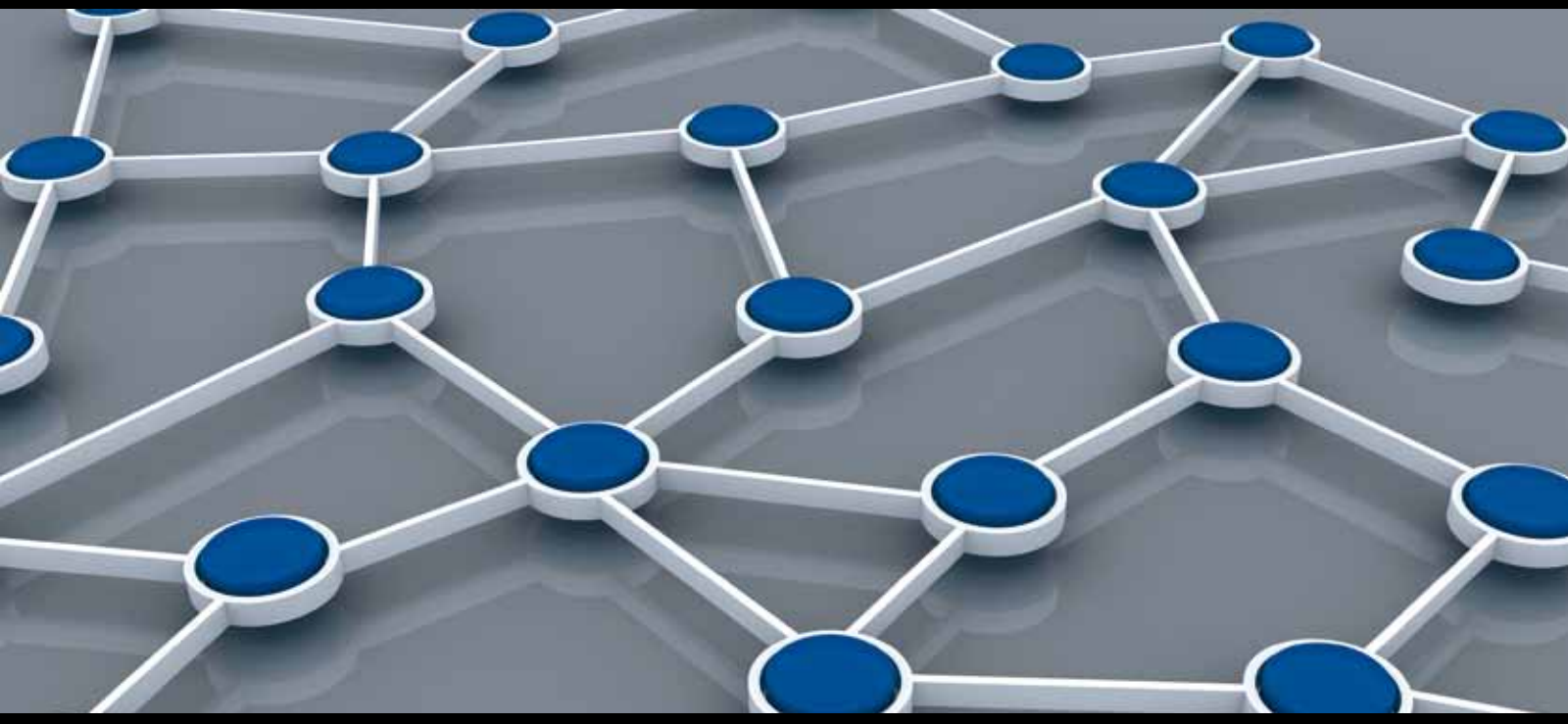
International Journal of Distributed Sensor Networks

Advanced Self-Organizing Technologies over Distributed Wireless Networks

Guest Editors: Mugen Peng, Zhiguo Ding, Yiqing Zhou, and Yonghui Li



Advanced Self-Organizing Technologies over Distributed Wireless Networks

International Journal of Distributed Sensor Networks

Advanced Self-Organizing Technologies over Distributed Wireless Networks

Guest Editors: Mugen Peng, Zhiguo Ding, Yiqing Zhou, and Yonghui Li

Editorial Board

Contents

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 821982, 2 pages
doi:10.1155/2012/821982

Editorial

Advanced Self-Organizing Technologies over Distributed Wireless Networks

Mugen Peng,[1] Zhiguo Ding,[2] Yiqing Zhou,[3] and Yonghui Li[4]

[1] *The Key Laboratory of Universal Wireless Communications for the Ministry of Education, Beijing University of Posts and Telecommunications, China*
[2] *The School of Electrical, Electronic, and Computer Engineering, Newcastle University, UK*
[3] *Institute of Computing Technology, Chinese Academy of Sciences, China*
[4] *School of Electrical and Information Engineering, The University of Sydney, Australia*

Correspondence should be addressed to Mugen Peng, pmg@bupt.edu.cn

Received 12 December 2012; Accepted 12 December 2012

Self-organizing Networks (SON) have experienced an explosive growth during the last few years. The principal objectives of introducing self-organizing network are to effectuate substantial operational and capital expenditure (O/CAPEX) reductions by diminishing human involvement in network operational tasks, and to optimize network capacity, coverage, and service quality. The general idea of SON is to integrate network planning, configuration, and optimization into a single, mostly automated process requiring minimal manual intervention. Most research work focused on the mobile cellular network. However, the SON over the distributed sensor network can enhance the network robustness, improve the energy efficiency, and improve the transmission performances. Main features in distributed sensor networks include scalability with respect to the number of sensor nodes, self-configuration, self-optimization, self-healing, energy saving, and a sufficient degree of connectivity among nodes.

In this special issue, the SON protocols and technical solutions over both the distributed and centralized wireless networks are focused on. We have received a total of 21 original submissions, out of which ten papers were accepted for publication after peer reviewing. We regret that we had to reject many good papers due to the limited number of papers that can be published with in this special issue. The accepted papers cover a broad area of SON-related topics, including the self-configuration, self-optimization, self-healing, self-routing and self-clustering.

The physical cell identity (PCI) self-configuration is a main feature for the advanced wireless network; a paper entitled "*Automatic distributing schemes of physical cell identity for self-organizing networks*" by Y. Wei et al., proposes an automatic distributing PCI scheme to optimize the PCI reuse distance and decrease the multiplexing interference. Simulation results illustrate that the proposal can achieve a significant performance gain even under the condition of severe PCI deficiency.

For the self-optimization related aspect, there are 4 accepted papers. In a paper entitled "*Self-organizing energy-saving management mechanism based on pilot power adjustment in cellular networks*" by Y. Peng et al., the authors present a regional self-organizing energy-saving management mechanism through pilot power adjustment in cellular network. Another paper entitled "*Self-optimization of coverage and capacity in LTE networks based on central control and decentralized fuzzy Q-learning*" by J. Li et al. proposes a central control mechanism that utilizes the fuzzy Q-learning algorithm, where each eNB is a learning agent that tries to optimize its antenna downtilt automatically using information from its own and its neighboring cells, and the initialization and the termination of the optimization processes of all agents are in the control of the central entity. A paper entitled "*Hybrid BS-cooperative power management scheme for self-organized sleep mode in virtual cell-based femto networks*" by W. Zheng et al. discusses a hybrid BS-cooperative power management scheme for self-organized sleep mode in virtual cell-based femto networks. In the fifth paper entitled "*Self-organized connectivity control and optimization subjected to dispersion of mobile ad hoc sensor networks*" by Z. Mi et al., a distributed link removal algorithm

to reduce redundant communication links while preserving global connectivity and a distributed connectivity control system to disperse a team of mobile sensors with guaranteed connectivity and collisions avoidance are presented.

For the self-healing related aspect, a paper entitled "*Centralized management mechanism for cell outage compensation in LTE networks*" by L. Wenjing et al. presents a centralized cell outage compensation management mechanism and a concrete autonomic particle swarm compensation algorithm.

A paper entitled "*A novel hybrid self-organizing clustering routing algorithm*" and another paper entitled "*Vascular network-based nonuniform hierarchical fault-tolerant routing algorithm for wireless sensor networks*" researched the self-routing related techniques. A hybrid self-organizing clustering routing protocol by combining energy sense and maximum connectedness is introduced. A vascular network based fault tolerant routing algorithm is presented by nonuniform hierarchical clustering.

To support the SON functionalities, the location technique is very urgent. In another paper entitled "*Cooperative group localization based on factor graph for next-generation networks*" the weighted factor graph-based cooperative group localization algorithm which incorporates the optimal weights based on the information reliability is presented. Meanwhile, to improve the energy efficiency by the modification of choosing cluster-heads formula and the steady-state phase in the distributed sensor network, a vice cluster head for each cluster during the communication process is formulated, which aims to diminish the energy consumption spent on the reclustering and prolong the time of being steady-state phase.

Acknowledgments

Before the end of this editorial, we would like to thank the anonymous reviewers for their great efforts in reviewing the submitted manuscripts, without which this special issue would not have been published as scheduled. We are also thankful to the editorial office manager, Nourhan Ashraf, for his supportive guidance during the whole process in the organization of this special issue.

Mugen Peng
Zhiguo Ding
Yiqing Zhou
Yonghui Li

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 649609, 6 pages
doi:10.1155/2012/649609

Research Article

Improved LEACH Routing Communication Protocol for a Wireless Sensor Network

Fuzhe Zhao, You Xu, and Ru Li

Department of Computer Science and Technology, Central China Normal University, Wuhan 430079, China

Correspondence should be addressed to Fuzhe Zhao, fzh_zhao@yahoo.com.cn

Received 7 June 2012; Revised 7 November 2012; Accepted 19 November 2012

Academic Editor: Zhiguo Ding

A WSN (wireless sensor network) consists of thousands of sensor nodes with limited energy, memory, and computation capability. The applications of WSN in some extreme environment make sensor nodes difficult to replace once they use up the resource. Hence, many researchers in this field focus on how to design a property routing protocol to prolong the life span of the network. The classical hierarchical protocols such as LEACH and LEACH-C have better performance in saving the energy consumption. However, the choosing formula neglecting the change of nodes' energy will make the nodes acting as cluster heads too many times die early owing to the consumption of too much energy. Also, the high frequency of reclustering wastes certain amount of energy. In order to make the energy distribute more evenly among different nodes, we improve the tradition equation used for selecting cluster heads with considering the dynamic change of nodes' energy. Meanwhile, we propose to establish a vice cluster head for each cluster during the communication process, which aims to diminish the energy consumption spent on the reclustering and prolong the time of being in a steady-state phase. Simulations show that our improved protocol performs better than the LEACH and the LEACH-C.

1. Introduction

WSN (wireless sensor network) [1] consists of more than hundreds of small sensor nodes which have limited power, memory, and computational capabilities. The application of the WSN involves many fields, such as the military battlefield, forest fire detection, and other extreme environments [2]. In these situations, it is difficult to replace the dead nodes caused by energy's depletion with new ones to supply energy for the system. Therefore, making sensor nodes working as long as possible is the main method to maximize the lifecycle of the system. Because the energy's consumption of sensor node mainly originates from the long distance transmission of data along the routing path, an efficient routing path formed by the routing protocol will have a great impact on the energy's consumption [3]. So how to design an energy-efficient routing protocol becomes the main goal for the wireless sensor network.

The conventional wire routing protocol cannot adapt to the wireless sensor network due to the inherent property of WSN. Hence, many protocols have been proposed to satisfy the need of WSN. The cluster idea present in [4, 5] has a significant reflection to the research of WSN routing protocols. It organizes the sensor nodes into different clusters, with a cluster head in each cluster executing the data collecting and transmitting tasks for other member nodes. The main idea introduced in the hierarchical routing protocols is to divide the whole network into two or more levels with each level performing different tasks. LEACH (low energy adaptive clustering hierarchy) [5] is a classical version in the hierarchical routing protocol family which divides the communication process into rounds with each round including a set-up phase and a steady-state phase. In the set-up phase, some sensor nodes are selected as cluster heads (CHs) according to certain rules and other nodes join in the clusters as member nodes. In the steady-state phase, the CHs collect and aggregate the data coming from their own cluster members and then transmit them to a base station (BS). Based on the LEACH, LEACH-C differs from LEACH in that it uses a central algorithm to finish the choice of CHs in each round, but it needs all the sensor nodes to transmit their information to remote BS [6]. No matter LEACH or LEACH-C, the existence of cluster head (CH) in a cluster avoids the long transmission distance to BS in a communication process for each sensor node. Meanwhile, the aggregation of the data

on the cluster heads reduces almost simultaneously a lot of redundancy data message coming from different member nodes [7]. However, due to the inherent characteristic of LEACH and LEACH-C, the unnecessary energy consumption caused by the unreasonable choosing formula and the high frequency of reclustering among sensor nodes will cause the uneven energy distribution and waste a certain amount of energy in the whole network. Based on the ideology of traditional LEACH, we modify the choosing formula for considering the dynamic change of sensor nodes' energy and change the process of choosing CHs to reduce the frequency of re-cluttering.

The remainder of the paper is organized as follows. In Section 2, we present LEACH and LEACH-C in detail. Section 3 develops our proposed protocol and provides its theoretical analysis. We evaluate the performances of three protocols via ns2 in Section 4. In Section 5, conclusion is derived based on the analysis and simulation.

2. LEACH and LEACH-C Protocols

LEACH is the most popular hierarchical cluster based routing protocol for a wireless sensor network [5]. In LEACH, the nodes in the deployed area are organized into local clusters and the communication process is divided into rounds with each round including set-up and steady-state phases. During the communication process, each cluster has a cluster head (CH) which is responsible for creating and manipulating a TDMA (time division multiple access) schedule table used by its member nodes to get when to transmit data packets. Once some emergency affairs happen in the monitor area, the sensor nodes are triggered to send data to their own cluster head instead of the remote BS by themselves. The cluster head mainly collects the data coming from different member nodes and does some aggregation to diminish the redundancy firstly and then transmits them to BS. In the whole process, the cluster head just works as a relay node to help member nodes shorten the transmission distance so as to save energy. As for the set-up and steady-state phases in a round, they can be described as follows.

2.1. Set-Up Phase. After finishing the deployment of sensor nodes, each node in the monitor field decides independently of other nodes whether it can become a cluster head in the current round. During the phase, each node generates a random number between 0 and 1 and then compares the threshold value with [8]

$$T(n) = \begin{cases} \dfrac{p}{1 - p * [r \bmod (1/p)]}, & n \in G, \\ 0, & n \notin G, \end{cases} \quad (1)$$

where p is the percentage of cluster heads over all nodes in the network, r is the number of rounds of selection, and G is the set of nodes that have not been selected as cluster heads in round $1/p$. The node whose number is larger than the threshold will select itself as a cluster head and then broadcasts the message to its surround sensor nodes. In this phase, a node may receive more than one broadcast message from different cluster heads, but the node can judge its distance to a cluster from the strength of received broadcast signal; the stronger the signal, the closer to a cluster. So the node whose number is smaller than threshold will only send request message containing its ID to the cluster which has the strongest signal strength for saving energy spent on the transmitting distance. Once the cluster head receives request message coming from one node, it records the node's ID and proclaims it as its member node. After the message exchanges between cluster heads and normal nodes, each CH gets its own member nodes' information about IDs and each normal node gets which cluster it belongs to. Based on the message it records, the CH creates a TDMA schedule table and broadcasts it to the cluster members. Therefore, all the member nodes get their idle slots for data transmission, and then the steady-state phase starts.

2.2. Steady-State Phase. The establishment of a cluster head in each cluster during the set-up phase provides a guarantee for the data transmission in a steady-state phase. In normal circumstances, member nodes can turn off their radio until they sense the necessary environment data. If there are some data in need to transmit, they will send the data to CH during the idle slots recorded in the TDMA schedule table. As for the CHs, they have to keep up communication status at all times so as to receive the data from different member nodes. After receiving all the data sent by their members, CHs will aggregate them firstly and then send them to BS. Because some sensor nodes may sense similar environment data, the aggregation on the cluster head can diminish unnecessary bandwidth cost and communication traffic, which has a positive reflection to the energy's consumption. Also, the data transmission distance becomes shorter comparing with transmitting to BS separately for each member node, which can save some energy for the member nodes. However, the heavy tasks executing on CH can lead to too much energy consumption. In order to avoid making the CHs die early and cause the cascade effect in the network, a new round begins and new clusters will be rebuilt in the whole network.

2.3. LEACH-C Protocol. Based on the LEACH, LEACH-C also organizes the sensor nodes into clusters with each cluster a cluster head and divides a round into set-up and steady-state phases. It differs from LEACH only in that it uses a high-energy base station to finish the choice of cluster heads. In the set-up phase of each round, every sensor node sends its information about energy to remote BS. Then the BS selects the cluster heads based on the energy information and broadcasts the IDs of cluster heads to other member nodes. This method can make the nodes with more energy and more chance to become the cluster head in the current round. But in this phase, every sensor node needs to send its ID and energy information to remote BS to compete for the role of cluster heads, which causes energy consumption on the long distance transition. Equation (2) has a good description of transition distance influence on the consumption of energy [3]:

$$E_{\text{trans}}(k, d) = \begin{cases} kE_{\text{elec}} + k\varepsilon_{fs}d^2 & (d < d_0), \\ kE_{\text{elec}} + k\varepsilon_{mp}d^4 & (d \geq d_0). \end{cases} \quad (2)$$

The transmission energy of transmitting a k-bit message over a distance d will cost E_{trans}; ε_{fs} and ε_{mp} are the power consumption of transferring 1 bit of data in different condition. We can obtain that there will be a certain amount of energy needed to be spent on the transition of energy information for each sensor node in every round, which cannot be neglected in communication; especially the BS locates far away from the monitor field and the network has a lot of sensor nodes.

Although LEACH and LEACH-C protocols act in a good manner, they also suffer from many drawbacks like the following.

(i) CHs' selection is random, which does not take into account the residual energy of every node or need the support of BS.

(ii) The high frequency of reclustering wastes a certain amount of energy.

(iii) It cannot cover a large area.

(iv) CHs are not uniformly distributed, where CHs can be located at the edge of the cluster.

3. The Improvement to the Cluster-Based Leach Protocol

Motivated by the original LEACH, LEACH-C and other improvement protocols [9, 10], we propose a modification to the cluster head selection process to reduce energy consumption. For a microsensor network, we first make the following assumptions.

(i) The base station (BS) is located far from the sensors and is immobile.

(ii) All nodes in the network are homogenous and have limited energy with an indentify ID.

(iii) All nodes are able to reach BS and can communicate with each other.

(iv) Cluster heads perform data compression and aggregation.

In the improvement, we also make use of the clustering ideology in hierarchical and divide a round into a set-up phase and steady-state phase. The set-up phase will use improved formula to select appropriate cluster heads (CHs) which are responsible for collecting data from their member nodes and transmitting them to BS. As the introduction in Section 3, CHs will consume more energy than member nodes because of the heavy tasks. In order to avoid making the CHs die early, LEACH and LEACH-C take the measure of beginning a new round and rebuilding the clusters. However, in this paper, we will make use of the member nodes' information dynamically achieved by cluster heads in the steady phase to choose the vice cluster heads (VCHs) which take over the role of cluster heads in the later period of steady phase. Comparing with the traditional LEACH and LEACH-C, the VCHs proposed will diminish the frequency of reclustering in the same interval and prolong the time of being in steady-state phase, which will prolong the lifecycle of the whole network.

3.1. Choosing Cluster Heads (CHs) in the Set-Up Phase. Based on the fact that LEACH does not take into account the residual energy of the nodes during the selection of cluster heads in the set-up phase, we develop the current energy and the times being selected CH or VCH which will be shown later in the paper. We first consider that the threshold $T(n)$ is modified to the following equation:

$$T(n) = \begin{cases} \dfrac{p}{1 - p * [r \bmod (1/p)]} \\ \times \left[\dfrac{E_{n\_\text{current}}}{E_{n\_\text{init}}} + \left(1 - \dfrac{E_{n\_\text{current}}}{E_{n\_\text{init}}}\right) \right. \\ \quad \left. \times \dfrac{p}{\text{CH\_times} + \text{VCH\_times} + 1} \right], & n \in G, \\ 0, & n \notin G, \end{cases} \tag{3}$$

where p is the percentage of cluster heads over all the nodes in the network, R is the number of rounds selection in current time. G is the set of nodes that have not been selected as cluster heads in round $1/p$. E_{n_current} is the residual energy of the node and E_{n_init} is the initial energy of every node. CH_times (VCH_times) is the times of being selected CH (VCH_times) once. Deducing from (3), we can obtain that the larger the E_{n_current}, the larger the $T(n)$. So we can infer that the node which has more energy will have a bigger probability to become the cluster head in the current round. At the same time, if a node acts as CH or VCH for too much time, the energy it consumes will be larger than other sensor nodes. However, the improved equation can make the probability of a node acting as too much time CH or VCHs to become CH again lower. We can observe that the improved formula adds some helpful determinacy factors in the selection of cluster heads, which is beneficial to the stabilization of clusters. The authors in [11] proposed that if there are too much cluster heads in the deployed area, it will cause some unnecessary consumption of energy. In order to limit the cluster heads' number to a reasonable range, we develop the simulated annealing algorithm to create appropriate numbers of cluster heads which is about 4%-5% of the total sensor nodes introduced in [12]. After finishing the selection of cluster heads in the set-up phase by using the improved equation and simulated annealing algorithm, the steady-state phase of a round begins.

3.2. Vice Cluster Heads' (VCHs') Establishment during the Steady-State Phase. In the steady-state phase of LEACH and LEACH-C protocols, the cluster heads will deplete more energy than member nodes because they have to take the responsibility of aggregating and relaying data to remote BS for their member nodes. In order to avoid making the cluster heads die early after undergoing certain of communication time, a new round begins to reorganize the nodes into clusters and reselect the cluster heads. So, all the nodes have to reappraise themselves and rebuild the cluster heads in order to campaign for new cluster heads. As a result, it consumes some energy spent on recompeting the cluster heads and shortens the total time of being in steady-state phase. In this paper, we propose a new scheme to prolong

the time of being in steady phase and diminish the frequency of recluster. The new scheme works as follows.

During the data communication in steady-state phase, because all member nodes send the data sensed from environment to their own cluster head, the cluster head can have the opportunity to learn the status information of its members. Based on this fact, the cluster head can record the information of different member nodes dynamically, the format of the information just like $\langle id, E\rangle$, which means the member node *id* has residual energy *E*. Through this way, the CH will have global energy information to its member nodes. In order to prolong the time of being in steady-state phase and delay a new round's coming, CH will appoint a member node which has the maximum energy in cluster to take over the role of it if consuming too much energy in the later steady-state phase of current round. We can call the member node which is appointed by the CH vice cluster (VCH). In order to make the rest of member nodes get the *id* of VCH, the CH will broadcast this message containing the VCHs *id* to other member nodes. After that, the CH itself will become a normal member node because of the too much energy consumption and the establishment of VCH in a cluster. Since then, all the member nodes will send their data to VCH, which compresses data firstly and then relays them to BS. We can observe that the establishment of VCH in cluster can prolong the communication time of being in steady-state phase and delay the coming of a new round. But after a certain time, the VCH will also consume more energy than the member nodes due to the heavy tasks undertaken as previous CH. To avoid making the VCH die early, a new round of selecting CHs in the set-up phase will start among all the nodes. So we can call the whole communication in our improved protocol as the cycle of "CH-VCH-CH." It can be described using the Figure 1.

In the proposed protocol, we take the measure of selecting a VCH for each cluster in the later period of the steady-state phase in a round by using the energy information achieved by CH, which can diminish the frequency of reclustering and prolong the time of being in steady-state phase. In the whole communication phase of a round, CH and VCH have the same role to undertake collecting data from member nodes and relaying them to BS. The difference is that the CH takes the responsibility in the earlier stage of the steady-state phase in a round, while VCH replaces the CH and works in the later stage of the steady phase. Also, the CHs selection originates the competition among all the nodes in the set-up phase. However, VCH is established directly by CH in the later stage of the steady-state phase in a round. We can obtain that the method of establishing VCH is simple and rapid comparing with the generation and cooperation of random numbers in the set-up phase. They all have a good benefit to the saving of energy in the whole network.

4. Simulation Result

In this section, we examine the improved protocol through NS2 [13]. A network of 100 nodes is deployed in an area of 100 m × 100 m with BS at (50, 175). The main parameters of the simulation experiments are described in Table 1.

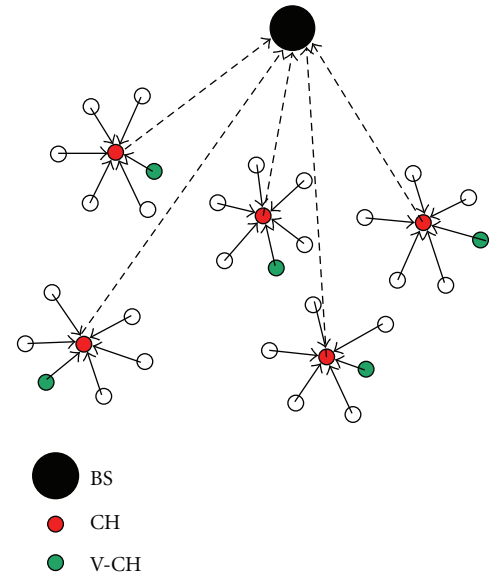


Figure 1: Improved hierarchical protocol working process.

Table 1: Summary of the parameters used in the simulation experiments.

Parameter	Value
Number of nodes	100
Network size	100 m × 100 m
BS location	(50, 175)
Radio speed	2 Mbps
Data size	25 bytes
Initial node power	2 *J*
Numbers of CH	5
Simulation time	600 secs

In order to compare the advantage of the improved protocol with the original LEACH and LEACH-C, we use three performance metrics for comparison: numbers of nodes alive over simulation time, the consumption of the whole network's energy over simulation rounds, and the message amounts created by the three different protocols. The simulation results are illustrated in Figures 2–4.

Observed from Figure 2, we can obtain that the numbers of nodes alive in improved protocol surpasses the nodes alive in LEACH and LEACH-C protocols at the same time. The network using LEACH routing protocol stops its life at about 450 seconds and LEACH-C can maintain the lifecycle to about 500 seconds, while the improved protocol can prolong its lifetime to 570 seconds.

Because of the introducing of VCH, some energy spent on the reclustering and recomputing among different nodes gets certain of economy. Observed from Figure 3, we can obtain that our improved routing protocol consumes less energy than LEACH and LEACH-C over the simulation rounds.

Due to our modification to the steady-state phase, the times for choosing the cluster heads and broadcasting the

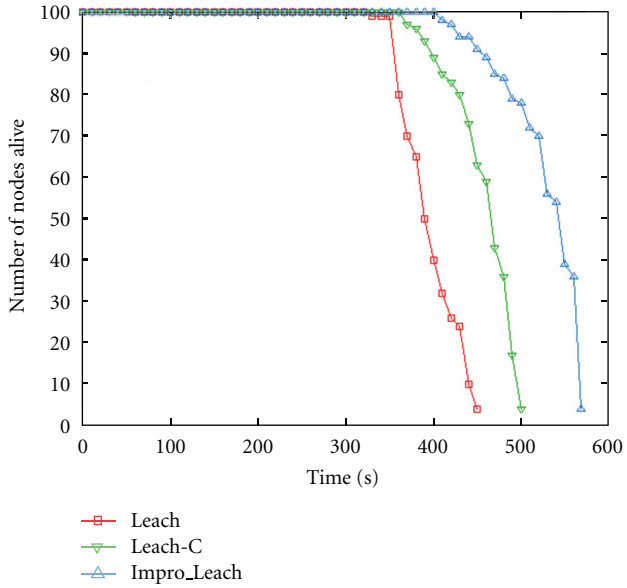


Figure 2: Number of nodes alive over simulation time.

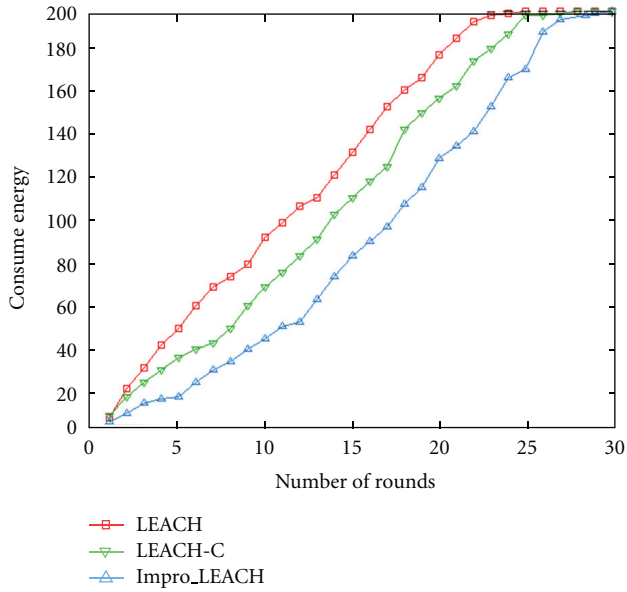


Figure 3: Energy consumption over simulation rounds.

notifications to each member node in the whole network become less as well as the message created by the nodes, which means the remaining energy of the network using our improved protocol exceeds that of the original LEACH and LEACH-C used. The result is validated through ns2 simulation in Figure 4.

5. Conclusion

In this paper, an overview of the original LEACH and LEACH-C protocols is presented and a new version of hierarchical protocol is proposed. The proposed protocol obtains energy efficiency by the modification to choosing of cluster heads formula and the steady-state phase.

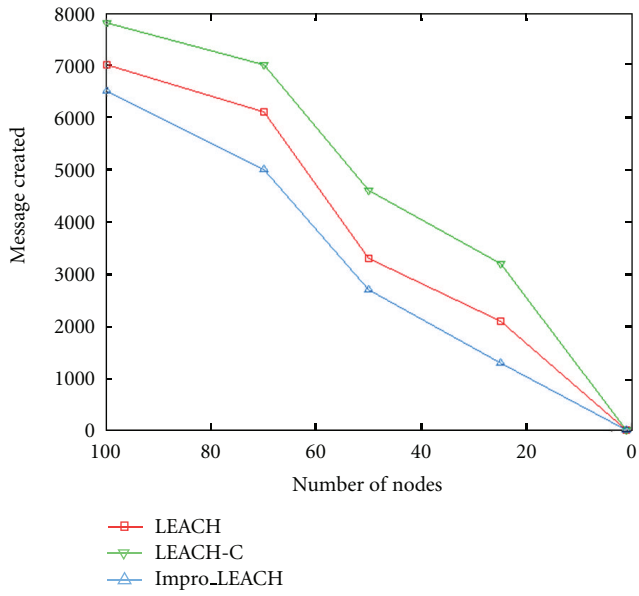


Figure 4: Message created over number of nodes.

The modification to the choosing of cluster heads formula makes the sensor nodes which have more energy and play less role in making the CH or VCH have more opportunity to act as CHs in the coming round. So the total energy of the whole network has more even distribution among different nodes. The introduction of VCH makes the frequency of reclustering more lowly and prolongs the time of being in steady-state phase; thus the energy used for calculating the formula on every nodes reduces. Through the modification and simulation, we can conclude that our proposed protocol performs better than LEACH and LEACH-C protocols.

Acknowledgments

The authors wish to thank the editor and reviewers for their valuable comments, corrections, and suggestions, which led to an improved version of the original paper. This research is a project partially supported by the National Natural Science Foundation of China (Grant no. 61070197).

References

[1] I. Akyildiz and M. C. Vuran, *Wireless Sensor Networks*, University Publishers, Tsinghua, China, 2010.

[2] J. Li and H. Gao, "Research advances in wireless sensor networks," *Journal of Computer Research and Advances*, vol. 45, no. 1, pp. 1–15, 2008.

[3] K. Khamforoosh and H. Khamforoush, "A new routing algorithm for energy reduction in wireless sensor networks," in *Proceedings of the 2nd IEEE International Conference on Computer Science and Information Technology (ICCSIT '09)*, pp. 505–509, August 2009.

[4] C. M. Liu and C. H. Lee, "Power efficient communication protocols for data gathering on mobile sensor networks," in *Proceedings of the 60th IEEE Vehicular Technology Conference (VTC '04)*, pp. 4635–4639, September 2004.

[5] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd Annual Hawaii International Conference on System Siences (HICSS '00)*, January 2000.

[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] J. Xu, G. Yang, Z.-Yu. Chen, L. Chen, and Z. Yang, "Performance analysis of data aggregation algorithms in wireless sensor networks," in *Proceedings of the International Conference on Electrical and Control Engineering (ICECE '11)*, September 2011.

[8] B. Deosarkar, N. Yaday, and R. P. Yadav, "Clusterhead selection in clustering algorithms for wireless sensor networks: a survey," in *Proceedings of the International Conference on Computing, Communication and Networking*, Tamilnadu, India, December 2008.

[9] G. Ran, H. Zhang, and S. Gong, "Improving on LEACH protocol of wireless sensor networks using fuzzy logic," *Journal of Information and Computational Science*, vol. 7, no. 3, pp. 767–775, 2010.

[10] M. B. Yassein, A. Al-zou'bi, Y. Khamayseh, and W. Mardini, "Improvement on LEACH protocol of wireless sensor network," *Journal of Digital Content Technology and Its Applications*, vol. 3, no. 2, pp. 260–264, 2009.

[11] H. Abusaimeh and S. H. Yang, "Dynamic cluster head for lifetime efficiency in WSN," *International Journal of Automation and Computing*, vol. 6, no. 1, pp. 48–54, 2009.

[12] T. Murata and H. Ishibuchi, "Performance evaluation of genetic algorithms for flowshop scheduling problems," in *Proceedings of the 1st IEEE Conference on Evolutionary Computation*, pp. 812–817, June 1994.

[13] "The Network Simulator −ns − 2," http://www.isi.edu/nsnam/ns/.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 981380, 14 pages
doi:10.1155/2012/981380

Research Article

A Vascular-Network-Based Nonuniform Hierarchical Fault-Tolerant Routing Algorithm for Wireless Sensor Networks

Hongbing Li,[1,2] Peng Gao,[1] Qingyu Xiong,[3] Weiren Shi,[1] and Qiang Chen[2]

[1] *School of Automation, Chongqing University, Chongqing 400030, China*
[2] *School of Computer Science and Engineering, Chongqing Three Gorges University, Chongqing 404100, China*
[3] *School of Software Engineering, Chongqing University, Chongqing 400030, China*

Correspondence should be addressed to Hongbing Li, sxxylhb@163.com

Received 10 April 2012; Revised 7 August 2012; Accepted 6 September 2012

Academic Editor: Mugen Peng

Fault tolerance is the key technology in wireless sensor networks which attracts many research interests. Aiming at the issue that the nodes' failures affect the network's stability and service quality, a vascular-network-based fault-tolerant routing algorithm is presented by nonuniform hierarchical clustering. According to the distribution characteristics of the vascular network and inspirations to the fault tolerance for wireless sensor networks, a mathematical model and network topology are, respectively, established. It applies the improved particle swarm optimization (IPSO) to the nonuniform hierarchical clustering, and multipaths are established between the neighbor hierarchical nodes based on the best-worst ant system (BWAS). It introduces the normalized values of the pheromone generated by the ants as the selection probabilities of transmitting paths to establish the hierarchical routing. Theoretical analysis and simulations show that the algorithm has higher packet receiving rates, lower average transmission delay, and balanced energy consumption. It has the good performance in fault tolerance and stability of data transmitting, and it avoids the hot issue in energy consumption and achieves the network load balance.

1. Introduction

Due to the characteristics of dynamic variation of the topology, no center ad hoc network and the constraints of the resources in wireless sensor networks (WSNs), as well as the unpredictability of the working scenario, such as the vibration and electromagnetic interference, it is prone to unbalance of energy consumption, poor quality of data transmission, and routing instability [1, 2]. So they have great influence on the key technologies, including average transmission delay, packet receiving rate, and average energy consumption, which closely correlate with the network's stability, accuracy, and reliability. They also weaken the reserved function and bring greater challenges to the existing network technologies, especially for the application in the complex scenario and necessity for strict performance, just like condition monitoring for key facilities and poisonous gas, emergency, and disaster releasing.

The network with good performance should have the integrity of the fault features, the accuracy of fault diagnosis, and fault recovery efficiency. It can perform well in the robustness of nodes' interconnection, accuracy of data transmission, and against the fast energy consumption as well as the malicious invasion. It also can make a timely diagnosis, find a reasonable control decision, and continue to provide high-credibility computational services by adaptively handling the anomalies of the network. Fault tolerance is the key technology in wireless sensor networks which has attracted great attention and becomes a hot issue that needs further study [3].

