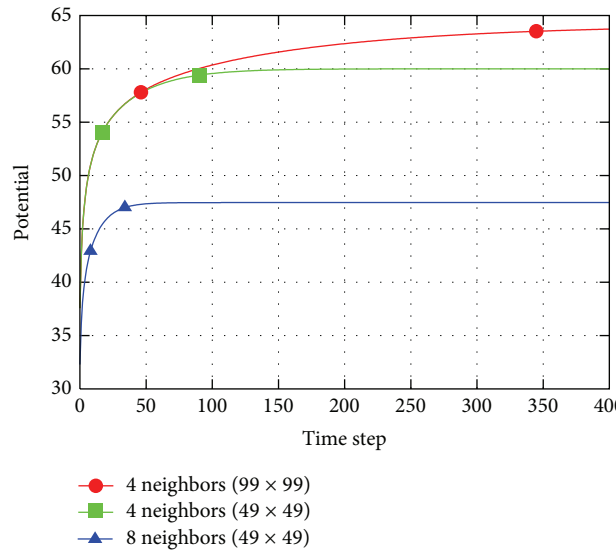



Figure 3: Potential convergence in grid networks.

For an example of the potential convergence, in Figure 3, the simulation results of the potential convergence in two grid networks (99 × 99 and 49 × 49) are shown. The center node is a potential source (potential is 100) and the outer circumferential nodes have potential of zero. Each node performs potential exchanges with its nearby four or eight nodes and updates its own potential according to (1) every time step. In the figure, the horizontal axis means the time step and the vertical axis is potential of a neighbor node of the center node. The symbols (circle, square, and triangle) mean 90% and 99% convergence from an initial value of zero in each result. From the results, convergence speed becomes quicker as a network size becomes small and as communication range increases. Figure 4 shows the potential convergence in a random network with 100 sensors. Due to the random deployment of sensor nodes, the convergence speed in this case is slower than that in grid cases.

## 5. Simulation Results

In this section, we evaluate the packet delivery ratio under the periodical environmental perturbations. We use an event-driven simulator written by C++ for evaluation. For a network model, we deploy 100 sensor nodes at random places over the square region 500 m on a side and install a sink node in a corner of the domain. Each sensor node generates one data every 500 s, and it is delivered to the sink node in a multihop manner. For a communication model, we utilize the disk model, and communication between two nodes within communication range is successful unless a message collision occurs or wireless channel condition between the nodes is

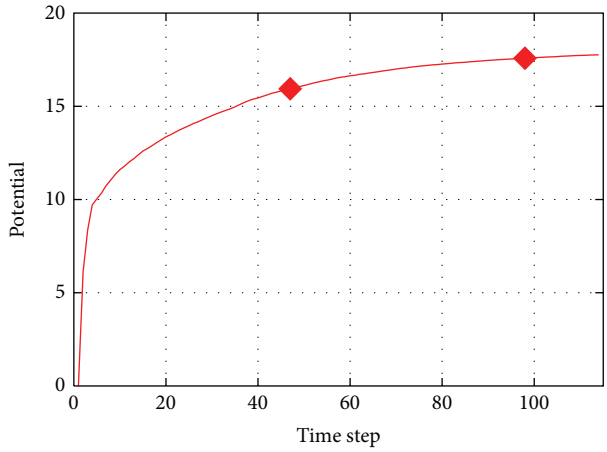



Figure 4: Potential convergence in random networks.

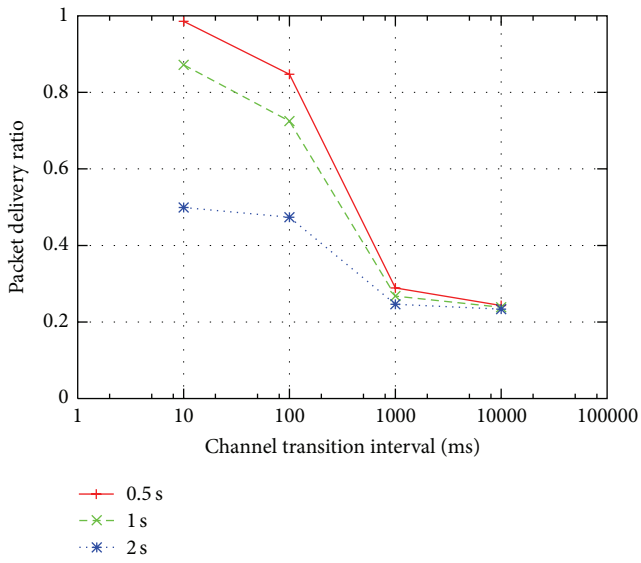



Figure 5: Packet delivery ratio against channel condition transition.

bad. The main parameters in a simulation are shown in Table 1.

*5.1. Transitions of Channel Conditions.* When the cycle of the sleep control in the MAC layer is set to 0.5 s, 1.0 s, and 2.0 s, respectively, the data delivery ratio against the periodic transition of the channel condition is shown in Figure 5. When the transition of the channel condition arises with the cycle of 10 ms and 100 ms, it turns out that shorter sleep control cycles are required for a high data delivery ratio. Since the MAC layer quickly responds to change in the channel conditions and the opportunity of the retransmission in the MAC layer increases as a sleep control cycle is shorter, even if there is no support in an underlying layer, perturbations with shorter cycle are absorbed. On the other hand, when perturbations occur with the cycle more than 1,000 ms, the delivery performance deteriorates greatly, and above the cycle, the MAC layer cannot handle perturbations. Therefore, it is essential to cope with such perturbations in a higher layer.

Table 1: Setting of parameters.

| Parameters | Value |
|---|---|
| Communication range | 100 m |
| Time to live (TTL) | 32 hops |
| $T_d$ | 5 s |
| Channel-condition transition probability (good to bad) | 30% |
| Channel-condition transition probability (bad to good) | 70% |
| Node speed | 4–6 km/h |
| Pause time | 250–350 s |
| Memory span for neighbor potential | 250 s |
| Update interval of potential | 50 s |

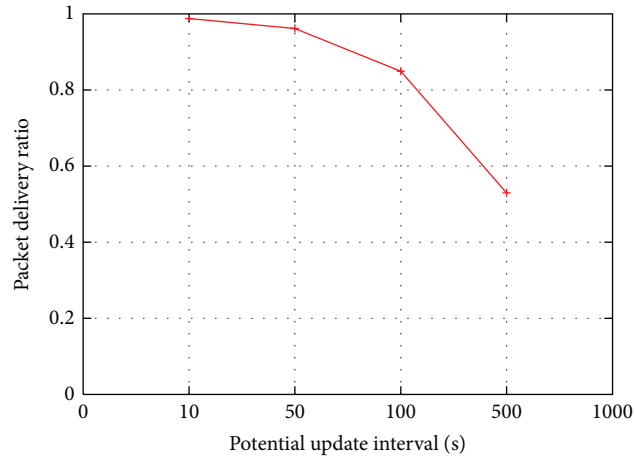



Figure 6: Packet delivery ratio against node mobility.

*5.2. Node Mobility.* Here, we set the same value to the cycles of potential advertisement and update. In order to eliminate the influence of perturbations other than the cycles, the number of the maximum relay has a sufficiently large value so that there may be no excess. Figure 6 shows that the delivery ratio is decreasing as the update cycle of potential becomes large. It is because as longer the update cycle is, the higher the possibility that a potential of a node is incorrect, and as a result, the node transmits data away from the sink node, or discards the data since there have already been no appropriate next-hop neighbor nodes.

Given the update cycle of potential is $T_i$ (s), when a data is generated at a node at a certain time, the elapsed time from the last update is $(1/2)T_i$ on average. Since we assume that nodes move at 4–6 kilometers per hour, the displacements of nodes from the last update are presumed to be $0.55T_i$–$0.83T_i$ (m). In our network model, where 100 sensor nodes and square domain with a 500 m side are assumed, the average distance with the nearest node is about 50 m, and therefore, connectivity with the nearest node can change with the cycle of a 10 s order. It is obvious that connectivity between other neighbor nodes can change with much shorter cycle. Therefore, in this network model, it is desirable that the value of the update cycle is at least shorter than 10 s, and when it is

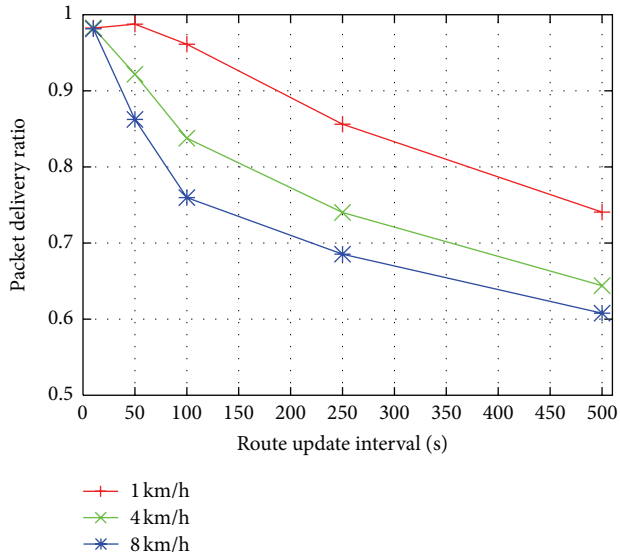


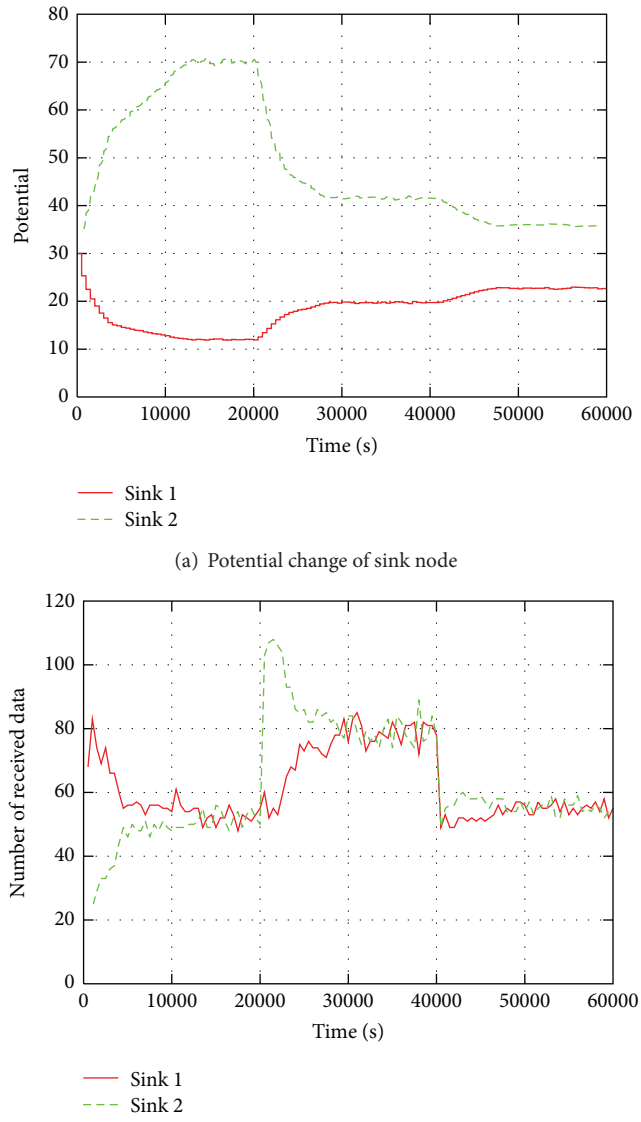

Figure 7: Packet delivery ratio against node mobility with constant velocity.

set to 10 s, from the simulation result, the delivery ratio more than 95% is obtained. Figure 7 also shows the packet delivery ratio when nodes moves at constant velocity and this result exhibits a similar relationship.

*5.3. Cross-Layer Interaction.* Here, we set two sink nodes at the center (sink 1) and a corner (sink 2) of the network. In Figures 8 and 9, we show the potential of two sink nodes (Figures 8(a) and 9(a)) and the number of received data by the sink nodes at every control interval (Figures 8(b) and 9(b)). In those results, 50 sensor nodes are added at a random position at time 20,000 s, and random 50 sensor nodes fail at 40,000 s. The potential update cycle is set to 50 s as shown in Table 1, and from the preliminarily experiment, it is found that simulation time of 500 s (100 s) can obtain the 99% (90%) convergence of the potential at each node.

The potential of two sink nodes is controlled by (2) so that the number of received data mutually becomes equal for every fixed cycle. At the beginning of the simulation, more data arrive at sink 1. Then, the control node makes potential of sink 1 up in order to reduce the number of received data by sink 1. Equalization of the received data by the sink nodes is attained at 12,000 s as shown in Figure 8(b) and their potential is also converged. Furthermore, equalization of the received data is attained right from the beginning as shown in Figure 9(b).

Meanwhile, some changes take place to the number of the received data immediately after 20,000 s when addition of 50 nodes occurs. In addition, in this case, convergence finishes within about 10,000 s (or immediately after the perturbation). Shortly after the failures of sensor nodes at 40,000 s, the number of received data decreases. It is because a convergence commences after nodes erase the potential of failed nodes. The time for erasing depends on the potential memory span shown in Table 1. It turns out that after failure, as well as the addition, potential converges.



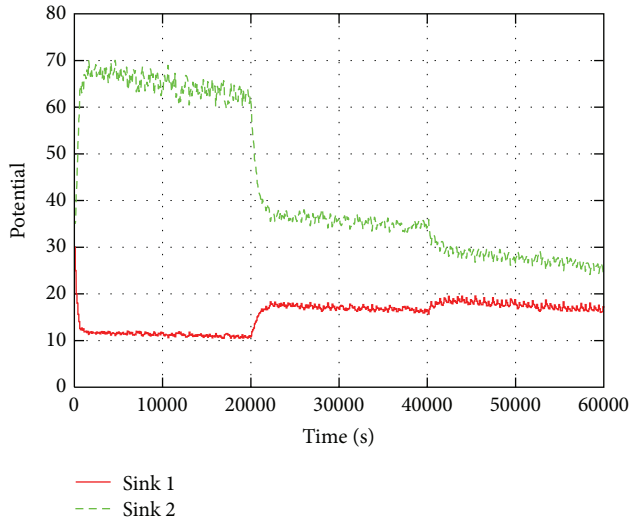

(a) Potential change of sink node

(b) Number of received data by sink node

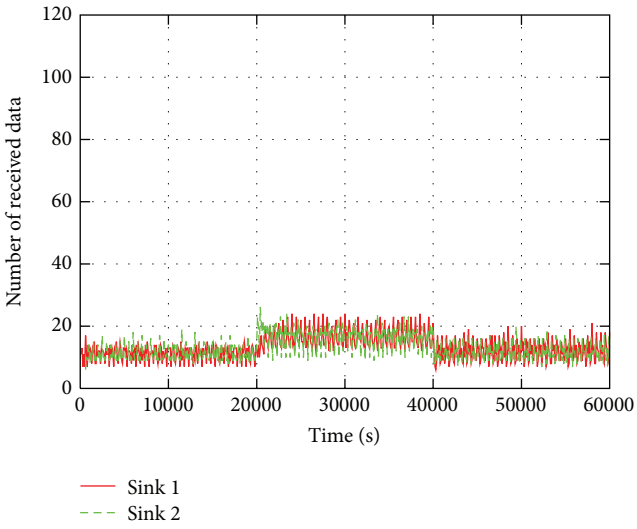
Figure 8: Potential control in case of 2 sinks (control interval 500 s).