Fault tolerance in wireless sensor networks is a complex system problem which includes fault prevention, detection, isolation, diagnosis, and recovery [3]. It mainly includes the optimization of each layer's protocols or algorithms and joints optimal control among multilayer and hardware fault tolerance [4–6]. Fault tolerance in the network layer [7] and

collaborative optimization among the layers are to be the important direction. Fault tolerance in the network layer is to explore the mechanism of multirouting data transmission [7], network coding [8], and so on. The mechanism of multirouting data transmission is to reestablish a path or select one from the multirouting table established in advance to realize the fault tolerance when the source node confirms that the fault exists in the transmitting routing, namely, the redundancy routing fault tolerance [9–11]. Network coding is to transmit the encoded data along multipaths and realize the fault tolerance by decoding and reconstructing the data at sink node, that is, the redundant data fault tolerance [12–14]. The combination of a variety of fault-tolerant methods or collaborative optimization with other layers will achieve the good result [7, 15]. The mechanism of multipath data transmission and network coding improves the stability of data transmission and network load balance. However, it increases the computational complexity, transmission delay, and energy consumption.

Fault tolerance based on the principle of bionics and modern intelligent bionic algorithms has recently been attracting growing research interests in the study of wireless sensor networks. They include biological immune system mechanism, the fuzzy diagnosis, the expert system, the artificial neural network, the particle swarm optimization, ant colony algorithm, and genetic immune algorithm. They provide good ideas and methods to the fault-tolerance and show better fault tolerant effect in wireless sensor networks. For example, the ant colony algorithm with the advantage of swarm intelligence is introduced into the routing establishment in the network layer; the particle swarm algorithm is introduced into the topology establishment with its characteristics of fast clustering convergence to improve the efficiency of clustering; artificial neural network or its combination with the fuzzy algorithm is applied to forecast the perception data to realize fault tolerance in the application layer; genetic algorithm can be applied to the data unit coding to improve fault tolerance of the data transmission in the link layer or network layer.

It has already shown the good performance and advantages when introducing the mechanism of biological immune system into the fault tolerance in wireless sensor networks [16–23]. Some basic research works have been done in this aspect. Bokareva et al. have presented a fault-tolerant architecture SASHA based on the biological immune system, in which the lymph mechanism is used to produce the detector for fault detection, and the thymus mechanism is used to complete the fault diagnosis. This structure cannot only identify the known fault model, but also provide a good adaptive ability of learning and evolving to the unknown fault model [16]. Jabbari and Lang apply the mechanisms of biological immune system to the security detection and fault tolerance by simulating the mechanisms of self-learning, self-organization, and memory and information processing in the biological immune system or the nerve immune system [17]. Atakan and Akan propose the new distributed node and rate selection method (DNRS), which is based on the principles of natural immune system to select the minimum number of sensor nodes based on the B-cell stimulation in immune system to transmit the data to the sink and select the appropriate reporting frequency of sensor nodes to achieve the minimum energy consumption [18]. Rui et al. propose an immune system inspired approach to locally discover and recover from losses of query messages at sensor nodes, similar to antibodies in an immune system [19]. Chen et al. propose a neighborhood node selection algorithm to decide if a sensor node is activated or not according to the B-cell and T-cell models of immune systems by comparing the similarity between artificial immune system and wireless sensor networks [20]. Hu et al. present a hybrid routing scheme and an immune cooperative PSO algorithm for fault-tolerant routing problem [21]. In [22], uniform immunization and temporary immunization are conducted on small worlds of tree-based wireless sensor networks to combat the sensor viruses. Salmon et al. propose an intrusion detection system (IDS) framework inspired by the human immune system to monitor their neighborhood and to identify an intruder [23]. Mechanism of biological immune system or immune algorithms can also be applied in routing optimization, intrusion detection, nodes' deployment, target coverage, and performance optimization.

This paper will explore establishing the nonuniform hierarchical clustering based fault-tolerant routing algorithm by the bionic intelligent algorithms according to the characteristics of vascular network and inspiration to fault tolerance based on the preliminary studies. It will carry out the fault tolerance studies by combining the routing establishment with the topology design to improve the network's performance of stability and reliability. The main work and contribution of this paper is: (1) to study the structure and characteristics of the vascular network and the inspirations to the fault tolerance and to establish the mathematical model and network topology; (2) to study the IPSO and BWAS, making these intelligent algorithms more optimal and more suitable for resource-restrained network; (3) to study the nonuniform hierarchical static clustering by applying the IPSO; (4) to study the routing establishment by applying the BWAS.

It innovatively establishes the hierarchical fault-tolerant topology based on the vascular network and uses the normalized values of the ants' pheromone as the selection probability of the transmission paths to establish the fault-tolerant routing by improved intelligent algorithms. Deference with previous studies in the methods of fault tolerance in the network layer is not to simultaneously establish multipaths or backup paths for each relay node, but to temporarily select the path with maximum normalized values of the pheromone between the two neighbor hierarchical nodes to establish the optimal transmission route. If this node fails, then it selects the other node as the substitute, in which the path is with the second maximum normalized values of the pheromone as the backup transmission path to realize the fault tolerance. The new reestablished transmission path is not from the source node to sink node but from the node ahead of the fault node to the sink node by the selection probability. Furthermore, the intelligent algorithms are improved to fit for the source-restrained network.

The rest of this paper is organized as follows. Section 2 studies the mathematical model of the vascular network and topology. Section 3 presents the nonuniform hierarchical clustering algorithm based on the improved particle swarm optimization (IPSO) and the fault-tolerant routing algorithm by the best-worst ant system (BWAS). Section 4 carries out the simulations and analysis to evaluate the performance. Finally, Section 5 concludes the paper and discusses the future work.

2. Mathematical Model and Network Topology

2.1. Mathematical Model. The vascular network is the fractal tree-like branching network. It has the complex structure with the characteristics of the fractal [24]. Great differences exist among the different sections of the vascular in the structures, physical characteristics, and functions. Different sections have different blood flowing rate. It has the properties of stability, flexibility, connectivity, and multiconnectivity. Blood flows in the directed way according to the pressure difference that exists between any two points. Blood flow has quantitative relationship with current velocity, cross-section, and the pressure.

The characteristics of the vascular network give an important inspiration to the establishment of fault-tolerant routing in WSN. The nodes are hierarchically marked to have the hierarchy differences according to distance to sink node to ensure the directed data transmission. Different clustering probabilities are adopted in different hierarchical areas. Finally, it forms the distribution with different density and scale of the clusters in different hierarchical areas, which has a similar topology with the vascular network model. The vascular-network-based topology has multipaths connectivity and uses the pheromone normalized values generated by ants in ant colony algorithm as the path selection probability to establish the optimal transmission routing to realize the fault tolerance.

The vascular-network-based mathematical model has the following characteristics. (1) Pressure differential network. Each node in WSN has the differential pressure value P', which reflects the distance to center node of the network. The ones closer to the center node have higher pressure values. So the pressure difference $\Delta P'$ exists between any two nodes, which reflects the directed data flow. (2) Connectivity of the weighted graph; it is the directed weighted graph and each side is assigned with the different weights $w(e)$. Any two points in the network are connected. (3) Clustering density. Different quantity and density of the clusters are presented in different hierarchies. Larger quantities and higher distribution density of the clusters are presented in the areas with smaller hierarchy values. The mathematical formulations are as follows:

$$P'_i > P'_j, \quad \text{if } \operatorname{dist}(i,s') < \operatorname{dist}(j,s'),\ i,j \in (1,2,\ldots,n),$$

$$\Delta P'_{i,j} = p'_i - p'_j, \qquad \vec{p'_i} = \sum \vec{p'_i},$$

l is the neighbor node of i,

$$p'_i = \frac{r}{\operatorname{dist}(i,s')},$$

$$a_{ij} = \begin{cases} w_{ij}, & (v_i, v_j) \in E, \\ 0, & i = j, \\ \infty, & (v_i, v_j) \notin E, \end{cases} \tag{1}$$

where P'_i is the pressure of node i, $\Delta P'_{i,j}$ is the pressure difference between the nodes i and j, $dist(i,s')$ represents the Euclidean distance from the node i to center node s', n is the quantity of the clustering nodes of network, r is a parameter value, $\vec{p'_i}$ represents the directed pressure, a_{ij} is the adjacent matrix element of directed weighted graph, and E is the set of side (i,j).

2.2. Network Topology

Definition 1. Hierarchy: nodes in WSN are marked with different hierarchies according to the distance from the average node to sink node. The nodes belonging to the same hierarchical area are marked with the equal hierarchy value, that is,

$$G_i = m \quad \text{if } \operatorname{dist}(s,i) \in ((m-1)r, mr), \tag{2}$$

where G_i represents the hierarchy value of the node i, $dist(s,i)$ represents the Euclidean distance from the node i to sink node s, and r is the initial distance value.

2.2.1. Nonuniform Hierarchical Clustering Topology. The topology is established based on the theory and characteristics of vascular network. (1) The nodes in the network are divided into two categories: the average nodes and clustering nodes. Average nodes are responsible for data collection and transmission. Besides this, clustering nodes have the functions of data receiving, forwarding, and routing addressing. The network routing is established among the clustering nodes. (2) The network is divided into different hierarchies by the distance to the sink node. The nodes in the same hierarchical area have the same hierarchy value, and the routing is established among the clustering nodes with different hierarchies. (3) It is the directed weighted graph with connectivity between any two nodes in the network. Each link has different weight; the data propagation and message broadcasting have directions according to the nodes' pressure gradient. (4) Different quantity and scale of the clusters are presented in different hierarchical areas. Larger quantity and smaller scale of the clusters are presented in the areas with smaller hierarchy values and so does the opposition in the areas with higher hierarchy values. (5) Multipaths will be established between the nodes that are the neighbor hierarchical within their transmission power coverage. The normalized values of the pheromone are used as the probabilities to select the transmitting path. The link with maximum probability is to be selected as the transmitting path. The established physical topology of network is shown in Figure 1.

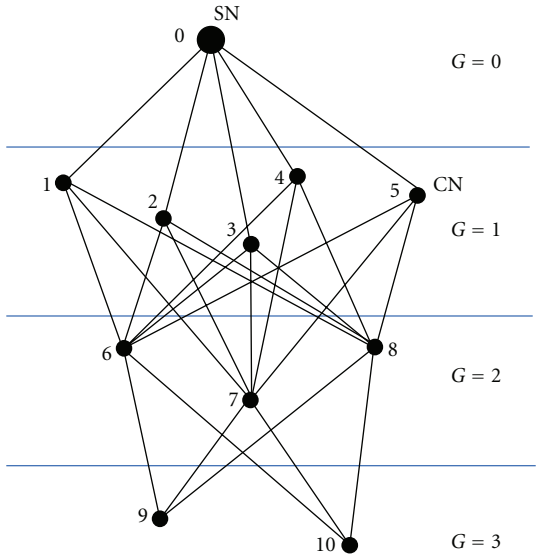


Figure 1: Physical topology of the network.

Figure 1 shows the network's topology based on the vascular network. It is divided into 3 hierarchies. Node 0 is the sink node, and nodes 1–10 are the clustering nodes. The nodes 1–5 in hierarchy 1 are responsible for the data fusion, receiving the data from the nodes in hierarchy 2 and retransmitting the data to the sink node. Nodes 6–8 in hierarchy 2 are responsible for data fusion in their respective clusters and retransmitting the data from nodes in hierarchy 3 to nodes in hierarchy 1. Nodes 9 and 10 in hierarchy 3 are only responsible for the data fusion and transmitting the data to the nodes in hierarchy 2. For example, node 7 in hierarchy 2 will establish the transmission path with nodes 1–5 in hierarchy 1 and nodes 9-10 in hierarchy 3 within its transmission power coverage excluding the nodes 6 and 8 which belong to the same hierarchy. And the value w_{ij} of each side is not equal.

The fault tolerance of the topology is owing to the normalized value of the ants' pheromone used as the selection probability of the transmission path, and it selects the path with the maximum probability as the transmission path. For example, the pheromone on the links connecting the node 7 in hierarchy 2 with the nodes 1–5 in hierarchy 1 is different, that is, $\tau_{7,3} \geq \tau_{7,2} \geq \tau_{7,4} \geq \tau_{7,1} \geq \tau_{7,5}$, where τ represents the normalized pheromone value. Therefore, the real transmitting path is $\text{Path}_{7,3}$. When this path fails, it will select the path with the second maximum probability as a substitute. If $\text{Path}_{7,3}$ fails, $\text{Path}_{7,2}$ is to be selected.

2.2.2. Establishing Rules. The topology will be established based on clustering model. The nodes are functionally divided into average nodes (ANs) and clustering nodes (CNs). The routing is established among the hierarchical clustering nodes. The clustering node is to establish the transmission paths only with neighbor hierarchical nodes, not the same hierarchical ones. It establishes the transmission path with all the neighbor hierarchical nodes within the transmitting power coverage. Only one path is to be selected as the actual transmission one, and the rest are the backup transmission paths. The establishing rule is shown in Algorithm 1.

2.2.3. To Determine the Clustering Probability. Different clustering probabilities are adopted in different hierarchical areas to form the different scale and quantity of the clusters, which means nonuniform clustering. Clustering node is dynamically selected according to the residual energy to balance the energy consumption in one cluster [25, 26].

The network is separated into different annular sections with the sink node as the center and $n \cdot R$ $(n \in 1, 2, \ldots)$ as the radius. The nodes in the same annular section have the same hierarchy value. It applies the larger clustering probability in hierarchy 1 where it is close to the sink node and smaller clustering probabilities in the hierarchical areas that are far from it. It eventually forms the distribution with large quantity and smaller scale of the clusters that are closer to the sink node, as shown in Figure 2.

This topology based on the vascular network model can balance the energy consumption and avoid the hot issue. However, how to determine the clustering probability in each hierarchical area becomes the key issue. The energy consumption model refers to [27], the radius of each area is $n \cdot R$, and the total number of the nodes is $N_{n \cdot R}$ that the network is divided into n areas. The clustering probability is p_n $(n \in 1, 2, \ldots)$ for the area where the radius is R. The number of the clusters is $N_{n \cdot R} \cdot p_n$ in each area. The number of the nodes in one cluster is $1/p_n - 1$ excluding the clustering node. Suppose each node only sends data k bit, so the energy consumption of the clustering node is as follows:

$$\begin{aligned} E_{R \cdot \text{clu}} &= E_{\text{fuse}}\left(\frac{1}{p_1} - 1\right)k + E_{\text{elec}}(n-1)k \\ &\quad + E_{\text{elec}}(k + (n-1)k) \\ &\quad + \varepsilon_{\text{amp}}(k + (n-1)k)d^2, \\ &\qquad\qquad i \in (1 \cdots (N_{n \cdot R} p_i)) \\ &= E_{\text{fuse}}\left(\frac{1}{p_1} - 1\right)k + E_{\text{elec}}(2n-1)k \\ &\quad + \varepsilon_{\text{amp}} nkd^2. \end{aligned} \tag{3}$$

The energy consumption of the nodes excluding the clustering nodes is as follows:

$$E_{R \cdot \text{ave}} = E_{\text{elec}}\left(\frac{1}{p_1} - 1\right)k. \tag{4}$$

The whole energy consumption in one cluster is as follows:

$$E_{R \cdot \text{total}} = \sum_{1}^{N_R p_1} (E_{R \cdot \text{clu}} + E_{R \cdot \text{ave}}). \tag{5}$$

In order to minimize the energy consumption, from $\partial E_{R \cdot \text{total}} / \partial p_1 = 0$, we can get p_1. According to $E_{R \cdot \text{total}} = E_{2R \cdot \text{total}}$ for the balance of energy consumption, we can get p_2, and so on.

Suppose $G_k = G_{k+1} = \cdots = G_{k+l} = m$, $i \neq i', i, i' \in (k, k+1, \ldots k+l)$;
$G_i = G_{i'} = m$, $i \not\rightleftharpoons i'$;
If $G_j = \{m-1, m+1\}$, $j \in (k', k'+1, \ldots k'+l')$;
  If Dist$(i, j) >$ Dist*RF;*
    $i \not\rightleftharpoons j$;
  else
    $i \rightleftharpoons j$;
  end
end

Where Dist(i, j) represents the Euclidean distance of node i and j, *DistRF* represents the nodes' RF distance. G_k represents the hierarchy of node k. $\not\rightleftharpoons$ shows that it cannot establish the transmission path.

Algorithm 1: Establishing rules of non-uniform hierarchical topology.

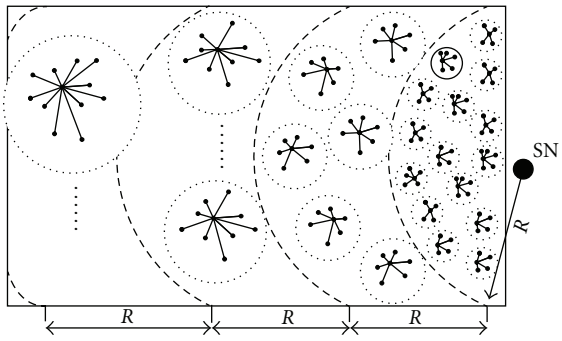


Figure 2: Nonuniform hierarchical static clustering.

3. Nonuniform Hierarchical Clustering Routing

3.1. Nonuniform Hierarchical Clustering Strategy. Particles, with equal quantity of the nodes, are randomly deployed in the active hierarchical area. It does the initial $N_{n \cdot R} \cdot p_n$ clustering by clustering probability and then introduces the IPSO to intelligent clustering. It modifies the particles' flying rules when the particles move to the location coinciding with the sensor nodes, the location of the particle becomes fixed and the velocity turns to zero. The particles no longer follow the rule that the particles' velocities and locations are updated by their own and swarms' optimum. The rest particles continue to fly by established rules until the whole particles' locations coinciding with sensor nodes. It forms the optimal clustering distribution and then turns to the next hierarchical area.

3.1.1. PSO Prototype. PSO is an optimization algorithm based on swarm intelligence which imitates birds' behavior. Swarm intelligent optimization searching is carried out by the cooperation and competition of the particles. Each solution of the generation has the characteristics of learning the optimum of itself and the group. Each particle is determined by its position and velocity vector. Each particle confirms them by the established direction as well as experiences from the optimal direction of itself and the swarm's. The velocity and location is updated by the following formulae:

$$v_i^{t+1} = w v_i^t + c_1 r_1 \left(p\text{best}_i^t - x_i^t \right) + c_2 r_2 \left(g\text{best}_j^t - x_i^t \right), \quad (6)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1}, \quad (7)$$

where the inertia weight coefficient $w = w_{\max} - (w_{\max} - w_{\min}) \cdot k/m$ is the linear descending function by time. It can make PSO explore the larger area at the beginning and roughly locate the position faster. With w descending gradually, the particles' speed gets slower and begins to locally search subtly. $w_{\max}$, $w_{\min}$ are, respectively, the initial and eventual weight, m is the maximum iterations, k is the current iterations, v_i^t is the particle's velocity vectors, and x_i^t is the particle's current position, r_1, r_2 are random numbers between 0 and 1 to keep the diversity of the group, c_1, c_2 are the learning factors which make it keep the ability of self-summary and learning from the optimal particles, and $p\text{best}_i^t$ and $g\text{best}_j^t$ are, respectively, the best positions of the particle itself and the swarm.

3.1.2. IPSO

(a) Modify the PSO formula with inertia weight. The particles $G = \{x_1, x_2, \ldots, x_m\}$ are divided into k clusters of particle swarms C_i $(i = 1, 2 \ldots, k)$. Modify the formula (6) into

$$V_i^{t+1} = w V_i^t + c_0 r_0 \left(p\text{best}_i^t - X_i^t \right) + \sum_{j=1}^{k} c_j r_j \left(g\text{best}_j^t - X_i^t \right). \quad (8)$$

It updates the particles' position and velocity by the formulae (7) and (8) in each iteration.

(b) Modify particles' flying rules. The particles no longer follow the rule that the particles' velocities and locations are updated by their own and swarms' optimum. In the improved way, when the particle moves

to the location coinciding with the sensor nodes, the particle's location becomes fixed and the velocity turns to zero. The rest particles continue to fly by established rules until the whole particles' locations coincide with sensor nodes. It forms the optimal clustering distribution, and then the clustering is to be carried out in the next hierarchical area.

3.1.3. Steps of the Algorithm

Step 1. Divide the network into different hierarchies by the distance from the nodes to sink node. It uses different clustering probabilities for different hierarchical areas. It selects the nodes in hierarchy 1 as the active nodes and makes others as dormant ones in order to avoid the particles flying into other hierarchical areas.

Step 2. Randomly generate the number of the particles $M = N_{n\cdot R}$ as same as nodes and do the initialization. Define the initial position x_i and flying velocity v_i of each particle, the learning factor c_0, c_j $(j = 1, 2, \ldots, k)$, the inertia weight $w_{\max}$, $w_{\min}$, the maximum speed $v_{\max}$, the maximum position $x_{\max}$, and the maximum m.

Step 3. Determine the number of the clusters $k = N_{n\cdot R} \cdot p_n$ and cluster center c_i in C_i. Each particle swarm has $D = M/k$ dimensions, and different individual particles have different positions.

Step 4. Set the target function to evaluate the fitness of the particles and the clustering quality by Euclidean distance:

$$S = \frac{1}{\sum_{i=1}^{K} \sum_{x \in C_j} dist(c_i, x)^2}, \tag{9}$$

where the *dist* is the normative Euclidean distance between two objects, K is the number of the clusters, x is the particle, and c_i is the center of cluster C_i.

Step 5. Calculate the particle's self-optimal position $p\text{best}_i^t$ in C_i and the optimal position $g\text{best}_j^t$ of C_i.

Step 6. Judge the particles' positions. When the particle moves to the sensor nodes' position, set $v_i = 0$, $x_i = x_{\text{sensor}}$, where x_{sensor} represents the sensor node's position.

Step 7. Update the rest particles' positions and velocities by formulae (7) and (8).

Step 8. Repeat Steps 3 to 7 until the last particle coincides with the last sensor node. The clustering is over in this hierarchy, and the nodes turn to the dormant state.

Step 9. Wake up the nodes in the hierarchical area $(n-1)R < \text{dist}(x_{\text{senor}}, x_{\sin k}) < nR$, $n \in (2, 3, \ldots, n)$ and do the clustering by Steps 2 to 8.

3.2. Routing Establishment Based on BWAS. BWAS is the improvement of the ant colony algorithm. It introduces the rewards-punishment mechanism to magnify the pheromone between the best and worst ants. It improves the ability of the optimal path searching and the convergence rate. And pheromone generated by ants is an important parameter reflecting the path's optimization. It considers the nodes' energy, inspiration of the nodes' distance, and the advantages of swarm intelligence.

Now suppose some artificial ants are set at each clustering node and they will die when finishing searching the paths from the clustering nodes to the sink node. The ants have the memory of the pheromone on the paths passed by.

Step 1. Initialize the parameters; to determine the clustering nodes, set ants' quantity at each clustering node and define their property.

Step 2. Select the path for each ant by formulae (10) and (11):

$$p_{ij}^k(t) = \begin{cases} \dfrac{\tau_{ij}^{\alpha}(t)\eta_{ij}^{\beta}(t)}{\sum_{s \in \text{allowed}_k} \tau_k^{\alpha}(t)\eta_{is}^{\beta}(t)}, & j \in \text{allowed}_k \\ 0, & \text{otherwise}, \end{cases} \tag{10}$$

$$\tau_{ij}(t+n) = \rho_1 \tau_{ij}(t) + \Delta\tau_{ij}(t, t+n), \tag{11}$$

$$\Delta\tau_{ij}(t, t+n) = \sum_{k=1}^{m} \Delta\tau_{ij}^k(t, t+n), \tag{12}$$

$$\Delta\tau_{ij}^k(t, t+n) = \begin{cases} \dfrac{Q}{L_k}, & \text{if ant } k \text{ passes } (i, j) \text{ in this ciculate}, \\ 0, & \text{otherwise}, \end{cases} \tag{13}$$

where p_{ij}^k is the transition probability of ant k, j is the unvisited node, τ_{ij} is the pheromone intension on side (i, j), η_{ij} is the visibility of the side (i, j) which reflects the inspiration transferring from node i to j, and allowed_k is the node set that the ant is allowed to visit. Formula (11) is to update the pheromone when the ant has finished establishing the whole path, and Q is a parameter.

Step 3. Update the pheromone on the generated paths by formula (14) after each ant finishes the task:

$$\tau_{rs} \longleftarrow (1-\rho)\tau_{rs} + \rho\Delta\tau_{rs}, \tag{14}$$

$$\Delta\tau_{rs} = (nL_{nn})^{-1}, \tag{15}$$

where τ_{rs} represents the pheromone value on the side (r, s), ρ is a parameter, $0 < \rho < 1$, n is the quantity of the nodes, and L_{nn} is the length of the path.

Step 4. Execute Step 2 and Step 3 until each ant generates a path and evaluate the best and worst ants according to the length of the paths passed by.

Step 5. Globally update the pheromone on the paths passed by the best ants by formula (16):

$$\tau_{rs} \longleftarrow (1-\alpha)\tau_{rs} + \alpha\Delta\tau_{rs}, \tag{16}$$

$$\Delta\tau_{rs} = \begin{cases} \left(L_{gb}\right)^{-1}, & \text{if } (r,s) \in \text{global best}, \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

Step 6. Globally update the pheromone on the paths passed by the worst ants by the following formula:

$$\tau(r,s) = (1-\rho)\tau(r,s) - \varepsilon\frac{L_{\text{worst}}}{L_{\text{best}}}, \tag{18}$$

where ε is a parameter, L_{best} and L_{worst} are the lengths of the paths, respectively, passed by the best and worst ant, L_{gb} is the length of global optimal path, and α is the volatilizing parameter of the pheromone, where $0 < \alpha < 1$.

Step 7. Execute Steps 2 to 7 for ants in the rest clustering nodes until all the ants finishes task and record the pheromone values on the paths.

After the hierarchical dividing and static clustering, it uses the normalized pheromone as the selection probability to establish the fault-tolerant routing. When the energy consumption of the clustering nodes reaches the threshold value, they report the information along the established route in order to reduce the energy consumption. When the transmission cycle is over, it selects the nodes with higher energy as the clustering nodes in the clusters. Then, the whole network calculates the new pheromone values of the paths based on BWAS to establish a new routing.

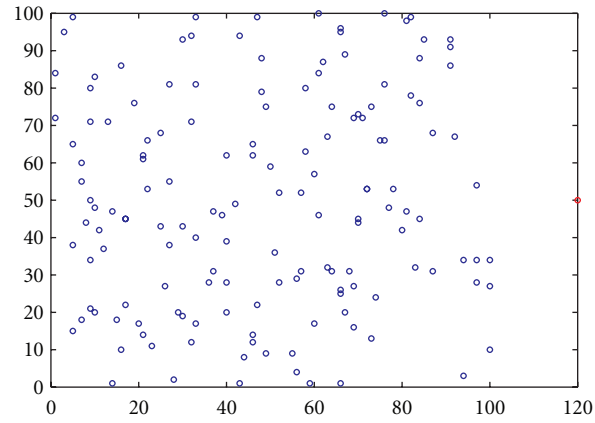


Figure 3: Nonhierarchical nodes in WSN.

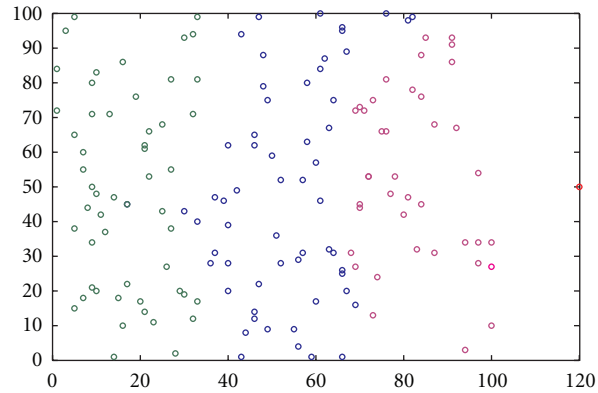


Figure 4: Hierarchical nodes in WSN.

4. Simulation and Analysis

4.1. Simulation on Nonuniform Hierarchical Clustering

4.1.1. The Distribution of Hierarchical Nodes. Assumptions: (1) the nodes in the network are stationary. (2) The locations of the nodes are known. (3) All nodes have the same status and parameters. (4) Each node works at full-duplex operation mode. (5) The nodes distribute uniformly in the rectangular areas, and the sink node is outside.

The simulation is based on PC with processor i3-2100, RAM with 4 G. 150 nodes are deployed in the area $[100, 100]$, the number of hierarchies is $n = 3$, and the initial coordinate of sink node is $(120, 50)$. The radius $R = 40$. The inertia weight coefficient w in IPSO is linearly down from 0.9 to 0.4. The learning factors $c_0 = 1.5$, $c_j = 2$; the quantity and scale of the particle swarm are determined by the clustering probability. The maximum iteration m is set by the experiments to ensure that the last particle's location coincides with the sensor node. r_0, r_j are the random values in $[0, 1]$, and $K = N_{n\cdot R} \cdot p_n$.

Figure 3 shows that 150 sensor nodes are deployed randomly in $[100, 100]$, which are labeled with blue color. These nodes are not hierarchically divided, and all have the equal hierarchical property. The sink node is deployed at $(120, 50)$ and is labeled with red color. The nodes in Figure 4 are hierarchically divided based on Figure 3. With the sink node as the center of the circle and the distance to sink node as the radius, 3 hierarchical areas are divided, respectively, marked by the colors red, blue, and green. The red nodes belong to hierarchy 1; the blue and green nodes, respectively, belong to hierarchies 2 and 3.

4.1.2. Nonuniform Hierarchical Clustering. The nodes are hierarchically clustered by the IPSO; different clustering probabilities are adopted in different hierarchical areas. Finally, it forms the cluster distribution with different scales and quantities, as shown in Figures 5 and 6.

Figure 5 shows the equal probability hierarchical clustering by IPSO with 3 hierarchies. 150 nodes are clustered by the same probability. The scales of clusters are the same, and there are about 2–4 nodes in each cluster. Figure 6 shows the nonequal probability hierarchical clustering by IPSO with 3 hierarchies. So it forms the distribution that different scales and quantity of clusters are in different hierarchical areas. It, at last, forms the distribution with smaller scales and larger amount of the clusters in the lower hierarchical level where it is near to sink node. And so does the opposition in hierarchy 3 where it has the larger scales of clusters and fewer nodes in each one. It reaches the expected clustering structure.

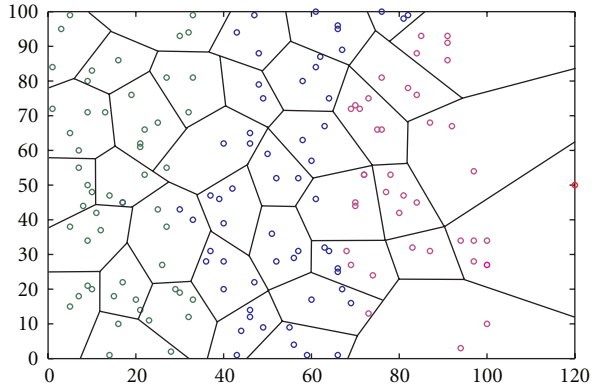


Figure 5: Equal-probability hierarchical clustering.

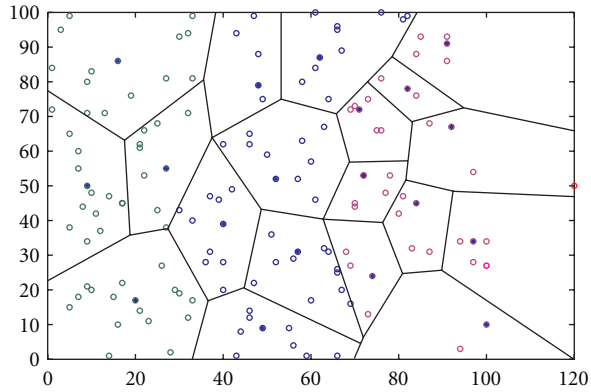


Figure 6: Nonequal-probability hierarchical clustering.

Table 1 shows the statistics of the nonuniform hierarchical clustering by IPSO. From it, the number of the nodes in hierarchy 1 is 38, the clustering probability is 0.237, and the number of the clustering is 9. According to the energy balance method presented above, we can calculate the probabilities of the second and third hierarchies; they are, respectively, 0.100 and 0.070. The simulation results show the corresponding values are 0.103 and 0.074. From the clustering probabilities, the number of the clusters, and member nodes, we can get the conclusion that the distribution of the clusters is the same with the topology previously supposed.

4.2. Simulation on Nonuniform Hierarchical Clustering Routing

4.2.1. Parameters Setting. After the hierarchical division and nonequal probability static clustering, to place the ants at each clustering node with the number $k = 30$, the initial pheromone value $\tau = 100$. Pheromone inspiration factor $\alpha = 1$; it indicates the relative importance of the pheromone when the ants select the forward way. Expect heuristic factor $\beta = 5$, indicating the relative importance of visibility and affecting the algorithm's convergence rate. The greater the value is, the closer it is to be greedy algorithm according to the state transition probability. Set the pheromone

Table 1: Statistics of the nonuniform hierarchical clustering.

Hierarchy	Hierarchy 1	Hierarchy 2	Hierarchy 3
Number of nodes	38	58	54
Number of clusters	9	6	4
Number of members	29	52	50
Probability of clustering	0.237	0.103	0.074

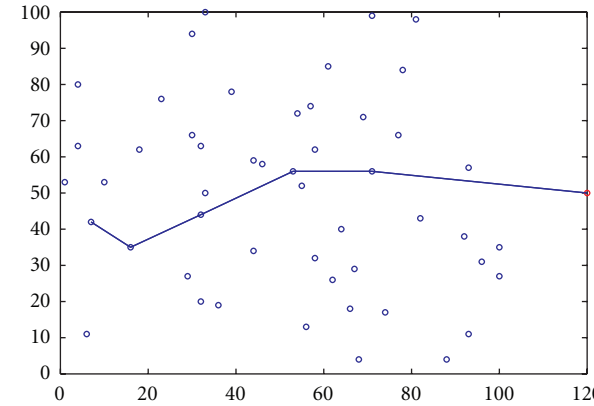


Figure 7: Transmission route from node (7, 42) to sink node.

evaporation coefficient $\rho = 0.1$, pheromone residual factor $1 - \rho$, and the maximum iterations $NC\_\max = 200$.

4.2.2. Process of Routing Establishment. Now we use A to E to identify hierarchies 1 to 5. (x_I, y_I) represents the coordinate of the nodes in hierarchy I, ρ' represents ant's pheromone normalized value. Table 2 shows the normalized values of the ants' pheromone by BWAS on the paths from the node $(7, 42)$ in E to sink node. The node $(7, 42)$ is randomly selected to establish a transmission path from it to sink node. $\rho'_{E \to D}$ represents the normalized values on the paths from the node $(7, 42)$ in E to nodes in D. Table 3 shows the nodes' coordinates corresponding with the normalized values in Table 2. $\rho' = 0.3034$ represents the normalized values of the pheromone on the path from the node $(7, 42)$ to $(x_D, y_D) = (16, 35)$.

From all the normalized values of the ants' pheromone on the paths from the node $(7, 42)$ in E to the nodes in D, $\max \rho' = 0.3034$. So the node $(x_D, y_D) = (16, 35)$ is to be selected as the next hop to establish the transmission route. If the node $(16, 35)$ fails, it will select the path with $\rho' = 0.1515$; therefore, $(x_D, y_D) = (18, 62)$ is to be selected as the next hop. So a routing for data transmission has been set up from node $(7, 42)$ in E to sink node. The hierarchy order of the established path is $E \to D \to C \to B \to A \to$ sink node, and the corresponding order of nodes' coordinates is $(7, 42) \to (16, 35) \to (32, 44) \to (53, 56) \to (71, 56) \to$ sink node, as shown in Table 3. Figure 7 shows that the actual transmission route is established among 50 clustering nodes based on Table 2.

4.2.3. Simulation on the Routing Establishment. Simulation is completed when all the ants reach the sink node by

Table 2: Normalized values of the ants' pheromone.

$\rho'_{E\to D}$	$\rho'_{D\to C}$	$\rho'_{C\to B}$	$\rho'_{B\to A}$
0.3034	0.1814	0.1285	0.1162
0.1515	0.1033	0.0873	0.0742
0.0920	0.1469	0.1277	0.1029
0.1299	0.1300	0.0796	0.1254
0.1041	0.0683	0.1086	0.0510
0.0608	0.1188	0.0983	0.0868
0.1039	0.1039	0.0619	0.0601
0.0544	0.0881	0.0889	0.0710
—	0.0729	0.0726	0.0360
—	—	0.0578	0.0525
—	—	0.0461	0.0375
—	—	0.0426	0.0564
—	—	—	0.0454
—	—	—	0.0409
—	—	—	0.0438

Table 3: Node's coordinates on the paths.

(x_E, y_E)	(x_D, y_D)	(x_C, y_C)	(x_B, y_B)	(x_A, y_A)
(1, 53)	**(16, 35)**	(32, 63)	**(53, 56)**	(64, 40)
(4, 63)	(18, 62)	**(32, 44)**	(54, 72)	(67, 29)
(4, 80)	(23, 76)	(33, 50)	(55, 52)	(69, 71)
(6, 11)	(29, 27)	(36, 19)	(57, 74)	**(71, 56)**
(7, 42)	(30, 66)	(39, 78)	(58, 32)	(74, 17)
(10, 53)	(30, 94)	(44, 34)	(58, 62)	(77, 66)
—	(32, 20)	(44, 59)	(61, 85)	(78, 84)
—	(33, 100)	(46, 58)	(62, 26)	(82, 43)
—	—	(56, 13)	(66, 18)	(88, 4)
—	—	—	(68, 4)	(92, 38)
—	—	—	(71, 99)	(93, 11)
—	—	—	(81, 98)	(93, 57)
—	—	—	—	(96, 31)
—	—	—	—	(100, 27)
—	—	—	—	(100, 35)

BWAS. The pheromone on each path stops evaporating and is recorded. Node visited by ants is not hierarchical, only by this can pheromone reflect the links' quality.

Figure 8 shows the distribution of the clustering nodes in three hierarchies. According to the nonequal probability clustering method, 19 clustering nodes are elected from total 150 nodes and marked with three different colors as for three hierarchies. Simulations in Figure 9 illustrate that all the possible transmission paths are established among the clustering nodes. Paths in different hierarchy are marked with different colors for distinction. The paths established in Figure 9 show the hierarchical characteristic clearly.

For fewer clustering nodes in WSN, we can more clearly understand the establishment of the transmission paths. However, better performance can be reflected for the large-scale network, as shown in Figures 10 and 11.

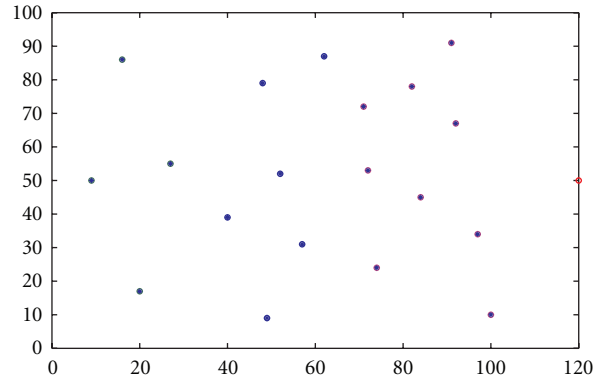


Figure 8: Clustering nodes in three hierarchies.

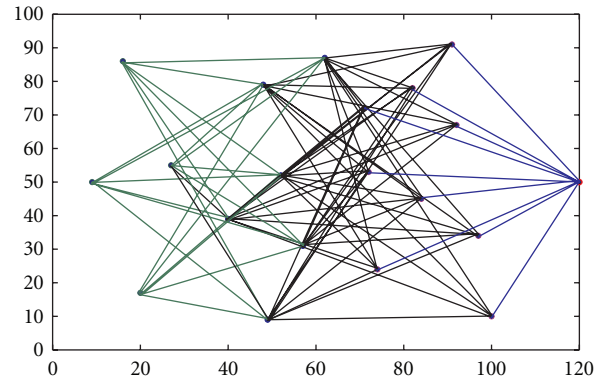


Figure 9: Paths among the clustering nodes in three hierarchies.

Figure 10 shows the distribution of 50 clustering nodes in five hierarchies. Figure 11 shows all the possible transmission paths among the clustering nodes. We also mark the hierarchical transmission paths with five different colors for distinction. There are five hierarchies for the established paths; the node only establishes the links to the nodes in the neighbor hierarchies. It has an important significance to the quick establishment and optimization of the transmission routing for large-scale network.

4.3. Performance and Complexity Analysis of the Fault Tolerance

4.3.1. Performance Analysis of the Fault Tolerance. (1) Pheromone, generated by ants during the process of searching the optimal path, is an important parameter reflecting the path's quality. The higher τ is, the greater probability the path is selected as transmitting path. (2) It uses the normalized pheromone values as the selectionprobabilities of the paths and selects one with maximum probability as the actual data transmitting path. Any clustering node will establish the possible paths to the nodes in neighbor hierarchies within its transmission power coverage. But only the path with maximum probability is to be selected as actual data transmitting path. If it fails, the path with second maximum probability is to be selected to realize the fault tolerance.

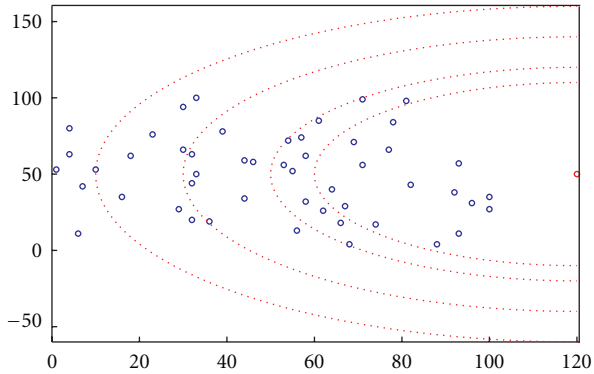


Figure 10: Clustering nodes in five hierarchies.

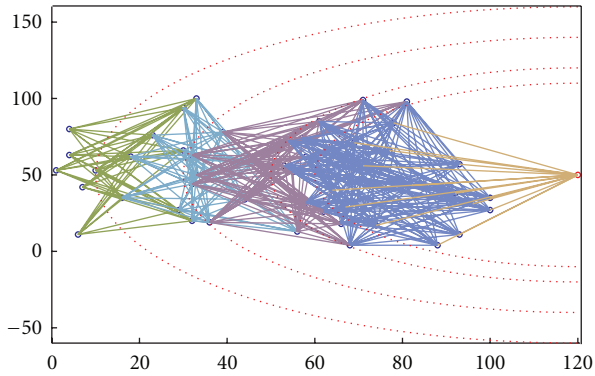


Figure 11: Paths among the clustering nodes in five hierarchies.

Node 7 in Figure 1, for example, has established the paths path_{7i} $(i = 1, 2, \ldots, 5)$ with the nodes 1–5 in the neighbor hierarchy. The normalized values of pheromone are not equal, like $\tau_{7i} \neq \tau_{7j}$ $(i \neq j, ij = 1, 2, \ldots, 5)$. And the respective selection probabilities are not equal, like $p_{7i} \neq p_{7j}$. So the actual transmission path is $\text{Path}_{7i} \rightarrow \max(p_{7i})$, $i = 1, 2, \ldots, 5$. If this path fails, it will select $\text{Path}_{7j} \rightarrow \max(p_{7j}, j \neq i, i, j = 1, 2, \ldots, 5)$ as the actual transmission path and so on. The direction of the gathered data transmitting to sink node is $G_3 \rightarrow G_2 \rightarrow G_1$ and opposite direction for the message broadcasting. (3) The method of multipaths establishment is different from others. When the fault occurs, the newly established transmission route reestablishes the route not from the source node to sink node but from the node ahead of fault node to the sink node by the selection probability. It saves the computational complexity and overhead. (4) The fault tolerance of the network is greatly related with the transmission range of the nodes. The larger the transmission range is, the better is the performance of the network; however, the more the energy consumption. The node's transmission range based on the nonuniform hierarchical clustering routing should be set in $[R, 2R]$, where R is the radius of the ring in Section 2.2.3. The node with transmission range in $[R, 2R]$ can cover some nodes in the neighbor hierarchical area and select the best node in the transmission rang coverage to establish the links. It considers the fault tolerance of the network as well as the energy consumption by reasonable transmission range.

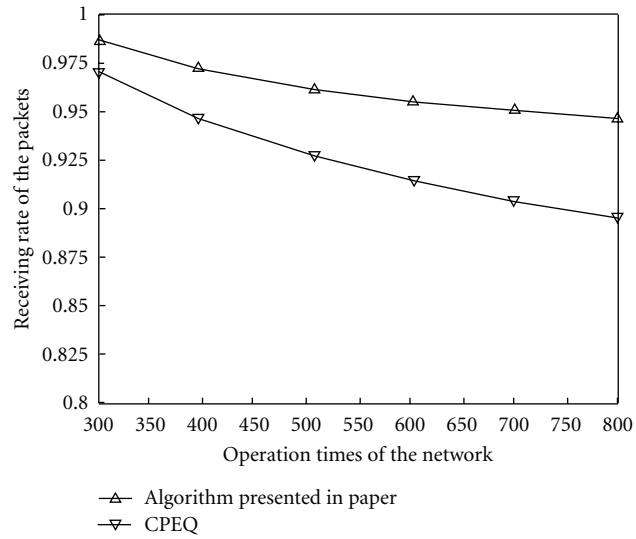


Figure 12: Comparison of packet receiving rate.

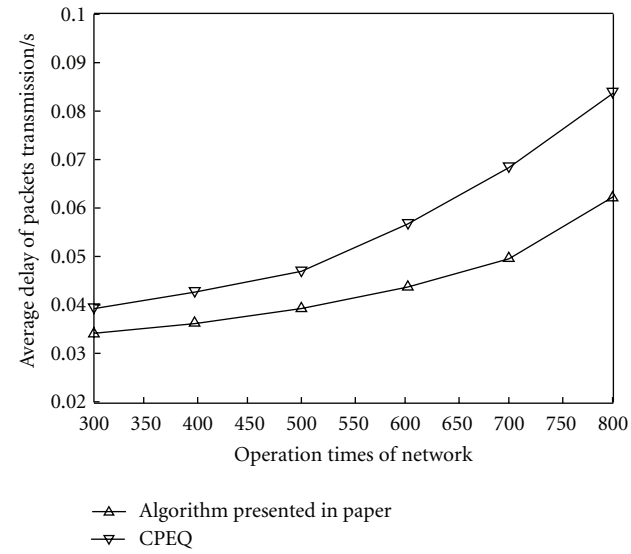


Figure 13: Comparison of average packet transmission delay.

4.3.2. Complexity Analysis of the Algorithm. The computational overhead of the nonuniform hierarchical clustering based fault tolerant routing algorithms is in terms of nonuniform hierarchical clustering by IPSO and computing the pheromone of each path by BWAS. k particle swarms are generated in the process of nonuniform hierarchical clustering by IPSO, and each particle swarm contains $D = M/K$ individual particles; so the computing complexity of the IPSO is $o((M/k)(M/k)d^2d^2k + d^2(M/k)k + k^2d^2) = o((M^2/k)d^4 + k^2d^2)$. When the positions of the particles are gradually becoming consistent with the sensors', the number of active particles becomes less during the process of the clustering. Its computational complexity presents exponential decreasing with the reduction of the number of individual particles, not keeping the same computational complexity from the beginning to the end of the clustering

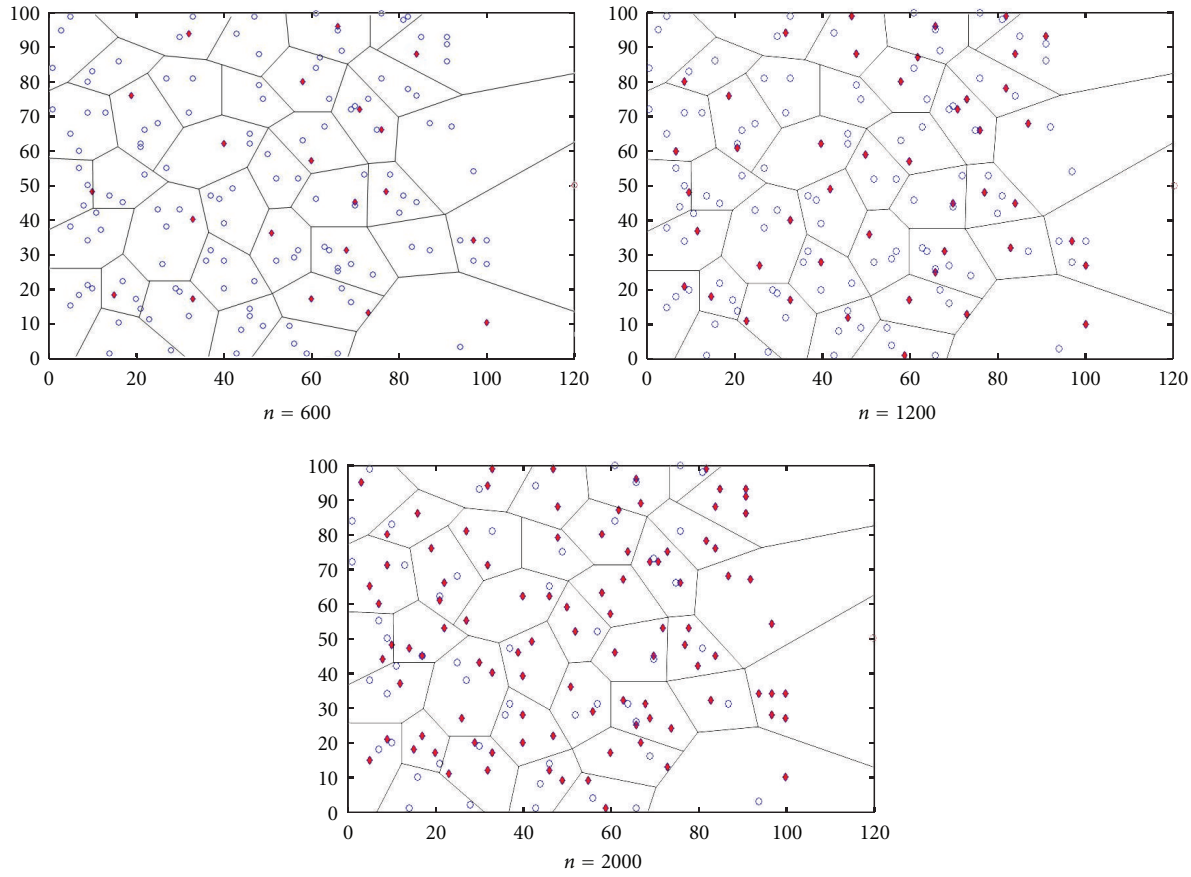


Figure 14: Energy consumption by equal probability and nonhierarchical clustering.

by the IPSO algorithm. This is much important to the resource-restricted wireless sensor networks and presents the advantages of the IPSO when using it to cluster in wireless sensor networks.

It has $\sum_{n=1}^{N} N_{nR}$ average nodes and $\sum_{n=1}^{N} N_{nR} p_n$ clustering nodes in wireless sensor networks; so the time complexity of establishing the routing and computing the pheromone of each path is $o(N(\sum_{n=1}^{N} N_{nR} p_n)^3)$ by the improved ant colony algorithm BWAS, where N is the number of iterations, and n is the number of the clustering nodes in each hierarchies. The computing and time complexities are determined by the scale of the network and the selection probability of cluster head nodes as well as the number of iterations.

4.4. Analysis of Receiving Rate and Average Packet Delay. It is important to select the high-quality communication link to improve the reliability of network data transmission. Two criterions, receiving rate and average delay, can best reflect the quality of links and performance of fault tolerance. The receiving rate of the packet is defined as the ratio of packets received at sink node and packets sent to source nodes. The closer to 1 the packet receiving rate is, the higher qualities of the links are. The average delay of the packet transmission is defined as average delay of all the packets received at sink node from the source nodes.

The fault nodes in network are supposed to be the ones in which the energy consumption reaches the initial threshold value, and the network is based on the model of timing acquisition transmission. CPEQ presented in [15] is the representative of the fault tolerance in multilayer joint optimization and shows the better performance by the comparison with DD (directed diffusion) routing algorithm. Therefore, we will make a comparison of the algorithm presented in this paper with CPEQ in [15] in receiving rate and average packet delay, as shown in Figures 12 and 13.

Figure 12 shows the comparison of the presented algorithm and CPEQ in the packets receiving rate. Fewer fault nodes have arisen before 300 times of the network operation. It cannot show the apparent influence on the performance of the network by them; so we choose the sampling data during the 300–800 times. Along with the increase of operation times, the fault nodes' number also increases. The proposed algorithm has a better performance in packet receiving rate and reflects the good stability in data transmission. Figure 13 shows the comparison of the average packet transmission delay during the operation 300–800 times between the presented algorithm and CPEQ. The newly established transmission route reestablishes the route not from the source node to sink node but from the node ahead of fault node to the sink node by the selection probability. Simulation

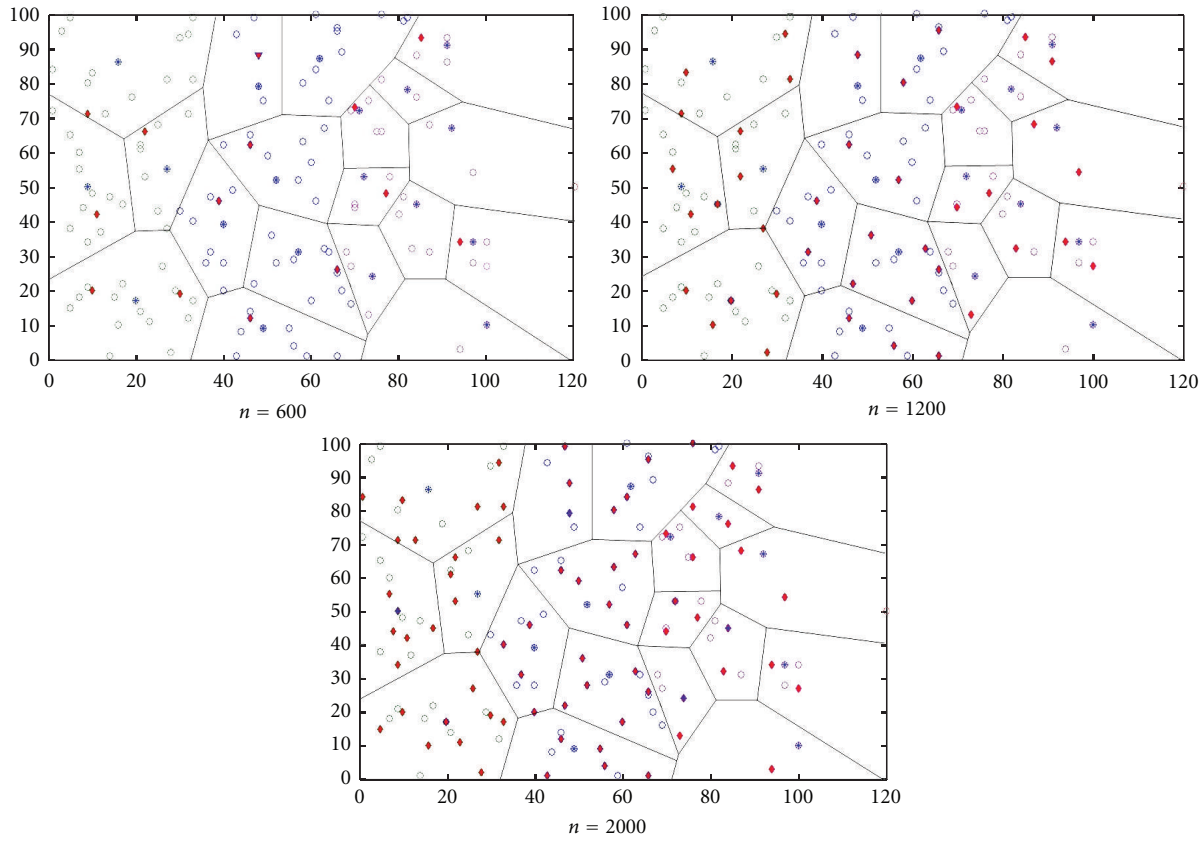


Figure 15: Energy consumption by nonequal probability and hierarchical clustering.

shows the better performance in average data transmission delay than CPEQ.

4.5. Analysis of Energy Consumption. The routing has been established based on BWAS according to two different clustering models. The first clustering model is equal probability clustering based on IPSO among the nodes with same hierarchy property. The second clustering model is the nonequal probability clustering based on the same BWAS in three hierarchies that different clustering probability is adopted in different hierarchy. Now we do the comparative analysis of energy consumption of the nodes, as shown in Figures 14 and 15.

Figure 14 shows the energy consumption by equal probability and nonhierarchical clustering. Figure 15 shows the energy consumption by nonequal probability and hierarchical clustering. There are total 150 nodes in the network. The red diamond point represents that the sensor node runs out of energy. We analyze the energy consumption at cycle $n = 600$, 1200, and 2000 times. From Figures 14 and 15, the number of dead nodes is, respectively, 21, 44, 96, and 14, 39, 72 when two models run at time $n = 600$, 1200, and 2000. The second model shows the relative smaller number of dead nodes. The dead nodes also represent randomly uniform distribution. It overcomes the "hot" issues in energy consumption caused by equal-probability clustering in WSN.

5. Conclusion and Prospect

In this paper, a fault-tolerant routing algorithm is presented based on the nonuniform hierarchical clustering inspired by the characteristics of vascular network. We build the mathematical model and the topology by marking the nodes with different hierarchies. Nonuniform hierarchical clustering is done based on IPSO with different clustering probabilities in different hierarchical areas. It establishes multiple transmission paths between the neighbor hierarchical nodes by the normalized values of the paths' pheromone generated in BWAS as the path selection probability. It selects the path with maximum probability as the actual data transmitting path to establish the fault-tolerant routing.

Theoretical analysis and simulations show that the topology based on the hierarchical nonuniform clustering can balance the nodes' energy consumption and avoid the hot issue. The routing has higher packets receiving rate and lower average transmission delay. It can avoid the data loss due to the fault of the nodes or links. So it has a good performance in the fault-tolerance and stability of the data transmission.

Our work introduces the characteristics of the vascular network into the fault-tolerant routing for WSN and carries out the studies within the framework of the biological mechanism. As a novel routing algorithm, in-depth study

is needed in the theoretical framework and practical application, especially when it applies to the fault detection and tolerance by the vascular model and the mechanism of the immune system and blood test.

Acknowledgments

This work was supported by the National Science and Technology Major Projects of China (no. 2009ZX07528-003-09), Specialized Research Fund for the Doctoral Program of Higher Education of China (no. 20100191110037), Chongqing Key Project of Science and Technology of China (no. CSCT2010AA2036) and (cstc2012gg-yyjs40008), Wanzhou District Science and Technology Planning Projects of China (no. 201203037) and ([2010]23), Youth Project of Chongqing Three Gorges University of China (no. 12QN14) and Science and Technology Project of Chongqing Three Gorges University of China ([2011]52).

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, 2002.

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

[3] L. Paradis and Q. Han, "A survey of fault management in wireless sensor networks," *Journal of Network and Systems Management*, vol. 15, no. 2, pp. 171–190, 2007.

[4] X. Wang, G. Xing, Y. Zhang, C. Lu, R. Pless, and C. Gill, "Integrated coverage and connectivity configuration in wireless sensor networks," in *Proceedings of the 1st International Conference on Embedded Networked Sensor Systems (SenSys '03)*, pp. 28–39, November 2003.

[5] R. A. F. Mini, A. A. F. Loureiro, and B. Nath, "The distinctive design characteristic of a wireless sensor network: the energy map," *Computer Communications*, vol. 27, no. 10, pp. 935–945, 2004.

[6] R. Badonnel, R. State, and O. Festor, "Management of mobile ad hoc networks: information model and probe-based architecture," *International Journal of Network Management*, vol. 15, no. 5, pp. 335–347, 2005.

[7] H. Alwan and A. Agarwal, "A survey on fault tolerant routing techniques in wireless sensor networks," in *Proceedings of the 3rd International Conference on Sensor Technologies and Applications (SENSORCOMM '09)*, pp. 366–371, June 2009.

[8] P. Djukic and S. Valaee, "Minimum energy fault tolerant sensor networks," in *Proceedings of the IEEE Global Telecommunications Conference Workshops (GLOBECOM '04)*, pp. 22–26, December 2004.

[9] N. M. Hoang and V. N. Son, "Disjoint and braided multipath routing for wireless sensor networks," in *Proceedings of the International symposium on Electrical & Electronics Engineering*, pp. 11–12, 2005.

[10] Y. Challal, A. Ouadjaout, N. Lasla, M. Bagaa, and A. Hadjidj, "Secure and efficient disjoint multipath construction for fault tolerant routing in wireless sensor networks," *Journal of Network and Computer Applications*, vol. 34, no. 4, pp. 1380–1397, 2011.

[11] T. A. Babbitt, C. Morrell, B. K. Szymanski, and J. W. Branch, "Self-selecting reliable paths for wireless sensor network routing," *Computer Communications*, vol. 31, no. 16, pp. 3799–3809, 2008.

[12] R. Luigi, "Effective erasure codes for reliable computer communication protocols," *Computer Communication Review*, vol. 27, no. 2, pp. 24–36, 1997.

[13] P. Djukic and S. Valaee, "Reliable packet transmissions in multipath routed wireless networks," *IEEE Transactions on Mobile Computing*, vol. 5, no. 5, pp. 548–559, 2006.

[14] S. Dulman, S. Nieberg, J. Wu, and P. Havinga, "Trade-off between traffic overhead and reliability in multipath routing for wireless sensor networks," *Wireless Communications and Networking*, no. 3, pp. 1918–1922, 2003.

[15] A. Boukerche, R. Werner Nelem Pazzi, and R. Borges Araujo, "Fault-tolerant wireless sensor network routing protocols for the supervision of context-aware physical environments," *Journal of Parallel and Distributed Computing*, vol. 66, no. 4, pp. 586–599, 2006.

[16] T. Bokareva, N. Bulusu, and S. Jha, "SASHA: toward a self-healing hybrid sensor network architecture," in *Proceedings of the 2nd IEEE Workshop on Embedded Networked Sensors (EmNetS-II '05)*, pp. 71–78, May 2005.

[17] A. Jabbari and W. Lang, "Advanced bio-inspired plausibility checking in a wireless sensor network using Neuro-immune systems: autonomous fault diagnosis in an intelligent transportation system," in *Proceedings of the 4th International Conference on Sensor Technologies and Applications (SENSORCOMM '10)*, pp. 108–114, July 2010.

[18] B. Atakan and O. B. Akan, "Immune system based distributed node and rate selection in wireless sensor networks," in *Proceedings of the 1st Bio-Inspired Models of Network, Information and Computing Systems (BIONETICS '06)*, December 2006.

[19] T. Rui, L. Kenji, and Z. Bing, "Immune system inspired reliable query dissemination in wireless sensor networks," in *Proceedings of the 10th International Conference on Artificial Immune Systems*, pp. 282–293, 2011.

[20] Y. J. Chen, S. F. Yuan, J. Wu, and Y. J. Zhang, "Performance optimization for wireless sensor networks based on immune system," *Systems Engineering and Electronics*, vol. 32, no. 5, pp. 1065–1069, 2010.

[21] Y. Hu, Y. Ding, and K. Hao, "An immune cooperative particle swarm optimization algorithm for fault-tolerant routing optimization in heterogeneous wireless sensor networks," *Mathematical Problems in Engineering*, vol. 2012, Article ID 743728, 19 pages, 2012.

[22] Q. Li, B.-H. Zhang, L.-G. Cui, Z. Fan, and A. V. Vasilakos, "Immunizations on small worlds of tree-based wireless sensor networks," *Chinese Physics B*, vol. 21, no. 5, Article ID 050205, 2012.

[23] H. M. Salmon, C. M. de Farias, P. Loureiro et al., "Intrusion detection system for wireless sensor networks using danger theory immune-inspired techniques," *International Journal of Wireless Information Networks*. In press.

[24] K.-I. Goh, G. Salvi, B. Kahng, and D. Kim, "Skeleton and fractal scaling in complex networks," *Physical Review Letters*, vol. 96, no. 1, 2006.

[25] I. Ahmed, M. Peng, and W. Wang, "Energy efficient cooperative nodes selection in wireless sensor networks," in *Proceedings of the International Conference on Parallel Processing Workshops (ICPPW '07)*, p. 50, September 2007.

[26] I. Ahmed, M. Peng, W. Wang, and S. I. Shah, "Joint rate and cooperative MIMO scheme optimization for uniform energy distribution in Wireless Sensor Networks," *Computer Communications*, vol. 32, no. 6, pp. 1072–1078, 2009.

[27] S. Lindsey, C. Raghavendra, and K. M. Sivalingam, "Data gathering algorithms in sensor networks using energy metrics," *IEEE Transactions on Parallel and Distributed Systems*, vol. 13, no. 9, pp. 924–935, 2002.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 170589, 8 pages
doi:10.1155/2012/170589

Research Article

Centralized Management Mechanism for Cell Outage Compensation in LTE Networks

Li Wenjing, Yu Peng, Jiang Zhengxin, and Li Zifan

State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to Li Wenjing, wjli@bupt.edu.cn

Received 3 May 2012; Revised 10 July 2012; Accepted 10 July 2012

Academic Editor: Mugen Peng

To mitigate the performance degradation induced by the cell outage, in this paper, a centralized cell outage compensation management mechanism and the corresponding workflow are proposed. Then a concrete algorithm named autonomic particle swarm compensation algorithm (APSCA) is proposed to generate the compensation scheme. The simulation results show that the proposed APSCA is effective to resolve the coverage problems induced by the outage. In the simulation scenario, almost 98.3% users that would otherwise be dropped will be served by the surrounding cells, and the block probabilities of the compensating cells are still kept on the accepted levels.

1. Introduction

Self-healing is one of the functionalities of self-organizing network (SON). The purpose of self-healing is to solve or mitigate the faults which could be solved automatically by triggering appropriate recovery actions [1]. Cell outage Management (COM) is the main task of self-healing. COM includes several phases, such as cell outage detection (COD) and cell outage compensation (COC) [2, 3]. Cell outage is the total loss of radio services in the coverage area of a cell. There are diverse reasons for the cell outage, including

(1) system hardware or software failures (e.g., circuit pack failure, channel processing error, or configuration error, etc.),

(2) transmission problems (e.g., connectivity failure, optical cable failure, etc.),

(3) external failures (e.g., power supply outage, lightning strike, etc.).

In practical network, some outage causes can be detected by the network itself. In such cases, alarms will be sent to the operation, administration, and maintenance (OAM) system or operators. Some causes can be detected by OAM through analyzing performance counters or other measurement data. Also some causes cannot be detected by network or OAM directly in real time. In such cases, automatic cell outage detection is needed to derive the probability of an outage [4]. The aim of COD is to automatically detect the occurrence, type, and scope of an outage.

When a cell is in the outage, the coverage gap is induced, and almost all the users cannot establish or maintain the radio bearers via that particular cell. The aim of COC is to mitigate the performance degradation induced by the outage to the utmost extent by automatically adjusting the related radio parameters of the neighboring cells. Adjusting such parameters may influence the users served by those cells, and the performance may be degraded in those cells, such as throughput decrease and interference increase. The trade-off between the coverage/capacity gain of outage cell and the performance degradation of the surrounding cells should be considered and achieved. The design of COC mechanism and algorithm is extremely necessary to meet the coverage, capacity, and quality requirements based on the operator's compensation policy.

In this paper, we mainly focus on the COC under the precondition that cell outage has already been detected. Firstly, a centralized COC management mechanism and the corresponding workflow are proposed. Then a concrete

algorithm named autonomic particle swarm compensation algorithm (APSCA) is proposed to generate the parameter adjustment scheme. APSCA is used to compensate the coverage gap induced by the outage while considering the balance between coverage/capacity and quality. At last, the achieved performance effects are evaluated.

The rest of this paper is organized as follows. Section 2 introduces the related works about COC. Section 3 proposes the centralized COC management mechanism and corresponding workflow. In Section 4, the COC algorithm APSCA is described and the complexity is analyzed, followed by the simulation results and discussion in Section 5. The paper is concluded in Section 6, where we conclude the current work and present a prospect of the future work.

2. Related Works

In 3GPP TS 32.541 [1], the general steps of automatic cell outage recovery and cell outage compensation are presented, but no any concrete COC algorithm is presented. The network parameters that can be adjusted by COC include cell reference signal transmission power P_{RS}, antenna tilt or azimuth, uplink target received power level P_0, and mobility parameters. In paper [5], a quantitative analysis of the compensation potential of different control parameters in mitigating outage-induced performance degradations in LTE networks is presented. In paper [6], a framework for automatic cell outage management and the key components necessary to detect and compensate the outages are presented. But no concrete compensation algorithms are proposed in [5, 6] yet.

In paper [7], an automated coverage optimization algorithm is proposed to optimize wireless downlink coverage in the RAN. However, the algorithm is based on the long-term measurements and analysis which is suitable for the normal network optimization scenario, but not suitable for the real-time cell outage compensation scenario. In papers [8–11], the algorithms, respectively, adjusting pilot power, target received power density, or antenna tilt of the surrounding cells to compensate the coverage gap are proposed. But these algorithms are all based on the stepwise adjustment methods in which the control parameters are consecutively adjusted step by step. After each step the system performance should be evaluated and then go to the next adjustment step until the compensation objective is reached. The convergence speed of those stepwise adjustment algorithms is comparatively fast, but the interim parameters may not be suitable and will usually bring instability to the system.

In this paper, we present a framework for centralized COC management mechanism and propose a heuristic COC algorithm for LTE networks, which is named Autonomic Particle Swarm Compensation Algorithm (APSCA). By APSCA, the final optimal parameter adjustment scheme can be automatically generated by the centralized node. APSCA can overcome shortcomings of the previous work, and the convergence speed is rapid enough to meet the real-time requirements of COC.

3. Centralized COC Management Mechanism

When a cell is out of the service, the control parameters of the surrounding cells need be adjusted to implement the coverage compensation and guarantee the quality of network service. Several factors should be considered when adjusting the control parameters, such as traffic load, amount of users, interference, downlink reference signal transmission power, uplink target received power, and antenna patterns. If the distributed COC management framework is adopted, large numbers of data need be exchanged among different eNodeBs, which will greatly increase the complexity of the COC algorithm and the difficulty of the implementation. As such, a centralized COC management framework is proposed in the paper. In the framework, it is the responsibility of OAM to collect and process all kinds of data and generate the final adjustment scheme. The centralized framework is more suitable for COC features.

To realize the COC management, a closed control loop is designed, which is shown in Figure 1.

The COC closed control loop, including monitor, analysis, plan, enforcement, and evaluation functionalities, is located in the OAM node, and the managed elements (eNodeBs in LTE networks) communicate with OAM through the management interface. There are two kinds of data exchanged through the management interface for COC. On one hand, eNodeBs report necessary measurement data to the OAM, and on the other hand, OAM delivers the control parameter adjustment request to the specified eNodeBs.

The workflow of the centralized COC management is illustrated in Figure 2.

The steps of the workflow are described as follows.

(1) Monitor function continuously collects the necessary data from eNodeB, UE, and OAM to determine the cell outage and the conditions of COC. The necessary data can be alarms, event notifications, KPIs, counters, and other measurement data. In [2, 6], a list of major data is presented.

(2) For cell outage detection, measurement data mentioned above is considered to determine whether an outage has occurred. As cell outage detection is not a major concern of this paper, there is no detailed discussion about it. For more information, please refer to [4]. If the cell outage is detected, then go to step 3 else continue step (1).

(3) When the cell outage is detected, the COC specified measurements will be triggered. Monitor function gathers more necessary information from UEs and eNodeBs in LTE, such as configuration, performance, and measurement data. Compared with the data monitored in step (1), these data are more detailed and particular.

(4) Analysis function analyzes and diagnoses above information and makes a conclusion.

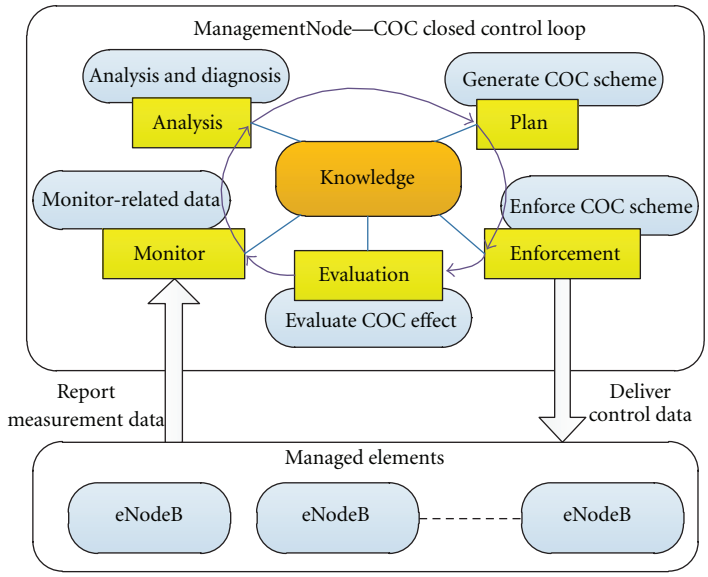


Figure 1: Centralized COC management framework.