## 6. Conclusion

In this paper, we discussed an approach for network design based on controlled self-organization. This approach is for future large-scale and complex networks. As an example of networks based on controlled self-organization, we focus on a wireless sensor network where a self-organized routing protocol and an external control mechanism are applied. In particular, our concern is on cyclic nature of the environmental perturbation. In order to obtain robustness of a system against environmental perturbations, multiple network layers should not handle them separately, but should cope with in a coordinated fashion. We show that our approach can deal with various perturbations by appropriately defining the control timescale of each layer. Further investigation on other perturbations and networks and validation with real experiments are our future work.




(a) Potential change of sink node




(b) Number of received data by sink node

FIGURE 9: Potential control in case of 2 sinks (control interval 100 s).

## Acknowledgment

This research was supported in part by "Grant-in-Aid for JPSP Fellows (24738)" of the Japan Society for the Promotion of Science (JSPS) in Japan.

## References

[1] C. Prehofer and C. Bettstetter, "Self-organization in communication networks: principles and design paradigms," *IEEE Communications Magazine*, vol. 43, no. 7, pp. 78–85, 2005.

[2] A. Banerjee, R. Agarwal, V. Gauthier, C. K. Yeo, H. Afifi, and F. Lee, "A self-organization framework for wireless ad hoc networks as small worlds," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 6, pp. 2659–2673, 2012.

[3] H. Zhang, J. Llorca, C. C. Davis, and S. D. Milner, "Nature-inspired self-organization, control, and optimization in heterogeneous wireless networks," *IEEE Transactions on Mobile Computing*, vol. 11, no. 7, pp. 1207–1222, 2012.

[4] C. Müller-Schloer, H. Schmeck, and T. Ungerer, *Organic Computing—A Paradigm Shift for Complex Systems*, Autonomic Systems, Springer, 2011.

[5] J. Branke, M. Mnif, C. Müller-Schloer et al., "Organic Computing—a ddressing complexity by controlled self-organization," in *Proceedings of the 2nd International Symposium on Leveraging Applications of Formal Methods, Verification and Validation (ISoLA '06)*, pp. 185–191, November 2006.

[6] F. Nafz, J. P. Steghöfer, H. Seebach, and W. Reif, "Formal modeling and verification of self-* systems based on observer/controller-architectures," in *Assurances for Self-Adaptive Systems*, pp. 80–111, Springer, 2013.

[7] F. Allerding, M. Premm, P. K. Shukla, and H. Schmeck, "Electrical load management in smart homes using evolutionary algorithms," in *Evolutionary Computation in Combinatorial Optimization*, pp. 99–110, Springer, 2012.

[8] B. M. Ali and Z. Nacereddine, "Web service-based emergence control in cooperative information systems," in *Proceedings of International Conference on Information Technology and e-Services (ICITeS '12)*, pp. 1–5, March 2012.

[9] J. Zheng and M. J. Lee, "A comprehensive performance study of IEEE 802.15.4," *Sensor Network Operations*, pp. 218–237, 2004.

[10] J. Polastre, J. Hill, and D. Culler, "Versatile low power media access for wireless sensor networks," in *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems (SenSys '04)*, pp. 95–107, November 2004.

[11] M. Buettner, G. V. Yee, E. Anderson, and R. Han, "X-MAC: a short preamble MAC protocol for duty-cycled wireless sensor networks," in *Proceedings of the 4th International Conference on Embedded Networked Sensor Systems (SenSys '06)*, pp. 307–320, November 2006.

[12] L. Tang, Y. Sun, O. Gurewitz, and D. B. Johnson, "PW-MAC: an energy-efficient predictive-wakeup MAC protocol for wireless sensor networks," in *Proceedings of the 30th IEEE International Conference on Computer Communications (INFOCOM '11)*, pp. 1305–1313, April 2011.

[13] K. Zeng, K. Ren, W. Lou, and P. J. Moran, "Energy aware efficient geographic routing in lossy wireless sensor networks with environmental energy supply," *Wireless Networks*, vol. 15, no. 1, pp. 39–51, 2009.

[14] J. Heo, J. Hong, and Y. Cho, "EARQ: energy aware routing for real-time and reliable communication in wireless industrial sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 5, no. 1, pp. 3–11, 2009.

[15] Y. M. Lu and V. W. S. Wong, "An energy-efficient multipath routing protocol for wireless sensor networks," *International Journal of Communication Systems*, vol. 20, no. 7, pp. 747–766, 2006.

[16] D. Kominami, M. Sugano, M. Murata, and T. Hatauchi, "Energy-efficient receiver-driven wireless mesh sensor networks," *Sensors*, vol. 11, no. 1, pp. 111–137, 2011.

[17] D. Kominami, M. Sugano, M. Murata, and T. Hatauchi, "Controlled potential-based routing for large-scale wireless sensor networks," in *Proceedings of The 14th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM '11)*, pp. 187–195, October 2011.

[18] T. Rusak and P. Levis, "Burstiness and scaling in the structure of low-power wireless links," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 13, no. 1, pp. 60–64, 2009.

[19] E. N. Gilbert, "Capacity of a burst-noise channel," *Bell System Technical Journal*, vol. 39, no. 9, pp. 1253–1265, 1960.

[20] D. B. Johnson and D. A. Maltz, "Dynamic source routing in Ad Hoc wireless networks," in *Mobile Computing*, pp. 153–181, 1996.

[21] A. Cunha, R. Teixeira, and L. Velho, "Discrete scale spaces via heat equation," in *Proceedings of the 14th Brazilian Symposium on Computer Graphics and Image Processing*, pp. 68–75, IEEE, October 2001.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 652051, 12 pages
http://dx.doi.org/10.1155/2013/652051

*Research Article*

# A Credible Routing Based on a Novel Trust Mechanism in Ad Hoc Networks

**Renjian Feng, Shenyun Che, Xiao Wang, and Ning Yu**

*School of Instrumentation Science and Opto-Electronics Engineering, Beihang University, Beijing 100191, China*

Correspondence should be addressed to Renjian Feng; rjfeng@buaa.edu.cn

Received 8 January 2013; Accepted 20 March 2013

Academic Editor: Adel Soudani

Copyright © 2013 Renjian 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.

Many existing routing protocols in Mobile Ad hoc Networks (MANETs) focus on finding paths in dynamic networks without considering security. In this paper, we propose a trust model which evaluates neighbours' direct trust by factors of encounter time, mobility, and successful cooperation frequency. The revised D-S evidence theory is used to combine multiple recommended pieces of evidence and obtain the recommended trust value. Then based on the novel trust mechanism, we propose a trusted routing protocol named TDS-AODV protocol by extending the AODV protocol. In this protocol, a node makes a routing decision according to the trust values of its neighbour nodes. Finally, two routes are built: the main route with highest route trust value in the candidate routes and a backup route. Simulation results reveal that TDS-AODV can eliminate malicious nodes effectively when building the route; furthermore, it also achieves better performance than TAODV and AODV in terms of throughput, packet delivery ratio, and average end to end delay.

## 1. Introduction

The past decade has witnessed tremendous research efforts devoted to Mobile Ad hoc Networks (MANETs). MANETs are temporary autonomous systems with the special characteristics of dynamic network topology, limited computational abilities, and continuously changing scale. Due to its flexibility, a MANET is attractive for applications, such as disaster relief, military service, and robot networks [1]. However, this flexibility also causes security problems. Routing security is one of the challenging issues in current research.

Traditional MANET routing protocols, such as destination-sequenced distance vector routing (DSDV) [2], dynamic source routing (DSR) [3], and ad hoc on-demand distance vector routing (AODV) [4], assume that all nodes in the network work in a benevolent manner and no predefined trust exists between communication partners. However, the fact is that malicious behavior among nodes exists; for example, selfish nodes deny relaying the packets of other nodes, and malicious nodes perform impersonation, fabrication, or modification attacks against the network traffic [5]. Hence, it is necessary to incorporate security mechanisms into MANET routing protocols to mitigate the impairment from malicious nodes. However, the security mechanism basing on the traditional cryptosystem is used to resist external attacks, but it cannot effectively solve the internal attacks by malicious nodes [6]. Therefore, the trust mechanism which is considered to be an effective measure to solve those questions has recently been studied.

In our trust mechanism, the successful cooperation frequency factor is considered in direct trust evaluation to guarantee the security of network. It is calculated according to its accumulated observations using the Bayesian inference which adopts Beta distribution. Unlike most trust mechanisms [7–12] that focus on trust evaluation without considering performance of the network, we take other two factors (factors of encounter time and mobility) into account. A good network performance can help save nodes' limited resources and prolong the network lifetime, which is very important in MANET. The network topology in MANET is dynamic; hence, the next hop of a node may not be its next hop the next moment. To create a relatively stable network topology as much as possible, we propose two factors, nodes' average encounter time and mobility, when calculating nodes' direct trust value. The two factors make the trust mechanism more suitable for resource-restricted MANET. D-S evidence

theory, which was first introduced by statistician of Dempster [13] and extended by Shafer [14], is used to calculate direct trust value, integrate indirect evidence, and obtain the overall trust value. We choose D-S evidence here because it does well in dealing with the uncertainty of trust value.

Based on the novel trust mechanism, we put forward a trusted routing protocol, by extending the AODV protocol in MANETs, TDS-AODV for short. In this protocol, a node evaluates its neighbours' trust value according to the trust model and selects reliable nodes as its next-hop nodes. A source can establish multiple reliable paths to a destination in one route discovery process. We consider the number of hops as well as the trust value of paths to the destination. A destination will respond with three shortest paths as candidates and the path trust will be calculated during the process of Route Reply (*RREP*) Message delivery. The one with maximum path trust will be selected as the forwarding route and the second reliable one will be regarded as the backup route. We perform some simulations to compare the performance of TAODV, AODV, and TDS-AODV on Matlab platform. Simulation results show that our method is practical to detect malicious nodes and outperform TAODV and AODV in throughput, packet delivery ratio, and average end to end delay.

The rest of this paper is organized as follows. Section 2 summarizes the related work on trust evaluation and trust-based routing protocols. Section 3 presents the novel trust evaluation mechanism. Section 4 describes TDS-AODV in detail. Section 5 provides the simulation studies. Finally, we conclude this paper in Section 6.

## 2. Related Works

Researchers are becoming more and more interested in integrating trust into a MANET and have proposed numerous works. In this section, we first focus our attention on trust evaluation models in MANETs and then discuss the trust based routing protocols in MANETs.

*2.1. Trust Evaluation.* Peng et al. [7] assessed the subjective trust of nodes through the Bayesian method, but they were not able to detect dishonest recommendations. Zouridaki et al. [8] chose to determine the node trustworthiness with respect to reliable packet forwarding by combining first-hand trust and second-hand trust information. However, the trust calculation in unsupervised ad hoc environment involved complex aspects such as availability and mobility. besides packet forwarding. Omar et al. [9] sought to establish a fully distributed trust model based on trust graphs and threshold cryptography.

At present, most of the trust evaluation literatures ignore the uncertainty of trust value. To deal with this problem, some researchers [10–12] resort to D-S evidence theory. D-S evidence theory has the capacity of expressing directly for "uncertain," which makes it suitable to calculate the trust value in MANETs. Xie et al. [10] proposed a trust model for MANETs based on D-S evidence theory. The model can be a good solution for the combination of pieces of evidence, but it failed in addressing the issues concerning conflicting recommendation pieces of evidence. In this paper, we adopt the revised D-S combination rule which includes a consistent intensity to calculate nodes' trust value.

*2.2. Trust-Based Routing Protocols.* Wang and Wu [15] introduced the trust metric which depended on network traffic statistics to evaluate the trust and then loaded the trust model on the previously proposed distance-based location-aided routing (LAR). The algorithm utilized direct trust and recommendation trust to prevent malicious nodes from joining the forwarding. Li et al. [16] built a simple trust model to evaluate neighbours' behaviours forwarding packets and proposed a trust-based reactive multipath routing protocol extending from AODV. Peng et al. [17] incorporated a new dynamic trust mechanism which was based on multiple constraints and collaborative filtering into the extending DSR. Narula et al. [18] selected soft encryption systems and implemented them in conjunction with a trust-based reputation system and a multipath routing to provide a secure routing scheme. The implementation of this trust-based approach using DSR was then discussed. Sirotheau and Sousa [19] proposed an evaluation mechanism that aimed to mitigate routing misbehavior and other network failures. Four attributes of the routes were considered: level of activity, trust, mobility, and number of hops.

When transmitting a packet to a given destination, a node may have two routes: one is short but incredible while the other is long but credible. One of our main aims is to design a rational strategy which involves both hop counts and trust values in making decisions. The detailed implementation of our scheme is a secure extension of the AODV. Because of its ability to cope with network dynamic changes and repair broken links in routes, AODV is one of the promising protocols for deployment in a MANET.

## 3. Trust Model Based on D-S Evidence Theory

Trust model essentially performs trust derivation, computation, and application [20]. Trust applications including trust-based route discovery and route selection will be discussed in the next section.

*3.1. D-S Evidence Theory.* D-S evidence theory is based on the identification frame $\Omega$ set comprised by basic propositions which are both exclusive and exhaustive. $2^{\Omega}$ is the power set of $\Omega$, that is, the set of all the possible propositions based on $\Omega$. Here we define $\Omega$ as $\{T, -T\}$, where $T$ and $-T$ represent two trust states, namely, credible and incredible. $2^{\Omega}$ is $\{\emptyset, \{T\}, \{-T\}, \{T, -T\}\}$, in which $\emptyset$, $\{T\}$,$\{-T\}$, and $\{T, -T\}$ represent the empty set, the propositions of nodes' "Trust", "Distrust", and "Uncertain", respectively. There are definitions of basic reliability function $m$ on $2^{\Omega}$: $2^{\Omega} \rightarrow [0, 1]$, Belief Bel: $2^{\Omega} \rightarrow [0, 1]$ and Plausibility Pl: $2^{\Omega} \rightarrow [0, 1]$, satisfying the following equations:

$$m(\emptyset) = 0,$$

$$\sum_{A \subseteq \Omega} m(A) = 1, \quad A \neq \emptyset,$$

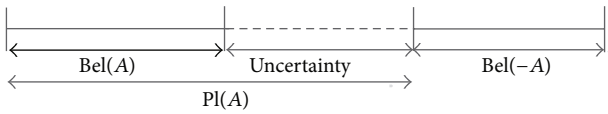



Figure 1: Belief (Bel) and Plausibility (Pl).

$$\begin{aligned} \mathrm{Bel}(A) &= \sum_{B \subseteq A} m(B), \quad \forall A \subseteq \Omega, \\ \mathrm{Pl}(A) &= 1 - \mathrm{Bel}\left(\overline{A}\right), \quad \forall A \subseteq \Omega, \end{aligned} \tag{1}$$

where $A$ is named focal element, $m(A) > 0$ is the basic confidence level of $A$, representing how much the evidence supports $A$ to happen.

The difference between Belief and Plausibility is referred to as Belief Interval. It is represented by the range of maximum uncertainty. The relationship of Belief and Plausibility is shown in Figure 1.

*3.2. Trust Factors.* The definition of "Trust" in this paper refers to the confidence that node $i$ has on node $j$ about the ability to forward packets successfully. Nodes tend to select the neighbour that has higher trust value as the intermediate node. In general, the trust between nodes only has some connection with malicious behaviors; however, we should consider more factors that depend on the interactions between neighbour nodes in a MANET due to its flexibility.