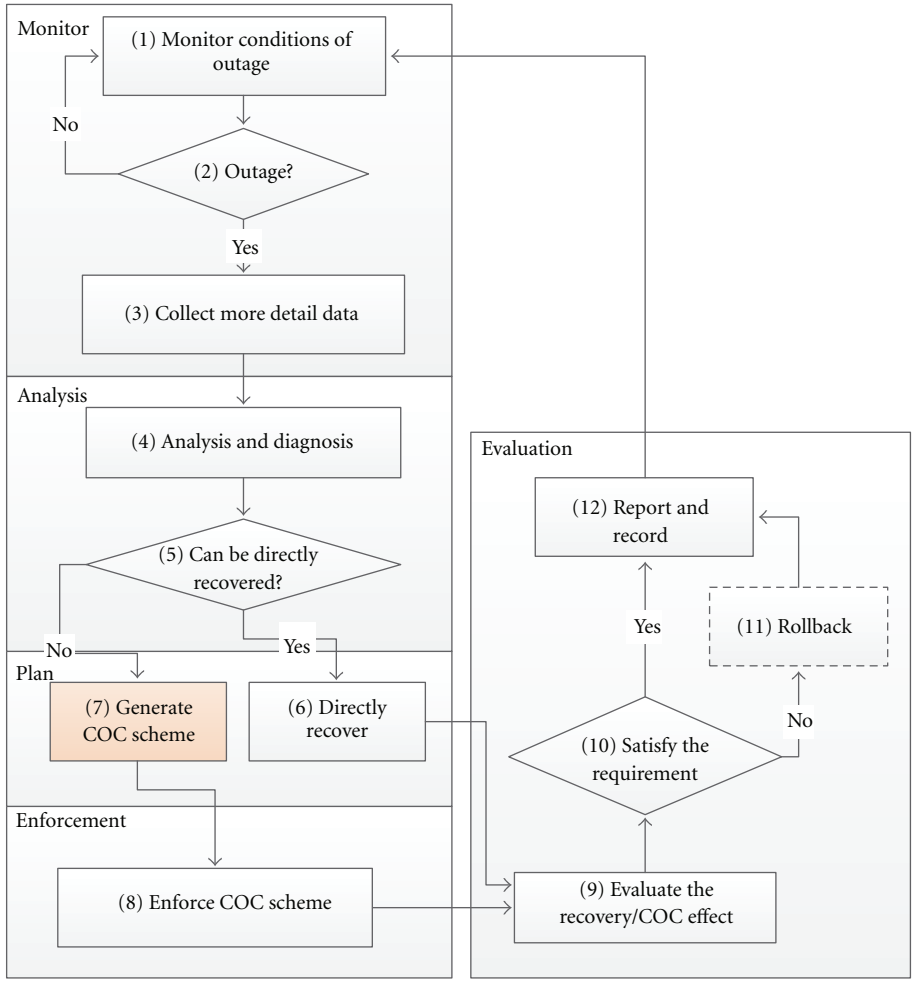


Figure 2: Workflow of centralized COC management.

(5) If the conclusion shows that the problem can be directly recovered, then go to step (6); else go to step (7).

(6) The network elements directly recover the problem which is not discussed in this paper. When the recovery action is finished, then go to step (9).

(7) To mitigate the influence of the cell outage and guarantee the QoS of users, COC will be triggered to take appropriate actions. A COC scheme generation algorithm named Autonomic Particle Swarm Compensation Algorithm (APSCA) is proposed, which will be discussed in detail in Section 4.

(8) After the COC scheme is generated, the enforcement function executes the COC scheme by delivering the parameter adjustment request to the specified eNodeBs. It should be noted that it is the eNodeBs that really adjust the control parameters according to the request.

(9) After recovery or COC, the evaluation function begins to assess the performance improvement brought by the recovery or COC. It is necessary to clarify the criteria to evaluate to what extent the compensation actions have reached in meeting the optimization goals. The COC evaluation criteria are given in Section 4.1.

(10) If the problems are recovered or the COC meets the requirements, then go to step (12), else go to step (11).

(11) If the problems are not recovered, then execute the rollback process when required.

(12) Report the COC enforcement result. OAM can record the necessary information according to the requirements. Then go to step (1) to continue monitoring the related data.

4. COC Algorithm

4.1. Coverage Evaluation Criteria. In LTE networks, the coverage effect can be evaluated by reference signal received power (RSRP) and reference signal received quality (RSRQ) measured by mobile users. The objective reference values of effective coverage of TD-LTE macrocell are as follows: firstly, the probability that the RSRP received by mobile users from the serving cell is stronger than $\text{RSRP}_{\text{minTh}}$ should exceed $\text{RSRP}_{\text{minRate}}$; secondly, the probability that the RSRQ received by mobile users from the serving cell is stronger than $\text{RSRQ}_{\text{minTh}}$ should exceed $\text{RSRQ}_{\text{minRate}}$.

Coverage problems can be represented as coverage gap and pilot pollution. In TD-LTE network, the coverage gap is defined as the area where RSRP < $\text{RSRP}_{\text{gapTh}}$, and the pilot pollution is defined by two criteria:

(a) number of cells whose RSRP > $\text{RSRP}_{\text{polTh}}$ exceeds K, and,

(b) the strongest RSRP−the Kth strongest RSRP ≤ R.

According to the practical network, the typical values of the above criteria are chosen as in Table 1.

Table 1: Coverage evaluation criteria.

Parameters	Value
$\text{RSRP}_{\text{minTh}}$	−105 dBm
$\text{RSRP}_{\text{minRate}}$	95%
$\text{RSRQ}_{\text{minTh}}$	−13.8 dB
$\text{RSRQ}_{\text{minRate}}$	95%
$\text{RSRP}_{\text{gapTh}}$	−119 dBm
$\text{RSRP}_{\text{polTh}}$	−90 dBm
K	3
R	6 dB

4.2. Model of COC Problem. In this paper, the cell reference signal transmission power P_{RS} is selected as COC adjustment parameter. The objective of COC is to serve the users, who would otherwise be dropped, by the surrounding cells as much as possible. The COC algorithm is to find out the optimal power adjustment scheme of the surrounding cells. P_{RS} can be calculated from RSRP according to link budget. For simplicity, in this paper, the P_{RS} adjustment objective is converted into RSRP adjustment objective, namely, a vector $P = [p_1, p_2, \ldots, p_m]$, where p_{i} represents the adjusted RSRP of each surrounding cell. Actual RSRP reported by the mobile user is denoted as $\mathbf{P}_i = [p_1^i, p_2^i, \ldots, p_m^i]$, where p_j^i is the reported RSRP by the ith mobile user from the jth neighboring cell and the unit is dBm.

Assume that outage cell is C_0, the amount of users in C_0 is n, and the number of neighboring cells is m. In this paper, two optimization objectives are considered.

(1) *Minimize the Coverage Gap.* The optimization objective can be described by the following expression:

$$\min P_{\text{gap}} = \frac{\sum_{i=1}^{n} \varepsilon\left(-1 \times \sum_{j=1}^{m} \varepsilon\left(p_j^i + p_j - P_{\text{th}}\right)\right)}{n}, \quad (1)$$

where p_j^i is the RSRP received by the ith mobile user from the jth surrounding cell. P_j represents the adjusted RSRP of the jth surrounding cell, P_{th} represents the threshold of the RSRP (see Table 1), and the unit of all these parameters is dBm. $\varepsilon(x)$ is step function. If $x \geq 0$, then $\varepsilon(x) = 1$; else $\varepsilon(x) = 0$. $(p_j^i + p_j - p_{\text{th}}) < 0$ denotes that user i cannot be served by cell j. When the user i cannot be served by all the m surrounding cells, then $\sum_{j=1}^{m} \varepsilon(p_j^i + p_j - p_{\text{th}}) = 0$, that is, $\varepsilon\{-1 \times \sum_{j=1}^{m} \varepsilon(p_j^i + p_j - p_{\text{th}})\} = 1$. The expression (1) is the percentage of uncovered users in n users, which can be represented as the probability of coverage gap.

(2) *Minimize the Pilot Pollution.* The optimization objective can be described by the following expression:

$$\min P_{\text{pollution}} = \frac{\sum_{i=1}^{n} \varepsilon\left(\left(\sum_{j=1}^{m} \varepsilon\left(p_j^i + p_j - P_{\text{th}}\right)\right) - K\right)}{n}. \quad (2)$$

When the dominant pilots received by the user exceed K (see Table 1), expressed as $\sum_{j=1}^{m} \varepsilon(p_j^i + p_j - p_{\text{th}}) \geq K$, that is,

$\varepsilon\{\{\sum_{j=1}^{m}\varepsilon(p_j^i + p_j - p_{\text{th}})\} - K\} = 1$. The expression (2) is the percentage of pilot pollution users in n users, which can be represented as the probability of pilot pollution.

According to the expressions (1) and (2), the optimization problem is a multiobjective nonlinear optimization problem. In the paper, an Autonomic Particle Swarm Compensation Algorithm (APSCA) is proposed to solve it. APSCA is suitable for multiobjective optimization problems and especially suitable for COC scenario because of the high speed of convergence.

4.3. APSCA Description. In APSCA, the surrounding cells are regarded as particle swarm. To find the optimal reference signal transmission powers of surrounding cells is equivalent to finding the optimal positions of particle swarm. Suppose the particle swarm includes m particles. APSCA defines each particle in the D-dimensional space. The location and velocity of the particles are, respectively, denoted as vector $X_i = (x_{i1}, x_{i2}, \ldots, x_{id})$ and $V_i = (v_{i1}, v_{i2}, \ldots, v_{id})$ where $i = 1, 2, \ldots, m$. The previous best position of a single particle is denoted as $P_i = (p_{i1}, p_{i2}, \ldots, p_{id})$, and the previous global best position of the particle swarm is denoted as $P_g = (p_{g1}, p_{g2}, \ldots, p_{gd})$.

Before finding the best position, the particles update their velocity and position according to the following formulae:

$$v_{id} = v_{id} \cdot w + c_1 \cdot r(p_{id} - x_{id}) + c_2 \cdot r\left(p_{gd} - x_{id}\right), \quad (3)$$

$$x_{id} = x_{id} + v_{id}, \quad (4)$$

where w is the constant inertia factor, c_1 and c_2 are acceleration coefficients, and r is a random number between (0, 1). We choose expressions (1) and (2) as fitness functions of particle i and choose the objective reference values of effective coverage listed in Table 1 as objective optimization criteria.

The process of APSCA is described in Algorithm 1.

4.4. Complexity Analysis of APSCA. According to the paper [12], the APSCA is convergent if the algorithm parameters ($c1, c2,$ and w) are given appropriate values.

In APSCA, the number of particles in the swarm is m which denotes the number of surrounding cells, and the dimension of each particle is D. The time complexity of APSCA is calculated as follows.

(1) The time complexity of initialization is $O(mD)$.

(2) In each iteration, the time complexity to update the location and velocity for all particles is $O(mD)$.

(3) In each iteration, the time complexity to calculate the fitness function is $O(nmD)$, where n denotes the number of original users located in the outage area.

(4) In each iteration, the time complexity to update the optimal particle is $O(m)$.

(5) In each iteration, the time complexity to update the global optimal particle swarm is $O(m)$.

(6) Suppose the iteration times is N; then the total time complexity is represented as $O[mD + N * (nmD + mD + 2m)]$ which can be approximately represented as $O(NnmD)$, where m is not more than 7 and D is not more than 2 in COC scenario.

Table 2: Key simulation system parameters.

Parameters	Value
Carrier frequency	2.6 GHz
Inter-eNodeB distance	500 m
System bandwidth	20 MHz (100 RB)
Antenna mode	3 GPP 3D Model
Antenna downtilt	15°
Channel mode	Typical urban
Maximum eNodeB output power	46 dBm
Maximum UE output power	23 dBm
Path loss from Macro to UE	$L = 128.1 + 37.6\log_{10}(R)$ (Unit of R is km)
Std deviation of shadow fading	−8 dB
Penetration loss	20 dB
UE density across macrocells	Uniform distribution
UE distribution within a macrocell	Uniform distribution
Service rate of edge users	128 kbps (UL), 512 kbps (DL)

Based on the previous analysis, the time efficiency of the algorithm is high. The APSCA is suitable for the centralized COC management scenario.

5. Simulation and Discussion

The compensation effect of the APSCA is validated in the TD-LTE scenario using Qualnet and Matlab. The simulation shows APSCA can effectively solve the coverage problem induced by cell outage while the performance degradation of compensating cells can be kept on the accepted levels.

5.1. Simulation Parameters Configuration. The simulation scenario is a TD-LTE urban macrocell region of 3.5 km × 3.5 km. Top view of simulation environment is shown in Figure 3 . In the simulation scenario, there are 7 eNodeBs with space about 500 m and about 300 users randomly scattered in the area which are not totally shown in the figure.

Refer to [13]; the key simulation system parameters are configured as Table 2.

5.2. Simulation Results and Discussion. Assume that eNodeB0 (and the corresponding cell C_0) which is located in the middle of the region is in the outage. Figure 4 shows the simulation results of cell outage compensation by APSCA.

By calculating the simulation data, we can learn that almost 98.3% of the coverage area of the cell C_0 is covered by the surrounding cells after compensation. Among the compensating cells, cells C_3 and C_6 are the most two cells to compensate the cell C_0, while other surrounding cells make the minor contribution to compensation.

–Initialization
Initialize number of particles, dimension of particle, maximum iterations, objective optimization criteria, $c1$, $c2$ and w, and so forth.
For each particle i
Initialize V_i and X_i;
Initialize P_i;
END
Initialize P_g;
–Search the Best Solution
Do
For each particle i
Update V_i according to Equation (3);
Update X_i according to Equation (4);
Evaluate the fitness of current X_i according to Equations (1) and (2);
If current X_i is better than P_i, **then**
$P_i = X_i$;
End
End
Update P_g using the best particle;
While max iterations or objective optimization criteria is reached.

Algorithm 1: Process of APSCA.

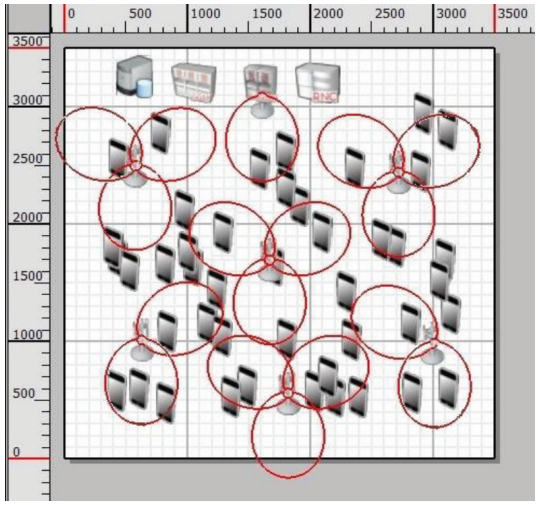


Figure 3: Simulation scenario.

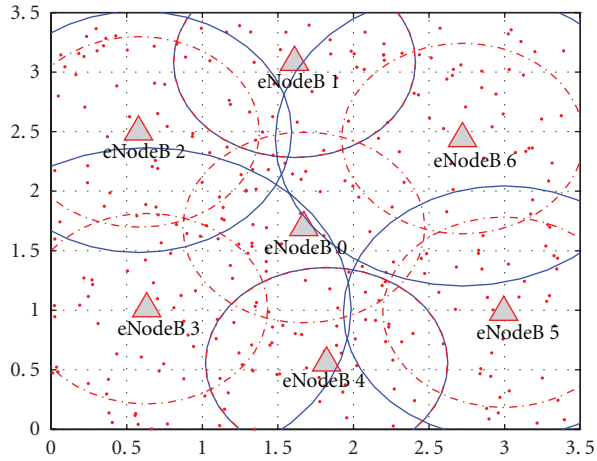


Figure 4: Compensation results.

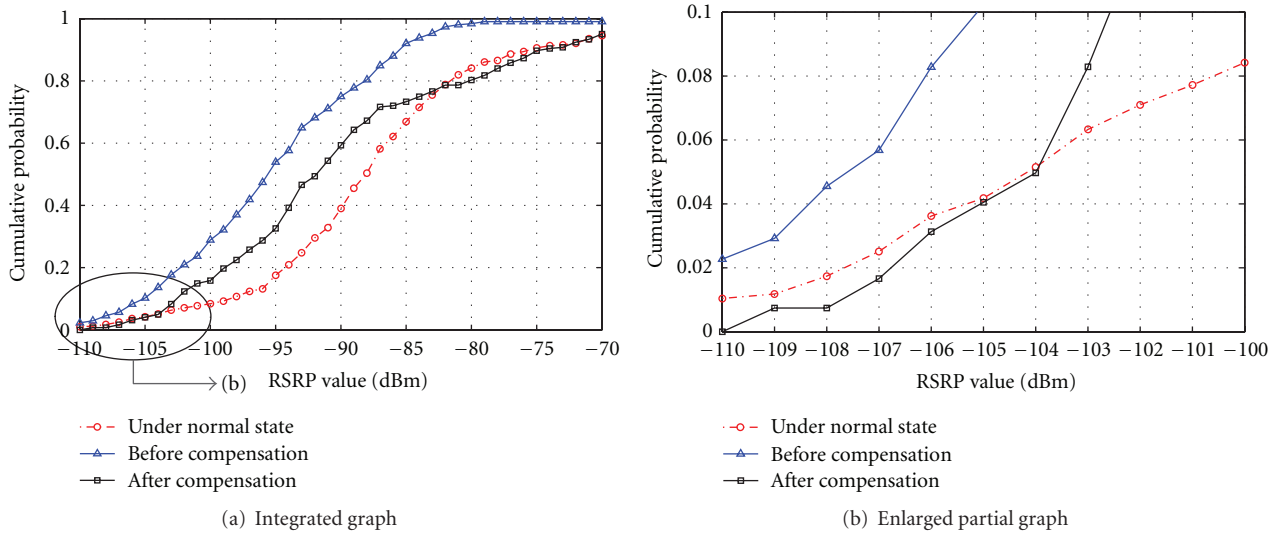
The coverage gains of the outage cell and the performance degradation of the surrounding cells induced by compensation will be discussed in detail.

(1) *Coverage Gains Analysis.* On one hand, from the viewpoint of the users in the outage cell, the received RSRP/RSRQ and their cumulative probabilities are shown in Figures 5 and 6,respectively.

As illustrated in Figure 5(a), compensation brings obvious coverage gains, especially for the weak coverage. As illustrated in Figure 5(b), before compensation, the cumulative probability that RSRP is lower than $\text{RSRP}_{\text{minTh}}$ (-105 dBm) exceeds 10% which is far worse than the objective reference value of effective coverage. After compensation, the cumulative probability that RSRP is lower than $\text{RSRP}_{\text{minTh}}$ (-105 dBm) equals to 3.9% which is within the scope of effective coverage.

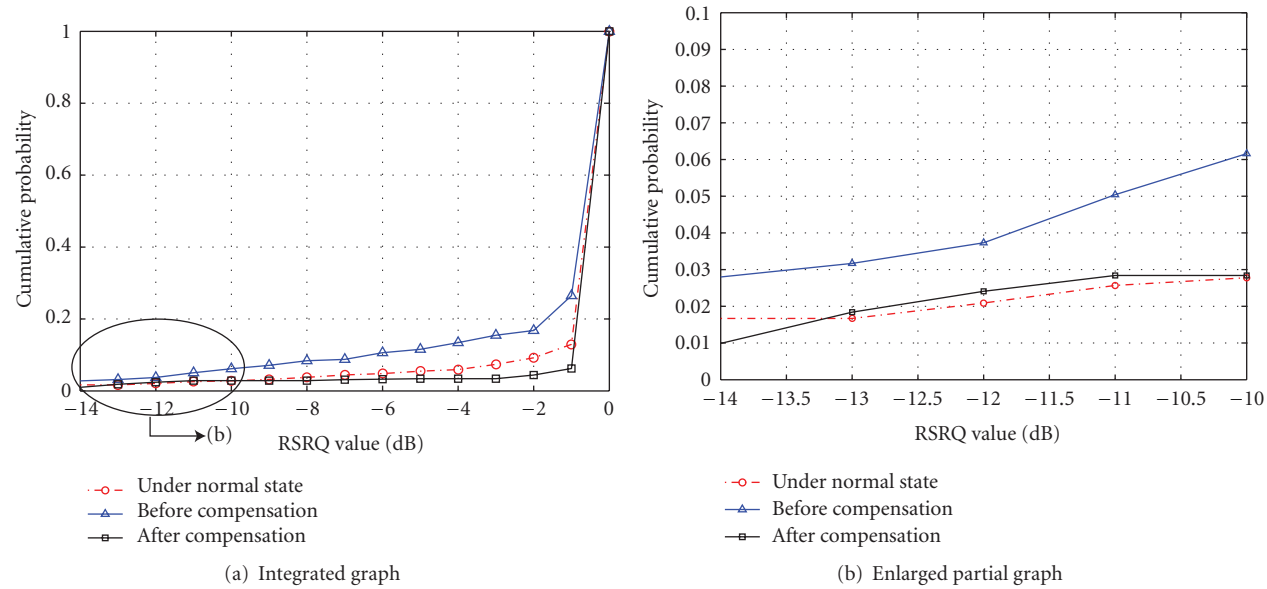
Furthermore, according to Figure 5(b), when the RSRP is below -105 dBm, the cumulative probability after compensation is lower than that of normal state, which means compensation is effective in decreasing the probability of weak coverage.

As illustrated in Figures 6(a) and 6(b), after compensation, the cumulative probability that RSRQ is stronger than -13.8 dB is 98.2%, which is within the scope of



(a) Integrated graph (b) Enlarged partial graph

Figure 5: RSRP and cumulative probability of outage cell.



(a) Integrated graph (b) Enlarged partial graph

Figure 6: RSRQ and cumulative probability of outage cell.

effective coverage, and is 11.33% higher than that of before compensation.

Figures 5 and 6 show that APSCA can effectively solve the coverage problems induced by cell outage.

(2) *Performance Impact Analysis.* On the other hand, from the viewpoint of users in the compensating cells, the service block probability can reflect the performance impact induced by compensation in some degrees. So we choose the cell C_3 to simulate the service block probability and take the HTTP service as an example. The simulation result is shown in Figure 7.

As shown in Figure 7, the cell C_3 begins to compensate the cell C_0 at 11:00 am; after that time, the block probability is a little higher than that of normal state. According to statistics, the service peak hours are during 13:00~17:00. After 13:00, the block probability rises obviously, and the peak block probability is 63% higher than that of normal state (no compensation). The reason is that many additional users from the cell C_0 are now served by C_3, which degrades the performance of C_3 in some ways. But the block probability is still below the threshold of 1% [14] which means the block probabilities after the compensation are still kept on the accepted levels.

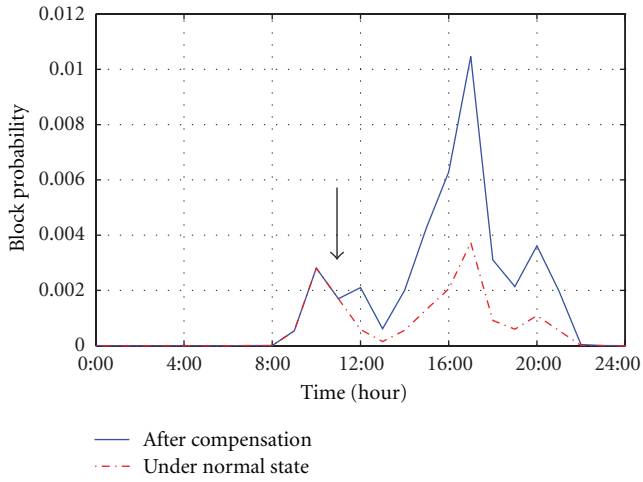


Figure 7: Service block probability of compensation cell.

6. Conclusion

To mitigate the performance degradation induced by the cell outage, a centralized cell outage compensation management mechanism and a concrete compensation scheme generation algorithm named APSCA are proposed in the paper. Simulation results show that the APSCA is effective in compensating the coverage problem induced by cell outage while the performance degradation of the compensating cells can be kept on the accepted levels.

Our future work will focus on the following three steps. First, we will investigate how to adjust other control parameters, such as antenna parameter and uplink target received power, to compensate the outage, and analyze which kind of parameters are more suitable for effective compensation in different scenarios. Second, the relations among these control parameters will be studied, and cooperative adjustment scheme generation algorithm will be designed. Lastly, more factors, such as load, interference and services, will be taken into consideration when designing the COC mechanism and algorithm.

Acknowledgments

This research is supported by the Funds for Creative Research Groups of China (61121061), National S&T Major Projects (2011ZX03003-002-01), NCET-10-0240, and Chinese Universities Scientific Fund (BUPT2012RC0608).

References

[1] 3GPP TS 32.541 V10.0.0, Telecommunications Management, Self-organizing Networks (SON), Self-healing concepts and requirements (Release 10), March 2011.

[2] INFSO-ICT-216284, SOCRATES D2.1, Use Cases for Self-Organising Networks, March 2008.

[3] NGMN Recommendation on SON and O&M Requirements, Next Generation Mobile Networks Recommendation on SON and O&M Requirements, December 2008.

[4] C. Mueller, M. Kaschub, C. Blankenhorn, and S. Wanke, "A cell outage detection algorithm using neighbor cell list reports," in *Proceedings of the International Workshop on Self-Organizing Systems*, pp. 218–229, 2008.

[5] M. Amirijoo, L. Jorguseski, R. Litjens, and R. Nascimento, "Effectiveness of cell outage compensation in LTE networks," in *Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC '11)*, pp. 642–647, January 2011.

[6] M. Amirijoo, L. Jorguseski, T. Kürner et al., "Cell outage management in LTE networks," in *Proceedings of the 6th International Symposium on Wireless Communication Systems (ISWCS '09)*, pp. 600–604, Siena, Italy, September 2009.

[7] Z. Heng, Q. Xue-song, M. Luo-ming, L. Wen-jing, and Z. Xi-dong, "An automated coverage optimization scheme used in radio access network[J]," *Journal of Beijing University of Posts and Telecommunications*, vol. 34, no. 5, pp. 29–32, 2011.

[8] M. Amirijoo, L. Jorguseski, R. Litjens, and L. C. Schmelz, "Cell outage compensation in LTE networks: algorithms and performance assessment," in *Proceedings of the IEEE 73rd Vehicular Technology Conference, VTC2011-Spring*, May 2011.

[9] L. Fuqiang, Q. Xuesong, L. Wenjing, and W. Honglin, "High load cell outage compensation method in TD-SCDMA wireless access network TD-SCDMA," *Journal of Beijing University of Posts and Telecommunications*, vol. 35, no. 1, 2012.

[10] L. Fu-qiang, Q. Xue-song, M. Luo-ming, and Z. Heng, "An algorithm of cell outage compensation in wireless access networks," *International Journal of Advancements in Computing Technology*, vol. 4, no. 1, pp. 404–414, 2012.

[11] L. Fu-qiang, Q. Xue-song, M. Luo-ming, and Z. Heng, "Achieving cell outage compensation in radio access network with automatic network management," in *Proceedings of the GlobeCom Workshops*, pp. 673–677, December 2011.

[12] F. Yangguang, Z. Chengping, and D. Mingyue, "Convergence analysis of standard particle swarm optimization algorithm," *Mathematica Applicata*, vol. 24, no. 1, pp. 187–194, 2011.

[13] 3GPP TR 36.814, Further advancements for E-UTRA physical layer aspects, V9.0.0, March 2010.

[14] L. Chiaraviglio, D. Ciullo, M. Meo, and M. A. Marsan, "Energy-efficient management of UMTS access networks," in *Proceedings of the 21st International Teletraffic Congress (ITC 21 '09)*, pp. 1–8, Paris, France, September 2009.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 785418, 12 pages
doi:10.1155/2012/785418

Research Article

Hybrid BS-Cooperative Power Management Scheme with Self-Organized Sleep Mode in Virtual Cell-Based Femto Networks

Wei Zheng,[1,2] Wei Li,[1,2] Yuanbao Xie,[1,2] and Xiangming Wen[1,2]

[1] *School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China*
[2] *Beijing Key Lab of Network System Architecture and Convergence, Beijing University of Posts and Telecommunications, Beijing 100876, China*

Correspondence should be addressed to Wei Zheng, zhengweius@gmail.com

Received 4 May 2012; Revised 19 September 2012; Accepted 21 September 2012

Academic Editor: Yiqing Zhou

A large-scale deployment of femtocell BSs (FBSs) causes substantial energy consumption. This paper proposes a hybrid BS-cooperative power management (HBCPM) scheme with self-organized sleep mode in virtual cell-based femto networks. Firstly, HBCPM builds a leader-member virtual cell framework, in which only one FBS is a FBS leader (FL) and others are FBS members (FMs). Then, the FL acts as an autonomous entity and is responsible for detecting active calls in the virtual cell coverage, while the FMs without active connections can entirely shut down pilot transmissions and the related processing all the time. So it is suited for self-organizing networks (SONs). Based on the proposed scheme, the state transition model is established by the semi-Markov stochastic process, and the analytic formulas of average cumulative delay and interference time as well as the energy consumption are derived. Meanwhile, three prominent-related schemes are also studied by the proposed model. With the practical long-term evolution (LTE) system parameters and three-dimensional femtocell network model, the numerical simulation and theoretical analysis match pretty well, and the tradeoff between the energy consumption and average cumulative delay is also manifested. Moreover, simulation results show that the proposed algorithm outperforms other three schemes in terms of average cumulative interference time and the energy consumption.

1. Introduction

1.1. The Motivation. The tremendous success of mobile cellular networks leads to wide proliferation and demand for ubiquitous heterogeneous broadband mobile wireless services, meanwhile, recent investigation confirms that most of voice services and data services take place in indoor scenarios [1]. Due to the high penetration loss and the long transmission distance, there is a huge cost for the macrocell to serve massive indoor users with extensive transmission demands. To address this problem, the applications of so called femtocell BS (FBS) have been investigated [1]. FBS is a short-range, low-power, low-cost, as well as customer-owned cellular base station deployed by the terminal consumers. It is connected to the existing broadband access network via fixed backhaul such as digital subscriber line or coaxial cable. FBS brings with it not only the enhanced indoor coverage without much additional cost, but also the other significant benefits such as increased spectral efficiency, offloaded macrocell burden, improved network capacity, and prolonged life for phone battery.

Although femtocell technique is deemed to be an emerging candidate solution of catering for the ever-growing demand for the wireless service, researchers and operators so far mainly focused on capacity or spectrum efficiency improvement by adding the femtocell into the macrocell [2–4] and do not pay much enough attention to the potential energy expenditure with more FBSs. ABI research [5] predicts that more than 36 million of femtocells are expected to be sold worldwide with 150 million of customers by the end of year 2012, and thus the energy consumption of the wireless cellular network might be drastically increased by the overwhelming FBSs. The growth of the energy consumption will cause an increase in global carbon dioxide emissions and impose more and more challenging operational costs for operators. Therefore, it is necessary to consider the energy

consumption of femtocell networks, and then design an energy saving scheme without sacrificing its core benefits and functions.

Contrasting with macrocells, the traffic load of femtocell networks has more significant fluctuations in space and time due to a number of factors such as user mobility and behavior, as well as supporting very fewer simultaneous users for each FBS. During daytime, traffic load is generally higher in office areas compared to residential areas, while it is the other way around during the night. Therefore, there will always be some femtocells under the low traffic load, while some others may be under the heavy traffic load. However, FBSs in current cellular network architecture are always required to be active on air interface by continuously transmitting pilot signals and doing some related processing. In addition, the wireless air interface consumes energy with the same rate in receive, transmit, or idle states. In turn, the less the wireless air interface is operating, the less energy is consumed as well as the fewer interference time is caused. Hence, the most important issue in developing networks which are energy-aware is to model the power consumption of the wireless air interface [6]. So, energy saving potential of FBSs needs to be exploited by designing protocols to enable to shut down the wireless air interface, or to go to the sleep mode in FBSs [6]. Then, it is also necessary to study a suitable scheme to control the sleep mode behavior for FBSs.

1.2. Related Work. With green communication emerging, energy efficiency has become an important system design parameter for the long-term economic evaluation indicator of cellular networks. Researchers and operators gradually focus on the energy efficiency in two-tier femtocell networks [7–9]. Khirallah and Thompson present a methodology for estimating the total energy consumption, taking into account the total operational power and embodied energy, and TCO (total-cost-of-ownership) of two-tier femtocell cellular networks [7], which also shows that macro-femtocell networks reduce the networks TCO at the expense of increased energy-consumption. The above study mainly analyze energy efficiency whether can be improved with the femtocell deployment. However, they do not consider how to reduce energy consumption for a large-scale deployment of femtocells. Cheng et al. [8] introduce a spectrum splitting strategy to minimize the downlink energy consumption while suppress cross-tier interference. Domenico et al. [9] propose two resource management schemes that can limit the overall interference per resource block (RB) from neighbor FBSs and reduce the transmission power in each RB as well.

Usually, the wireless interface consumes energy with the same rate in receive, transmit, or idle states. In turn, the less the wireless interface is operating, the less energy is consumed. Therefore, the best strategy to minimize the energy consumption is to shut down the wireless interface, or to go to energy saving mode [6]. By contrast, there are relative less literature on designing protocols to enable sleep modes in FBSs and controlling the sleep mode behavior of FBSs for energy efficient improvement. The LTE standard introduces power saving protocols such as discontinuous transmission (DTX) and discontinuous reception (DRX) mode for both mobile stations and base stations [10]. DTX and DRX are methods to momentarily power down the device by switching off the transceivers whenever there is no need to transmit or receive. A timer-based energy-efficient solution is presented in IEEE 802.16 m standard [11] that the FBS periodically transmits pilot signals in low traffic scenario. The transmission period of the pilot signal is divided into two parts: fixed available and unavailable intervals, where the FBS switches off pilot transmissions and enters into the sleep mode. To reduce unnecessary available intervals, Widiarti et al. propose a dynamic shutting down pilot transmissions mechanism based on FBSs grouping [12]. However, the above schemes will not be efficient as long as the sleep/idle users are in the FBS coverage. Regarding this, the FBS does not need to turn on its pilot transmissions and related processing via user activity detection, irrespective of the location of registered users [13]. Nevertheless, this scheme only studies how to reduce the energy consumption from single FBS aspect, not suitable for the large-scale femtocell deployments. It also does not analyze the impact of various key parameters on the proposed scheme such as session arrival rate and so on.

1.3. The Contributions. In this paper, a novel hybrid (centralized/distributed) BS-cooperative power management (HBCPM) scheme is proposed for self-organized sleep mode in virtual cell-based femto networks. Firstly, according to the femtocell network topology, BCHPM builds a leader-member virtual cell framework, that is, femtocell networks are classified into a number of virtual cells based on one-hop neighbor list, the size of which is the sum of the sizes of FBSs. Within each virtual cell, the FBS is elected as either a FBS leader (FL) or a FBS member (FM) based on the list degree and the sum of received pilot signal power. Moreover, because of the time variation characteristic of the femtocell network topology, the role assignment algorithm for the newly installed FBS is introduced. Secondly, the FL is responsible for detecting active calls in the virtual cell coverage, while the FMs without the active user can completely switch off the pilot transmission and the associated processing all the time. Once the active registered user is detected by the FL, FL cooperates with the target FM, and then the target FM becomes active on the wireless air interface. Thirdly, the performance of HBCPM is analyzed theoretically in terms of average cumulative delay and the energy consumption and average cumulative interference time from the following two aspects: (i) state transition model is established; (ii) the analytic formulas of average cumulative delay and interference time and the energy consumption are derived. Finally, according to theoretical analysis of the proposed scheme, we further research three-prominent related schemes, and then, we assess the effectiveness of the proposed scheme by comparing its performance with them. With a three dimensional (3D) femtocell network model and practical LTE parameters, simulation results show that the proposed scheme outperforms three existing schemes in terms of average accumulative interference time and the energy consumption. Furthermore, the results provide some guidelines for deploying energy efficient femtocell networks.

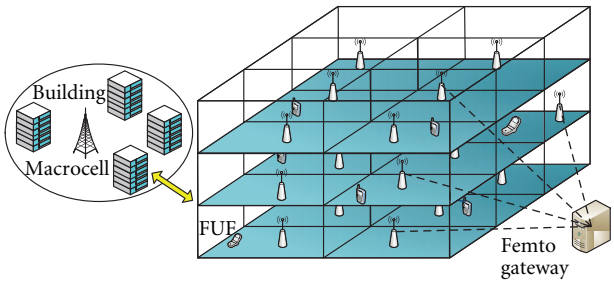


Figure 1: The two-tier femtocell cellular network deployment.

The rest of this paper is organized as follows. Section 2 presents a 3D femtocell network deployment model, and describes a novel hybrid BS-cooperative power management scheme for the self-organized sleep mode in virtual cell-based femto networks. Section 3 analyzes the proposed scheme theoretically from the aspects of the energy saving and average cumulative delay and interference time. Meanwhile, three-prominent related schemes are investigated in theory as well. Section 4 provides simulation results and discussions. Finally, Section 5 concludes the paper and future work.

2. Approach: HBCPM

2.1. Femtocell Cellular Network Model. Figure 1 shows three dimensional topology of two-tier femtocell cellular networks, also called macrocell-femtocell overlay networks. A large number of FBSs with a small coverage area are randomly distributed on every floor of office blocks or residential buildings. The buildings are in the coverage of single overlay macrocell BS (MBS), and all FBSs in the building are under the centralized control of a femtocell server referred as the femto gateway (F-GW), which manages operation and maintenance (OAM) information such as the femtocell location and the identification through a backhaul link [14]. In addition, closed subscriber group (CSG) mode is adopted in this paper, where only authorized users and a list of invited users are allowed to access the given femtocell and occupy the resources for data communications.

2.2. Proposed Scheme. This section presents the HBCPM approach, which is based on the following two main components: (i) leader-member virtual cell construction, (ii) hybrid FBS-cooperative power management.

2.2.1. Leader-Member Virtual Cell Construction. The objective of this part is to describe the virtual cell construction procedure which divides the femtocell network into different virtual cells. Considering the uneven distribution of the FBSs in femtocell networks, a virtual cell construction algorithm is proposed on the basis of the femtocell distribution density and the received pilot signal power strength. According to the third generation partnership project (3GPP) standard [15], when a FBS is installed or switched on, it firstly enters into an initialization state. The FBS in this state needs to scan the radio environment to search for neighbor FBSs for getting necessary information. During this scanning process, the one-hop neighbor list including the received pilot signal power indicator from neighbors can be obtained. Note that the received pilot signal power strength must be above the predefined threshold to ensure that they are very close.

Utilizing the one-hop list, each FBS can calculate the sum of the received pilot signal power and the number of the FBSs in the list (i.e., list degree), and then sends these information to the F-GW. Note that the greater the sum of the received pilot energy strength or the list degree is, the more FBSs are located around this FBS. In subsequent frame, the F-GW elects the one which has the highest list degree among all FBSs as the first FL. If more than one FBSs are chosen as the FLs, the one with the maximal is chosen as the first FL. If the sums of the received pilot signal power are equal, break the tie randomly. After the first FL is determined, each FBS in the first FL's one-hop list decides whether it can be a FM of the first FL. If one FBS calculates that the pilot signal power received from the first FL is maximal in its one-hop list, it serves as FMs and attached to the elected FL. Otherwise, wait for the next role assignment. After the first virtual cell is generated, the second virtual cell is initiated with the FBS having the highest sum of the received pilot signal power among FBSs excluding the ones belonging to the first virtual cell and is updated as the first virtual cell. Besides, the rest of virtual cells are also constructed in the same way.

More formally, the whole algorithm to construct the virtual cell is summarized in Algorithm 1, where the set of operational FBSs in the femtocell network is given by $\varphi = \{1, 2, \ldots, N\}$. ψ_n is the set of FBSs which are in the one-hop list of the FBS n, and $|\psi_n|$ denotes the list degree of the FBS n. Define the received pilot signal power measured by the FBS n, from the FBS m as p_{nm} $(n, m \in \varphi)$, and p_n represents the sum of the received pilot signal power computed by the FBS n.

Furthermore, a role assignment algorithm is studied for the newly added FBS. When the new FBS F_n is installed, it scans the surrounding area to search for neighbor FBSs and sends the one-hop list to the F-GW. Since the F-GW has the global knowledge about the topology of the femtocell networks, it knows each virtual cell configuration, including the FL and the FMs based on the identification. Then, using this information, the F-GW will send a message to notify the new FBS what role it should play. If there is only one FL in the one-hop list, the new FBS serves as a FM belonging to this FL; if there are more than one FLs in the list, compare the pilot power received from these FLs and choose the one with the largest pilot power as the FL of the new FBS (if the received pilot power is equal, break the tie randomly). Otherwise, the new FBS forms a new virtual cell by itself and be a FL. The role assignment algorithm for the new FBS is described by the pseudocode in Algorithm 2.

2.2.2. Hybrid FBS-Cooperative Power Management. The object of this part is to present the FBS-cooperative power management algorithm in each virtual cell. We need to introduce a low power consumption device installed with each FBS, called a sniffer. It is used to detect the energy strength of

Initialize: $p_n = 0$, $p_{nm} = 0$, $\psi_n = \phi$, $\varphi = \{1, 2, \ldots, N\}$, $|\psi_n| = 0$, $n, m = 1, 2, \ldots, N$
(1) Each FBS n creates the one-hop list and calculates the list degrees $|\psi_n|$ and the sum of the received pilot signal power $p_n = \sum_{m \in \psi_n} p_{nm}$;
(2) Elect the FL n^* from set φ, where $n^* = \arg_{n \in \varphi} \max |\psi_n|$. If there is more than one maximum, choose the FBS n^{**} as the FL, where $n^{**} = \arg_{n^*} \max p_n$. If sums of the pilot signal power are equal, break the tie randomly;
(3) Update the lth virtual cell as $C_l = \{n^{**}\}$;
(4) If $\psi_{n^{**}} \neq \phi$, the FBS $m^* \in \psi_{n^{**}} \cap \varphi$ acts as a FM of the lth virtual cell, which should meet the following condition: $p_{m^* n^{**}} \geq p_{m^* n'}, \forall n' \in \psi_{m^*}$; otherwise go to step 6;
(5) update C_l as: $C_l \leftarrow C_l \cup \{m^* \mid p_{m^* n^{**}} \geq p_{m^* n'}, \forall n' \in \psi_{m^*}\}$;
(6) Update φ to exclude the FBSs belonging to the first virtual cell, that is, $\varphi \leftarrow \varphi \setminus C_l$, and then recalculate $|\psi_n|$ and p_n, $(n \in \varphi)$;
(7) Repeat steps 2~6 until the set φ is empty.

Algorithm 1: Virtual cell construction algorithm.

(1) F_n scans the surrounding radio environments and creates the one-hop neighbor list;
(2) F_n sends the one-hop list to the F-GW, and then the F-GW notifies F_n what role it should serve as;
(3) **if** F_n has only one FL in its one-hop neighbor list, **then**
(4) the F-GW elects this FBS as the FL of F_n, and F_n acts as a FM;
(5) **else**
(6) **if** F_n has more than one FLs in its one-hop list, **then**
(7) compare the pilot power strength of F_n received from these FLs, and choose the one with the largest pilot power strength as the FL of F_n;
(8) **else**
(9) F_n forms a new virtual cell by itself and be a FL
(10) **end if**
(11) **end if**

Algorithm 2: Role assignment algorithm.

an active call. In each virtual cell, the sniffer of the FL needs to keep alive all the time. When there is an active call in the virtual cell coverage, the sniffer of FL detects an energy rise of the received signal on the uplink frequency band. If the received signal energy strength is higher than the predefined threshold, the registered user with an active call is considered to be close enough to the virtual cell.

Figure 2 shows the operational flowchart of cooperative FBS power management procedure. Assume that the FL can exchange information with the FMs in the same virtual cell, similar to the information interaction among MBSs through X2 interface in 3GPP [14]. Take a virtual cell for example. In the initial state, the FBSs without the active user in the virtual cell disable pilot transmissions and the associated processing and the sniffers, while the sniffer of the FL is busy performing measurements on the macrocell uplink band. When the sniffer of the FL detects an active registered user in the vicinity of the virtual cell and the received energy strength p_{MF} exceeds the threshold level p_{TH}^1, the FL sends a message to notify the target FM to open the sniffer. Here, we assume that the FL knows which FM the registered user can access. Meanwhile, the time counter of the target FM is started and the sniffer begins to detect the energy rise in the received signal. If the received energy strength is lower than the threshold p_{TH}^2 within the timer range T_0, the sniffer switches off and the time counter resets immediately, then the target FBS returns to the origin state and sends message to the FL. Otherwise, the target FM switches on pilot transmissions and the associated processing. Then, the registered user can detect the target FM in the next scanning and establish a connection with it. After the active call is finished, the target FM turns off the pilot transmission and the processing and sends message to notify the FL.

Specially, when the sniffer of the FL detects the registered user making an active call in the target FM coverage area, the FL directly notifies the target FM to activate the sniffer and pilot transmissions. By doing so, the active user can detect the target FM and perform handover to it with a slight delay.

3. Theoretical Analysis

This section establishes a state transition model based on the proposed scheme, and then closed form expressions of the energy consumption, average cumulative delay, and interference time are derived.

3.1. State Transition Model. To analyze the proposed scheme, we build up the state transition model including the following three states as shown in Figure 3. For the ease of description, define these states as state 1, state 2, and state 3, respectively. In each state, the FL needs to keep its sniffer on for the active call detection. In state 1, not only the pilot

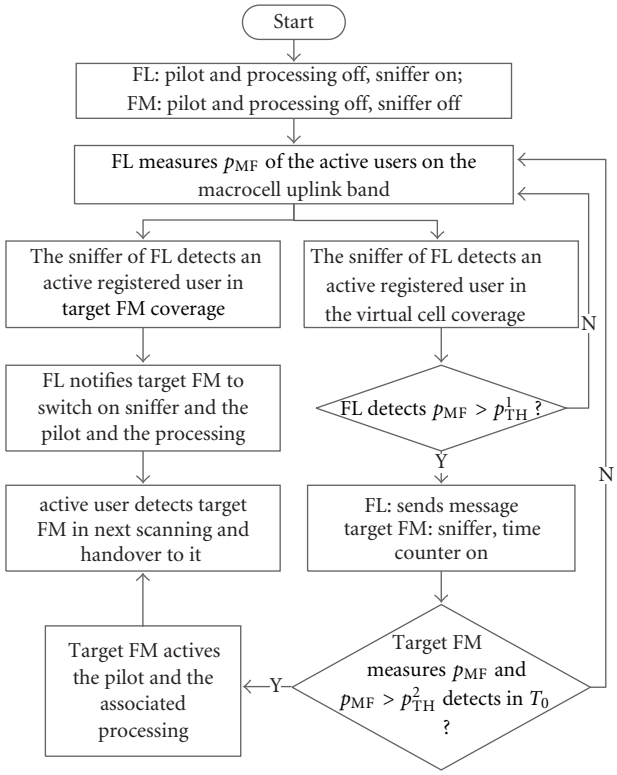


Figure 2: Flowchart of hybrid FBS-Cooperative Power Management.

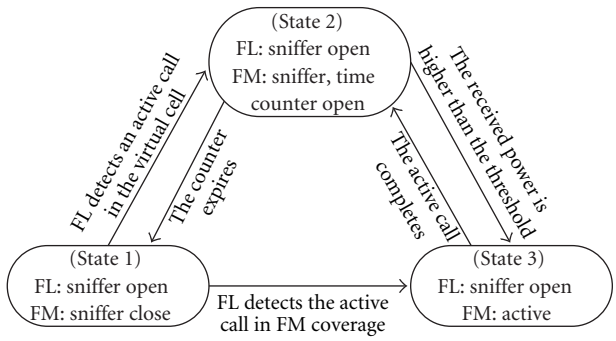


Figure 3: State transition model for HBCPM.

transmission and the associated processing of the FMs are off, but also the sniffer. When the sniffer of the FL detects an active registered user in the vicinity of the virtual cell and the received signal energy strength exceeds the threshold level, the FM moves into state 2. In state 2, both the sniffer and the time counter of the FM are on. If there is no session arrival in the timer range, the time counter is reset and the sniffer is switched off immediately, and the FM returns back to the state 1. Otherwise, the FM enters into the state 3. In state 3, the pilot transmission and the processing of the FM are both on, which provides services with active users. If the active call completes, the FM disables its pilot transmission and moves into state 2. Specially, when the sniffer of the FL detects an active registered user in the target FM coverage, the target FM directly enters into state 3.

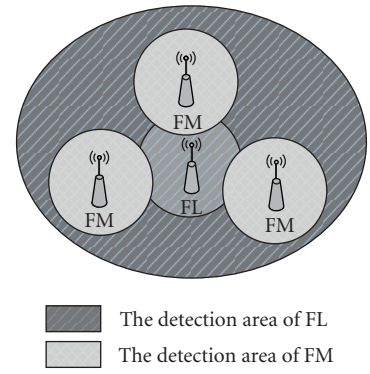


Figure 4: simplified model of femtocell deployments.

3.2. Analysis of the Proposed Scheme. Our objective in this subsection is to derive the analytic formulas of the energy consumption, average cumulative delay, and interference time. For simplicity, we consider a single virtual cell which consists of a FL and a FM as illustrated in Figure 4. The detection area of the FL is the coverage of the virtual cell, while the detection area of the FM is its own coverage. First, we shall make the following assumptions for the rest of the work.

AS 1. At the initial condition, no user exists in the FM coverage or the users under the coverage of the FM are in sleep/idle mode.

AS 2. The session arrival in the virtual cell coverage and the FM coverage follow Poisson distribution with the rate λ_c and λ_m, respectively.

AS 3. The time duration in state 1 follows a $M/M/\infty$ queuing model, and the service time for each session follows negative exponential distribution with the average service rate ρ, ($\rho > 0$).

AS 4. The timer range of the time counter of the FBS is constant and set as T_0.

Let the waiting time of the nth session arrival be defined as T_n. Based on the assumption above, the distribution function $M_{T_n}(t)$ and probability density function $m_{T_n}(t)$ of T_n can be given by

$$M_{T_n}(t) = P(T_n \le t) = 1 - e^{-\lambda_m t}, \quad t > 0,$$
$$m_{T_n}(t) = \frac{dM_{T_n}(t)}{dt} = \lambda_m e^{-\lambda_m t}, \quad t > 0. \tag{1}$$

According to the state transition model, define the energy consumption E of the FBS per hour, average cumulative delay D_{ave}, and interference time T_{int}, respectively, as

$$\begin{aligned} E &= \sum_{i=1}^{3} E_i P_i, \quad i = 1,2,3, \\ D_{\text{ave}} &= \frac{P_1 \times \overline{t_1} \times \lambda_m \times (D_{12} + D_{23}) + P_2 \times \overline{t_2} \times \lambda_m \times D_{23}}{(P_1 \times \overline{t_1} + P_2 \times \overline{t_2} + P_3 \times \overline{t_3}) \times \lambda_m} \\ &= \frac{P_1 \times \overline{t_1} \times (D_{12} + D_{23}) + P_2 \times \overline{t_2} \times D_{23}}{P_1 \times \overline{t_1} + P_2 \times \overline{t_2} + P_3 \times \overline{t_3}}, \\ T_{\text{int}} &= P_3 T_H, \end{aligned} \tag{2}$$

where E_i denotes the energy consumption of the FBS per hour in state i and P_i represents the steady state probability of each state. $\overline{t_i}$ is the mean residence time of the FBS in state i. D_{12} and D_{23} are average cumulative delay needed to initiate a new session from state 1 to state 2 and from state 2 to state 3, respectively. T_H is a statistical period. Since the residence time of the FBS in each state does not follow the exponential distribution, a semi-Markov process model is introduced to analyze the state transition behavior of the FBS [16]. Thus, the steady state probability of each state can be represented as

$$P_i = \frac{\pi_i \overline{t_i}}{\sum_{j=1}^{3} \pi_j \overline{t_j}}, \quad i = 1,2,3, \tag{3}$$

where π_i represents the stationary probability of the state i, which can be obtained by solving the following equations:

$$\begin{aligned} \pi_i &= \sum_{j=1}^{3} \pi_j P_{ij}, \quad j = 1,2,3, \\ 1 &= \sum_{j=1}^{3} \pi_j, \end{aligned} \tag{4}$$

where P_{ij} is the transition probability from the state i to the state j. According to the state transition model, the state transition probability matrix H can be directly given by

$$H = \left(P_{ij}\right) = \begin{pmatrix} 0 & P_{12} & P_{13} \\ P_{21} & 0 & P_{23} \\ 0 & P_{32} & 0 \end{pmatrix}. \tag{5}$$

Utilizing the time distribution T_{ij} from the state i to the state j, the state transition probability can be derived. Since there is only one event for the FBS moving away from state 3, we have $P_{32} = 1$. The event for the FBS exiting from state 2 is that the received signal energy strength whether satisfies the threshold level within the timer range T_0. Therefore, the transition probability P_{23} is

$$\begin{aligned} P_{23} &= \int_0^{\infty} m_{n_T}(t) P(T_{23} > t) dt \\ &= \int_0^{\infty} \lambda_m e^{-\lambda_m t} U(T_0 - t) dt = 1 - e^{-\lambda_m T_0}. \end{aligned} \tag{6}$$

Due to $P_{21} + P_{23} = 1$, then we have $P_{21} = 1 - P_{23} = e^{-\lambda_m T_0}$. Similarly, the event for the FBS exiting from the state 1 to state 2 is that the user initiates the session in the virtual cell coverage excluding the target FM coverage and the received signal energy strength meets the predefined threshold. Hence, the transition probability P_{12} can be obtained as

$$P_{12} = \frac{\lambda_c}{(\lambda_c + \lambda_m)} = 1 - \frac{\lambda_m}{\lambda_c}. \tag{7}$$

Due to $P_{12} + P_{13} = 1$, then we have $P_{13} = 1 - P_{12} = \lambda_m / \lambda_c$. From (3)–(7), we can derive the stationary probabilities given as follows:

$$\begin{aligned} \pi_1 &= \frac{P_{21}}{(2 + P_{21} P_{13})} = \frac{\lambda_c e^{-\lambda_m T_0}}{2\lambda_c + \lambda_m e^{-\lambda_m T_0}}, \\ \pi_2 &= \frac{1}{(2 + P_{21} P_{13})} = \frac{\lambda_c}{(2\lambda_c + \lambda_m e^{-\lambda_m T_0})}, \\ \pi_3 &= \frac{\lambda_m e^{-\lambda_m T_0} + \lambda_c \left(1 - e^{-\lambda_m T_0}\right)}{2\lambda_c + \lambda_m e^{-\lambda_m T_0}}. \end{aligned} \tag{8}$$

The mean residence time $\overline{t_1}$ and $\overline{t_2}$ can be easily derived as:

$$\overline{t_1} = \frac{1}{\lambda_c}, \tag{9}$$

$$\overline{t_2} = E[t_2] = E[\min\{T_{21}, T_{23}\}] = \frac{\left(1 - e^{-\lambda_m T_0}\right)}{\lambda_m}. \tag{10}$$

In state 3, the time duration is a busy period following $M/M/\infty$ queuing model, and thus the mean residence time $\overline{t_3}$ can be given by [17]

$$\overline{t_3} = \frac{\left(1 - e^{-\lambda_m / \rho}\right)}{\lambda_m}. \tag{11}$$

According to the stationary probability π_k and the mean residence time $\overline{t_k}$, the steady state probability P_k in each state can be obtained based on (3), and we have

$$\begin{aligned} P_1 &= \frac{\lambda_m e^{-\lambda_m T_0}}{2\lambda_m e^{-\lambda_m T_0} - \lambda_m e^{-\lambda_m (T_0 + 1/\rho)} + \lambda_c (1 - e^{-\lambda_m T_0})(1 - e^{-\lambda_m / \rho})}, \\ P_2 &= \frac{\lambda_c \left(1 - e^{-\lambda_m T_0}\right)}{2\lambda_m e^{-\lambda_m T_0} - \lambda_m e^{-\lambda_m (T_0 + 1/\rho)} + \lambda_c (1 - e^{-\lambda_m T_0})(2 - e^{-\lambda_m / \rho})}, \\ P_3 &= \frac{1 - e^{-\lambda_m / \rho}}{2 - e^{-\lambda_m / \rho}}. \end{aligned} \tag{12}$$

The partial derivative of the steady state probabilities with respect to the timer range T_0 are, respectively,

$$\frac{\partial P_1}{\partial T_0} = \frac{\lambda_m e^{-\lambda_m T_0}\left(e^{-\lambda_m T_0} - 1\right)}{\lambda_m e^{-\lambda_m T_0}\left(2 - e^{-\lambda_m/\rho}\right) + \lambda_c\left(1 - e^{-\lambda_m T_0}\right)\left(e^{-\lambda_m/\rho}\right)} < 0,$$

$$\frac{\partial P_2}{\partial T_0} = \frac{\lambda_m \lambda_c e^{-\lambda_m T_0}\left(2 - e^{-\lambda_m T_0}\right)}{\lambda_m e^{-\lambda_m T_0}\left(2 - e^{-\lambda_m/\rho}\right) + \lambda_c\left(1 - e^{-\lambda_m T_0}\right)\left(e^{-\lambda_m/\rho}\right)} > 0,$$

$$\frac{\partial P_3}{\partial T_0} = 0. \tag{13}$$

Therefore, the steady state probability P_1 and P_2 are the monotone decreasing function and the monotone increasing function with the increasing of T_0 respectively, while the steady state probability P_3 has nothing relationship with the timer range T_0. To simplify the description, we define:

$$f(\lambda_c, T_0) = 2\lambda_m e^{-\lambda_m T_0} - \lambda_m e^{-\lambda_m (T_0 + (1/\rho))} + \lambda_c\left(1 - e^{-\lambda_m T_0}\right)\left(2 - e^{-\lambda_m/\rho}\right). \tag{14}$$

In the subsequent, the analytic formulas of the energy consumption and average cumulative delay and interference time are expressed as:

$$E = \frac{\lambda_m e^{-\lambda_m T_0} E_1 + \lambda_c\left(1 - e^{-\lambda_m T_0}\right) E_2}{f(\lambda_m, \lambda_c, \rho, T_0)} + \frac{1 - e^{-\lambda_m/\rho}}{2 - e^{-\lambda_m/\rho}} E_3, \tag{15}$$

$$D_{\text{ave}} = \frac{\lambda_m^2 e^{-\lambda_m T_0}(D_{12} + D_{23}) + \lambda_c^2\left(1 - e^{-\lambda_m T_0}\right)^2 D_{23}}{\lambda_m^2 e^{-\lambda_m T_0} + \lambda_c^2 (1 - e^{-\lambda_m T_0})^2 + \lambda_c (1 - e^{-\lambda_m/\rho})^2 f(\lambda_m, \lambda_c, \rho, T_0)/(2 - e^{-\lambda_m/\rho})},$$

$$T_{\text{int}} = \frac{1 - e^{-\lambda_m/\rho}}{2 - e^{-\lambda_m/\rho}} T_H. \tag{16}$$

3.3. Analysis of the Compared Scheme. In this subsection, we analyze the existing prominent schemes theoretically and make a comparison among these schemes. IEEE 802.16 m [11] proposes a simple and practical scheme, called as periodical discontinuous transmission (PDTX), in which the FBS without active user changes alternately between available interval (AI) and unavailable interval (UAI). During the UAI, the FBS switches off the pilot transmission and the related processing. During the AI, the FBS becomes active on the air interface for synchronization and signaling purposes. In PDTX, Let τ be the low duty cycle in each period between AI and transmission period, the energy consumption and the cumulative interference time can be derived as

$$\begin{aligned} E_{\text{PDTX}} &= (1 - P_3)(1 - \tau)E_1 + (1 - P_3)\tau(E_3 - E_s) \\ &\quad + P_3(E_3 - E_s) \\ &= \tau(E_3 - E_1) + (1 - \tau)[P_3(E_3 - E_s - E_1) + E_1] \\ &= \tau(E_3 - E_1) + (1 - \tau) \\ &\quad \times \left[\frac{\left(1 - e^{-\lambda_m/\rho}\right)(1 - \tau)}{2 - e^{-\lambda_m/\rho}}(E_3 - E_s - E_1) + E_1\right], \end{aligned} \tag{17}$$

$$\begin{aligned} T_{\text{PDTX}} &= P_3 T_H + (1 - P_3)\tau T_H = [P_3(1 - \tau) + \tau] T_H \\ &= \left[\frac{\left(1 - e^{-\lambda_m/\rho}\right)(1 - \tau)}{2 - e^{-\lambda_m/\rho}} + \tau\right] T_H. \end{aligned} \tag{18}$$

Widiarti et al. [12] present a novel sleep mode for the FBS based on a grouping methodology, called as GPDTX. This scheme is closely related to the user location. The FBS in the absence of users switches off the pilot transmission, while the FBS with a nonactive user periodically transmit the pilot signal. Suppose that the user lies indoor with a probability P_{in}, and $P_{\text{out}} = 1 - P_{\text{in}}$. The energy consumption and the cumulative interference time can be derived as

$$\begin{aligned} E_{\text{GPDTX}} &= \left[\frac{P_{\text{in}}(1 - P_3)^2 \tau}{P_{\text{out}} + P_{\text{in}}(1 - P_3)} + P_3\right](E_3 - E_s) \\ &\quad + \left[\frac{P_{\text{in}}(1 - P_3)^2(1 - \tau)}{P_{\text{out}} + P_{\text{in}}(1 - P_3)} + \frac{P_{\text{out}}(1 - P_3)}{P_{\text{out}} + P_{\text{in}}(1 - P_3)}\right] E_1, \\ T_{\text{GPDTX}} &= \frac{P_{\text{in}}(1 - P_3)^2 \tau}{P_{\text{out}} + P_{\text{in}}(1 - P_3)} T_H + P_3 T_H. \end{aligned} \tag{19}$$

The partial derivatives of the energy consumption and interference time with respect to P_{in} are

$$\frac{\partial E_{\text{GPDTX}}}{\partial P_{\text{in}}} = \frac{(1 - P_3)^2 \tau (E_3 - E_s - E_1)}{(1 - P_3 P_{\text{in}})^2} > 0,$$

$$\frac{\partial T_{\text{GPDTX}}}{\partial P_{\text{in}}} = \frac{P_{\text{in}}(1 - P_3)^2 \tau}{(1 - P_3 P_{\text{in}})^2} T_H > 0. \tag{20}$$

Therefore, the energy consumption and cumulative interference time are the monotone increasing function with the increasing of P_{in}, respectively. It means the energy

efficiency of this scheme is directly relative to P_{in}. When $P_{\text{out}} = 0$, we have

$$E_{\text{GPDTX}} = [(1 - P_3)\tau + P_3](E_3 - E_s) + (1 - P_3)(1 - \tau)E_1, \quad (21)$$

$$T_{\text{GPDTX}} = (1 - P_3)\tau T_H + P_3 T_H. \quad (22)$$

Compare (17) and (21), (18) and (22), they are the same with each other. It implies that when the users stay at home all the time, GPDTX has the same energy-saving effect as PDTX. However, the signaling overhead of GPDTX is more serious than that of PDTX.

Reference [13] investigates the behavior of the sleep mode for the FBS based on user activity detection, called UDTX. The FBS switches off the pilot transmission until it detects the active user in its coverage. After completing the service for the active call, it immediately moves into a sleep mode. Hence, the energy consumption and interference time can be derived as

$$E_{\text{UDTX}} = (1 - P_3)E_1 + P_3 E_3 = E_1 + (E_3 - E_1)\frac{1 - e^{-\lambda_m/\rho}}{2 - e^{-\lambda_m/\rho}},$$
$$T_{\text{UDTX}} = P_3 T_H. \quad (23)$$

4. Performance Evaluation

In order to evaluate the performance of the proposed scheme, Monte Carlo simulation analysis is employed. A dual-strip model is chosen to simulate the multifloor based 3D femtocell cellular networks [15]. In this model, each stripe has 2 by 10 apartments and 6 floors. Each apartment is of size 10 m × 10 m. There is a street between the two stripes of apartments, with a width of 10 m. Each femto block is of 120 m × 70 m. Two critical parameters are introduced: the deployment rate ω and the activation rate ν. The deployment rate is used to simulate that each apartment has a FBS with a probability ω. The activation rate is defined as the active proportion of FBSs. If the FBS is active, it will transmit the user data in the traffic channel, or else it will transmit the pilot signal in the control channel. These two parameters determine operational femtocell density μ, that is, $\mu = \omega\nu$. Assume that femto users (FUEs) are randomly distributed on each floor of the building, and the session arrival of each FUE in the virtual cell coverage and the FM coverage follows Poisson distribution with the rate λ_c and λ_m.

The channel propagation model [15] includes path loss and log-normal shadowing. The path loss consists of free space loss, indoor path loss, indoor wall penetration loss, floor penetration loss, and outdoor wall penetration loss. If the FBS and the FUE (or another FBS) are inside the different apartments, the path loss is expressed as (24), or else (25). Consider

$$PL(\text{dB}) = \max\left(2.7 + 42.8\log_{10}R, 38.46 + 20\log_{10}R\right) + L_{ow,1} + 18.3 n_0^{((n_0+2)/(n_0+1)-0.46)} + q \times L_{iw} + L_{ow,2}0.7d_2, \quad (24)$$

$$PL(\text{dB}) = 38.46 + 20\log_{10}R + 0.7d_{2D,\text{indoor}} + 18.3\, n_0^{((n_0+2)/(n_0+1)-0.46)} + q \times L_{iw}, \quad (25)$$

where R denotes the distance between the FBS and the FUE (or another FBS), and $d_{2D,\text{indoor}}$ takes account of penetration loss due to walls inside an apartment. n_0 is the number of penetrated floors, q represents the number of walls separating apartments between the FBS and the FUE (or another FBS). L_{iw} is the penetration loss of the wall separating apartments, and $L_{ow,1}$ and $L_{ow,2}$ represent the penetration losses of outdoor wall for two houses, respectively.

Table 1: Simulation parameter.

Parameter	Value
System parameters	
Carrier frequency/bandwidth	2 GHz/10 MHz
Subcarrier bandwidth	15 kHz
Log-normal shadowing standard deviation	8 dB
Macrocell parameters	
Antenna pattern	Omni-directional
Macrocell radius	288 m
MBS height	32 m
Maximum MBS TX power	46 dBm
Femtocell parameters	
Maximum FBS TX power	20 dBm
User per femtocell	1
Deployment rate ω	80%
Activation rate ν	50%
Session arrival rate λ_m, $\lambda_c = 2\lambda_m$	5/h, 10/h, 20/h
The average service rate ρ	100/h

In addition, the simulation parameters are mostly referenced from [15] as summarized in Table 1. The simulation analysis is divided into the following two aspects. Firstly, with realistic LTE system parameters and 3D femtocell network model, we verify the rationality of the theoretical analysis and investigate the impact of the key parameters of the proposed scheme on the performance of femtocell networks. Secondly, we assess the effectiveness of the proposed scheme by comparing its performance with other three-prominent schemes.

4.1. Parameter Analysis of the Proposed Scheme. Figure 5 depicts steady state probability of the proposed scheme according to the timer range of the FBS. The close agreement between theoretical and numerical simulations indicates that analytic expressions (12) are the accurate approximations

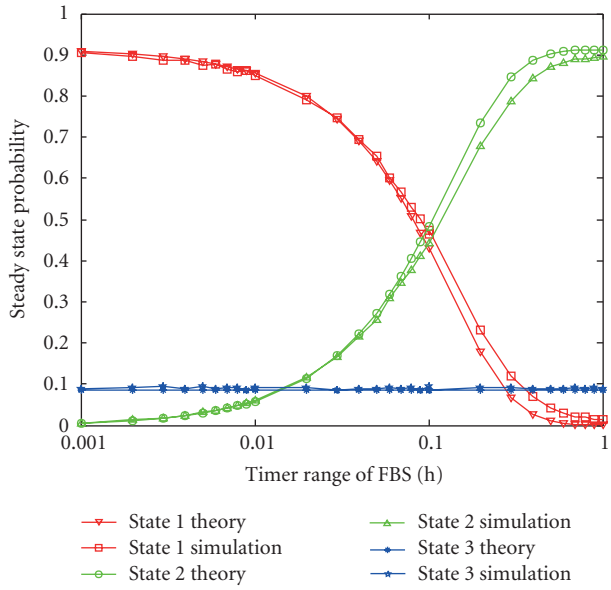


Figure 5: Steady state probability for both theory and simulation.

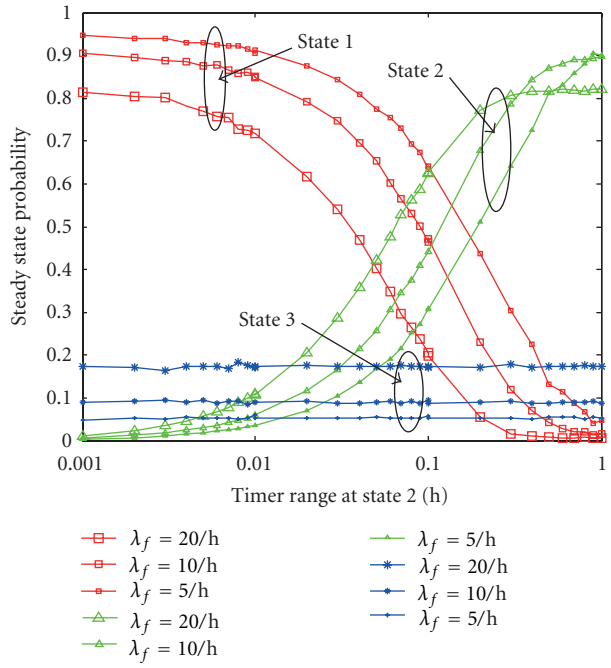


Figure 6: Steady state probability for numerical analysis.

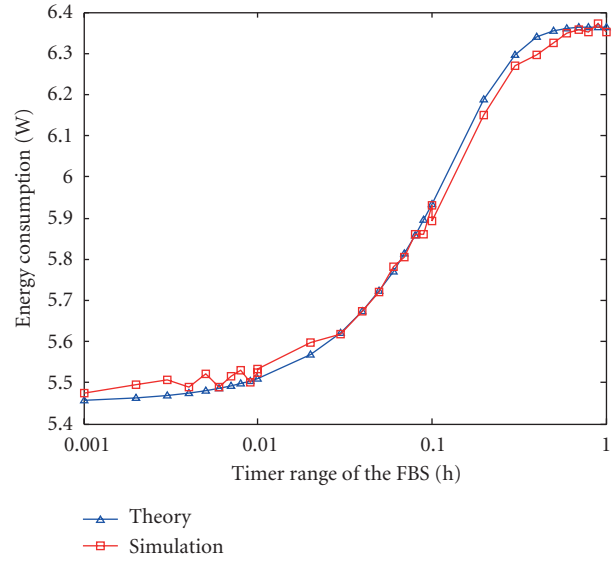


Figure 7: Energy consumption for both theory and simulation.

for practical steady state probability in the dense femtocell environments. Moreover, it can be seen that as the timer range of the FBS increases, the steady state probability of the state 1 decreases, the steady state probability of the state 2 increases, and the steady state probability of the state 3 has no correlation with the timer range and remains unchanged. That is also in accordance with the theoretical analysis.

Figure 6 shows the plots of the steady state probabilities of the proposed scheme from the simulation analysis. The steady state probabilities are compared with different session arrival rate. Moreover, for a fixed timer range of the FBS, as the session arrival rate increases, the steady state probability of state 1 decreases, while the steady state probabilities of both state 2 and state 3 increase. That is mainly due to the fact that the higher the session arrival rate, the greater the probability of the activation for the FM.

Figure 7 shows the energy consumption of the proposed scheme according to the timer range of the FBS. It can be seen that the theoretical analysis and numerical simulations get matched pretty well, which indicates that the analytic expression (15) is the accurate approximation for the practical energy consumption in dense femtocell environments. Moreover, we also can know that as the timer range of the FBS increases, the energy consumption increases. That is mainly due to the fact that the longer the timer range is, the greater the steady state probability of the state 3 is, which consumes the most energy of the all states.

Figure 8 depicts the relation between the energy consumption of simulation analysis and the timer range, respectively. For the same value of timer range, as the session arrival rate increases, the energy consumption increases. For fixed session arrival rate, the energy consumption increases with the increasing of the timer range.

Figure 9 depicts the average cumulative delay of the proposed scheme according to the timer range. It can be seen that as the timer range increases, the average cumulative delay decreases. That is mainly due to the fact that the greater the timer range, the more the session arrival will directly enter into the state 3, which will avoid the delay from state 1 to state 2. In addition, as the session arrival rate increases, the average cumulative delay decreases.