*3.2.1. Factor of Average Encounter Time $ACF_{i,j}$.* The concept of average encounter time does well in quantifying node's encounter history record. Encounter means that two nodes enter each other's wireless transmission range. The larger the $\mathrm{ACF}_{i,j}$ is, the more possibly node $i$ chooses node $j$ as the next hop. The $\mathrm{ACF}_{i,j}$ during period $T$ is calculated by the following equation

$$\mathrm{ACF}_{i,j} = \frac{\sum_{t=0}^{t=T} \delta_{i,j}}{T}. \tag{2}$$

If two nodes enter each other's wireless transmission range $\delta_{i,j} = 1$, else $\delta_{i,j} = 0$. For example, in Figure 2, node $i$ and node $j$ encounter three times during period $T$; the $\mathrm{ACF}_{i,j}$ is:

$$\mathrm{ACF}_{i,j} = \frac{\sum_{t=0}^{t=T} \delta_{i,j}}{T} = \frac{T_2 - T_1 + T_4 - T_3 + T_6 - T_5}{T}. \tag{3}$$

*3.2.2. Factor of Mobility $MOL_{i,j}$.* The topology of MANET is dynamic due to the node movement; hence, in order to establish a more stable routing, it is necessary to take the node mobility into account when a node selects its cooperative nodes. The factor $\mathrm{MOL}_{i,j}$ is constructed as

$$\begin{aligned} \mathrm{MOL}_{i,j} &= \left| \frac{d_{i,j}(t+T) - d_{i,j}(t)}{d_{i,j}(t+T) + d_{i,j}(t)} \right|, \\ d_{i,j}(t) &= \sqrt{\left(x_i(t) - x_j(t)\right)^2 + \left(y_i(t) - y_j(t)\right)^2}, \end{aligned} \tag{4}$$

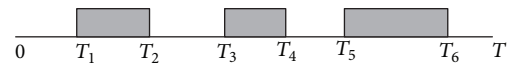



Figure 2: Average encounter time.

where $d_{i,j}$ denotes the distance between node $i$ and node $j$ at time $t$, $(x_i(t), y_i(t))$ and $(x_j(t), y_j(t))$ are the coordinates of node $i$ and node $j$ at time $t$, respectively.

*3.2.3. Factor of Successful Cooperation Frequency $SCF_{i,j}$.* Node $i$ has a detection mechanism to obtain its interaction results $\mathrm{record}_{i,j} = (\alpha_{i,j}, \beta_{i,j})$ with node $j$. $\alpha_{i,j}$ and $\beta_{i,j}$, respectively, denote the number of successful cooperation and unsuccessful cooperation about node $j$ observed by node $i$. Suppose $\mathrm{SCF}_{i,j}$ can be easily expressed by beta distribution, that is, $\mathrm{SCF}_{i,j} \sim \mathrm{Beta}(\alpha_{i,j}, \beta_{i,j})$. The factor $\mathrm{SCF}_{i,j}$ is constructed as

$$\mathrm{SCF}_{i,j} = \frac{\alpha_{i,j}}{\alpha_{i,j} + \beta_{i,j}}. \tag{5}$$

Once node $j$ behaves badly, $\beta_{i,j}$ will increase and $\mathrm{SCF}_{i,j}$ will decrease, which leads to the decrease of the possibility that node $i$ chooses node $j$ as the next hop.

*3.3. Direct Trust.* Subject node $i$ monitors the behaviors of object node $j$ in one cycle and acquires the current trust value $\mathrm{CDT}_{i,j} = (m_{i,j}^C(\{T\}), m_{i,j}^C(\{T, -T\}), m_{i,j}^C(\{-T\}))$ based on the following expression:

$$\begin{aligned} m_{i,j}^C(\{T\}) &= \frac{\left(\omega_1 * \mathrm{ACF}_{i,j} + \omega_2 * \left(1 - \mathrm{MOL}_{i,j}\right) + \omega_3 * \mathrm{SCF}_{i,j}\right)}{\left(\sum_{k=1}^{3} \omega_k\right)}, \\ m_{i,j}^C(\{-T\}) &= \frac{\left(\omega_1 * \left(1 - \mathrm{ACF}_{i,j}\right) + \omega_2 * \mathrm{MOL}_{i,j} + \omega_3 * \left(1 - \mathrm{SCF}_{i,j}\right)\right)}{\left(\sum_{k=1}^{3} \omega_k\right)}, \\ m_{i,j}^C(\{T, -T\}) &= 1 - m_{i,j}^C(\{T\}) - m_{i,j}^C(\{-T\}), \end{aligned} \tag{6}$$

where $0 < \omega_k < 1$, $k = 1, 2, 3$, $\omega_k$ are determined by specific application environment, usually $\omega_3 > \omega_1$, $\omega_3 > \omega_2$ as security is more important.

Furthermore, the direct trust value is recalculated in accordance with history records. Assuming the direct trust value of latest cycle is $\mathrm{HDT}_{i,j}$, the update of direct trust value is calculated as follows:

$$\mathrm{DT}_{i,j} = \gamma \times \mathrm{HDT}_{i,j} + (1 - \gamma) \times \mathrm{CDT}_{i,j}, \tag{7}$$

where $\mathrm{DT}_{i,j}$ is the direct trust value of subject node $i$ on object node $j$ in current cycle, parameter $\gamma$ is the adaptive time factor

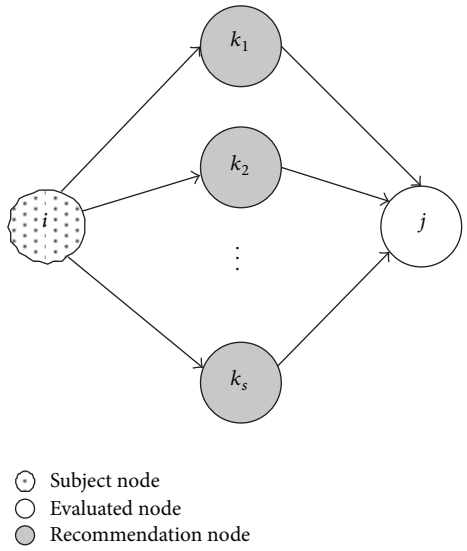



Figure 3: Recommendation relationship between subject node $i$ and object node $j$.

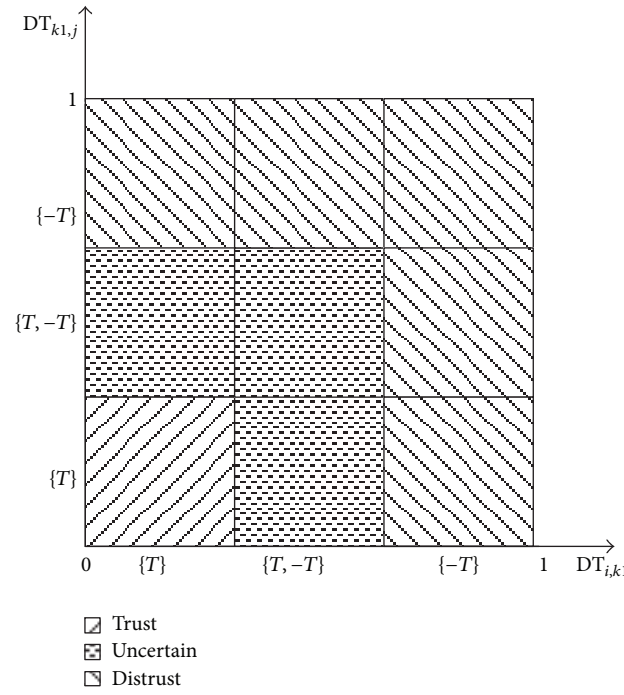



Figure 4: The process of trust transitivity.

used to weigh history experience against current information. To keep $\gamma$ preferably dynamic characteristic, it is satisfied as

$$\gamma = \begin{cases} \gamma_s, & m_{i,j}^H(\{T\}) \geq m_{i,j}^C(\{T\}), \\ \gamma_l, & m_{i,j}^H(\{T\}) < m_{i,j}^C(\{T\}), \end{cases} \tag{8}$$

where $0 < \gamma_s < \gamma_l < 1$, the parameter $m_{i,j}^C(\{T\})$ and $m_{i,j}^H(\{T\})$ represent the trust components of $\mathrm{CDT}_{i,j}$ and $\mathrm{HDT}_{i,j}$, respectively.

### 3.4. Recommendation Trust Evaluation

*3.4.1. Trust Transitivity.* Suppose the recommended trust value of node $i$ on node $j$ can be obtained through $s$ different paths, and the number of recommendation paths $s$ depends on nodes' distribution and communication radius. In order to avoid trust recycle recursion and decrease network communication payload, the recommendation values are confined to direct trust value of the common neighbours owned by both node $i$ and node $j$. As shown in Figure 3, node $i$ can get the trust recommendation of node $j$ from $k_1, k_2, k_3, \ldots, k_s$.

$\mathrm{RT}_{i,j}^1$ denotes the recommended trust value of node $i$ on node $j$ through recommendation path pt1 = $\{k1\}$. The vector forms of $\mathrm{RT}_{i,j}^1$, $\mathrm{DT}_{i,k_1}$, $\mathrm{DT}_{k_1,j}$ are as follows:

$$\begin{aligned} \mathrm{RT}_{i,j}^1 &= \left(m_{i,j}^1(\{T\}), m_{i,j}^1(\{T,-T\}), m_{i,j}^1(\{-T\})\right), \\ \mathrm{DT}_{i,k_1} &= \left(m_{i,k_1}^D(\{T\}), m_{i,k_1}^D(\{T,-T\}), m_{i,k_1}^D(\{-T\})\right), \\ \mathrm{DT}_{k_1,j} &= \left(m_{k_1,j}^D(\{T\}), m_{k_1,j}^D(\{T,-T\}), m_{k_1,j}^D(\{-T\})\right). \end{aligned} \tag{9}$$

Let us set $\Theta = \{\{T\}, \{T,-T\}, \{-T\}\}$, $A$, $E$ and $F \subseteq \Theta$. Then, the $\mathrm{RT}_{i,j}^1$ is calculated as follows:

$$m_{i,j}^1(A) = \begin{cases} m_{i,k_1}^D(A) \times m_{k_1,j}^D(A), & A = \{T\}, \\ \sum\limits_{E=A \text{ or } F=A} m_{i,k_1}^D(E) \times m_{k_1,j}^D(F), & A = \{-T\}, \\ 1 - m_{i,j}^1(\{T\}) - m_{i,j}^1(\{-T\}), & A = \{T,-T\}. \end{cases} \tag{10}$$

Using the symbol $\otimes$ to denote this operation, we get

$$\mathrm{RT}_{i,j}^1 = \mathrm{DT}_{i,k_1} \otimes \mathrm{DT}_{k_1,j}. \tag{11}$$

To vividly show the process of trust transitivity, we resort to Figure 4. It is obvious to see that as long as one of $\mathrm{DT}_{i,k_1}$ and $\mathrm{DT}_{k_1,j}$ is distrust, then $\mathrm{RT}_{i,j}^1$ is distrust.

Extending the above transitivity to multihop, we can get recommended trust through complex recommendation paths with many middle nodes

$$\mathrm{RT}_{i,j}^1 = \mathrm{DT}_{i,\bullet} \otimes \cdots \otimes \mathrm{DT}_{\bullet,j}, \tag{12}$$

where the symbol $\bullet$ indicates anonymous nodes in recommendation path.

*3.4.2. Dynamic Aggregation of Recommended Trust.* On the basis of trust transitivity, node $i$ obtains recommended trust values on node $j$ through $s$ recommendation paths, namely,

$$\begin{aligned} \mathrm{RT}_{i,j}^{1} &= \left(m_{i,j}^{1}(\{T\}), m_{i,j}^{1}(\{T,-T\}), m_{i,j}^{1}(\{-T\})\right) \\ \mathrm{RT}_{i,j}^{2} &= \left(m_{i,j}^{2}(\{T\}), m_{i,j}^{2}(\{T,-T\}), m_{i,j}^{2}(\{-T\})\right) \\ &\vdots \\ \mathrm{RT}_{i,j}^{s} &= \left(m_{i,j}^{s}(\{T\}), m_{i,j}^{s}(\{T,-T\}), m_{i,j}^{s}(\{-T\})\right). \end{aligned} \tag{13}$$

Then, node $i$ would aggregate these pieces of evidence to get a consensus on node $j$. Due to the existence of malicious nodes that may offer false recommendation, we introduce the revised D-S combination rule which adopts a consistent intensity to adjust weights of recommended trust values. The integration process is described in detail as follows.

Firstly, we compute the corresponding average weight denoted as $I_u$. The consistent intensity between $\mathrm{RT}_{i,j}^{u}$ and $\mathrm{RT}_{i,j}^{v}$ is defined as follows [21]:

$$I_{u,v} = 1 - \sqrt{\frac{1}{2}\left(\left\|\vec{m}_{i,j}^{v}\right\|^2 + \left\|\vec{m}_{i,j}^{u}\right\|^2 - 2\left\langle \vec{m}_{i,j}^{v}, \vec{m}_{i,j}^{u}\right\rangle\right)}, \quad v = 1,2,\ldots,s;\ u = 1,2,\ldots,s, \tag{14}$$

where $\|\vec{m}_{i,j}^{v}\|^2 = \langle \vec{m}_{i,j}^{v}, \vec{m}_{i,j}^{v}\rangle$, $\|\vec{m}_{i,j}^{u}\|^2 = \langle \vec{m}_{i,j}^{u}, \vec{m}_{i,j}^{u}\rangle$, $\langle \vec{m}_{i,j}^{v}, \vec{m}_{i,j}^{u}\rangle$ is the inner product of $\vec{m}_{i,j}^{v}$ and $\vec{m}_{i,j}^{u}$.

The difference between two pieces of recommended trust evidence increases with the reduction of consistent intensity. The lower the consistent intensity is, the more probably false trust recommendation may occur.

Furthermore, the matrix of consistent intensity composed of all the recommended trust values is defined as follows:

$$I_{s\times s} = \begin{bmatrix} 1 & I_{1,2} & \cdots & I_{1,s} \\ I_{2,1} & 1 & \cdots & I_{2,s} \\ \vdots & \vdots & \ddots & \vdots \\ I_{s,1} & I_{s,2} & \cdots & 1 \end{bmatrix}. \tag{15}$$

Through summation in row and normalization, the totally consistent intensity of recommended trust $\mathrm{RT}_{i,j}^{u}$, which is equal to the average weight $I_u$, is computed by

$$I_u = \frac{\sum_{v=1,\ v\neq u}^{s} I_{u,v}}{\mathrm{Max}\left(\sum_{v=1,\ v\neq w}^{s} I_{w,v}\right)_{1\le w\le s}}. \tag{16}$$

Then, the basic reliability function $m$ of every recommended trust evidence is amended by $I_u$ as follows:

$$\begin{aligned} m_{i,j}^{u\prime}(\{T\}) &= I_u \times m_{i,j}^{u}(\{T\}), \\ m_{i,j}^{u\prime}(\{-T\}) &= I_u \times m_{i,j}^{u}(\{-T\}), \\ m_{i,j}^{u\prime}(\{T,-T\}) &= 1 - m_{i,j}^{u\prime}(\{T\}) - m_{i,j}^{u\prime}(\{-T\}), \\ & u = 1,2,\ldots,s. \end{aligned} \tag{17}$$

Next, we apply the amended basic trust reliability function $m$ to D-S combination rule. Assume that $\mathrm{Bel}_1$ and $\mathrm{Bel}_2$ are two trust degree functions that are on the same identification frame $\Omega$; their basic reliability degree functions are $m_1$ and $m_2$. And $m$, the basic trust reliability function of Bel, can be expressed as follows:

$$\begin{aligned} m(A) &= m_1(A) \oplus m_2(A) = \frac{\sum_{X\cap Y=A} m_1(X)\times m_2(Y)}{1-K}, \quad A\neq\emptyset,\ A\subseteq\Omega, \\ m(\emptyset) &= 0, \\ K &= \sum_{X\cap Y=\emptyset} m_1(X)\times m_2(Y), \end{aligned} \tag{18}$$

where $\oplus$ is called "Direct Sum," representing the combinatorial operation between pieces of evidence.

Extending to $s$ independent pieces of evidence which belongs to the same identification frame $\Omega$, we can get

$$m(A) = \left(\left(m_1(A)\oplus m_2(A)\right)\oplus\cdots\right)\oplus m_s(A), \quad m(\emptyset)=0, \quad A\neq\emptyset,\ A\subseteq\Omega. \tag{19}$$

At last, the consistent recommended trust $\mathrm{RT}'_{i,j}$ is obtained.

*3.5. Overall Trust Value Synthesis.* Through the observation and recommendation from neighbour nodes, subject node $i$ computes $\mathrm{DT}_{i,j}(t)$ and $\mathrm{RT}_{i,j}(t)$. The D-S evidence theory can combine conflicting and uncertain information to make a correct decision and accelerate converge rate of trust calculation. Consequently, it is used to synthesize $\mathrm{DT}_{i,j}(t)$ and $\mathrm{RT}_{i,j}(t)$ for the overall trust value $\mathrm{OT}_{i,j}(t)$:

$$\begin{aligned} m_{i,j}^{O}(A) &= m_{i,j}^{D}(A)\oplus m_{i,j}^{R}(A) \\ &= \frac{1}{N}\sum_{E\cap F=A} m_{i,j}^{D}(E)\times m_{i,j}^{R}(F), \quad A\subseteq\theta, \\ N &= \sum_{E\cap F\neq\emptyset} m_{i,j}^{D}(E)\times m_{i,j}^{R}(F) > 0. \end{aligned} \tag{20}$$

Algorithm 1 shows the process that subject node $i$ judges whether node $j$ is "Trust", "Distrust" or "Uncertain". The threshold values $\eta$ and $\xi$ are determined by specific application environment; here, we define $\eta = 0.4$ and $\xi = 0.1$. If the trust component is the biggest and the uncertain component is smaller than $\eta$, node $i$ regards node $j$ as "Trust." If the distrust component is the biggest and the uncertain component is smaller than $\eta$, node $i$ regards node $j$ as "Distrust." Otherwise, node $i$ regards node $j$ as "Uncertain".

## 4. Trust-Based Routing Protocol

In this section, we extend the AODV protocol to which can establish trusted route with minimum hops and maximum

(1) **if** $m^o_{i,j}(\{T, -T\}) > \eta$ **then**
(2) node $i$ regard node $j$ as "Uncertain";
(3) **else if** $m^o_{i,j}(\{T\}) - m^o_{i,j}(\{-T\}) > \xi$ && $m^o_{i,j}(\{T\}) > m^o_{i,j}(\{T, -T\})$ **then**
(4) node $i$ regard node $j$ as "Trust";
(5) **else if** $m^o_{i,j}(\{-T\}) - m^o_{i,j}(\{T\}) > \xi$ && $m^o_{i,j}(\{-T\}) > m^o_{i,j}(\{T, -T\})$ **then**
(6) node $i$ regard node $j$ as "Distrust";
(7) **else**
(8) node $i$ regard node $j$ as "Uncertain";
(9) **end if**

Algorithm 1: Process of judging node $j$'s style.

path trust based on trust mechanism denoted by TDS-AODV. The differences between AODV and TDS-AODV are listed as follows.

(1) We append the model of trust computation and fields including $\mathrm{ACF}_{i,j}$, $\mathrm{MOL}_{i,j}$, $\mathrm{SCF}_{i,j}$, and $\mathrm{OT}_{i,j}$ in the neighbour table of each node.
(2) Every node maintains a local black list.
(3) We append $T_{\mathrm{route}}$ field in the route reply message and $T_{\mathrm{route}}$ denotes the accumulated route trust.
(4) We set backup route to avoid initiating the route discovery frequently.

*4.1. Route Discovery.* During the process of route discovery, when node $i$ chooses another node $j$ to forward a packet, node $i$ may suffer some attacks from node $j$, such as black hole attack. Thus, it is important to choose a reliable next hop node. The process of judging whether node $j$ can be the next hop of node $i$ is as follows.

*Step 1.* Node $i$ checks whether it has the trust value of node $j$ ($\mathrm{OT}_{i,j}$); if it has, turn to Step 5, else turn to Step 2.

*Step 2.* Node $i$ computes $D_{i,j}$ according to (6)–(8) and broadcasts a *Recommendation_Query* message to the common neighbours denoted as node $k$.

*Step 3.* After receiving the *Recommendation_Query* message, node $k$ sends $D_{k,j}$ to node $i$ if $m_{k,j}(\{T, -T\}) < \eta$.

*Step 4.* Node $i$ calculates $\mathrm{RT}_{i,j}$ based on (13)–(18) and $\mathrm{OT}_{i,j}$ based on (19).

*Step 5.* Whether node $j$ is reliable can be estimated using Algorithm 1. If node $j$ is trusted, node $i$ will update $\mathrm{OT}_{i,j}$ and regards node $j$ as its credible next hop node, else node $i$ will not choose node $j$ to transmit packets and move node $j$ into its local black list as a malicious node.

Once a node is in a black list, it will neither receive packets from its neighbour nor have its packets forwarded. That is, a malicious node in a black list is excluded by its neighbours. When a node exists in the black lists of all its neighbours, it will be excluded from the local network.

Sending packets by the trusted route will decrease the probability of malicious attacks and improve the survivability of MANETs. We evaluate the trustworthiness of a route by the trust value of nodes along the route, denoted by $T_{\mathrm{route}}$ [16]

$$T_{\mathrm{route}} = \prod m_{i,k}(\{T\}), \quad n_i \in \mathrm{route}, \quad n_k \in \mathrm{route}, \quad n_i \rightarrow n_k, \quad n_k \neq n_d, \tag{21}$$

where $n_i$ and $n_k$ are any two adjacent nodes among the route; $n_d$ is the destination node in the route; $n_i \rightarrow n_k$ means that $n_k$ is the next hop node of $n_i$; $n_k \neq n_d$ means that the destination node $n_d$ should not forward the packets for itself and $m_{i,d}(\{T\})$ is not used to calculate the path trust to node $n_d$.

As shown in Figure 5, the trust value of path *P(A, B, C, D)* is equal to 0.68 (i.e. $T_{A,B,C,D} = m_{A,B}(\{T\}) \times m_{B,C}(\{T\}) = 0.85 \times 0.8 = 0.68$). Figure 6 shows an example of a multiple path. Among the three paths from $A$ to $H$, path *P(A, F, G, H)* is the most credible path.

In our trusted routing mechanism, the route discovery includes three processes: (i) Route Request (*RREQ*) Message Delivery; (ii) Route Reply (*RREP*) Message Delivery; and (iii) route selection.

*4.1.1. RREQ Delivery.* An *RREQ* packet contains the following fields: ⟨*SourceAddr*, *SourceSequenceNo*, *BroadcastID*, *DestAddr*, *DestSequenceNo*, *HopCounter*⟩.

When the source node $S$ needs to send data to the destination node $D$, it first checks whether there is a feasible path found between $S$ and $D$. If so, $S$ sends the data to $D$; otherwise, $S$ will broadcast a *RREQ* to start a route discovery.

When any reliable intermediate node $K$ whose authentication process was discussed before receives a *RREQ* packet from a neighbour $J$, it deals with the request according to the following steps.

*Step 1.* It checks whether one copy of the same *RREQ* has been received according to the *BroadcastID*. If so and the later copy has greater *HopCounter*, the *RREQ* will be discarded and the procedure ends; otherwise, go to Step 2.

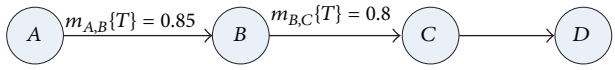



Figure 5: Path trust computation of a single path.

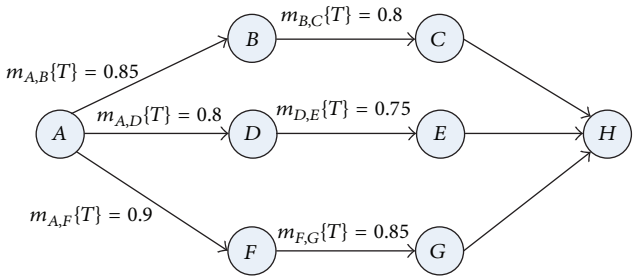



Figure 6: Path trust computation of a multiple path.

*Step 2.* If node $J$ is not the source, node $K$ creates a reverse route to $S$ using the previous hop (node $J$) of the *RREQ* as the next hop.

*Step 3.* $K$ checks whether there is a valid route to the destination. If so and the *DestSequenceNo* of the route is greater than that in the *RREQ*, $K$ unicasts a Route Replay (*RREP*) message to $S$ via $J$ through the reverse route; otherwise, go to Step 4.

*Step 4.* $K$ increases *HopCounter* by one and propagates the *RREQ* to all its neighbours.

The pseudocode of the *RREQ* is shown in Algorithm 2.

*4.1.2. RREP Delivery.* An *RREP* packet contains the following information: ⟨*SourceAddr*, *SourceSequenceNo*, *DestAddr*, *DestSequenceNo*, *HopCounter*, *LifeTime*, *PathTrust*⟩. When $D$ receives the *RREQ* packet, it deals with the request according to the following steps.

*Step 1.* If it is the first time for $D$ to receive a *RREQ* packet, then $D$ sets a timer window $t_D$ and records the route of *RREQ* in its cache and go to Step 6, otherwise go to Step 2.

*Step 2.* If $t_D$ expires, it discards the follow-up *RREQ* packets, otherwise go to Step 3.

*Step 3.* If there are less than three routes in the cache of $D$, then add the new route in its cache and go to Step 6, otherwise go to Step 4.

*Step 4.* $D$ compares the hop count of the new route with that of the route which owns the maximum hop count in its cache (denoted as route $X$). If the former is more than or equal to the latter, $D$ discards the new *RREQ*, otherwise turn to Step 5.

*Step 5.* $D$ uses the new route to substitute route $X$ and then turns to Step 6.

*Step 6.* $D$ sets $T_{\text{route}}$ and then unicasts the *RREP* packets with $T_{\text{route}}$ to the intermediate node.

After receiving a *RREP* packet, the intermediate node computes $T_{\text{route}}$ according to (21) and updates the field of $T_{\text{route}}$ then it forwards the *RREP* packet with $T_{\text{route}}$. The pseudo code of *RREP* delivery algorithm is shown in Algorithm 3.

*4.1.3. Route Selection.* When $S$ receives the *RREP* packet, if the timer window $t_D$ does not expire, it needs to update the $T_{\text{route}}$ field of this message according to (21). Otherwise, $S$ discards follow-up *RREP* packets and picks the one with largest $T_{\text{route}}$ as its main route. The route with second largest $T_{\text{route}}$ is regarded as backup route which aims at avoiding initiating the route discovery frequently. The pseudo code of route selection algorithm is shown in Algorithm 4.

*4.2. Route Maintenance.* After each successful route discovery takes place, $S$ can deliver its data to $D$ through a route. However, the route may break at any time instant due to the mobility of nodes or attacks. In order to maintain a stable and secure network connection, route maintenance is necessary to ensure the system survivability. AODV protocol designed two types of route maintenance mode one is a local repair mechanism and the other is that $S$ reestablishes the route. Detailed process is discussed as follows.

Once the route is found, each node along the route periodically sends *HELLO* messages to its neighbour node for link failure detection. Link failure occurs when the neighbour node does not reply to the *HELLO* messages after a period of time. When a node $N$ detects a link failure, it first sends a Route Error (*RERR*) message to $S$. $S$ checks whether there is a backup route; if a backup route is found, $S$ replaces the failure route with the backup and sends a *FoundBackup* message to $N$. Otherwise $S$ sends a *NonBackup* message to $N$ and then $N$ starts a local repair mechanism. $N$ broadcasts a *RREQ* message to find an alternative route between $N$ and $D$. If no route is found, the system resorts back to another mechanism of sending a *RERR* message upstream to $S$, starting a new route discovery.

In TDS-AODV, besides link failure, if $T_{\text{route}} < T_{\text{thr}}$, $S$ will also perform route maintenance which works as follows. During the transmission, if $S$ finds the trust of a route has decreased, it sends a route check message along the route to check the route status and sets a timeout period to wait for the route check message from $D$. When $S$ receives the reply, it will update the $T_{\text{route}}$ and judge whether $T_{\text{route}}$ is larger than $T_{\text{thr}}$. If $T_{\text{route}} < T_{\text{thr}}$, $S$ resorts to the backup route and updates the path trust of the backup route (denoted as $T_{\text{rb}}$). If $T_{\text{rb}} > T_{\text{thr}}$, $S$ discards the main route and uses the backup route to send packets. Otherwise, a new route discovery is triggered.

## 5. Simulation Studies

To evaluate the performance of TDS-AODV, we use the simulation tool MATLAB. In our simulation, fifty nodes at first are randomly placed in a specific field (100 m ×

(1) **To source node**:
(2) **if** there is a feasible path found between $S$ and $D$ **then**
(3) $S$ sends data to $D$;
(4) **else**
(5) broadcasts the *RREQ* to start a route discovery;
(6) **end if**
(7) **To a reliable intermediate node**:
(8) checks whether one copy of the same *RREQ* has been received;
(9) **if** so and the later copy has greater *HopCounter* **then**
(10) discards *RREQ* and the procedure ends;
(11) **else**
(12) creates a reverse route to $S$ using the previous hop of the *RREQ* as the next hop;
checks whether there is a valid route to the destination;
(13) **if** so and the *DestSequenceNo* of the route is greater than that in the *RREQ* **then**
(14) unicasts a Route Replay (*RREP*) message to $S$ via $J$ through the reverse route;
(15) **else**
(16) increases *HopCounter* by one;
propagates the *RREQ* to all its neighbours;
(17) **end if**
(18) **To destination node**:
(19) calls the process of route reply;
(20) **end if**

Algorithm 2: The *RREQ* delivery algorithm.

(1) **To destination node**:
(2) sets $T_{\text{route}} = 1$;
(3) **if** received the first *RREQ* packet **then**
(4) sets a timer window $t_D$;
increases the destination sequence number by 1;
records the route of *RREQ* in its cache;
sends the *RREP* with $T_{\text{route}}$ along the path to the intermediate node;
(5) **else if** $t_D$ expires **then**
(6) discards the follow-up *RREQ* packets;
(7) **else if** there are less than three routes in its cache **then**
(8) adds the new route in the cache;
sends the *RREP* with $T_{\text{route}}$ along the path to the intermediate node;
(9) **else if** the hop count of the new route is more than or equal to that of route $X$ **then**
(10) discards the new *RREQ*;
(11) **else**
(12) uses the new route to substitute route $X$;
sends the *RREP* with $T_{\text{route}}$ along the path to the intermediate node;
(13) **end if**
(14) **To a reliable intermediate node**:
(15) updates $T_{\text{route}}$ according to (21);
forwards the *RREP*;
(16) **To source node**:
(17) updates $T_{\text{route}}$ according to (21);
calls route selection;

Algorithm 3: The *RREP* delivery algorithm.