From the analysis above, we can conclude that the timer range has a significant effect on the energy consumption and the average cumulative delay and cumulative interference time. Hence, we should choose the proper timer range for the energy efficient FBS. Furthermore, as the timer range of

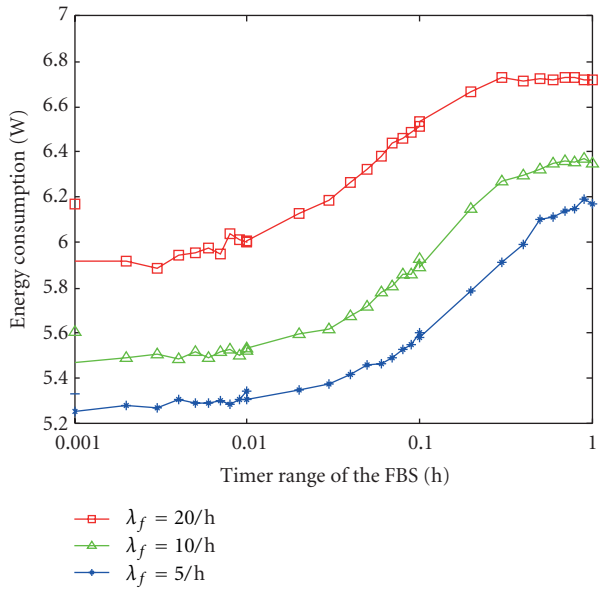


Figure 8: Energy consumption of simulation analysis versus timer range.

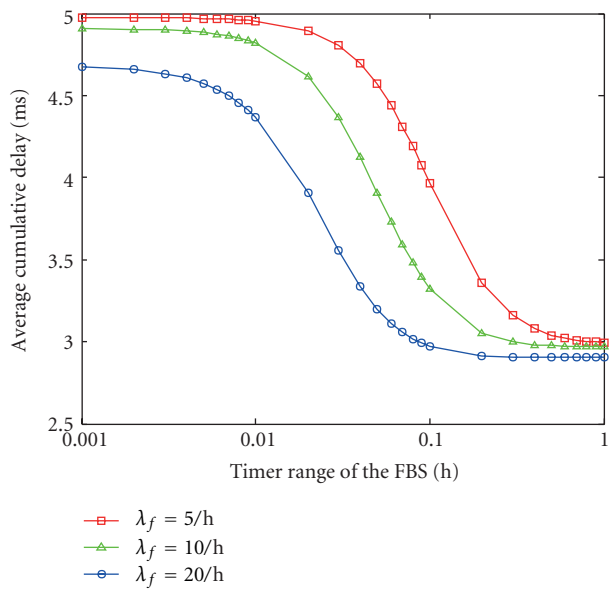


Figure 9: Average cumulative delay versus timer range.

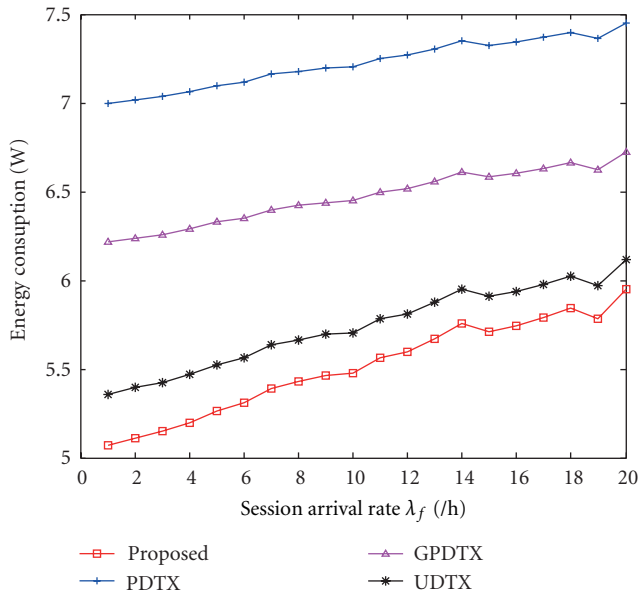


Figure 10: Energy consumption for different schemes.

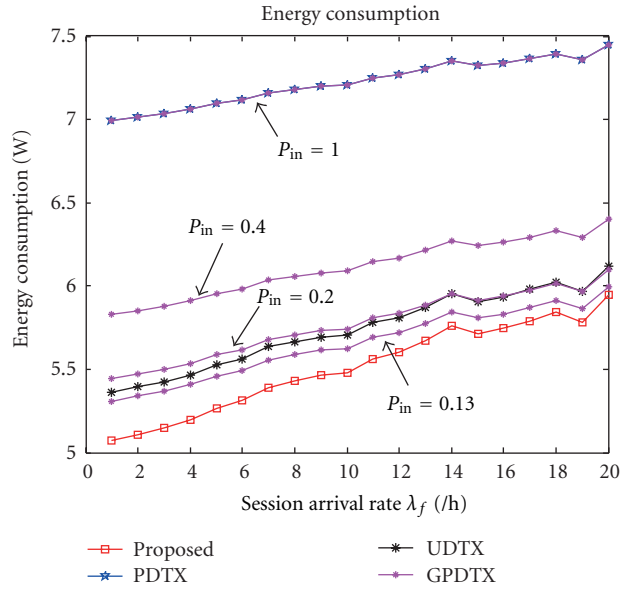


Figure 11: Energy consumption for different schemes with different P_{in}.

the FBS increases, curve trends of energy consumption and average cumulative delay are just contrary. Hence, we should elect the proper operation to balance them.

4.2. Performance Comparison with Different Schemes. Figure 10 and Figure 11 depict the energy consumption of different schemes with various session arrival rates. It is assumed that low duty cycle τ is 0.4 for two figures, and $P_{in} = 0.6$ for Figure 10. From these two figures, it can be seen that the energy consumption of all these schemes increases with the increasing of the session arrival rate. From Figure 10, for a fixed session arrival rate, the proposed scheme costs the lowest energy consumption of all schemes, while the traditional PDTX scheme spends the greatest energy consumption of all schemes. The energy consumption of femtocells using GPDTX is higher than using UDTX.

However, it can be seen from Figure 11 that the energy consumption of GPDTX decreases with the increasing of the probability P_{in}. When $P_{in} = 1$, the energy consumption curve of GPDTX is almost coincided with that of PDTX. When $P_{in} = 0.2$, the energy consumption curve of GPDTX is almost coincided with that of UDTX. When $P_{in} = 0.13$, the energy

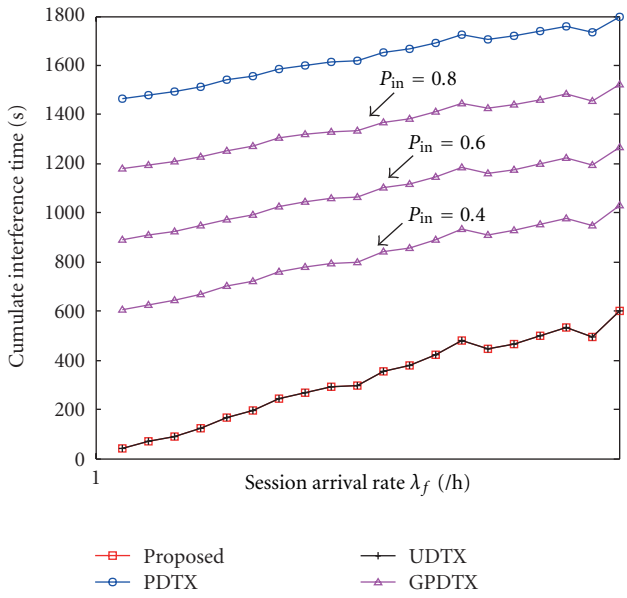


Figure 12: Interference time for different schemes with different P_{in}.

consumption of UDTX is higher than that of GPDTX. The reason is that when P_{in} is small, more FMs in the GPDTX scheme than in UDTX can completely switch off the pilot transmission due to the FBS grouping mechanism; when P_{in} is high, the UDTX scheme does not need to transmit the pilot signal with the nonactive user, but GPDTX scheme cannot do this. Finally, though our scheme is best in terms of the energy saving, it is achieved at the expense of more signaling overhead due to FBSs cooperation within each virtual cell.

Figure 12 shows the cumulative interference time for different scheme according to the session arrival rate. It can be seen that as the session arrival rate increases, the cumulative interference time for various schemes increase. For a fixed session arrival rate, the cumulative interference time using PDTX is the worst of all, while the curves of the proposed scheme and UDTX are nearly coincided with each other, and best of all. The cumulative interference time using GPDTX decreases with the decreasing of P_{in}.

5. Conclusion

This work proposes a novel hybrid BS-cooperative power management scheme for self-organized sleep mode in virtual cell-based femto networks. Firstly, according to the femtocell network topology, HBCPM builds a leader-member virtual cell framework. Within each virtual cell, the FBS is elected as either a FL or a FM based on the list degree and the sum of received pilot signal power. Moreover, because of the time variation characteristic of the femtocell network topology, the role assignment algorithm for the newly installed FBS is introduced. Secondly, the FL is responsible for detecting active calls in the virtual cell coverage, while the FMs in the absence of the active user can completely switch off the pilot transmission and the associated processing all the time. Thirdly, utilizing the semi-Markov process, state transition model of HBCPM is established, and then the analytic formulas of energy consumption and average cumulative delay and interference time are derived as well as three-prominent related schemes. Simulation results verify that the proposed scheme outperforms three existing schemes in terms of interference time and energy consumption. Further, the timer range of the proposed scheme has a significant effect on the energy consumption and the average delay.

In addition, this scheme can be easily applied to orthogonal frequency division multiple access (OFDMA) networks [18, 19], or orthogonal frequency and code division multiplexing (OFCDM) systems [20]. In the future, we plan to incorporate the user mobility model to study its effect on the performance of the proposed scheme and evaluate the energy consumption with the prominent related schemes. Further, we also plan to design and implement the FBS cooperative power management algorithm in official blocks.

Acknowledgments

This work is supported by the National S&T Major Project of China (no. 2010ZX03003-001-01, 2011ZX03003-002-01), and the Cobuilding Project of Beijing Municipal Education Commission, the National Natural Science Foundation of China (61101109), and the Youth Research and Innovation Project of Beijing University of Posts and Telecommunications.

References

[1] J. G. Andrews, H. Claussen, M. Dohler et al., "Femtocell past present and future," *IEEE Journal on Selected Areas in Communications*, vol. 30, pp. 497–508, 2012.

[2] W. C. Cheung, T. Q. S. Quek, and M. Kountouris, "Throughput optimization, spectrum allocation, and access control in two-tier femtocell networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, pp. 561–574, 2012.

[3] J. H. Yun and K. G. Shin, "Adaptive interference management of OFDMA femtocells for co-channel deployment," *IEEE Journal on Selected Areas in Communications*, vol. 29, pp. 1225–1241, 2011.

[4] J. W. Huang and V. Krishnamurthy, "Cognitive base stations in LTE/3GPP femtocells: a correlated equilibrium game-theoretic approach," *IEEE Transactions on Communications*, vol. 59, pp. 3485–3493, 2011.

[5] S. Carlaw and C. Wheelock, "Femtocell market challenges and opportunities," *ABI Research, Research Report*, vol. 23, 2007.

[6] Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green cellular networks: a survey, some research issues and challenges," *IEEE Communications Surveys & Tutorials*, vol. 13, pp. 524–540, 2011.

[7] C. Khirallah and J. S. Thompson, "Energy and cost impacts of relay and femtocell deployments in long-term-evolution advanced," *IET Communications*, vol. 5, pp. 2617–2628, 2011.

[8] W. Cheng, H. Zhang, L. Zhao, and Y. Li, "Energy efficient spectrum allocation for green radio in two-tier cellular networks," in *Proceedings of the 53rd IEEE Global Communications Conference (GLOBECOM '10)*, pp. 1–5, December 2010.

[9] A. D. Domenico, S. Emilio Calvanese, and D. Andrzej, "Ghost femtocells: an energy-efficient radio resource management

scheme for two-tier cellular networks," in *Proceedings of the 11th European Wireless Conference, Sustainable Wireless Technologies*, pp. 1–8, 2011.

[10] 3GPP, "TS36. 300 v11. 1. 0-E-UTRA and E-UTRAN Overall description Stage 2," 2012, http://www.3gpp.org/.

[11] IEEE P802.16m, "Part 16: Air Interface For Fixed And Mobile Broadband Wireless Access Systems-Advanced Air Interface," 2010.

[12] H. Widiarti, S. Y. Pyun, and D. H. Cho, "Interference mitigation based on femtocells grouping in low duty operation," in *Proceedings of the IEEE 72nd Vehicular Technology Conference Fall (VTC-Fall '10)*, pp. 1–5, September 2010.

[13] I. Ashraf, L. T. W. Ho, and H. Claussen, "Improving energy efficiency of femtocell base stations via user activity detection," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '10)*, pp. 1–5, April 2010.

[14] 3GPP, "TS 22. 220 v11. 4. 0 Service requirements for Home Node B, (HNB) and Home eNode B, (HeNB)," 2011, http://www.3gpp.org/.

[15] 3GPP, "TR 36. 814 v9. 0. 0-Further advancements for E-UTRAN physical layer aspects," 2010, http://www.3gpp.org/.

[16] O. C. lbe, *Markov Processes for Stochastic Modeling*, Academic Press, 2009.

[17] D. Gross, J. F Shortle, J. M. Thompson, and C. M. Harris, *Fundamentals of Queueing Theory*, Wiley, 2008.

[18] H. Zhu and J. Wang, "Chunk-based resource allocation in OFDMA systems—part I: chunk allocation," *IEEE Transactions on Communications*, vol. 57, no. 9, pp. 2734–2744, 2009.

[19] Z. Huiling and W. Jiangzhou, "Chunk-based resource allocation in OFDMA systems—part II: joint chunk, power and bit allocation," *IEEE Transactions on Communications*, vol. 60, pp. 499–509, 2009.

[20] Y. Zhou, J. Wang, and M. Sawahashi, "Downlink transmission of broadband OFCDM systems—part I: hybrid detection," *IEEE Transactions on Communications*, vol. 53, no. 4, pp. 718–729, 2005.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 672436, 15 pages
doi:10.1155/2012/672436

Research Article

Self-Organized Connectivity Control and Optimization Subjected to Dispersion of Mobile Ad Hoc Sensor Networks

Zhenqiang Mi, Yang Yang, and Hao Ding

School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China

Correspondence should be addressed to Zhenqiang Mi, mizq@ustb.edu.cn

Received 3 May 2012; Revised 11 July 2012; Accepted 29 July 2012

Academic Editor: Zhiguo Ding

This paper addresses the problem of the connectivity control and the self-organized deployment/dispersion of a team of mobile ad hoc sensor nodes. First, to reduce redundant communication links while preserving global connectivity, a distributed link removal algorithm is developed that only requires local information of no more than two-hop neighbors. Secondly, for the purpose of preserving essential links while avoiding collisions, a combined piecewise-continuous motion controller is designed to regulate the motion of mobile sensors between two consecutive switches. The proposed hybrid control system can autonomously disperse a team of mobile sensors towards their final configuration with guaranteed connectivity and collision avoidance. Theoretical analysis and computer simulations have confirmed the efficiency and scalability of the proposed schemes.

1. Introduction

The self-organized connectivity control and dispersion of mobile ad hoc sensor networks (MASNs) has been extensively investigated due to its promising potential applications in various fields, such as remote supervision of hazard substances, exploration of unknown fields, and cooperative sensing. Enhancing the coverage of the MASNs could be beneficial, if not critical, for a variety of missions, such as environmental monitoring and disaster management. Moreover, from the wireless communication point of view, the sparse network structure resulted from the enhanced coverage can effectively reduce radio interferences, which is critical for the elimination of excessive message overheads and the reduction of the latency [1]. Furthermore, in energy critical wireless sensor networks, the reduction of message overheads as well as the complexity of the self-organizing algorithms can extend the service lifetime of the systems.

In a mobile sensor network, the mobility of sensor nodes provides the possibility for the nodes to deploy to a configuration with desired properties from an arbitrary initial distribution in a self-organized manner. The aforementioned concept of MASNs has attracted numerous research efforts. Among various topics that have been covered, emphasis is placed on the *consensus* of a group of mobile sensors, such as *flocking* and *rendez-vous*, with a particular interest in coverage and connectivity control. *Self-organized dispersion*, on the other hand, is another fundamentally essential aspect of the system and requires more research efforts.

Self-organized dispersion of MASNs can be loosely defined as maximizing coverage with a minimum number of mobile sensors [2]. Significant application potentials of self-organized dispersion have recently led to a surge of research attentions. Techniques including inverse agreement control law [3], clique-intensity algorithm [2], and "artificial physics" frame-work [4] have been developed to regulate the evolution of the underlying networks. Coverage control of mobile networks is another promising research area that is closely related to self-organized dispersion, and numerous efforts have concentrated on the deployment of mobile sensors [5–8], coverage control in stationary WASN [9], and coverage control of autonomous agents [10]. Commonly used models include Voronoi diagrams [7, 8] and potential functions [10, 11]. Voronoi diagrams method can partition the field into many subareas dedicated to each mobile sensor, allowing sensors to move to maximize coverage in its own subarea. However, the assumption that mobile sensors can easily detect most of its Voronoi neighbors through local

communication may not be satisfied in a real network, due to the limited communication range of the mobile sensors, which may not be sufficient for covering all Voronoi neighbors. Potential function, on the other hand, is able to fulfill the self-organized coverage control of mobile sensor nodes in a more local and distributed manner. The concept of this approach is to imitate the behavior of electromagnetic particles: whenever two electromagnetic particles are too close in proximity, a repulsive force pushes them apart. In mobile ad hoc sensor networks, this method can help move sensors from high density to low density areas, thereby dispersing the team to improving the overall network coverage.

Connectivity control of MASNs is rapidly becoming a hot research topic in the field of multiagent systems, and various strategies have been developed in recent years, including both centralized [12, 13] and decentralized [14–20] approaches. A common way to maintain connectivity in mobile networks is to shrink the communication links between any two neighbors whenever they tend to break [21]. Most studies have attempted to maximize *Fiedler value* [13] and to add communication links to increase connectivity [22]. However, these approaches often result in a tight network structure with dense communication links, which may largely restrict the mobility of mobile sensors and jeopardize the coverage and cooperation efficiency. More importantly, from the wireless communication point of view, high network density can severely aggravate radio interferences [1]. To deal with these problems, Zavlanos and Pappas [15] proposed a distributed market-based control strategy, which is able to reduce redundant communication links based on local estimation of *spanning subgraph*. This strategy, however, requires full knowledge of the network structure, which may cause large delay in dense and large-scale networks [23]. In [17], a connectivity control strategy is presented with a particular consideration of the relationship between the communication range and the sensing range, that is, radio range is at least twice the sensing range. The proposed control scheme can guaranteed a connected network with optimized sensing coverage. How to disperse the networked sensors to achieve a sparse topology is not yet discussed. It is also worthwhile to mention that certain topology control methods [24, 25] in ad hoc senor networks have been developed that similarly deal with the issue of removing redundant communication links. However, these works are mainly based on stationary network structure. The important issues of combing global connectivity and dispersion of mobile sensors with respect to mobility control strategy have not been addressed properly.

Contributions. Aiming at bridging the gap between dispersion and connectivity, we first present a distributed link removal algorithm (DLRA) to reduce redundant communication links while preserving global connectivity. The proposed algorithm is fully distributed and only requires local information of no more than two-hop neighbors. Then, by integrating DLRA with a novel combined piecewise-continuous potential function, a distributed connectivity control system is developed to disperse a team of mobile sensors with guaranteed connectivity and collisions avoidance. The present work provides answers to the following questions.

(1) How to remove communication links with respect to global connectivity based only on local information of neighbor status?

(2) How to disperse a team of sensors with connectivity preservation, collision avoidance, and limited actuation?

(3) Is the generated network structure sparse enough to fulfill dispersion requirement?

The rest of the paper is organized as follows: Section 2 provides necessary terminologies and notations. Section 3 presents the DLRA algorithm that is developed to address the link removal problem with respect to global connectivity. Section 4 introduces a distributed combined piecewise-continuous potential function to control the motion of mobile sensors. Computer simulations are included in Section 5, and this paper is concluded in Section 6 with a discussion of future research works.

2. Problem Formulation

Let the dynamic graph $\mathcal{G}(t) = (\mathcal{V}, E(t))$ denote a mobile network of n mobile sensor nodes with integrated wireless communication capabilities, where $\mathcal{V} = (1, \ldots, n)$ denotes the set of vertices indexed by the set of mobile sensors and $E(t) = \{i, j \mid f_{ij}(t) \geq \varepsilon,\ i, j \in \mathcal{V}\}$, $0 < \varepsilon < 1$ denotes the time-variant set of communication links. We define $0 \leq f_{ij}(t) \leq 1$ to be a normalized nonnegative weighting function symmetric in its arguments, that is, $f_{ij}(t) = f_{ji}(t)$, and assume that $f_{ij}(t) \neq f_{ik}(t)$, $j \neq k$. In this case, the adjacency matrix $\mathcal{A}(t) = (a_{ij}(t)) \in \mathbb{R}^{n\times n}$ (we define $a_{ii}(t) \equiv 0$ for all i, thus there are no self-loops in the network) can be defined as follows:

$$a_{ij}(t) = \begin{cases} a_{ji}(t) = f_{ij}(t), & (i, j) \in E(t), \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Note that *any normalized nonnegative function* can be treated as a weighting function. Nevertheless, to associate with the link quality, it is rather a natural choice that $f_{ij}(t)$ denotes received signal strength. Moreover, as indicated in radio propagation theory [1], in obstacle-free environment, signal fading can be treated as an exponential decay function of the Euclidean distance between i and j. In particular, we have the following definitions.

Definition 1. For any node i in $\mathcal{G}(t) = (\mathcal{V}, E(t))$, the normalized received signal strength from node j is set to be:

$$f_{ij}(t) \triangleq \begin{cases} 1, & \left\|\xi_i(t) - \xi_j(t)\right\| < a_0, \\ \exp\left[-\beta\left(\left\|\xi_i(t) - \xi_j(t)\right\| - a_0\right)\right], & \text{otherwise,} \end{cases} \tag{2}$$

where β (the value of β depends on the radio frequency, antenna gain, system loss, and environmental factors, etc.)

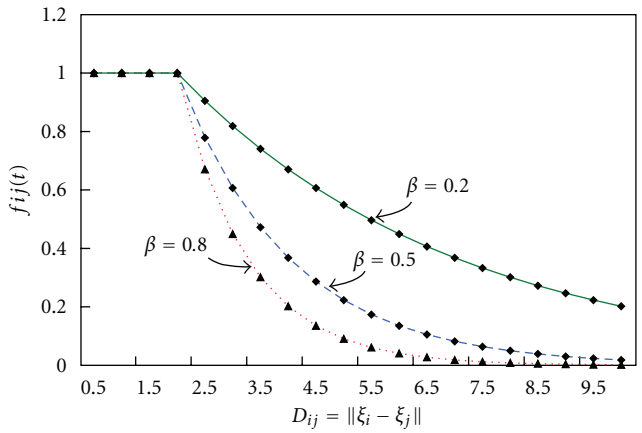


FIGURE 1: Modeling normalized receive signal fading with Euclidean distance, where $\mathcal{D} = \|\xi_i(t) - \xi_j(t)\|$, and $a_0 = 2$.

is positive path-loss value, and a_0 is reference distance (see Figure 1).

Definition 2. An undirected weighted dynamic graph $\mathcal{G}_s(t) = (\mathcal{V}, E(t))$ is defined as connected at time t if and only if there is at least one communicative path between any two vertices within it.

Furthermore, consider the vertices in $\mathcal{G}_s(t)$ as n mobile sensors with single-integrator kinematic model given by the following:

$$\dot{\xi}_i(t) = u_i(t), \tag{3}$$

where $\xi_i(t) \in \mathbb{R}^N$ and $u_i(t) \in \mathbb{R}^N$ denotes the position and velocity of node i, respectively.

Now the main objectives of this paper can be described as follows:

(1) for any nodes i and j in $\mathcal{G}_s(t) = (\mathcal{V}, E(t))$, determine whether link (i, j) is an essential communication link for connectivity maintenance, and then remove redundant communication links. The generated subgraph $\mathcal{G}_s(t) = (\mathcal{V}, E_s(t))$ contains only essential communication links, that is, $(i, j) \in E_s(t)$;

(2) for the corresponding kinematic system $\dot{\xi}_i(t) = u_i(t)$, derive a dynamic motion controller $u_i(t) \in \mathbb{R}^N$ for all sensors i between any two consecutive switches, so that (1) for any $(i, j) \in E_s(t)$, it is guaranteed that $\varepsilon \le f_{ij}(t) < \varphi$; (2) for any $(i, j) \notin E_s(t)$, nodes i and j are dispersed into the set $\{i, j \mid f_{ij} \le \varepsilon, i, j \in \mathcal{V}\}$ (see Figure 2).

3. Distributed Link Removal Algorithm with Respect to Global Connectivity

The objective of this section is to develop a distributed local connectivity control algorithm, aiming at removing redundant communication links to facilitate the self-organized dispersion of mobile sensors with respect to global connectivity.

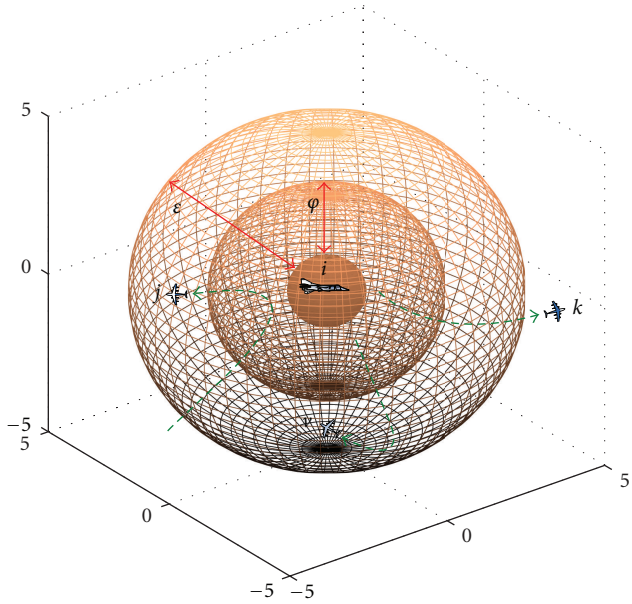


FIGURE 2: The trajectories of node j and v denote the possible movements where $(i, j), (i, v) \in E_s(t)$, and that of node k depicts the situation where $(i, k) \notin E_s(t)$, that is, be removed from $\mathcal{G}_s(t)$ and dispersed from each other.

3.1. Principle of DLRA. Let us first introduce the following definitions which categorize neighboring sensors into two distinctive neighbor relationship sets, namely, physical neighbor set and logical neighbor set.

Definition 3. Node j is in the physical neighbor set of i, denoted $j \in \mathcal{N}_p{}^{[i]}(t)$, if and only if $(i, j) \in E(t)$.

Definition 4. Node j is the logical neighbor set of i, denoted $j \in \mathcal{N}_l{}^{[i]}(t)$, if and only if $(i, j) \in E_s(t)$.

Definition 5. For any node i in the network, we denote $\mathcal{N}_r{}^{[i]}(t) = (a_{jk}(t))$ as the neighbor relationship matrix (NRM) at time t, where $j \in \mathcal{N}_l{}^{[i]}(t)$ and $k \in \mathcal{N}_l{}^{[j]}(t)$.

Based on the aforementioned definitions, the proposed DLRA is described as follows.

Information Exchange. The information that sensor i requires for the link removal algorithm is obtained by periodically receiving *Heartbeat* messages from all its neighbors in $\mathcal{N}_p{}^{[i]}(t)$. The periodical messages sent by neighbor j should at least contain the following information: node ID, position, logical neighbor set $\mathcal{N}_l{}^{[j]}(t)$, and link weight with all its logical neighbors. After exchanging information with all its neighbors, each node is only aware of the status of its one-hop physical neighbors and two-hop logical neighbors.

Selecting Candidate Links to Be Removed. Upon receiving *Heartbeat* message from each individual physical neighbor, node i updates $\mathcal{N}_p{}^{[i]}(t)$. As it is assumed that the initial

network is connected, DLRA starts at time $t = t_0$. Then, the process is started with initial condition that $\mathcal{N}_p{}^{[i]}(t_0) = \mathcal{N}_l{}^{[i]}(t_0)$, and $\mathcal{N}_r{}^{[i]}(t_0)$ is constructed for each i. The selection of candidate links is as follows.

Step 1. Look up neighbor relationship matrix $\mathcal{N}_r{}^{[i]}(t_0)$. If $\exists a_{kj}(t_0) \in \mathcal{N}_r{}^{[i]}(t_0)$, where $j \in \mathcal{N}_l{}^{[i]}(t)$, there exists a triangle consist of sensor i, j, k. If and only if $a_{ij}(t_0) = \max\{a_{ij}(t_0), a_{ik}(t_0), a_{jk}(t_0)\}$, then denote link (i, j) as candidate link to be removed from $\mathcal{N}_l{}^{[i]}(t)$, put sensor j into candidate removal set $\mathcal{C}_d{}^{[i]}(t)$, and remove corresponding row vector in $\mathcal{N}_r{}^{[i]}(t)$ (cf. line 2–7 in Algorithm 1).

Step 2. Update $\mathcal{N}_r{}^{[i]}(t)$, so that there is no triangle in the updated topology. Let $\mathcal{N}_r{}^{[ik]}(t) = [a_{jk}(t), a_{lk}(t), \ldots a_{nk}(t)]$ be the column vector of $\mathcal{N}_r{}^{[i]}(t)$, where $(j, l \ldots n) \in \mathcal{N}_l{}^{[i]}(t)$. Check $\mathcal{N}_r{}^{[ik]}(t)$ for every $k \in \mathcal{N}_r{}^{[i]}(t)$, if $\exists a_{jk}(t)$, and $a_{lk}(t)$, $j \neq l$ and $a_{jk}(t) \times a_{lk}(t) \neq 0$, compare link weights $a_{ij}(t)$, $a_{il}(t)$, $a_{jk}(t)$, and $a_{lk}(t)$, if $a_{ij}(t) = \max\{a_{ij}(t), a_{il}(t), a_{jk}(t), a_{lk}(t)\}$, then denote link (i, j) as candidate link to be removed from $\mathcal{N}_l{}^{[i]}(t)$, and put j into $\mathcal{C}_d{}^{[i]}(t)$ and remove corresponding row vector in $\mathcal{N}_r{}^{[i]}(t)$ (cf. line 8–15 in Algorithm 1 (Δt is the execution time for node i in every edge removal procedure, we assume Δt is constant for any node).

*Synchronized Local Coordination for DLRA (*Algorithm 2*).* A major challenge when dealing with dynamic network is that, the network topology may be changed during any consecutive switches, due to the movement and leaving/joining of the mobile sensors. Meanwhile, an unpredictable message delay may occur at any time. Therefore, a synchronization process is required for such situations. In particularly, a request and acknowledge mechanism is utilized to dynamically synchronize the disconnection of any redundant communication links between the corresponding vertices. That is, upon the determination of redundant communication link, for example, $j \in \mathcal{C}_d{}^{[i]}(t)$, a request message Req$D(i, j)$ will be sent from i to j. Upon receiving Req$D(i, j)$ sensor j will remove sensor i from the logical neighbor set $\mathcal{N}_l{}^{[j]}(t)$ and send back a acknowledge message Ack$D(j, i)$ based on the condition that $i \in \mathcal{C}_d{}^{[j]}(t)$. The network structure will be constructed based on the subgraph $\mathcal{G}_s(t)$, which is generated from the proposed local coverage enhancement algorithm (cf. Algorithm 2).

3.2. Capabilities and Message Complexity of DLRA. We first prove that the connectivity is guaranteed while the redundant links are removed from the underlying network.

Theorem 6. *Given a network $\mathcal{G}(t_0) = (\mathcal{V}, E(t_0))$ with initial connectedness and controllable links, the subgraph $\mathcal{G}_s(t) = (\mathcal{V}, E_s(t))$ generated from DLRA is connected all the time.*

Proof. Since neighbor discovery procedure and adding redundant links into $\mathcal{N}_p{}^{[i]}(t)$ do not violate network connectivity, we only need to consider the link removal mechanism for connectivity preservation.

Consider two nodes i_0, i_n where i_0, i_n are initially connected in $\mathcal{G}(t_0)$ through a series of vertices, for example, $i_0 \leftrightarrow i_1 \leftrightarrow i_2 \ldots i_{n-1} \leftrightarrow i_n$. Suppose that $(i_k, i_{k+1}) \notin E_s(t)$, and i_k, i_{k+1} are disconnected in $\mathcal{G}_s(t_s)$, we have the following.

Case 1. $\exists v$ where $v \in \mathcal{N}_l{}^{[i_k]}(t_0)$ and $v \in \mathcal{N}_l{}^{[i_{k+1}]}(t_0)$, since $(i_k, i_{k+1}) \notin E_s(t)$. According to step 2 in Algorithm 1, we have $a_{i_k i_{k+1}}(t_0) = \max\{a_{i_k i_{k+1}}(t_0), a_{i_k v}(t_0), a_{i_{k+1} v}(t_0)\}$, so $\exists i_k \leftrightarrow v \leftrightarrow i_{k+1}$, contradiction reached.

Case 2. $\exists v, u$ where $u, v \in \mathcal{N}_l{}^{[i_k]}(t)$, and $u, v \notin \mathcal{N}_l{}^{[i_{k+1}]}(t)$, since $(i_k, i_{k+1}) \notin E_s(t)$, $a_{i_k i_{k+1}}(t) = \max\{a_{i_k i_{k+1}}(t), a_{i_k v}(t), a_{u i_{k+1}}(t), a_{uv}(t)\}$, so we have $\exists i_k \leftrightarrow v \leftrightarrow u \leftrightarrow i_{k+1}$, contradiction reached.

The proof of Theorem 6 is completed. □

In addition, for Case 1, suppose that there exists a path $i \leftrightarrow v \leftrightarrow u \leftrightarrow i$, and $a_{iv}(t) = \max\{a_{uv}(t), a_{iv}(t), a_{iu}(t)\}$. From Algorithm 1 it can be seen that $(i, v) \notin \mathcal{G}_s(t)$, so contradiction reached. Similarly, for Case 2, suppose there exists $i \leftrightarrow v \leftrightarrow u \leftrightarrow j \leftrightarrow i$, same contradiction can also be met, which leads to the following corollary.

Corollary 7. *For any generated subgraph $\mathcal{G}_s(t) = (\mathcal{V}, E_s(t))$, the shortest cycle is 5, that is, there are no such links as $i \leftrightarrow v \leftrightarrow u \leftrightarrow i$ or $i \leftrightarrow v \leftrightarrow u \leftrightarrow j \leftrightarrow i$.*

Furthermore, through removing redundant links, node degree in $\mathcal{G}_s(t)$ can be constrained within a certain spectrum. In particular, one has the following.

Corollary 8. *For a generated subgraph $\mathcal{G}_s(t) = (\mathcal{V}, E_s(t))$ where $\xi_i(t) \in \mathbb{R}^2$, the upper bound of node degree is 5.*

Proof. To see contradiction, it is assumed that there are two links (i, j) and (i, k) in $\mathcal{G}_s(t)$, where (i, j) and (i, k) enclose an angle $\alpha \leq \pi/3$ at node i. Furthermore, suppose that $a_{ij}(t) < a_{ik}(t)$ and $j, k \in \mathcal{N}_l{}^{[i]}(t)$, from Definition 1 we know that link weight is a monotonic function of vertices' Euclid distance, for example, $a_{ij}(t) < a_{ik}(t) \Leftrightarrow \|\xi_i(t) - \xi_j(t)\| < \|\xi_i(t) - \xi_k(t)\|$. Since $a_{ij}(t) < a_{ik}(t)$ and $\alpha \leq \pi/3$, it is straightforward that $\|\xi_j(t) - \xi_k(t)\| < \|\xi_i(t) - \xi_k(t)\|$, see Figure 3, which leads to $a_{ik}(t) = \max\{a_{ij}(t), a_{jk}(t), a_{ik}(t)\}$. Consequently, we can conclude from Algorithm 1 that $k \notin \mathcal{N}_l{}^{[i]}(t)$, for example, $(i, k) \notin \mathcal{G}_s(t)$, thus we reach a contradiction to the assumption that $(i, k) \in \mathcal{G}_s(t)$. This proved that no adjacent links in $\mathcal{G}_s(t)$ enclose an angle equal or less that $\pi/3$, from which the corollary follows. □

In addition, the request and acknowledge mechanism in Algorithm 2 guarantees that, despite of trivial one-hop delays of the acknowledge information, the communication links in $\mathcal{G}_s(t)$ remain symmetric all the time. Moreover, DLRA depends only on local information of no more than two-hop neighbors, so that the information exchange between mobile sensors will not invite large delay and communication overhead. In particular, we have the following.