(1) when source node receives the *RREP*, checks the $t_s$;
(2) **if** $t_s$ does not expire **then**
(3) updates the $T_{\text{route}}$
(4) **else**
(5) discards the follow-up *RREP*;
selects the route with the largest as its main route;
picks the route with second largest $T_{\text{route}}$ as its backup routes;
(6) **end if**

Algorithm 4: The route selection algorithm.

100 m) and move to another random position with a speed chosen between 0 to 30 m/s. The malicious nodes randomly drop data packets based on their trust value. The simulation parameters are listed in Table 1.

*5.1. Performance Metrics.* To measure the performance of our proposed TDS-AODV, we identify three metrics: (i) throughput: the number of packets transmitted per unit time from the source node to the destination node; (ii) packet delivery ratio: the ratio of the number of packets received to the total number of packets; and (iii) average end to end delay: the average delay between the sending of the packets by the source node and its receipt at the destination node.

The network topology of TDS-AODV was compared with that of TAODV [22] and AODV in this paper. We also carried out three simulations in terms of the maximum node speed and the proportion of malicious nodes to compare the above three performances of two protocol.

*5.2. Simulation Results and Analysis.* Figures 7 and 8 are the network topology of TDS-AODV and AODV with 20% malicious nodes. It is obvious to see that our method can avoid malicious nodes becoming the next hop effectively while in AODV malicious nodes can be selected as the next hop. The reason is that TDS-AODV takes nodes' trust value into account.

Figure 9 shows the average routing hop of TDS-AODV and AODV with different numbers of malicious nodes. when the number of malicious nodes accounts for a certain proportion of the number of total nodes, the average route hop of TDS-AODV is a little higher than that of AODV, because nodes would rather choose a relative longer path than choose malicious nodes as the next hop nodes in TDS-AODV. Although the path of TDS-AODV may be a little longer, the performance of TDS-AODV is still better than that of AODV as it eliminates malicious nodes out of the routing paths, which will be proven by the following simulation experiments.

Figures 10 and 11 depict the throughput of TDS-AODV, TAODV, and AODV. The routing throughput of TDS-AODV is averagely 29.60% lower than that of AODV and 21.27% lower than that of TAODV in Figure 10. This is because that our method can detect malicious nodes effectively and thus prevent the channel congestion. The throughput changes little

Table 1: Simulation parameters.

| Parameters | Value |
|---|---|
| Simulation time | 100 s |
| Number of nodes | 50 |
| Source node | Node 1 |
| Destination node | Node 50 |
| Area size | 100 m × 100 m |
| Transmission radius | 25 m |
| Max speed | 0–30 m/s |
| Number of malicious node | 0–20 |

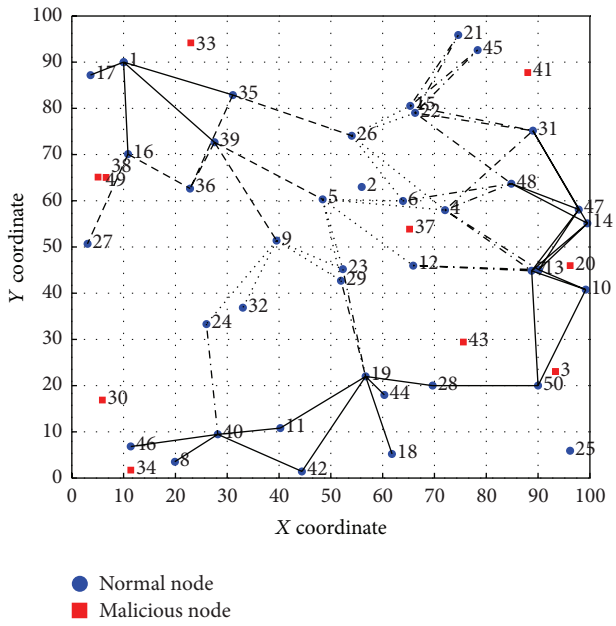



Figure 7: Network topology of TDS-AODV.

at different maximum speed which indicates our method has excellent dynamic. As shown in Figure 11, the throughput rises slowly with the increase in the number of malicious nodes. Besides, TDS-AODV rises more slowly than TAODV and AODV as it prevents the malicious nodes from becoming the next hop and affects less by malicious nodes.

The packet delivery ratio of TDS-AODV, TAODV, and AODV is shown in Figures 12 and 13. It can be observed

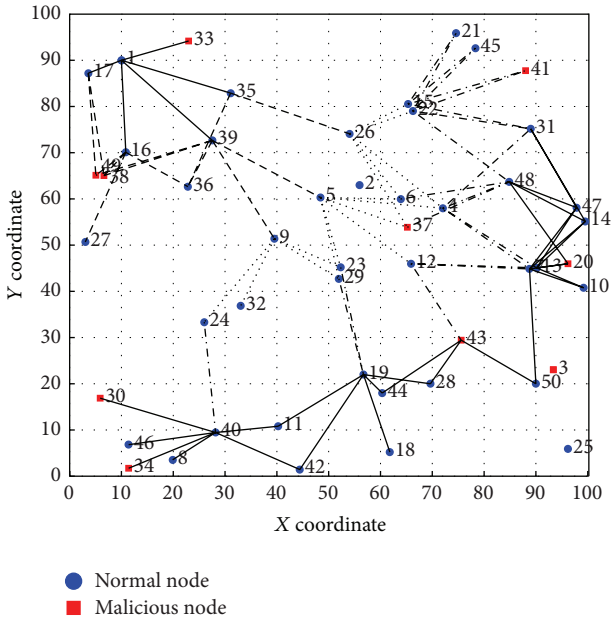



Figure 8: Network topology of AODV.

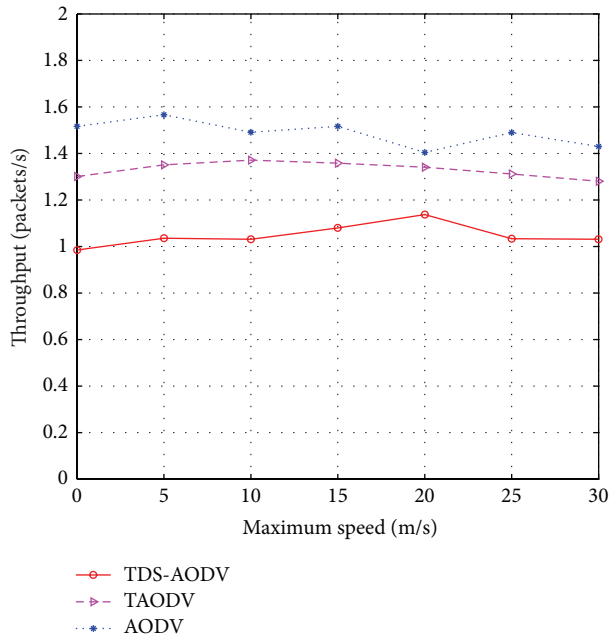



Figure 10: Performance of network throughput at different maximum speed.

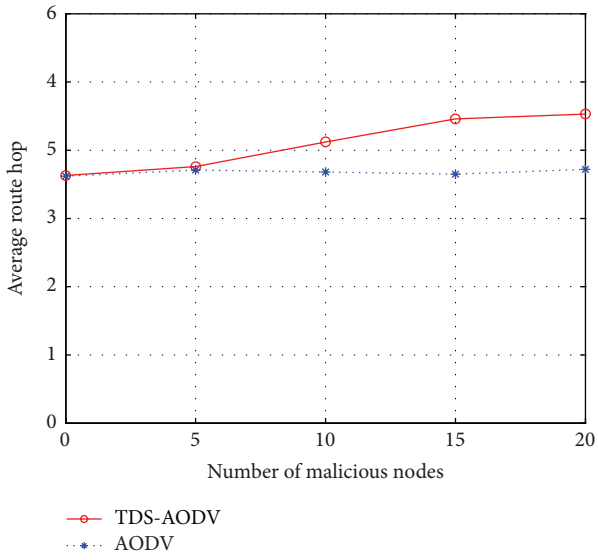



Figure 9: Everage route hop of TDS-AODV and AODV with different numbers of malicious nodes.

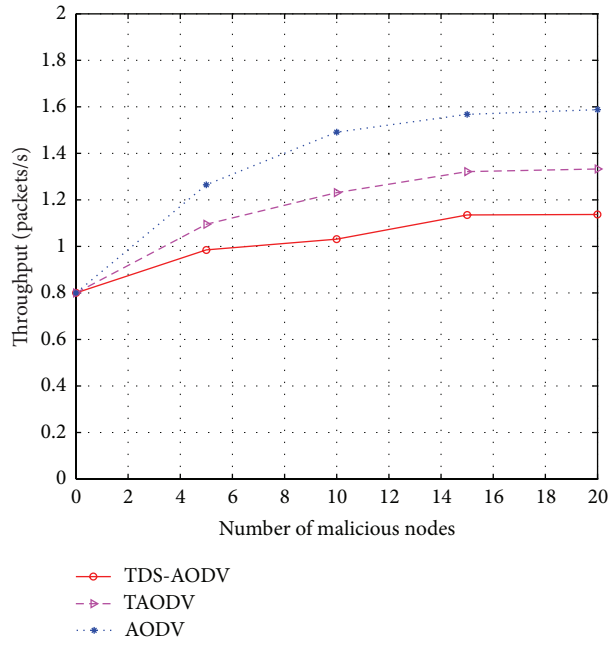



Figure 11: Performance of network throughput with different number of malicious nodes.

that TDS-AODV outperforms TAODV and AODV in the packet delivery ratio because of the fact that in TDS-AODV intermediate nodes make routing selection considering hop count and trust value. It shows the packet delivery ratio of TDS-AODV is averagely 46.24% higher than that of AODV and 17.18% higher than that of TAODV in Figure 12. Figure 13 indicates that TDS-AODV has better fault tolerance as its packet delivery ratio declines slowly with the increase in the number of malicious nodes.

We give the average end to end delay comparisons of TDS-AODV, TAODV, and AODV in Figures 14 and 15. As shown in Figure 14, the average end-to-end delay of three schemes rises very slowly with the increase in the maximum speed. However, the average delay of AODV is 18.73% higher than that of TDS-AODV and the average delay of TAODV is 7.74% higher than that of TDS-AODV in Figure 14 due to the

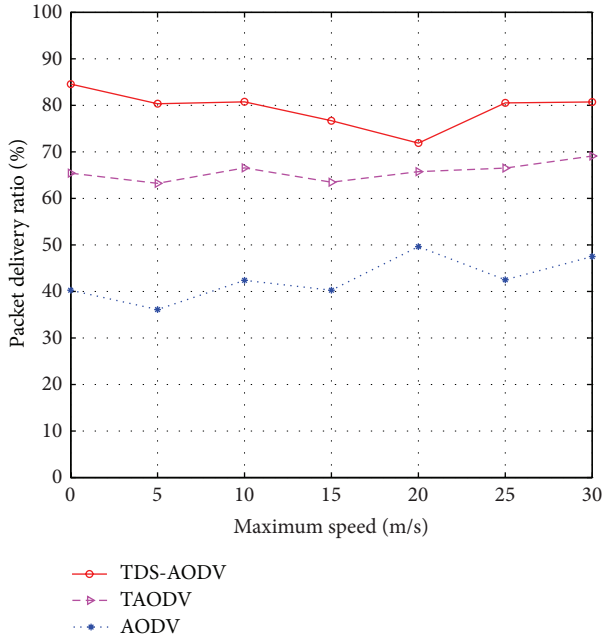



Figure 12: Performance of packet delivery ratio at different maximum speed.

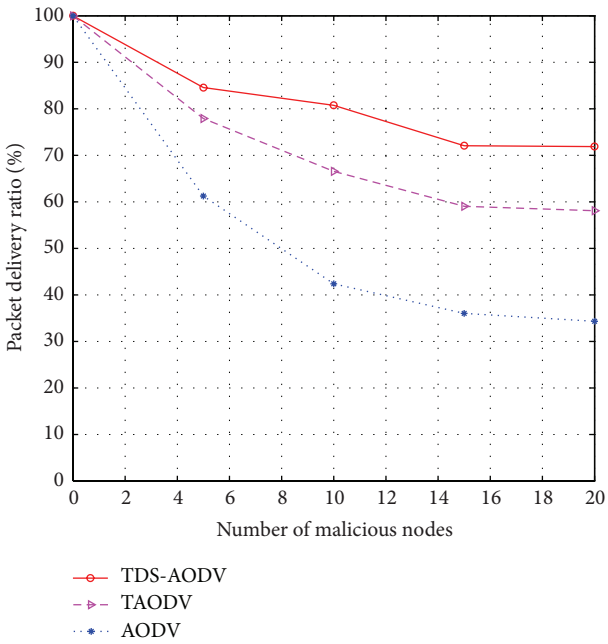



Figure 13: Performance of packet delivery ratio with different number of malicious nodes.

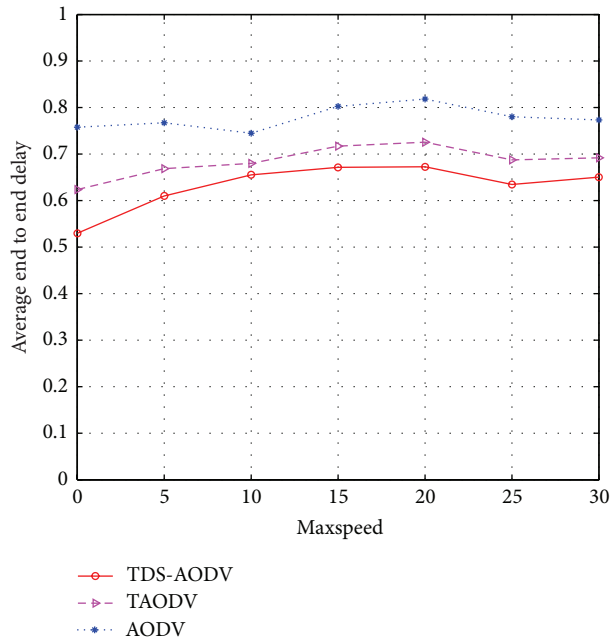



Figure 14: Performance of average delay at different maximum speed.

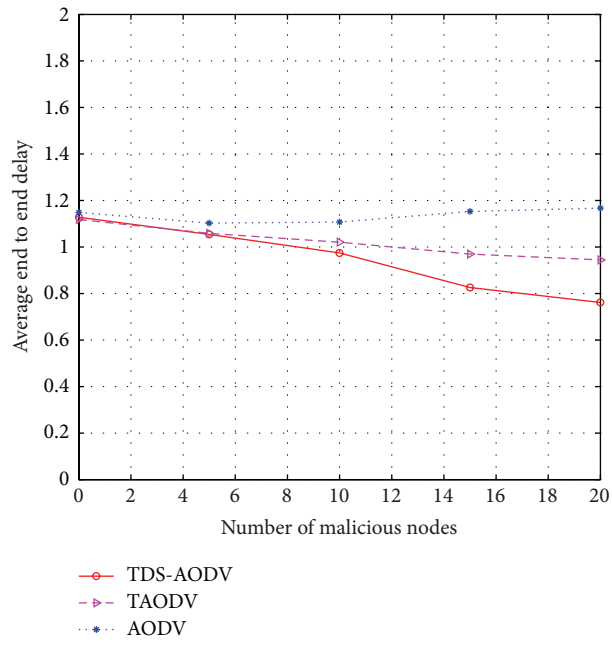



Figure 15: Performance of average delay with different number of malicious nodes.

lack of consideration of dynamic topology. Figure 15 depicts the performance of average delay with different number of malicious nodes. The average end to end delay of TDS-AODV declines faster than AODV. Because when intermediate nodes choose the next hop they will not consider the malicious nodes and thus save the time.

## 6. Conclusions

In this paper, we propose a novel trust mechanism after investigating on trust models of ad hoc networks and routing in current researches. In this trust mechanism, direct trust value on each neighbour node is calculated by using trust factors of average encounter time, mobility, and successful cooperation frequency, which are defined according to node

behaviors. Meanwhile, the revised D-S evidence theory is used to combine multiple recommended pieces of evidence and obtain the recommended trust value. Then, a trusted routing protocol based on the novel trust mechanism, by extending the AODV protocol is presented. In this protocol, a source establishes a main path and a backup path which are evaluated by two aspects: hop counts and trust values. At last, we validate the correctness and effectiveness of TDS-AODV by comparing its performance with TAODV and AODV on Matlab platform. Simulation results show that TDS-AODV is able to eliminate malicious nodes effectively when building the route and achieves an improvement in throughput, packet delivery ratio, and average end-to-end delay.

In our future work, we will conduct extensively simulation and rigorous analysis to quantify and evaluate the tradeoff between the security and the nodes' energy consumption. In addition, a comprehensive performance evaluation will be conducted to compare TDS-AODV with other routing protocols (e.g., DSR).

## Acknowledgments

The authors are grateful to the anonymous reviewers for their insightful comments. This work is supported by the National Natural Science Foundation of China under Grants no. 61201317 and no. 61001138.

## References

[1] B. Wang, C. H. Huang, L. Y. Li, and W. Z. Yang, "Trustbased trustbased minimum cost opportunistic routing for Ad Hoc networks," *Journal of Systems and Software*, vol. 84, no. 12, pp. 2107–2122, 2011.

[2] C. E. Perkins and B. Highly, "Dynamic destination-sequenced distance-vector routing (DSDV) for mobile computers," in *Proceedings of the Conference on Communications Architectures, Protocols and Applications (SIGCOMM '94)*, pp. 234–244, 1994.

[3] D. B. Johnson and D. A. Maltz, "Dynamic source routing in Ad Hoc wireless networks," *Mobile Computing*, vol. 353, pp. 153–181, 1996.

[4] C. E. Perkins and E. M. Royer, "Ad-Hoc on-demand distance vector routing," in *Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications (WMCSA '99)*, pp. 90–100, New Orleans, LA, USA, February 1999.

[5] A. A. Pirzada, A. Datta, and C. S. McDonald, "Trust-based routing for Ad-Hoc wireless networks," in *Proceedings of the 12th IEEE International Conference on Networks (ICON '04)*, pp. 326–330, November 2004.

[6] H. Xia, Z. P. Jia, X. Li, and F. Zhang, "A subjective trust management model based on AHP for MANETs," in *Proceedings of the International Conference on Network Computing and Information Security (NCIS '11)*, vol. 1, pp. 363–368, Guilin, China, May 2011.

[7] S. C. Peng, W. J. Jia, and G. J. Wang, "Voting-based clustering algorithm with subjective trust and stability in mobile Ad-Hoc networks," in *Proceedings of the 5th International Conference on Embedded and Ubiquitous Computing (EUC '08)*, vol. 2, pp. 3–9, Shanghai, China, December 2008.

[8] C. Zouridaki, B. L. Mark, M. Hejmo, and R. K. Thomas, "E-hermes: a robust cooperative trust establishment scheme for mobile Ad Hoc networks," *Ad Hoc Networks*, vol. 7, no. 6, pp. 1156–1168, 2009.

[9] M. Omar, Y. Challal, and A. Bouabdallah, "Reliable and fully distributed trust model for mobile Ad Hoc networks," *Computers and Security*, vol. 28, no. 3-4, pp. 199–214, 2009.

[10] H. Xie, J. F. Ma, L. Yang, and X. W. Dong, "A trust method of MANETs based on D-S evidence theory," *International Journal of Advancements in Computing Technology*, vol. 4, no. 2, pp. 247–257, 2012.

[11] L. D. Huang, G. Xue, X. L. He, and H. L. Zhuang, "A trust model based on evidence theory for P2P systems," *Applied Mechanics and Materials*, vol. 20, no. 23, pp. 99–104, 2010.

[12] L. M. Jiang, J. Xu, K. Zhang, and H. Zhang, "A new evidential trust model for open distributed systems," *Expert Systems with Applications*, vol. 39, no. 3, pp. 3772–3782, 2012.

[13] A. P. Dempster, "Upper and lower probabilities induced by a multi-valued mapping," *The Annals of Mathematical Statistics*, vol. 38, no. 2, pp. 325–339, 1967.

[14] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, New Jersey, NJ, USA, 1976.

[15] K. Wang and M. Wu, "Improved secure trust-based location-aided routing model for MANETs," *China Communications*, vol. 8, no. 3, pp. 154–162, 2011.

[16] X. Li, Z. P. Jia, P. Zhang, and H. Y. Wang, "Trust-based on-demand multipath routing in mobile Ad Hoc networks," *IET Information Security*, vol. 4, no. 4, pp. 212–232, 2010.

[17] S. C. Peng, W. J. Jia, G. J. Wang, J. Wu, and M. Y. Guo, "Trusted routing based on dynamic trust mechanism in mobile Ad-Hoc networks," *IEICE Transactions on Information and Systems*, vol. E93.D, no. 3, pp. 510–517, 2010.

[18] P. Narula, S. K. Dhurandher, S. Misra, and I. Woungang, "Security in mobile Ad-Hoc networks using soft encryption and trust-based multi-path routing," *Computer Communications*, vol. 31, no. 4, pp. 760–769, 2008.

[19] S. L. F. Sirotheau and R. T. D. Sousa, "Evaluating trust in Ad Hoc network routing by induction of decision trees," *IEEE Latin America Transactions*, vol. 10, no. 1, pp. 1332–1343, 2012.

[20] A. A. Pirzada, C. McDonald, and A. Datta, "Performance comparison of trust-based reactive routing protocols," *IEEE Transactions on Mobile Computing*, vol. 5, no. 6, pp. 695–710, 2006.

[21] A. L. Jousselme, D. Grenier, and E. Bosse, "A new distance between two bodies of evidence," *Information Fusion*, vol. 2, no. 2, pp. 91–101, 2001.

[22] X. Q. Li, M. R. Lyu, and J. C. Liu, "A trust model based routing protocol for secure Ad Hoc networks," in *Proceedings of the IEEE Aerospace Conference Proceedings*, vol. 2, pp. 1286–1295, March 2004.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 730819, 8 pages
http://dx.doi.org/10.1155/2013/730819

*Research Article*

# A Reliable Data Collection Protocol Based on Erasure-Resilient Code in Asymmetric Wireless Sensor Networks

**Jian-Jun Lei,[1] Taehyun Park,[2] and Gu-In Kwon[2]**

[1] *School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China*
[2] *School of Computer and Information Engineering, Inha University, Incheon 402-751, Republic of Korea*

Correspondence should be addressed to Gu-In Kwon; gikwon@inha.ac.kr

Received 23 December 2012; Revised 14 March 2013; Accepted 29 March 2013

Academic Editor: Adel Soudani

Copyright © 2013 Jian-Jun Lei 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 RECPE, a reliable collection protocol for aggregating data packets from all the sensor nodes to the sink in a large-scale WSN (wireless sensor network). Unlike some well-known reliable data collection protocols such as CTP (Collection Tree Protocol) that uses ETX (expected transmission count) as the routing metric, RECPE exploits ETF (expected transmission count over forward links) to construct a one-way collection tree, which avoids missing some good routes and reduces the effect of asymmetric link in the network. Crucially, RECPE guarantees the reliability by erasure-resilient codes in the application layer without retransmission required by other reliable protocols. Therefore, some lower layers such as data link layer only need to conduct best-effort data delivery. Meanwhile, to improve efficiency, RECPE also exploits Trickle algorithm to reduce routing beacons and pipeline data delivery to prevent self-interference. We evaluate the performance of RECPE via TOSSIM simulations, and our results show that, in comparison with CTP (the de facto data collection protocol for TinyOS), RECPE can obtain significant performance in terms of delivery cost, latency, and packet loss rate for reliable data collection especially in asymmetric link networks.

## 1. Introduction

Many-to-one data collection is an important issue for WSN applications and protocols. Collection tree provides efficient approach for the higher layer routing and transport protocol. Most metrics employed in constructing collection tree from sensors to sink node in recent years tend to rely on the minimum hop-count [1, 2] or ETX [3, 4].

In high-power stable wireless networks, data loss is infrequent and hop-count can adequately capture the cost of packet delivery to destination. However, with lossy link, as found more often in WSN, each hop may require one or more retransmissions to compensate for the lossy channel. The real cost of packet delivery can be much larger than the hop distance. Therefore, the minimum hop-count ignores the possibility that a longer path might offer more efficient route and higher throughput. Additionally, the minimum hop-count also chooses arbitrarily among the different paths of the same minimum length, regardless of the often large difference in link quality among these routes. One alternative metric to fix this problem is to use ETX, which is calculated as $1/(Q_f \times Q_r)$, where $Q_f$ and $Q_r$ are forward and backward link quality, respectively [3]. ETX-based routing protocols tend to avoid asymmetric links. However, many experiments [5–7] have shown that wireless links are complex and asymmetric. In general, the reverse link has less effect on data transmission, especially in some applications, a majority of data only need to be delivered from sensors to sink node. Also in the other real-time applications, retransmission data is almost useless. Therefore, path selecting based on bidirectional link quality often misses some good routes and discounts the potential resources.

In [5], Sang et al. proposed a reformative metric called ETF aiming at asymmetric network. They investigated the link asymmetric in WSN and demonstrated the improvement by exploiting ETF as the routing metric over ETX. ETF-based protocols choose the routes for data packet delivery solely according to its forward link quality toward the destination. Like other reliable protocols, ETF protocol also exploits retransmission mechanism to compensate for the packet loss.

However, in such protocols, the acknowledgement packets (ACKs or NACKs) that are transmitted in the reverse links usually are more important than data packets delivered in forwarding link. Even the synchronous ACKs exploited in [5] are demonstrated to obtain higher reliability; their loss will still be not tolerant for the reliable data delivery.

In this paper, we presented a protocol that uses erasure-resilient codes in application layer to guarantee the reliability of data delivery. Therefore, the reliable transmission is transferred into the best-effort data delivery according to the feature of erasure-resilient codes. Consequently, it can simplify the function for reliability requirement in data link layer and network layer. Our protocol can find high fidelity route by broadcasting asynchronous discovering beacons and form a single-direction collection tree, which avoids missing some good routes and reduces the effect of asymmetric link in the network. Meanwhile, in forwarder node, large buffer required by potential retransmission request also can be avoided. Additionally, in our protocol, some sophisticated strategies like Trickle algorithm [8] and pipeline mechanism are also exploited to reduce the control information and improve the efficiency of data delivery. This protocol is referred to as RECPE in this paper, which can benefit some applications significantly like real-time video delivery where retransmitting packets are useless. RECPE strives to provide the reliability for some applications in which bulk data needed to be delivered to the sink. The sender and sink have powerful processors, and the relaying nodes are some traditional resource-constrained nodes. Image collection and monitoring in the wild zone is a good example of application. We evaluate the performance of RECPE via TOSSIM, in comparison with CTP [9] protocol, RECPE significantly reduces the delivery cost, latency, and packet loss rate for reliable data collection especially in asymmetric link networks.

The remainder of the paper is organized as follows. In Section 2, we briefly review some data delivery approach in asymmetric network and application of erasure-resilient codes in WSN. In Section 3, we present the protocol details. Then, we provide some simulation results that show the protocol performance in Section 4. Finally, we make a conclusion in Section 5.

## 2. Related Work

In this section, we review briefly two related works: the data delivery in asymmetric network and the application of erasure-resilient codes in WSN.

Some earlier works indicated that link asymmetry is a real issue in low-power WSN. Zhou et al. [10] reported that about 30% links were asymmetric in their deployed system. Zamalloa and Krishnamachari [11] provided a comprehensive analysis of the root causes of link unreliability and asymmetry. Kannan et al. [12] found that most intermediate links are bursty and they shift between poor and good delivery. Based on these facts, some efficient mechanisms were proposed to tackle the asymmetry of link. ETF [5] has been proposed as a routing metric and shown to perform well in a variety of asymmetric wireless networks. Diversely, DEAL [13] proposed an approach to discover and manage the asymmetric link, which can reduce the expected packet count when they are exploited in the data collection application. Not just focusing on the routing metric, EERDC [14] proposed an energy-efficient reliable data collection protocol which uses the implicit ACK to cope with the asymmetry of network. Thus, our newly proposed RECPE protocol focuses on the reliable data delivery that is achieved by essentially combining the ETF metric and the packet-level erasure-resilient codes to avoid the retransmission. Unlike the DEAL and EERDC that need the acknowledgment packet and only provide the *statistical reliability*, RECPE is a package solution that does not require the reverse link and can guarantee the data delivery with *absolute reliability*. Additionally, in these proposed reliable schemes, the sender node and relaying nodes along the path must keep all transmitting packets in the buffer until they are acknowledged. Therefore, large buffer is required by each relay node, especially those that are close to the sink due to the need to relay packets for plenty of senders. Thus, these schemes would make buffer overflow at the relay node and sender on the large-scale WSN. Moreover, the use of a retransmission also introduces the duplicates and interference, even such as ETF scheme in which the dynamic retransmission threshold strategy is exploited. In contrast with these proposed protocols, RECPE can improve the throughput and reduce the delivery cost greatly.

The packet-level erasure-resilient codes first are presented as a mechanism for reliable and efficient multicast for file distribution system [15]. Recently, researchers started to use them to provide reliability for data delivery in WSN. Rateless Deluge [16] enhances the Deluge [17] protocol through a hybrid ARQ technique based on erasure-resilient codes. AdapCode [18] performs the encoding at each node during data dissemination; also the coding aggressiveness is adaptively changed according to the link qualities and the number of neighbors. SYNAPSE++ [19] disseminates reliably the data by implementing the hybrid ARQ similar to that in [16] but adds full support for pipelining through a joint design of MAC and Fountain codes. In [20], Cataldi et al. also exploited the erasure-resilient codes to disseminate the data in wireless vehicular network. As mentioned, all these applications focus on the reliable data dissemination in WSN.

In this paper, we present the reliable data collection approach which exploits the erasure-resilient codes to provide the reliability and avoids the requirement in the feedback channel. Furthermore, RECPE improves the efficiency of data delivery by constructing one-way collection tree using ETF metric. Since it is designed as a whole solution, the details of algorithm and frame format also are included in this paper. Furthermore, by modularization, our protocol also can be embedded into any other collection protocols to substantially improve their performance. To characterize the impact of our protocol, we use CTP protocol as the benchmark in our simulation.

## 3. Protocol Design

In this section, we present the motivations and demonstrate an example to show the improvement of exploiting ETF as

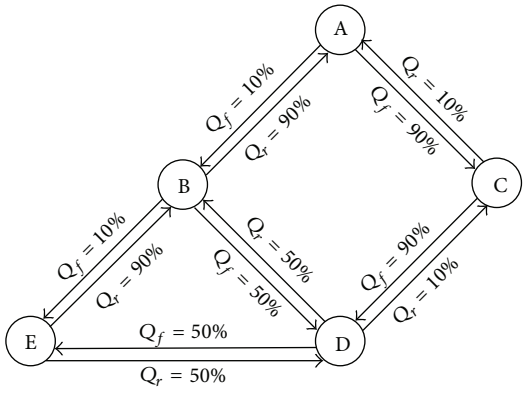



Figure 1: An example of route choosing under asymmetric link.

the metric in asymmetric network and then focus on the protocol design issues including exploiting efficient routing beacon mechanism to form a one-way collection tree for reliable data delivery. Our work puts heavy weight on optimization for function realization in transport and network layers.