Theorem 9. *The worst-case message complexity of DLRA is $O(n^2/2)$, where n is the number of sensors.*

Proof. Since DLRA depends only on two-hop neighbor information to determine redundant communication links, this

Upon receiving *Hello* message from j, update $\mathcal{N}_p{}^{[i]}(t)$, send *Hello* message back. Upon finishing neighbor discovery, set time $t = t_0$ and $\mathcal{N}_l{}^{[i]}(t_0) = \mathcal{N}_p{}^{[i]}(t_0)$. Construct $\mathcal{N}_r{}^{[i]}(t_0)$, Initialized.

(1) **For** $i \in \mathcal{V}$
(2) **while** $\exists a_{kj}(t_0) \in \mathcal{N}_r{}^{[i]}(t_0)$ and $j \in \mathcal{N}_l{}^{[i]}(t_0)$ **do**
(3) $\forall k \in \mathcal{N}_r{}^{[i]}(t_0)$
(4) **if** $a_{ij}(t_0) = \max\{a_{ij}(t_0), a_{ik}(t_0), a_{jk}(t_0)\}$
(5) $\mathcal{C}_d{}^{[i]}(t_0 + \Delta t) := \mathcal{C}_d{}^{[i]}(t_0) \cup \{j\}$, send Req$D(i, j)$ to j
(6) **end if**
(7) **end while**
(8) Update $\mathcal{N}_r{}^{[i]}(t), \mathcal{C}_d{}^{[i]}(t)$, hold
(9) Construct $\mathcal{N}_r{}^{[ik]}(t) = [a_{jk}(t), a_{lk}(t), \ldots a_{nk}(t)]$
(10) $\forall k \in \mathcal{N}_r{}^{[i]}(t)$
(11) **while** $a_{jk}(t) \times a_{lk}(t) \neq 0$ **do**
(12) **if** $a_{ij}(t) = \max\{a_{ij}(t), a_{il}(t), a_{jk}(t), a_{lk}(t)\}$
(13) $\mathcal{C}_d{}^{[i]}(t + \Delta t) := \mathcal{C}_d{}^{[i]}(t) \cup \{j\}$, send Req$D(i, j)$ to j
(14) **end if**
(15) **end while**
(16) **End**

Algorithm 1: Link removal mechanism for DLRA.

Upon receiving Req$D(i, j)$ from j
(1) **For** $i \in v$
(2) **if** $j \in \mathcal{C}_d{}^{[i]}(t)$
(3) $\mathcal{N}_l{}^{[i]}(t + \Delta t) := \mathcal{N}_l{}^{[i]} \setminus \{j\}$, send Ack$D(i, j) = 1$ to j
(4) **else if** $\exists \mathcal{N}_r{}^{[ij]}(t)$
(5) $\mathcal{N}_l{}^{[i]}(t + \Delta t) := \mathcal{N}_l{}^{[i]}(t) \setminus \{j\}$, send Ack$D(i, j) = 1$ to j
(6) **else** send Ack$D(i, j) = 0$ to j
(7) **end if**
(8) **else** send Ack$D(i, j) = 0$ to j
(9) **end if**
(10) Upon receiving Ack$D(i, j)$
(11) **if** Ack$D(i, j) = 1$
(12) $\mathcal{N}_l{}^{[i]}(t + \Delta t) := \mathcal{N}_l{}^{[i]}(t) \setminus \{j\}$
(13) **else**
(14) $\mathcal{C}_d{}^{[i]}(t + \Delta t) := \mathcal{C}_d{}^{[i]}(t) \setminus \{j\}$
(15) **end if**
(16) **End**

Algorithm 2: Synchronized coordination for DLRA.

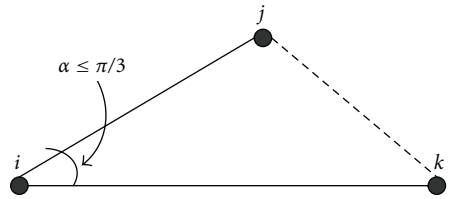


Figure 3: As indicated in proof of Corollary 8, $\|\xi_j(t) - \xi_k(t)\| < \|\xi_i(t) - \xi_k(t)\|$ if and only if $a_{ij}(t) < a_{ik}(t)$ and $\alpha \leq \pi/3$.

requires two messages per sensor, and total $2n$ messages will be sent. In addition, request and acknowledge mechanism requires two messages to remove each redundant link. In the worst case where $\mathcal{G}(t_0)$ is a undirected complete graph (UCG), and $E_s(t) = n - 1$, the maximum redundant communication links to be removed is $n(n-1)/2 - (n-1) = (n-1)(n-2)/2$, so the total messages to be sent is $(n-1)(n-2)/2 + 2n = (n^2 + n + 2)/2$. In conclusion, DLRA has a total message complexity of $O(n^2/2)$. □

Note that for large scale mobile networks with limited communication range, it is an unlikely case that the initial configuration is a UCG. Therefore, the message complexity can be reduced dramatically, which makes DLRA scalable to MASNs with a large number of mobile sensors.

4. Dispersion of Mobile Sensors

The invariance of network structure between any two consecutive switches in $\mathcal{G}(t)$ is prerequisite for the realization of

connectivity preserving dispersion of mobile sensors. The aim of this section is to develop a distributed motion controller that regulates the dispersion of mobile sensors in continuous space. First, we utilize a combined potential function to ensure that all essential links are remained connected and interagent collisions are avoided all the time during the self-organized dispersion of mobile sensors between any two consecutive topology updates.

We first introduce a repulsive potential function to deal with the collision avoidance between sensor i and all its physical neighbors $j \in \mathcal{N}_p{}^{[i]}(t)$, defined as ${\psi_{ij}}^r(t)$, where ${\psi_i}^r(t) = \sum_{j\in\mathcal{N}_p{}^{[i]}(t)} {\psi_{ij}}^r(t)$. In particular, we have

$$ {\psi_{ij}}^r(t) \triangleq \left(1 - \frac{f_{ij}(t) - \varepsilon}{1-\varepsilon}\right)^{-\rho}, \tag{4} $$

where $\rho \in \mathbb{R}^+$, and in this paper, associated with $f_{ij}(t)$ in (2), it yields

$$ {\psi_{ij}}^r(t) \triangleq \left(1 - \frac{\exp\left[-\beta\left(\left\|\xi_i(t) - \xi_j(t)\right\| - a_0\right)\right] - \varepsilon}{1-\varepsilon}\right)^{-\rho}. \tag{5} $$

It is straight forward that, whenever two mobile sensors approach to each other, the repulsive potential ${\psi_{ij}}^r(t)$ will continuously grow until $\varepsilon \to 1$, where the potential ${\psi_{ij}}^r(t) \to \infty$, so that intersensor collision is avoided. Similarly, to restrict all essential communication links within $f_{ij}(t) \geq \varepsilon$ for all $j \in \mathcal{N}_l{}^{[i]}(t)$, we introduce an attractive potential function, denote as ${\psi_{ij}}^a(t)$, where ${\psi_i}^a(t) = \sum_{j\in\mathcal{N}_l{}^{[i]}(t)} {\psi_{ij}}^a(t)$. Specifically, we have

$$ \begin{aligned} {\psi_{ij}}^a(t) &\triangleq \left(1 - \frac{f_{ij}(t) - \varphi}{\varepsilon - \varphi}\right)^{-\rho} \\ &= \left(1 - \frac{\exp\left[-\beta\left(\left\|\xi_i(t) - \xi_j(t)\right\| - a_0\right)\right] - \varphi}{\varepsilon - \varphi}\right)^{-\rho}. \end{aligned} \tag{6} $$

It can be seen from (6) that whenever two mobile sensors with critical communication links move away from each other, the attractive potential ${\psi_{ij}}^a(t)$ will continuously grow until $\varepsilon \to \varphi$, where the potential ${\psi_{ij}}^a(t) \to \infty$, so that the intersensor connection is guaranteed.

The proposed potential functions entitle us to assign each node i in the network a distributed control law, which is given as the combination of the negative gradients of the two potentials in the $\xi_i(t)$ direction

$$ \begin{aligned} u_i'(t) \triangleq &-K_r \nabla_{\xi_i(t)} \sum_{j\in\mathcal{N}_p{}^{[i]}(t)} {\psi_{ij}}^r(t) \\ &- K_a \nabla_{\xi_i(t)} \sum_{j\in\mathcal{N}_l{}^{[i]}(t)} {\psi_{ij}}^a(t). \end{aligned} \tag{7} $$

It is easy to observed from (7) that, the potential force can become infinite between pairs of sensors whenever $f_{ij}(t) \to 1$ or $f_{ij}(t) \to \varepsilon$. In practice, such unbounded actuation is unrealistic for most mobile sensor systems. In this paper, a piecewise-continuous function method is proposed with respect to sensors' actuation to approach the objective of *bounded* input.

Denote $u_i^m(t)$ as the upper bound of velocity for node i, in particular, for homogeneous systems, we have $u_i^m(t) = u_j^m(t)$, where $i \neq j \,\&\, i, j \in \mathcal{V}$.

Definition 10. Given a mobile sensor networks with fixed underlying network structure $\mathcal{G}^*(t)$, a distributed piecewise-continuous control law is assigned to each mobile sensor as follows:

$$ u_i(t) \triangleq \begin{cases} u_i'(t), & u_i'(t) \leq u_i^m(t), \\ u_i^m(t) \times \hat{u}_i'(t), & u_i'(t) > u_i^m(t), \end{cases} \tag{8} $$

where $\hat{u}_i'(t) = u_i'(t)/\|u_i'(t)\|$ represents the unit vector of $u_i'(t)$.

The definition of piecewise-continuous control law gives rise to the development of hybrid connectivity control system $\mathbb{C}$, as is described in Figure 4. In the connectivity control system, the proposed DLRA algorithm will first take every updated location of the mobile sensors as its input matrix to calculate the topology of the network. Therefore, after every interaction process, the redundant communication links will be determined and updated. Meanwhile, the output value of network topology from DLRA will in turn be utilized by the motion controller in every mobile sensor to assign different control functions to every neighboring sensor according to their neighboring types, that is, physical or logical neighbor. And the final output of the system is sensors' velocity and moving direction in the next step. The bounded velocity is implemented by comparing the output velocity from the controller with the designated maximum speed in the mobile sensor systems.

Now we can conclude the main result of this paper.

Theorem 11. *Given a MASNs with initially connected underlying network $\mathcal{G}_s(t_0)$, the distributed connectivity control system $\mathbb{C}$ guarantees that the connectivity of $\mathcal{G}_s(t)$ is maintained, interagent collisions are avoided and sensors' velocities are bounded for all time $t > t_0$.*

Proof. See Appendix. □

Furthermore, preserving connectivity without removing redundant communication links often results in a restricted network structure, as indicated in Figure 5(a). We argue that the restriction can be avoided by utilizing DLRA. In particular, from Corollary 7 we proved that the generated core structure (core structure is a subgraph of $\mathcal{G}_s(t)$ that contains only the logical neighbor links) of $\mathcal{G}_s(t)$ is a substructure of *α-Lattice* [4, 26] (pentagon), and the absence of cycle shorter than 5 guarantees that no mutual influences exist between any distinctive links. Therefore, for any two logical neighbors i and j, we have

$$ u_{ij}'(t) \triangleq -K_r \nabla_{\xi_i(t)} {\psi_{ij}}^r(t) - K_a \nabla_{\xi_i(t)} {\psi_{ij}}^a(t), \tag{9} $$

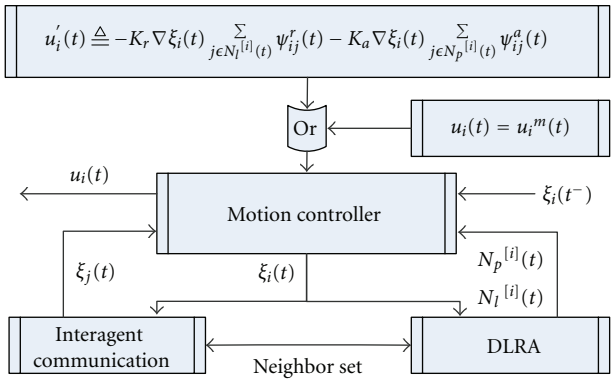


Figure 4: Control system $\mathbb{C}$ for the dispersion of MASNs.

Table 1: Simulation parameters.

Symbol	Quantity	Value
Δt	Execution time slot	0.01 s
ΔT	Interval of consecutive switches	0.1 s
ε	Minimum sensing capability of normalized RSSI	10%
φ	Optimized sensing value of normalized RSSI	40%
a_0	Restricted radium (Reference distance)	2 m
β	Path loss (outdoor, 802.11b)	0.02
ρ	Exponential gain	0.7
K_r	Repulsive potential gain	2500
K_a	Attractive potential gain	200
K_c	Restoration potential gain	10
$u^m(t)$	Maximum sensor velocity	5 m/s^{-1}

deriving the equilibrium distance $\mathcal{D}_e = \|\xi_i - \xi_j\|$ between i and j from the denotation that $u_{ij}'(t) = 0$, we have

$$\begin{aligned} &K_r \nabla_{\xi_i(t)} {\psi_{ij}}^r(t) + K_a \nabla_{\xi_i(t)} {\psi_{ij}}^a(t) \\ &\quad = \rho K_r \left(1 - \frac{f_{ij}(t) - \varepsilon}{1 - \varepsilon}\right)^{-\rho-1} \cdot \frac{\nabla_{\xi_i(t)} f_{ij}(t)}{1 - \varepsilon} \\ &\qquad - \rho K_a \left(1 - \frac{f_{ij}(t) - \varphi}{\varepsilon - \varphi}\right)^{-\rho-1} \cdot \frac{\nabla_{\xi_i(t)} f_{ij}(t)}{\varphi - \varepsilon} = 0 \\ &\quad \overset{\nabla_{\xi_i(t)} f_{ij}(t) \neq 0,\ \varepsilon < f_{ij} < 1}{\Longrightarrow} \\ &\frac{K_r \left(1 - \left(\left(f_{ij}(t) - \varepsilon\right)/(1 - \varepsilon)\right)\right)^{-\rho-1}}{1 - \varepsilon} \\ &\quad = \frac{K_a \left(1 - \left(\left(f_{ij}(t) - \varphi\right)/(\varepsilon - \varphi)\right)\right)^{-\rho-1}}{\varphi - \varepsilon}. \end{aligned} \tag{10}$$

Integrating (3) with the aforementioned equation and after some algebra manipulation, it yields

$$\mathcal{D}_e = a_0 - \frac{1}{\beta} \ln \frac{1 + \varepsilon \cdot \Phi}{1 + \Phi}, \tag{11}$$

where $\Phi = ((\varepsilon - 1)/(\varepsilon - \varphi))^{\rho+2} \cdot (K_a/K_r)^{\rho+1}$.

The aforementioned ratiocinations bring us the following result.

Corollary 12. *Considering a MASNs with initially connected underlying network $\mathcal{G}(t_0)$, where the position of vertices $\xi_i(t) \in \mathbb{R}^2$. The evolution of the system under the control system $\mathbb{C}$ can reach a set of condition $\mathcal{K}(\mathcal{D}_{ij}(t)) \triangleq \{\mathcal{D}_{ij}(t) \mid \mathcal{D}_{ij}(t) \geq a_0 - \ln((1 + \varepsilon \cdot \Phi)/(1 + \Phi))/\beta,\ i, j \in \mathcal{V}\}$ within finite time $t < \infty$.*

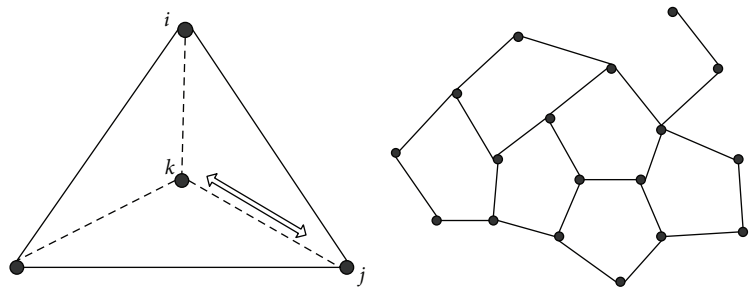
5. Simulations

To evaluate the performance of the proposed control system for the self-organized dispersion of mobile sensor networks, a variety of simulations have been conducted, and the results are presented in this section. The simulations are focused on the connectivity control and dispersion of mobile sensor networks in $\mathbb{R}^2$ and $\mathbb{R}^3$ spaces. General parameters are set as in Table 1.

First, to evaluate the correctness and efficiency of the proposed DLRA, a static simulation is conducted. In the first scenario, 15 static sensors are initially deployed within an area of 20 m $\times$ 20 m, as can be seen from Figure 6(a), where dense communication links are observed, after the execution of DLRA, the topology of the wireless sensor network is finally simplified with only 15 communication links (cf. Figure 6(b)). In case of large sensor networks, a 100 sensors scenario is then performed, the initial configuration of the network topology can be found in Figure 6(c). The final results from DLRA is shown in Figure 6(d), where the sensor network becomes extremely sparse with only a few redundancy.

To evaluate the proposed connectivity control system $\mathbb{C}$. We first simulate the dispersion of 10 sensors initially located within 20 m $\times$ 20 m $\mathbb{R}^2$ space. The evaluation of mobile sensor network is shown as in Figures 7(a)–7(d). For the purpose of illustration, the diamonds represent mobile sensors, and the circles represent the semidistance (semidistance represents a value that is 5% shorter than half of the communication range. The use of it here is for the sake of facilitating the observation) of communication range. Sensors are under the communication range of each other if and only if the two corresponding circles overlap. The total trajectories of mobile sensors are shown in Figure 7(e), where the solid diamonds denote the initial position of the sensors, final position is represented as hollow ones, and the trajectory is shown as the dotted lines. Furthermore, upon removal of redundant communication links, sensors are dispersed into $\mathbb{R}^2$ space, and the network reaches its final configuration within 30 seconds (see Figure 8(a)). Moreover, the average velocity of the sensors is decreased from the peak that is 5 m s^{-1} (see Figure 8(c)), the bounded velocity is guaranteed.

We further conduct a simulation for 100 sensors to verify the scalability of the proposed approach. Similarly to the case of 10 sensors, sensors are initially location in a 200 m $\times$ 200 m $\mathbb{R}^2$ space, the trajectories of networked sensors evolve as in Figures 6(f)–6(i). It is noticeable that the average degree of node in both simulations are between



(a) Restricted Structure (b) *α-Lattice* (pentagon)

FIGURE 5: Illustration of network structure (a) restricted link (j,k) is observed without removal of redundant links; (b) a typical *α-Lattice* (pentagon) structure generated by DLRA, absence of external constrains.

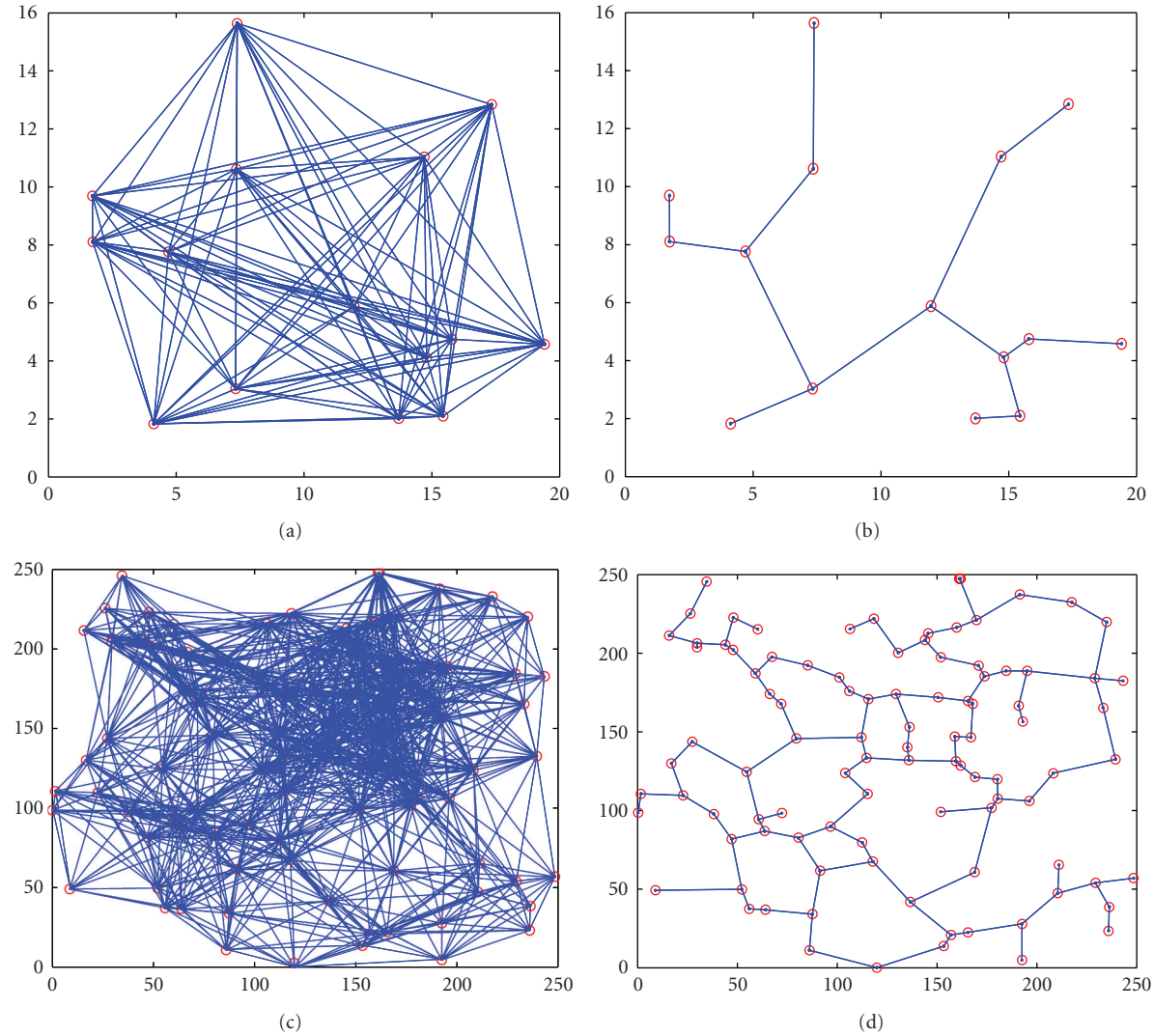


FIGURE 6: Simulation results for DLRA. (a) Original topology for 15 nodes; (b) final topology for 15 nodes; (c) original topology for 100 nodes; (d) final topology for 100 nodes.

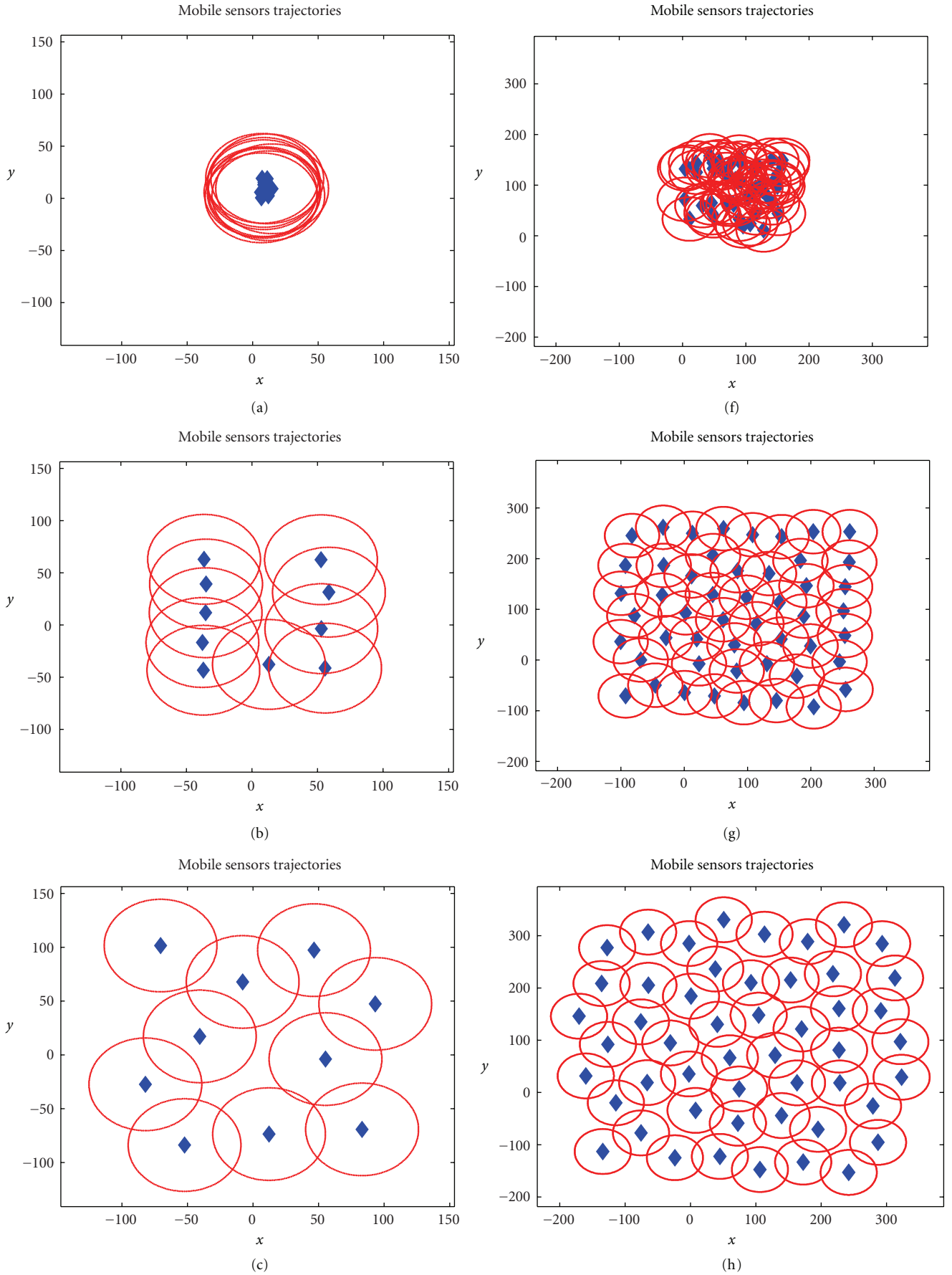


Figure 7: Continued.

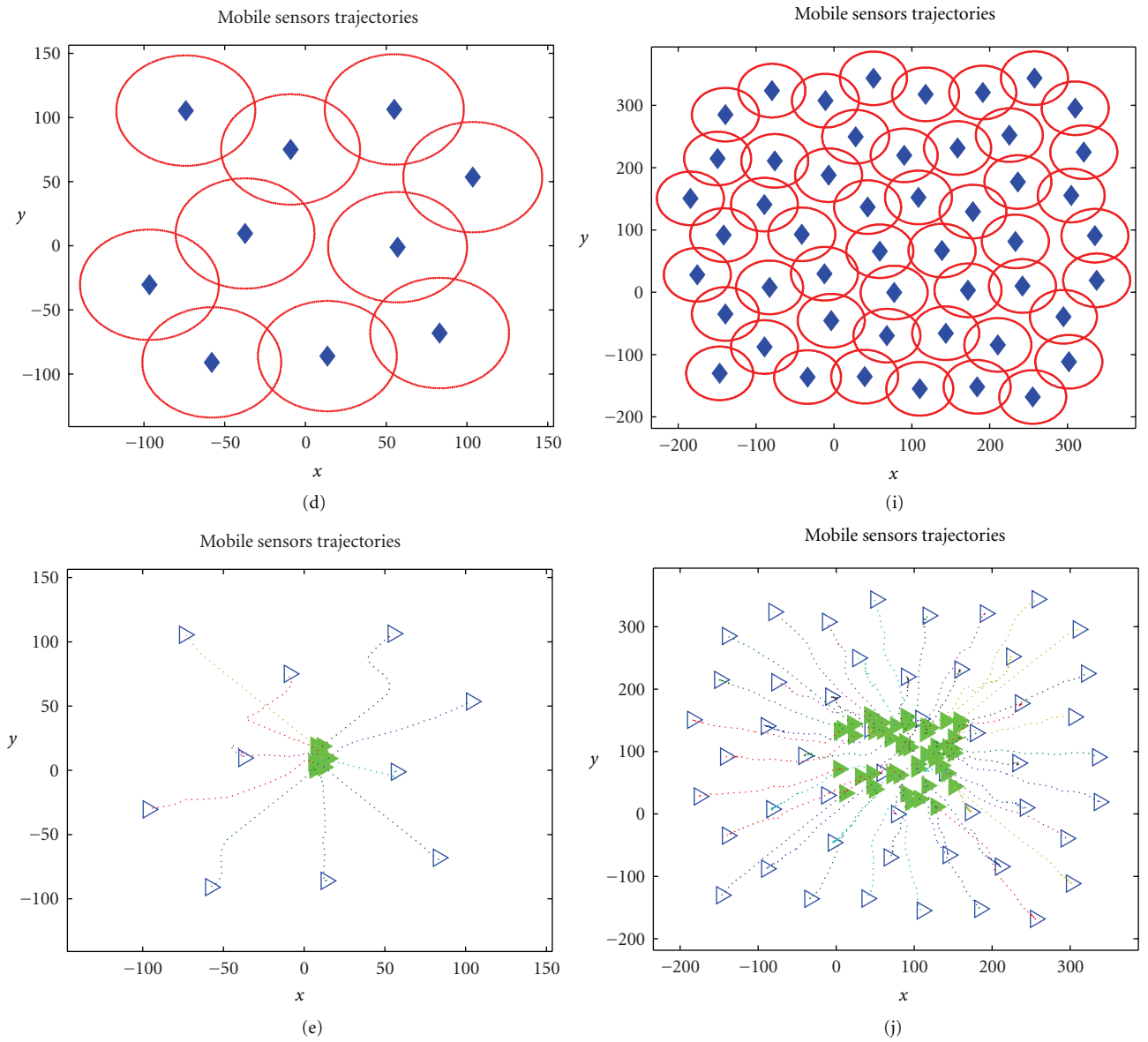


FIGURE 7: Simulation results for 2D distributed mobile multiagent dispersion. (a)–(d) Evolution snapshots for 10 nodes; (f)–(i) evolution snapshots for 100 nodes; (e) mobile sensor's trajectories in 10 nodes case; (d) mobile sensor's trajectories in 100 nodes case.

2 and 3, this phenomenon indicates that an asymptotically 1-connectivity (graph is 1-connected if it contains only one path between any pair of distinct vertices) graph structure is generated during the evolution of the network, such that the final configurations are sparse with only trivial signal conflictions.

The self-organized dispersion of MASNs in $\mathbb{R}^3$ space is also evaluated with similar simulation methodology. Sensors locate within 95% communication range of each other is connected by a blue line (see Figure 9), networks with 10 and 100 sensors are all investigated, and in both case, the initial locations are within $20\,\text{m} \times 20\,\text{m} \times 20\,\text{m}\,\mathbb{R}^3$ space. Note that Corollary 12 may not hold in the case of 3-dimensional space coverage optimization, this is due to fact that the lower bound of the node degree cannot be obtained and a certain freedom of mobile sensors may not exist with respect to the uncertain structure of the networks. Therefore, we then conducted a simulation to study the coverage efficiency of proposed system $\mathbb{C}$ with respect to different number of sensors. As it can be seen from Figure 10(c), with the aid of the proposed coverage enhancement system, 7 mobile sensors can cover an area up to 25000 m^2 in size (nearly the size of total 4 standard football fields), under the constrain of limited communication range ($R = 100\,\text{m}$). The interagent distances between logical neighbors are converged into a constant value, which is above the minimum distance derived from Corollary 12, and the distances between nonlogical neighboring sensors are strictly above the communication range (see either Figures 10(a) or 10(b)). The results verified the proposed control system in

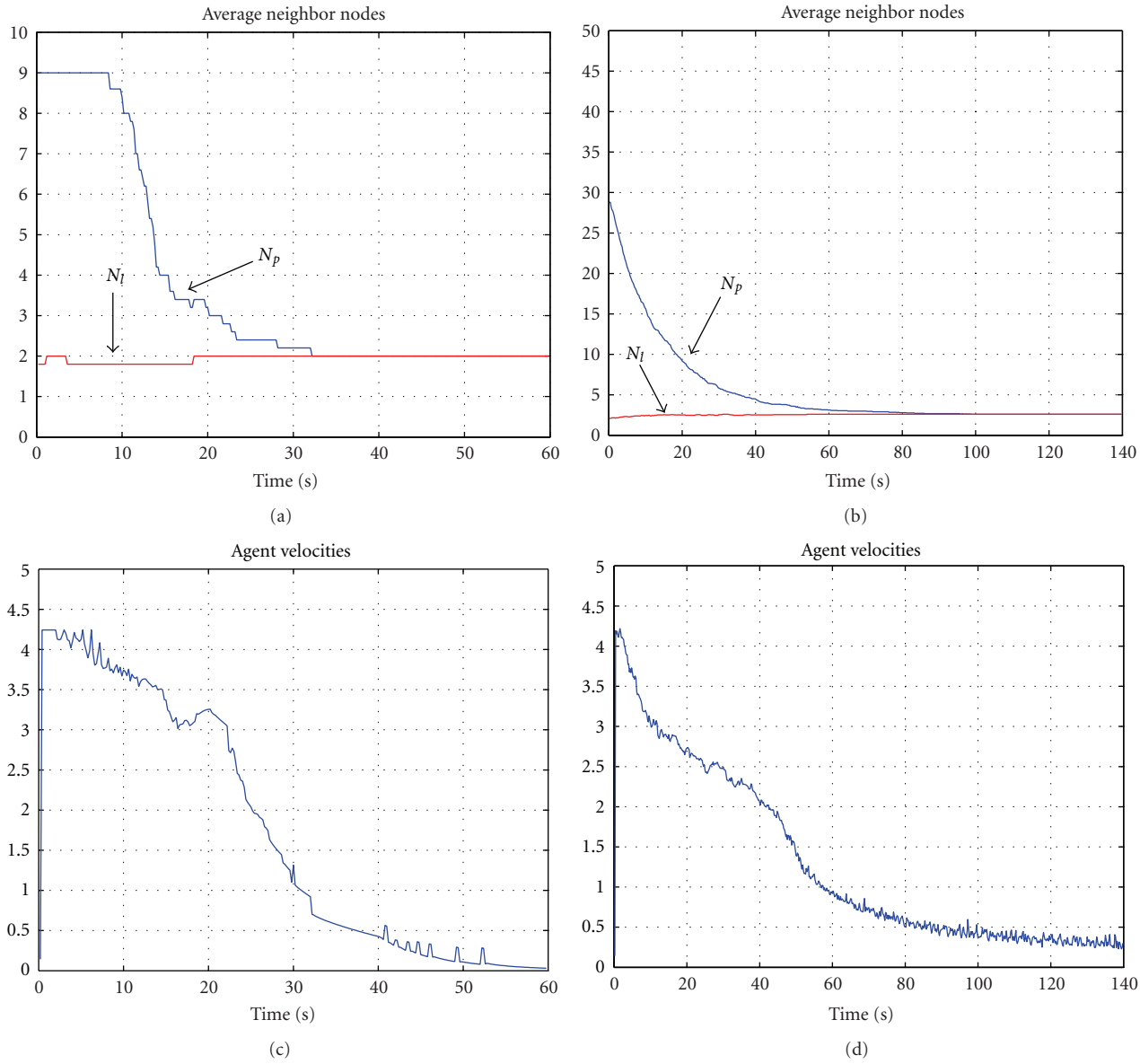


Figure 8: Simulation results for 2D distributed mobile multiagent dispersion. (a) The converge of average number of physical neighbors and logical neighbors during simulation in 10 nodes case; (b) the converge of average number of physical neighbors and logical neighbors during simulation in 100 nodes case; (c) average velocity of mobile sensors during simulation in 10 nodes case; (d) average velocity of mobile sensors during simulation in 100 nodes case.

its capability of coverage enhancement. Moreover, as it can be seen from Figures 10(a) and 10(b), no interagent distance whatsoever can reach zero, which indicates that the proposed method can also effectively handle the issue of collision avoidance. Moreover, the simulation results unfolded in Figure 10 suggest that the lower bound of the neighbors' distance can also be guaranteed in a network of 10 mobile sensors. The question of *whether this is a common case in 3-dimensional space coverage enhancement* will be investigated in our future research. Nonetheless, the capability of the coverage enhancement system in 3-dimensional space, for example, can cover up to 1.4×10^6 m^3 with 7 mobile sensors, is still fairly convincing.