*3.1. Motivation.* This work is mainly motivated from some experiments. During data delivery, substantial wireless links are asymmetric in manipulating a real-world system. Additionally, in data collection scenario, only forward links are needed to deliver the data packet to the sink. In our proposed protocol, due to the reliability guaranteed by the erasure-resilient codes, the reverse links required by the traditional reliable protocols are trivial. Therefore, we can construct the optimal one-way data collection tree for data delivery by only exploiting forwarding link quality. Particularly, ETX and minimum hop-count metrics cannot choose a better route when the deployment scenario experiences a great deal of asymmetry. An example in Figure 1 can demonstrate this situation. While node A delivers a packet to node E, Table 1 lists possible routes and their determination under different metrics. Obviously, it only takes 4.2 transmissions on average to deliver a data packet through route A → C → D → E chosen by metric ETF, while more than 3 times packets need to be delivered when other two metrics are used for routing.

*3.2. Design Overview.* RECPE is expected to aggregate data packets from all the nodes in the network to one sink efficiently. To this end, RECPE builds and maintains a collection tree that spans the entire network, in which each node has a one-way route to the sink. Meanwhile, the node attempts to deliver the packet to the best forwarder, which means that each node only needs to maintain a local routing table. RECPE finds a path with the minimum ETF obtained by adding all the ETF values of the links in the path. Some periodic beacons are exchanged between neighbors to estimate the ETF to neighbors and the whole ETF to sink. Consecutive sequence numbers are added to these beacons, so that neighbors can determine which beacons are lost during transmission. The fraction of lost beacons forms an estimation of the link quality from the neighbors. To reduce the control information, routing beacons are transmitted according to a variable timer, controlled by Trickle algorithm. Trickle will decrease the beacon frequency exponentially when the topology is stable. Once the inconsistency of topology is detected, the beacon frequency will be reset to the minimum value. This approach ensures the agile response to topology change while the control traffic overhead is minimized. Once the nodes have link estimation, they can build the routing tree. The routing beacons contain a node's ETF to the sink. Upon receiving the beacon, a node can choose a parent that yields the smallest routing cost, which is the sum of the parent's ETF and the ETF to its parent. Then, RECPE forwards the packet along the one-way collection tree. The reliability is guaranteed using erasure-resilient in the application layer, which can avoid the retransmission requirement of the data link layer. Therefore, the low layers of protocol only need to provide the best-effort one-way data delivery. Meanwhile, to prevent self-interference, RECPE uses a method like pipeline delivery to forward packets: after transmitting a packet the forwarder waits at least 2 packet transmitting time before delivering the next one.

*3.3. One-Way Link Estimation and Construction of Collection Tree.* Generally, there are two approaches to estimate link quality. One is hardware-assisted link quality estimation which uses signal strength provided by the radio as the estimation of link quality. This approach usually cannot give the accurate value for link selection and only can provide a qualitative indication about the link: low quality and high quality. Another approach broadcasts beacons periodically at the link layer, and the neighborhood nodes can estimate the link quality according to the sequence number of the beacons. Obviously, a high beacon rate can lead to a more agile response to the network dynamic, simultaneously introduce high control overhead. However, low beacon rate results in slow scalability for dynamic network, forming the loss of a lot of packets when the network changes frequently.

RECPE exploits beacons controlled by Trickle algorithm to get link estimation. Trickle obtains good tradeoff between agility and efficiency. Trickle broadcasts routing beacons using a variable timer which varies between 100 ms and 1 hour. When the timer expires, RECPE doubles it until the maximum value. Whenever the RECPE detects an event that indicates the topology changes, it resets the timer to the minimum value. If a node has a timer value $\tau$, it will choose a random time within the interval of $[\tau/2, \tau]$ to broadcast beacon, which can prevent the collision among nodes whose timers synchronize. This strategy uses a few beacons to maintain collection tree without sacrificing agileness.

In RECPE, three events can reset the timer to the minimum value:

(i) a node receives a beacon packet with a "P" bit set.

(ii) A node is asked to forward a data packet whose routing cost is lower than its own.

Table 1: Route result under different metrics.

| Route | A → B → E | A → B → D → E | A → C → D → E | Route result |
|---|---|---|---|---|
| Minimum hop | 2 | 3 | 3 | A → B → E |
| ETX | 22.2 | 19.1 | 26.2 | A → B → D → E |
| ETF | 20 | 14 | 4.2 | A → C → D → E |

(iii) The routing cost of a node decreases or increases significantly.

When a new node joins the network, it will broadcast a beacon with a "P" bit set, which enables the new node to join the network rapidly. The last two events are detected by the data frame and routing frame, respectively, which indicates that the topology changes significantly because of the node moving or dying. The network will evaluate the link qualities and create new routing table, which can avoid the loss of plenty of packets when route inconsistency emerges in the network.

Figure 2 shows the format of routing beacon, which contains several inbound link estimations and some control information. The beacon advertises the routing cost of the current node to the sink and the inbound link quality of its several closest neighbors. Then, the neighbor node that receives the beacon can know the routing cost if it chooses the broadcasting node as its parent. By comparison in the routing table, it can also choose the node with minimum ETF as its routing forwarder. At the same time, the neighbor node can calculate the inbound link estimation from the broadcasting node by the routing beacon. We calculate the reception probability by using a windowed moving average. The ETF estimation is $(Seq_{i+k} - Seq_i)/k$. When broadcasting the beacon, the node will choose several newest (updated newly) neighbors' ETF to insert into the routing frame. The number of link information entries can be different in each routing beacon, which is indicated in the field "Num_entry."

There are also some control bits in the routing frame. "Beacon_flag" differentiates the routing beacon to the data frame. The pull bit "P" is set when a new node joins networks, so that all the neighbors can broadcast routing beacon as soon as possible and the new node can get the inbound link estimation quickly. The bit "C" indicates the relay node is encountering congestion; thus, it will not be chosen as the forwarder by its neighbor. The beacon with "F" set only originated from the sink but can be forwarded by any other nodes, which means that a file with the sequence number "File_ID" is decoded successfully in the sink. Thus, the source node will stop to inject the encoding packets into network, and the forwarders also will give up delivering the data packets with the same file sequence number.

*3.4. Data Delivery.* In RECPE, the node transmits all received packets and encoding packets generated locally to the parent selected by the link estimation algorithm. Figure 3 shows the format of the data frame, whose size is eight bytes excluding data payload. One bit "Data_flag" can distinguish the data frame from the beacon frame. The "Origin_ID" in a data packet indicates the source node that generates

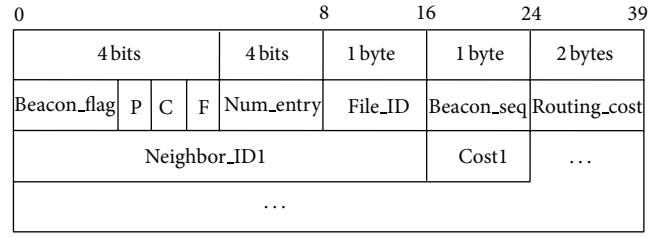



Figure 2: Routing Beacon frame format.

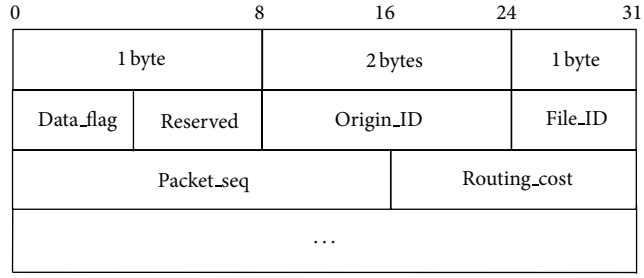



Figure 3: Data frame format.

the data packet. The field "File_ID" acts as an application dispatch identifier, which will dispatch the receiving packets to the corresponding decoder, and all packets produced by an application have their unique "File_ID" in the data frame. The "Packet_seq" can identify the unique packet generated by the same node. Particularly, the data frame also provides routing control field "Routing_cost" for dynamic route detecting, which can find link inconsistency quickly and reduce the control beacon by avoiding frequent routing frame exchanging.

The packets can be heard a few hops up and down the path. Therefore, to avoid self-interference, the node typically should wait at least two packets time when forwarding the data packets continuously. For example, considering the linear network shown in Figure 4, nodes A and D can simultaneously transmit the packets to their downstream neighbors without interference. The enforced delay depends on the packet rate of the radio and is empirically established by the following linear data delivery experiment.

A packet forwarding experiment shows the effect of a varying transmission wait time on a single node flow in the linear MICAz testbed. In the experiment, node A transmits some packets to node F without the end-to-end reliable control strategy; so the accumulative PLR (packet loss rate) can be defined as the ratio of the number of receiving packets in node F to the number of all generating packets in node A. Figure 5 shows the effect of interval time on packet loss

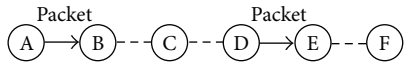



Figure 4: Pipeline delivery.

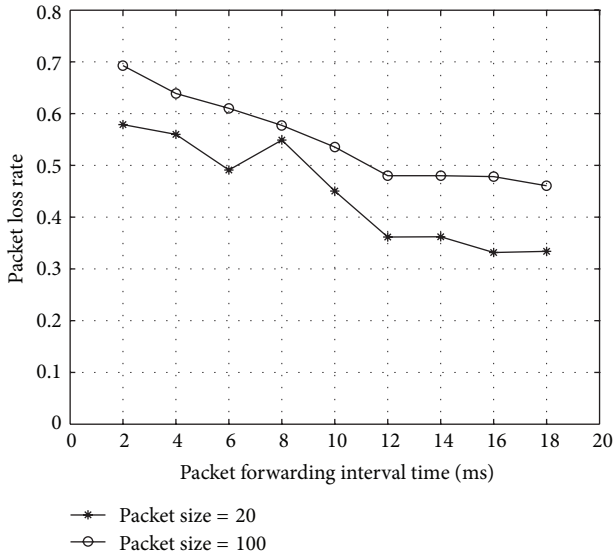



Figure 5: The effect of interval time on PLR during packet forwarding in a linear network.

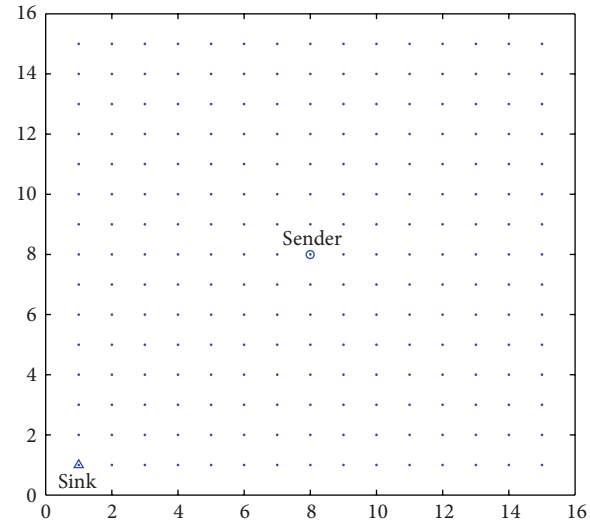



Figure 6: Grid network topology for simulation.

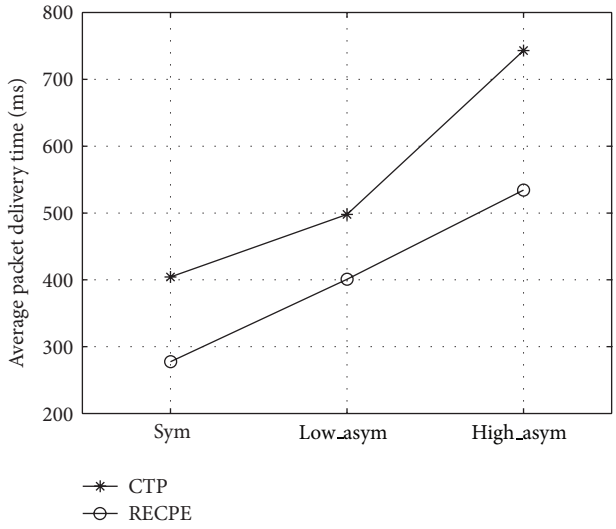



Figure 7: Packet delivery time of CTP and RECPE under symmetric, low and high asymmetric networks.

rate under different packet sizes during packet forwarding. The transmitting timers range from 2 ms to 18 ms. When the value is below 12 ms, PLR always decreases due to the self-interference between neighbor nodes. However, the PLR hardly changes when the transmitting interval time is more than 12 ms. If the interval time is too long, the network can be idle and the channel will be wasted. According to this experiment, the optimal waiting time is shown about 12 ms. It is reasonable that this optimal forwarding interval time is chosen based on the linear network since a lot of data packets often are generated in a burst by one node in WSN. This empirical interval time will be exploited in RECPE without considering the packet size.

*3.5. Reliability Design.* In RECPE, the reliability is guaranteed by encoding all packets in the application layer. Erasure-resilient codes can provide resilience to packet-level losses.

In the source node, the encoder produces a sequence of encoding packets from the set of input packets. For the erasure-resilient codes we use, each encoding packet is simply bitwise XOR of a specific subset of the input packets. In the sink, the decoder attempts to recover the original content from the encoding packets. Some simulation demonstrated that well-designed degree distribution can recover all original packets and only requires a few percent (less than 5%) of encoding packets beyond the number of original packets.

Provably good degree distributions for sparse parity check codes were first developed and analyzed in [21]. However, these codes are rate fixed, which means that only a predetermined number of encoding symbols are generated. Thus, in our application, it can lead to inefficiencies for decoding in the sink, as the source node will eventually be forced to transmit more encoding packets. Newer codes, called rateless codes, can avoid this limitation and allow unbounded numbers of encoding symbols to be generated on demand. Two examples of rateless codes, along with further discussion of the merits of ratelessness, may be found in [22, 23]. Both of these codes also have strong probabilistic decoding guarantees, along with low decoding overhead and average degrees. In our experiments, we use LT codes [22], and the degree distribution and importance sampling approach are described in [24]. The packets with the same file sequence number "File_ID" are forwarded to the same decoder of the application layer in the sink. The sink will broadcast a special beacon frame with "F" bit set when it succeeds to decode all original packets. This special noticing

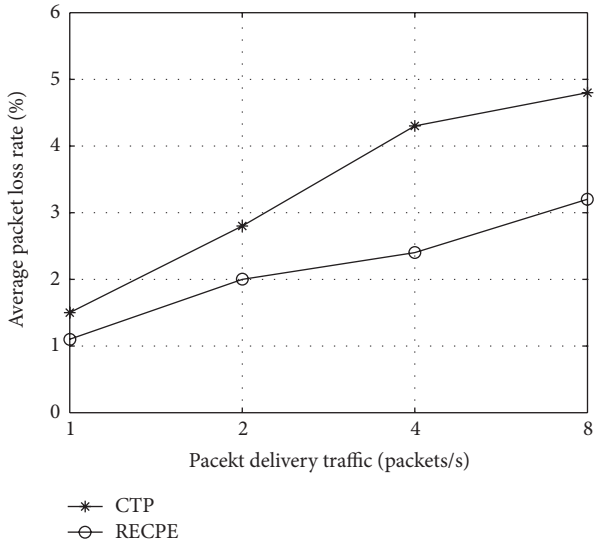



Figure 8: Packet loss rate of CTP and RECPE under different packet delivery traffic.

Table 2: Radio parameters for simulation.

| Radio parameters | $S_{11}$ | $S_{12}$ | $S_{21}$ | $S_{22}$ | N_F | W_G_N |
|---|---|---|---|---|---|---|
| High asymmetry | 3.7 | −3.3 | −3.3 | 6.0 | −106.0 | 4 |
| Low asymmetry | 0.9 | −0.7 | −0.7 | 1.2 | −105.0 | 4 |
| Symmetric link | 0 | × | × | 0 | −105.0 | 4 |

beacon will be scattered rapidly in the entire network with the highest precedence.

## 4. Experimental Evaluation

*4.1. Experimental Methodology.* To compare RECPE with CTP in large network topology, we use the TOSSIM [25] simulation tool to evaluate the relative performance in a variety of network densities. Except specified especially in the following simulation, the codes for CTP and RECPE were implemented based on the default parameters of TinyOS2.x. We generated a network topology in terms of power level gain by applying a theoretical propagation model. This model is configured by parameters in three aspects: channel, radio, and topology, which are described carefully in [26]. Table 2 presents the radio parameters which can generate three kinds of networks for our simulations: symmetric, low asymmetric, and high asymmetric link networks. To improve the quality of radio simulation, CPM [27] algorithm is also used to simulate the RF noise and interference. We utilize the noise trace from the Meyer library, which generates a statistical model in CPM.

Figure 6 plots the network topology, in which 225 nodes are placed in grid with different space (the distance between neighbors). Since plenty of packets often are generated in a burst by one node in some application such as event monitoring, in our simulations only node 112 located at (8, 8) generates the data packets and transmits them to sink node (0, 0) after encoding by erasure-resilient codes.

Table 3: Packet loss number and network stable speed under symmetric, low and high asymmetric networks.

| | Symmetric | Low asymmetric | High asymmetric |
|---|---|---|---|
| Packet loss num. | | | |
| CTP | 28 | 61 | 59 |
| RECPE | 20 | 58 | 41 |
| Network stable seq. | | | |
| CTP | 42 | 103 | 185 |
| RECPE | 27 | 94 | 176 |

*4.2. Simulation Result.* Firstly, we conducted a simulation to compare the average packet delivery time between CTP and RECPE when delivering 1000 packets from the sender to the sink. The packet traffic generated in the sender is 10 packets per second. The traveling time of each packet on the route is calculated from the sender to the sink, which is referred to as packet delivery time. Figure 7 shows the simulation result when the space is 2 meters.

Figure 7 showed that RECPE can obtain better latency when delivering data packet from the sender to the sink under symmetric, low and high asymmetric networks due to better route chosen and no retransmission. That indicates that RECPE also would spend less time to deliver a file under such networks.

At the same simulation, we counted the packet loss number and network stable speed. After delivering 1000 packet, RECPE only lost 20, 58, and 41 packets in symmetric, low and high asymmetric networks. However, CTP is 28, 61, and 59 packets, respectively, which are shown in Table 3. Simultaneously, we also observed the routing stable speed of protocol (when plenty of packets are generated in a burst in the resource node, CTP and RECPE usually spend time to form a stable route; before that, lots of packet are lost due to the variable route). CTP can find stable route under delivering 42, 103, and 185 packets, respectively, under three kinds of networks, while RECPE is faster, respectively, after 27, 94, and 176 packets delivery.

We also measured the effect of packet traffic on packet loss rate in the low asymmetric network when the space is 2 meters. In this simulation, 1000 packets are delivered from the sender to the sink. Figure 8 showed that RECPE always got lower packet loss rate than CTP, particularly, when the network had high traffic rate. The improved performance is contributed by less radio interference during packet transmission due to the pipeline data delivery and no retransmission requirement in RECPE.

We further investigated the efficiency of data delivery in terms of the delivery cost which accounts for all data and control overhead. Delivery cost is defined as the ratio of the total bytes transmitted or forwarded by all the nodes (including the source node) to the size of file delivered from the sender to the sink. Note that the overhead also includes the frame header and control information in each packet. Lower delivery cost means consuming less energy when delivering the same content in the network. In all simulation

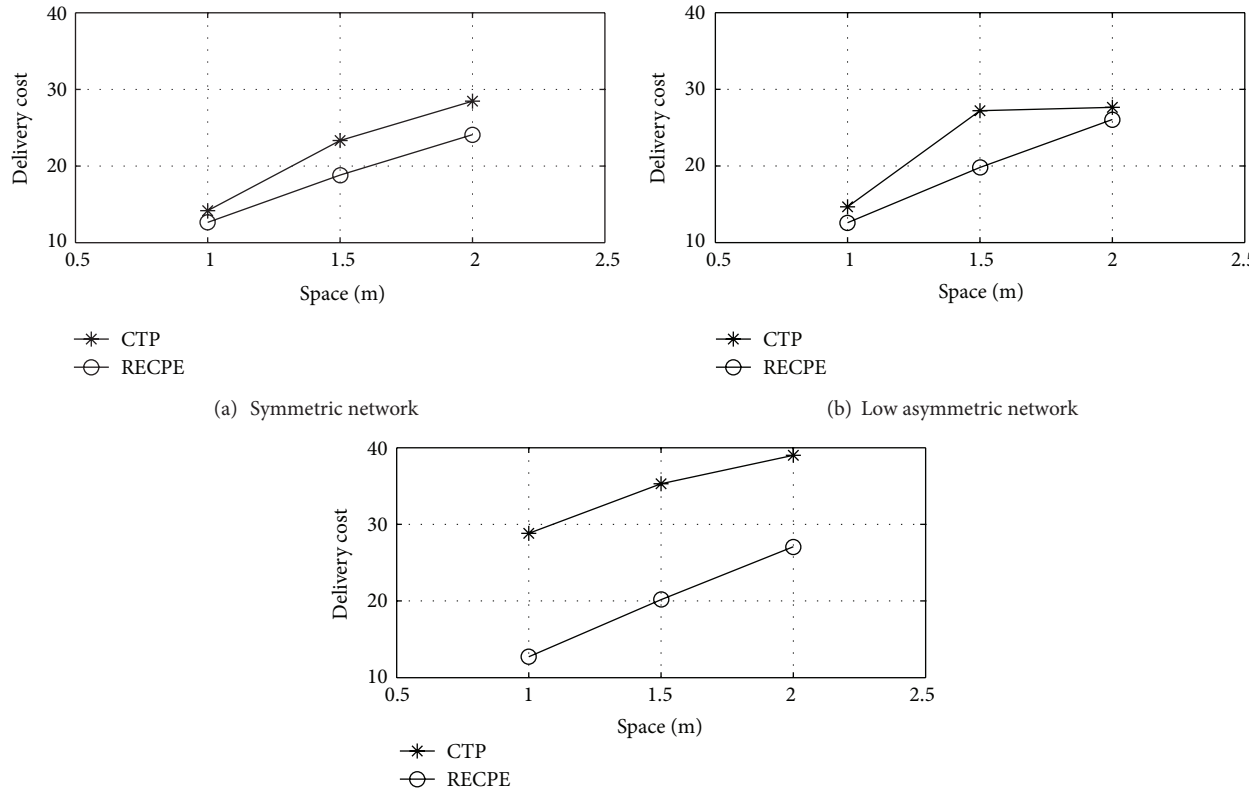



Figure 9: Delivery cost of CTP and RECPE under symmetric, low and high asymmetric networks.

scenarios, the file of 30k bytes packaged to 1000 packets is transmitted from the sender to the sink. In RECPE, these packets are encoded into enough packets by erasure-resilient codes in the sender and delivered into the network until the sink can recover all the original packets. Figure 9 plots the delivery cost of CTP and RECPE protocols in three kinds of networks.

As shown in Figure 9, in symmetric link network, RECPE obtained less delivery cost than CTP and kept the tendency under the other space simulations. In low and high asymmetric networks, the similar simulation results are obtained too. Another observation is that the delivery cost in RECPE almost increases linearly, while CTP fluctuates drastically. Totally, the performance of RECPE is hardly affected by link asymmetry of networks, which implies that RECPE will perform stably in almost all complex deployment scenarios.

## 5. Conclusion

In this paper, we present RECPE, a reliable bulk data collection protocol for large-scale multihop WSN. RECPE exploits beacons controlled by Trickle to obtain link estimation and constructs a one-way collection tree. The data packets are delivered through a pipeline approach. We also present the formats of beacon frame and data frame. RECPE differs from the state-of-the-art CTP protocol in its use of erasure-resilient codes to guarantee the delivery reliability, which avoids the requirement of retransmission in data link layer. RECPE can obtain significant performance in terms of energy efficiency and scalability in various density networks.

We evaluated the delivery cost of CTP and RECPE via TOSSIM simulation. Our results show that

(i) RECPE can get lower average packet delivery time than CTP under symmetric, low and high asymmetric networks. Also, RECPE obtain better packet loss rate than CTP under various packet delivery traffic.

(ii) The delivery costs of CTP and RECPE are comparable in symmetric and low asymmetric networks. However, RECPE performs much better in high asymmetric network.

(iii) RECPE also obtains significant stabilization in various asymmetric and different density networks.

## Acknowledgment

This work was supported by Inha University, Republic of Korea.

## References

[1] C. E. Perkins and P. Bhagwat, "Highly dynamic Destination-Sequenced Distance-Vector routing (DSDV) for mobile computers," *ACM SIGCOMM Computer Communication Review*, vol. 24, pp. 234–244, 1994.

[2] C. E. Perkins and E. M. Royer, "Ad-hoc on-demand distance vector routing," in *Proceedings of the 2nd IEEE Workshop on Mobile Computing Systems and Applications (WMCSA '99)*, pp. 90–100, New Orleans, La, USA, February 1999.

[3] D. S. J. De Couto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multi-hop wireless routing," *Wireless Networks*, vol. 11, no. 4, pp. 419–434, 2005.

[4] U. Ashraf, S. Abdellatif, and G. Juanole, "An interference and link-quality aware routing metric for wireless mesh networks," in *Proceedings of the 68th Semi-Annual IEEE Vehicular Technology (VTC Fall '08)*, September 2008.

[5] L. Sang, A. Arora, and H. Zhang, "On link asymmetry and one-way estimation in wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 6, no. 2, article 12, 2010.

[6] L. Sang, A. Arora, and H. Zhang, "On exploiting asymmetric wireless links via one-way estimation," in *Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '07)*, pp. 11–21, ACM, Montreal, Canada, September 2007.

[7] J. J. Lei and G. I. Kwon, "Reliable data transmission based on erasure-resilient code in wireless sensor networks," *KSII Transactions on Internet and Information Systems*, vol. 4, no. 1, pp. 62–77, 2010.

[8] L. Philip, P. Neil, C. David, and S. Scott, "Trickle: a self-regulating algorithm for code propagation and maintenance in wireless sensor networks," in *Proceedings of the 1st Conference on Symposium on Networked Systems Design and Implementation*, vol. 1, USENIX Association, San Francisco, Calif, USA, 2004.

[9] O. Gnawali, R. Fonseca, K. Jamieson, and P. Levis, "CTP: robust and efficient collection through control and data plane integration," Stanford Information Networks Group SING-08-02, 200, 2008.

[10] G. Zhou, T. He, S. Krishnamurthy, and J. A. Stankovic, "Models and solutions for radio irregularity in wireless sensor networks," *ACM Transactions on Sensor Networks*, vol. 2, no. 2, pp. 221–262, 2006.

[11] M. Z. Zamalloa and B. Krishnamachari, "An analysis of unreliability and asymmetry in low-power wireless links," *ACM Transactions on Sensor Networks*, vol. 3, no. 2, Article ID 1240227, 2007.

[12] S. Kannan, A. K. Maria, A. Saatvik, and L. Philip, "The $\beta$-factor: measuring wireless link burstiness," in *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems*, ACM, Raleigh, NC, USA, 2008.

[13] C. Bin Bin, H. Shuai, Z. Mingze, C. Mun Choon, and A. L. Ananda, "DEAL: discover and exploit asymmetric links in dense wireless sensor networks," in *Proceedings of the 6th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON '09)*, IEEE Press, Rome, Italy, June 2009.

[14] R. P. Liu, Z. Rosberg, I. B. Collings, C. Wilson, A. Y. Dong, and S. Jha, "Energy efficient reliable data collection in wireless sensor networks with asymmetric links," *International Journal of Wireless Information Networks*, vol. 16, no. 3, pp. 131–141, 2009.

[15] J. W. Byers, M. Luby, and M. Mitzenmacher, "A digital fountain approach to asynchronous reliable multicast," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 8, pp. 1528–1540, 2002.

[16] H. Andrew, S. David, and T. TAri, "Over-the-air programming of wireless sensor networks using random linear codes," in *Proceedings of the 7th International Conference on Information Processing in Sensor Networks*, pp. 457–466, IEEE Computer Society, April 2008.

[17] W. H. Jonathan and C. David, "The dynamic behavior of a data dissemination protocol for network programming at scale," in *Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems (SenSys '04)*, pp. 81–94, ACM, Baltimore, Md, USA, November 2004.

[18] I. H. Hou, Y. E. Tsai, T. F. Abdelzaher, and I. Gupta, "AdapCode: adaptive network coding for code updates in wireless sensor networks," in *Proceedings of the 27th IEEE Communications Society Conference on Computer Communications (INFOCOM '08)*, pp. 2189–2197, Phoenix, Ariz, USA, April 2008.

[19] M. Rossi, N. Bui, G. Zanca, L. Stabellini, R. Crepaldi, and M. Zorzi, "SYNAPSE++: code dissemination in wireless sensor networks using fountain codes," *IEEE Transactions on Mobile Computing*, vol. 9, no. 12, pp. 1749–1765, 2010.

[20] P. Cataldi, A. Tomatis, G. Grilli, and M. Gerla, "A novel data dissemination method for vehicular networks with rateless codes," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '09)*, April 2009.

[21] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, "Efficient erasure correcting codes," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 569–584, 2001.

[22] M. Luby, "LT codes," in *Proceedings of the 34rd Annual IEEE Symposium on Foundations of Computer Science*, pp. 271–280, November 2002.

[23] P. Maymounkov and D. Mazières, "Rateless codes and big downloads," in *Proceedings of the International Workshop on Peer-to-Peer Systems (IPTPS '03)*, pp. 247–255, 2003.

[24] E. Hyytiä, T. Tirronen, and J. Virtamo, "Optimizing the degree distribution of LT codes with an importance sampling approach," in *Proceedings of the 6th International Workshop on Rare Event Simulation*, Bamberg, Germany, 2006.

[25] L. Philip, L. Nelson, W. Matt, and C. David, "TOSSIM: accurate and scalable simulation of entire TinyOS applications," in *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems (SenSys '03)*, pp. 126–137, ACM, Los Angeles, Calif, USA, November 2003.

[26] M. Zuniga, "Building a network topology for tossim," USC Technical Report, 2011.

[27] H. Lee, A. Cerpa, and P. Levis, "Improving wireless simulation through noise modeling," in *Proceedings of the 6th International Symposium on Information Processing in Sensor Networks (IPSN '07)*, pp. 21–30, ACM, Cambridge, Mass, USA, April 2007.