It is worthwhile to notice that simulation results in $\mathbb{R}^3$ space are analogous to that in the case of $\mathbb{R}^2$ (see Figure 7), which entitles us to obtain the following conclusions.

Correctness. A variety of simulations in $\mathbb{R}^2$ and $\mathbb{R}^3$ spaces verified the correctness of the distributed control system for self-organized sensor dispersion, so that objectives can be achieved.

Sensibility. The required time for reaching the final configuration is distinctive in $\mathbb{R}^2$ and $\mathbb{R}^3$ spaces. The dimension-sensible property shows a better performance in spatial dispersion.

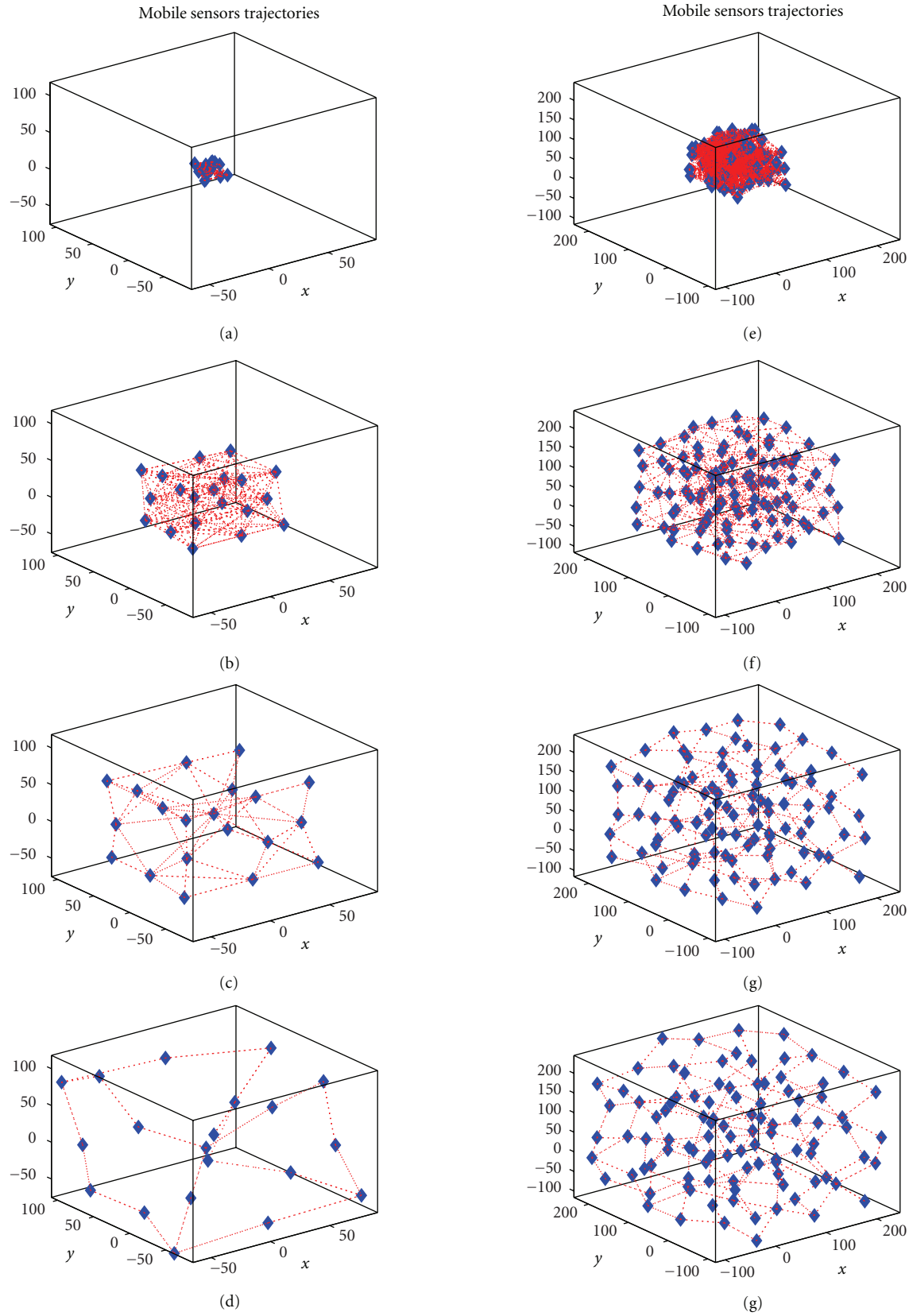


Figure 9: Evolution trajectory (snapshot) of 3D distributed dispersion of mobile sensors in different scales. (a) $n = 20$, $t = 1$; (b) $n = 20$, $t = 5$; (c) $n = 20$, $t = 15$; (d) $n = 20$, $t = 30$; (e) $n = 100$, $t = 1$; (f) $n = 100$, $t = 15$; (g) $n = 100$, $t = 40$; (h) $n = 100$, $t = 80$.

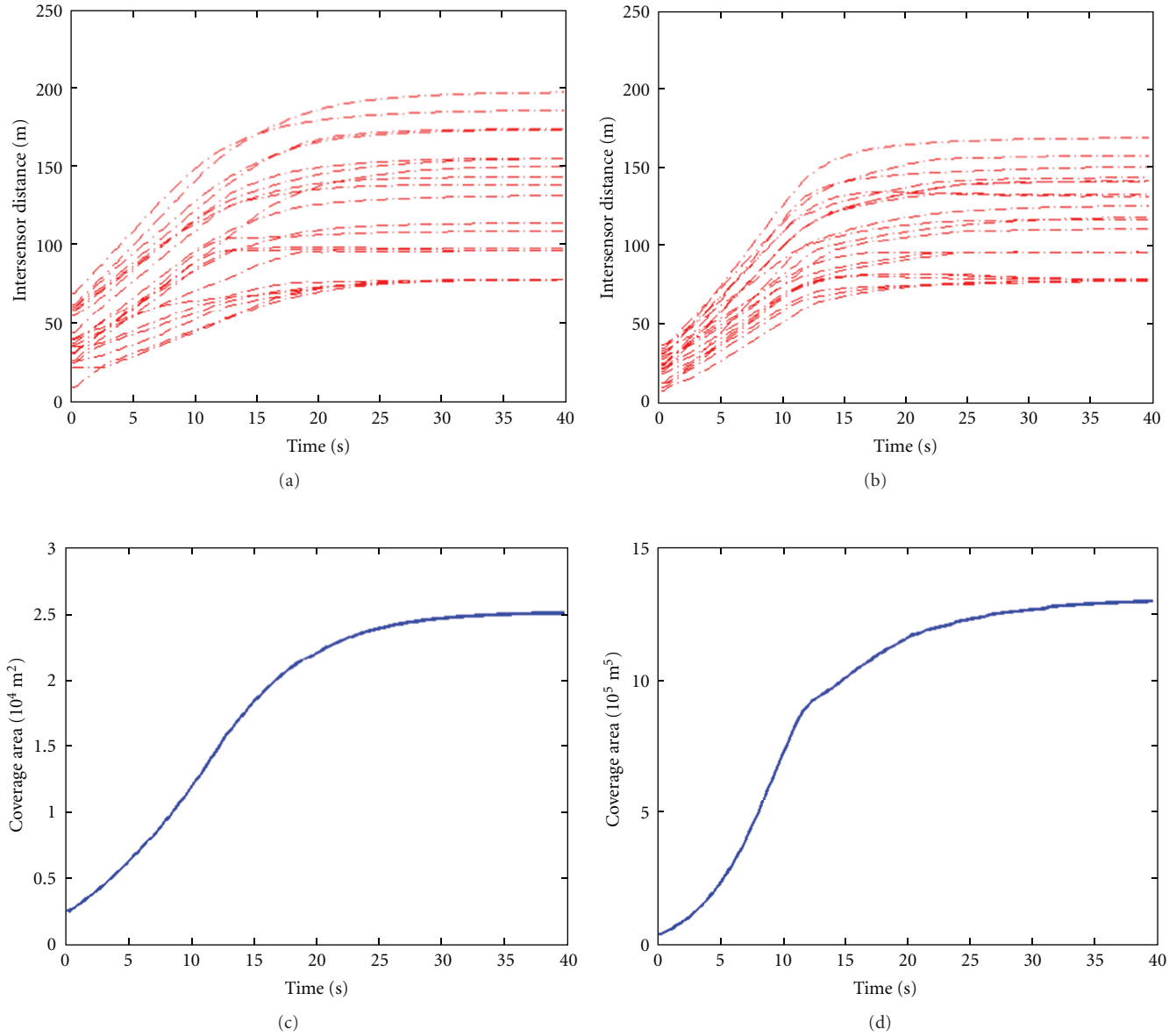


Figure 10: The efficiency of the proposed connectivity control system $\mathbb{C}$. (a) Interagents distances in case of 2-dimensional coverage optimization; (b) interagents distances in case of 3-dimensional coverage optimization; (c) coverage area in case of 2-dimensional coverage optimization; (d) coverage space in case of 3-dimensional coverage optimization.

Sparseness. Despite of dimension and network capacity, the final configurations of the networks are always sparse with small node degrees.

6. Conclusions

Self-organized dispersion of mobile ad hoc sensor networks with guaranteed connectivity and collision avoidance is studied in this paper. A distributed self-organized control system combined with a local dynamic link removal algorithm, namely, DLRA, is proposed to reduce redundant communication links and regulate the dispersion of mobile sensors. The proposed control system can disperse a team of mobile sensors with guaranteed connectivity and collision avoidance, and the DLRA does not require access to the whole topology information, and thus, the proposed approach is scalable and will not invite excessive delays and communication overhead in large scale networks.

Several extensions to this presented work will be studied in future research: to apply the mobile sensor networks into indoor environment, we plan to study the dispersion of networked mobile sensors in nonconvex environments; the impact of wireless communication on the connectivity of multiagent system is another promising topic. The proposed

approach can also be extended for dispersion in leader-follower system and potential applications in hazardous and battlefield environments.

Appendix

Proof of Theorem 11

To facilitate the proof of Theorem 11, we first unfold the following lemmas.

Lemma A.1. *For a mobile sensor network driven by the distributed control law defined by* (2) *and* (8) *with a set of initial condition* $\mathcal{I}(\xi(t_0)) \triangleq \{\xi(t_0) \mid f_{ij}(t_0) < 1,\ i, j \in \mathcal{V},\ i \in \mathcal{N}_p^{[j]}(t)\}$, *then* $\mathcal{I}(\xi(t))$ *is invariant for the trajectory of the closed-loop system.*

Proof. Suppose that $\mathcal{I}(\xi(t))$ varied at time $t = t_v$, furthermore, assume the variant was irritated by $f_{ij}(t_v) = 1$, thus there exist a small positive constant $\sigma | 0 < \sigma < 1 - \varepsilon$, that $\widehat{u}_i'(t) \cdot (\xi_i(t) - \xi_j(t)) \geq 0$ in every $(1 - \sigma, 1]$. Whereas, from the definition of (7), we know that $\lim_{f_{ij}(t) \to 1} \psi_{ij}{}^r(t) = \infty$, thus there exist a small positive constant $\theta | 0 < \theta < 1 - \varepsilon$, that $\widehat{u}_i'(t) \cdot (\xi_i(t) - \xi_j(t)) \leq 0$ in every $(1 - \theta, 1]$, contradiction is reached, this completes the proof. □

Lemma A.1 established collision avoidance between any pairs of sensors within each other's physical neighbor set. Now we investigate the connectivity preservation capability of the motion controller in $\mathcal{G}^*(t)$.

Lemma A.2. *For a MASNs driven by the distributed control law defined by* (2) *and* (8) *with a set of initial condition* $\mathcal{J}(\xi(t_0)) \triangleq \{\xi(t_0) \mid f_{ij}(t_0) > \varepsilon,\ i, j \in \mathcal{V},\ i \in \mathcal{N}_l^{[j]}(t)\}$, *then* $\mathcal{J}(\xi(t))$ *is invariant for the trajectory of the closed-loop system.*

Proof. Following the same methodology in the proof of Lemma A.1, we assume a variant of $\mathcal{J}(\xi(t))$ was irritated by $f_{ij}(t_v) = \varepsilon$, thus there exist a small positive constant $\sigma \mid 0 < \sigma < \varphi - \varepsilon$, that $\widehat{u}_i'(t) \cdot (\xi_i(t) - \xi_j(t)) \leq 0$ in every $[\varepsilon, \varphi)$. Whereas, from the definition of (8), we know that $\lim_{f_{ij}(t) \to \varepsilon} \psi_{ij}{}^a(t) = \infty$ and $\lim_{f_{ij}(t) \to \varepsilon} \psi_{ij}{}^r(t) = 0$, thus there exist a small positive constant $\theta \mid 0 < \theta < \varphi - \varepsilon$, that $\widehat{u}_i'(t) \cdot (\xi_i(t) - \xi_j(t)) \leq 0$ in every $[\varepsilon, \varphi)$. We choose $\sigma \geq \theta$, contradiction is reached, which completes the proof. □

Lemmas A.1 and A.2 guaranteed that link connections are all preserved and interagent collisions are avoided during the dispersion between any two consecutive topology updates. On the other hand, the synchronization mechanism in DLRA causes delays in both link addition and link removal procedures in $\mathcal{G}_f(t)$. We now show that the ineluctable delays cannot jeopardize network connectivity. Assume the interval time between two consecutive switches is ΔT, where $\Delta T = n \cdot \Delta t$, $n \in \mathbb{N}^+$, we have the following cases.

Case 1. $j \in \mathcal{N}_l^{[i]}(t)$ at time t, considering a small positive time-interval $\tau \mid 0 < \tau < \Delta T$, assume $j \notin \mathcal{N}_l^{[i]}(t + \tau)$, thus $\mathcal{J}(\xi(t))$ is invariant for $[t + \tau, t + \Delta T)$, where the communication links are rather unnecessarily maintained than disconnected.

Case 2. $j \notin \mathcal{N}_l^{[i]}(t)$ and $j \in \mathcal{N}_p{}^{[i]}(t)$ at time t, assume that $j \in \mathcal{N}_l^{[i]}(t + \tau)$ and $j \notin \mathcal{N}_p{}^{[i]}(t + \Delta T)$, thus $j \notin \mathcal{N}_l^{[i]}(t + \Delta T)$. Since candidate link (i, j) for link addition is redundant, connectivity is not violated. This result also hold for the case that $j \notin \mathcal{N}_l^{[i]}(t)$ and $j \notin \mathcal{N}_p{}^{[i]}(t)$, while hypostatically we assuming that $j \in \mathcal{N}_l^{[i]}(t + \tau)$ and $j \notin \mathcal{N}_p{}^{[i]}(t + \Delta T)$.

Derived directly from the aforementioned analysis and Theorem 6, we conclude the proof of Theorem 11.

Acknowledgments

The work is cosponsored by Nature Science Foundation of China (Grant nos. 61272432, 61202508), and China Postdoctoral Science Foundation (Grant no. 2011M500243). The authors would like to gratefully acknowledge the constructive and thorough comments of the reviewers, which is very helpful for improving the quality of the paper. The authors appreciate the constructive suggestions in the simulation methodology from Professor Michael M. Zavlanos. They also thank Dr. Yan C. Wang for the enlightening discussions in topology control algorithms.

References

[1] T. Rappaport, *Wireless Communications, Principles and Practice*, Prentice Hall, Upper Saddle River, NJ, USA, 2002.

[2] L. Ludwig and M. Gini, "Robotic swarm dispersion using wireless intensity signals," in *Proceedings of the 22nd Int'l Symposium on Distributed Autonomous Robotic Systems*, pp. 135–144, July 2006.

[3] D. V. Dimarogonas and K. J. Kyriakopoulos, "Inverse agreement protocols with application to distributed multi-agent dispersion," *IEEE Transactions on Automatic Control*, vol. 54, no. 3, pp. 657–663, 2009.

[4] W. M. Spears, D. F. Spears, J. C. Hamann, and R. Heil, "Distributed, physics-based control of swarms of vehicles," *Autonomous Robots*, vol. 17, no. 2-3, pp. 137–162, 2004.

[5] A. Chen, S. Kumar, and T. H. Lai, "Local barrier coverage in wireless sensor networks," *IEEE Transactions on Mobile Computing*, vol. 9, no. 4, pp. 491–504, 2010.

[6] A. Tahbaz-Salehi and A. Jadbabaie, "Distributed coverage verification in sensor networks without location information," *IEEE Transactions on Automatic Control*, vol. 55, no. 8, pp. 1837–1849, 2010.

[7] Z. Tu, Q. Wang, H. Qi, and Y. Shen, "Flocking based distributed self-deployment algorithms in mobile sensor networks," *Journal of Parallel and Distributed Computing*, vol. 72, no. 3, pp. 437–449, 2012.

[8] J. Cortes, "Coverage optimization and spatial load balancing by robotic sensor networks," *IEEE Transactions on Automatic Control*, vol. 55, no. 3, pp. 749–754, 2010.

[9] H. Zhou, T. Liang, C. Xu, and J. Xie, "Multiobjective coverage control strategy for energy-efficient wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2012, Article ID 720734, 10 pages, 2012.

[10] M. Zhong and C. G. Cassandras, "Distributed coverage control and data collection with mobile sensor networks," *IEEE*

Transactions on Automatic Control, vol. 56, no. 10, pp. 2445–2455, 2011.

[11] Z. Mi and Y. Yang, "Topology control and coverage enhancement of dynamic networks based on the controllable movement of mobile agents," in *Proceedings of the IEEE International Conference on Communication*, pp. 1–5, Kyoto, Japan, June 2011.

[12] M. M. Zavlanos and G. J. Pappas, "Potential fields for maintaining connectivity of mobile networks," *IEEE Transactions on Robotics*, vol. 23, no. 4, pp. 812–816, 2007.

[13] Y. Kim and M. Mesbahi, "On maximizing the second smallest eigenvalue of a state-dependent graph Laplacian," *IEEE Transactions on Automatic Control*, vol. 51, no. 1, pp. 116–120, 2006.

[14] T. Razafindralambo and D. Simplot-Ryl, "Connectivity preservation and coverage schemes for wireless sensor networks," *IEEE Transactions on Automatic Control*, vol. 56, pp. 2418–2428, 2011.

[15] M. M. Zavlanos and G. J. Pappas, "Distributed connectivity control of mobile networks," *IEEE Transactions on Robotics*, vol. 24, no. 6, pp. 1416–1428, 2008.

[16] P. Yang, R. A. Freeman, G. J. Gordon, K. M. Lynch, S. S. Srinivasa, and R. Sukthankar, "Decentralized estimation and control of graph connectivity for mobile sensor networks," *Automatica*, vol. 46, no. 2, pp. 390–396, 2010.

[17] Y. Stergiopoulos, Y. Kantaros, and A. Tzes, "Distributed control of mobile sensor networks under RF connectivity constraints," *International Journal of Distributed Sensor Networks*, vol. 2012, Article ID 741821, 7 pages, 2012.

[18] G. Notarstefano, K. Savla, F. Bullo, and A. Jadbabaie, "Maintaining limited-range connectivity among second-order agents," in *Proceedings of the American Control Conference*, pp. 2124–2129, Minneapolis, Minn, USA, June 2006.

[19] Z. Mi, Y. Yang, and G. Liu, "Connectivity control of mobile multi-robot networks," in *Proceedings of the IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pp. 123–128, Montreal, Canada, July 2010.

[20] X. Li and Y. Xi, "Distributed cooperative coverage and connectivity maintenance for mobile sensing devices," in *Proceedings of the 48th IEEE Conference on Decision and Control*, pp. 7891–7896, Shanghai, China, December 2009.

[21] A. Ajorlou, A. Momeni, and A. G. Aghdam, "Connectivity preservation in a network of single integrator agents," in *Proceedings of the 48th IEEE Conference on Decision and Control*, pp. 7061–7067, Shanghai, China, December 2009.

[22] G. Xing, T. Wang, Z. Xie, and W. Jia, "Rendezvous planning in wireless sensor networks with mobile elements," *IEEE Transactions on Mobile Computing*, vol. 7, no. 12, pp. 1430–1443, 2008.

[23] Y. Yang, Z. Mi, J. Y. Yang, and G. Liu, "A model based connectivity improvement strategy for vehicular ad hoc networks," in *Proceedings of the 76th IEEE Vehicular Technology Conference*, Ottawa, Canada, September 2010.

[24] R. Wattenhofer and A. Zollinger, "XTC: a practical topology control algorithm for ad-hoc networks," in *Proceedings of the 18th IEEE International Parallel & Distributed Processing Symposium*, pp. 216–222, Santa Ana, Calif, USA, 2004.

[25] A. Jawahar, S. Radha, and R. Sharath Kumar, "Capacity-preserved, energy-enhanced hybrid topology management scheme in wireless sensor networks for hazardous applications," *International Journal of Distributed Sensor Networks*, vol. 2012, Article ID 280932, 12 pages, 2012.

[26] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: algorithms and theory," *IEEE Transactions on Automatic Control*, vol. 51, no. 3, pp. 401–420, 2006.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 973713, 10 pages
doi:10.1155/2012/973713

Research Article

Automatic Distributing Schemes of Physical Cell Identity for Self-Organizing Networks

Yao Wei, Mugen Peng, Wenbo Wang, Shijun Min, Jia Mo Jiang, and Yu Huang

Wireless Signal Processing and Network Laboratory, Key Laboratory of Universal Wireless Communications of Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to Yao Wei, wybby60@gmail.com

Received 3 May 2012; Revised 19 August 2012; Accepted 16 September 2012

Academic Editor: Yiqing Zhou

This paper presents and puts forward an optimal automatic distributing of physical cell identity (ADPCI) scheme for the self-organizing network (SON). Considering the high number and the layered structure of the evolved node B (eNodeB, eNB) in the initial rollout phase, the assigning of PCI for cells would be quite complex. The PCI self-distributing problem is mapped to the well-known minimum spanning tree (MST) problem in order to optimize the PCI reuse distance and decrease the multiplexing interference. The correlation property of PCI is analyzed and taken into consideration in the assigning phase. Moreover, a suboptimal algorithm (SADPCI) is presented as it performs approximately to ADPCI but the computational complexity is lower. To demonstrate the proposal validity, performances of ADPCI and SADPCI are evaluated. Simulation results illustrate that these schemes can achieve significantly higher performance even under the condition of severe PCI deficiency.

1. Introduction

There is a strong momentum for self-organizing features in wireless communication networks, self-configuration and self-optimization are identified as two mechanisms to facilitate operation and manage the long term evolution (LTE) network [1, 2]. In parallel, the self-distribution of PCI is included in the self-configuration use cases defined by the Next Generation Mobile Networks (NGMN) Alliance [3].

The objective of this use case is to assign a PCI to each cell so that mobile terminals can identify neighboring cells without ambiguity. PCI indicates the primary and secondary synchronization signals which help user equipment (UE) to acquire frequency and time synchronization during the cell search phase. However, considering the high number and layered structure of eNodeBs, finite number of IDs need to be reused across the network. Therefore, PCI allocation must to be planned carefully to avoid multiplexing interference and conflict. Methods for managing other radio resources [4, 5] can also be applied in PCI allocation.

Some proposals are discussed in [6–10]. One suggests that eNB scans its own radio environment, especially in terms of reception of the down-link transmission band of neighboring radio cells to acquire PCI information. The scanning method helps eNB to identify unavailable PCI and thus avoid collision to a certain degree. However, in order to perform a confusion free selection of PCI, the PCI assigning information of neighbors' neighbors have to be measured, but the measurement result depends on the signal quality which cannot be guaranteed. Another method has performed a distributed solution that relies on the use of a temporary PCI. The new eNB chooses a random PCI from a predefined set provided by operation administration and Maintenance (OAM) and starts to operate. The automatic neighbor relationship (ANR) function is utilized to acquire neighbors' PCI information supported by UEs. The method performs more effectively to address both collision and confusion requirements than the previous method. However, it relies on the proper location of UEs to identify each of all the neighbors; furthermore, PCI reconfiguration causes UE dropping and brings too much overhead to the system. A centralized approach [9] has mapped the PCI assignment problem to the well-known and well-understood problem of graph coloring. The coloring algorithm is used and

simply extended in the approach, and it provides an efficient initial assignment even for complex networks. The scheme has analyzed the properties of the colored graph that is used for extending the network with new cells, and the results show that only minimal interruptions have occurred while still retaining the properties of a colored graph. Another centralized approach (HCPCI) [10] has introduced a hyper graph coloring PCI assigning scheme. The neighbor's relationship degree N is regarded to be the reuse distance; only when $N \geq 7$, PCI are PCI could be reused.

This paper proposes an automatic distributing PCI scheme. At first, the correlations between different PCIs are analyzed and the interinfluence IDs are classified into different groups. The PCI distributing problem is mapped to the well-known MST problem and the reuse distance is taken into account in order to decrease the multiplexing interference throughout the entire network. A suboptimal scheme is proposed as a second solution which has a good performance approximated to ADPCI but with much lower computational complexity.

The remainder of this paper is organized as follows. Section 2 contains a detailed description of scenarios and the framework of PCI management, the distributing principles as well. Section 3 analyzes the correlation property of IDs and represents the proposed schemes of ADPCI and SAD-PCI. The performance evaluation is presented in Section 4, followed by a conclusion in Section 5.

2. PCI Assigning Framework and Principles

2.1. PCI Distributing Scenario and Framework. Two typical scenarios are defined in the SON description specification [6]. In a macrocell deployment, a large number of eNBs are deployed and requiring PCI configuration during the initial network establishment. However, the number of PCIs is limited and insufficient to guarantee that each cell gets an individual PCI. Thus, PCI has to be reused in the network which brings the multiplexing interference inevitably. Another scenario is the assignment for individual eNB during the network growth.

To minimize the human interventions and decrease the planning, deployment, optimization, and maintenance activities, the newly deployed eNBs are configured by automatic installation procedures by OAM to get basic parameters and download necessary software for operation. The support for PCI assignment is translated into concrete functionalities, interfaces, and procedures as shown in Figure 1.

(i) Centralized SON function: according to centralized SON function, the newly deployed eNBs obtains relevant PCI allocating information, including antenna location, cell identity, cell radius, down tilt, and height of antenna. The geographic coordinates can be obtained with the help of the global position system (GPS) and other information is acquired during the base station establishment phase. The information of eNBs are classified and provided to upper layer OAM through Itf-N (interface-north).

(ii) Data processing/configuration interface: this module has the responsibility to process the data obtained from the network, like cell type, cell state, and the neighbor relations are classified and transmitted to the database. When the PCI allocating decision is made, it indicates eNBs to update the results.

(iii) Policy management base: this module indicates the PCI self-configuring policy. Operators can provide specific groups of IDs which meet specific requirements for example to ensure uniqueness in border regions. The policy can be created, edited, and modified by the operators or through the PCI allocation feedback information.

(iv) PCI algorithm execution: based on the information from database, PCI algorithm analyzes the allocating policy and executes for PCI distribution.

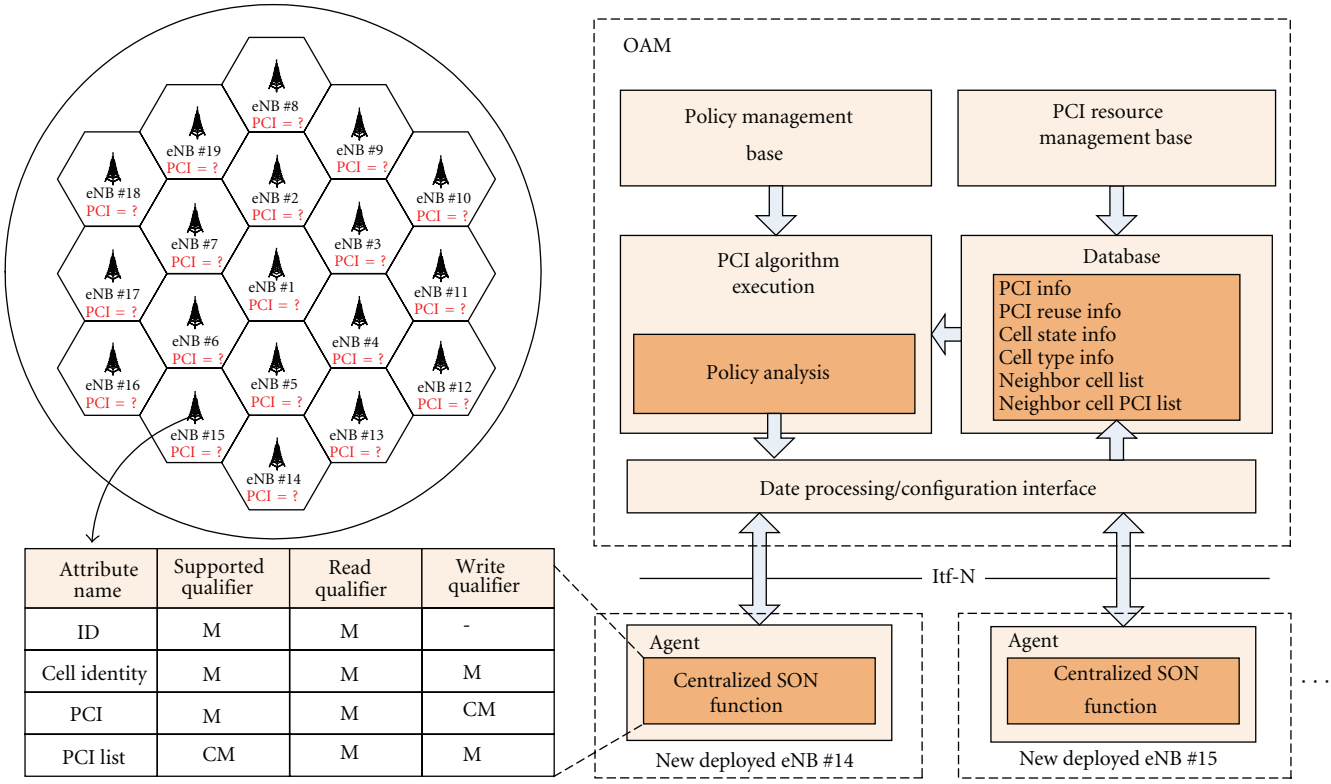
(v) PCI resources management bases: PCI usage status is stored in resource bases, including PCI reuse frequency, the PCI number for macro cell, and heterogeneous nodes.

(vi) Database: mainly contains related information that obtained from lower layer or indicated from other modules. It provides indispensible information for the algorithm, for example, cell state information, cell type information, and neighbor list.

2.2. Allocation Issues and Principles. Despite the existence of 504 different IDs, the actual available identities are limited to a smaller number. IDs are grouped into 168 groups and three sequential PCIs are generally corresponding to three sectors in an eNodeB. Furthermore, IDs are divided into subsets for macro-, micro-, and femtocells in order to simplify the new eNB introduction in one layer without impacting on other layers. The available PCI number for macro cell is not as redundant as expected, therefore PCI reuse is inevitable. The optimization of PCI assigning relies on the location and basic orientation of eNBs and the reuse distance should be taken into account. The objective of ADPCI is to optimize the reuse distance and minimize interference to achieve global optimum.

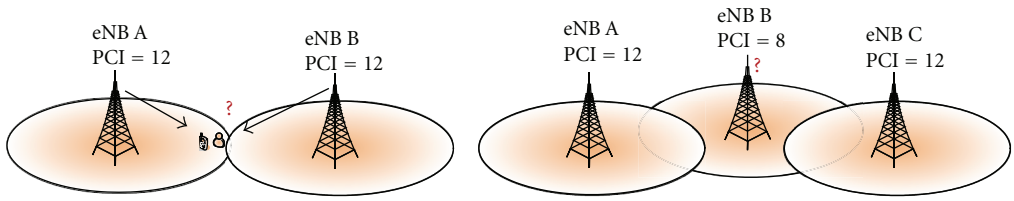
In addition, each cell identity corresponds to a unique combination of an primary synchronization signal (PSS) sequence and an secondary synchronization signal (SSS) sequence, each of which comprises of a sequence of length 62 symbols and perforce a great correlation property. However, a number of SSS sequences are high correlated to others, which affects UE to recognize the target cell. The cross-correlation properties are analyzed in the next section. On the other hand, the automatic configuration is specified to meet the requirements of collision and confusion free [6]. The former means two neighbor cells cannot use identical ID; the latter one implies that one cell cannot have any two neighbor cells that assigned with identical ID, thus allowing for ID reuse by 3rd degree neighbors, as shown in Figure 2.

Based on the above factors, the principles of PCI assigning mainly include the following.



Attribute name	Supported qualifier	Read qualifier	Write qualifier
ID	M	M	-
Cell identity	M	M	M
PCI	M	M	CM
PCI list	CM	M	M

Figure 1: The framework of PCI self-configuration.



(a) PCI collision (b) PCI confusion

Figure 2: Cell identity collision and confusion.

(i) To satisfy the collision free and confusion free, identical ID reuse is allowed by 3rd degree neighbors.

(ii) SSS sequences with high cross-correlation should be widely separated.

(iii) The reuse distance should be regarded to be as far as possible.

3. PCI Self-Distributing Algorithm

3.1. Synchronization Signals Sequences Correlation Analysis. In this section, the property of synchronization code is analyzed and the result will be provided. Highly correlated IDs are grouped together based on the results, and will be distributed separately. As described in [11], each ID can be expressed by the following equation:

$$N_{\mathrm{ID}}^{\mathrm{cell}} = 3N_{\mathrm{ID}}^{(1)} + 3N_{\mathrm{ID}}^{(2)}, \quad (1)$$

where $N_{\mathrm{ID}}^{(1)}$ is indicated by SSS sequences in the range of 0–167, it represents the ID groups. $N_{\mathrm{ID}}^{(2)}$ is indicated by PSS sequences in the range of 0–2, defining the actual cell identity within a group. The SSS sequences are specially designed. It is based on M-sequences, which can be created by cycling through every possible state of a shift register. Two codes are two different cyclic shifts of a single length-31 M-sequence and alternated between slot 0 and slot 10 in each radio frame, which enables the UE to determine the 10 ms radio frame timing from a single observation of a SSS in the synchronization procedure. The code is one-to-one mapped to the physical layer identity within the group corresponding to the target eNodeB. Details of the scrambling operations are given in [11].

Based on the understanding of SSS sequences, all SSS sequences are generated and the correlations between any two of codes are estimated. Figure 3 illustrates the cross-correlation of sequence *index* = 84 in frequency domain.

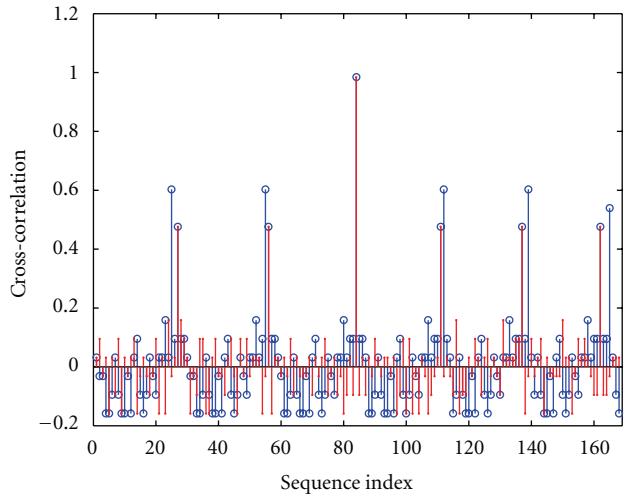


Figure 3: Cross-correlation of a pair of SSS sequences.

The blue and red stems represent the cross-correlation value in slot 0 and slot 10, respectively. The results show that most of the sequences have good properties (less than 0.2), but a few pairs are highly correlated (near 0.5). All the SSS sequences in one slot are correlated reciprocally, but not between different slots. However, both slots in a frame should be considered due to the synchronization process for cell search. As shown in the simulation result, the sequence $index = 84$ are highly correlated with sequences $\{25, 55, 84, 112, 139\}$ in slot 0 and $\{27, 56, 111, 137, 162\}$ in slot 10, respectively. Let $CRset(n)$ be the combination of sets that contains both slots for ID $index = n$, thus $CRset(84) = \{25, 27, 55, 56, 84, 111, 112, 137, 139, 162\}$. Identities from the same $CRset(n)$ are considered to be separated. The result of PCI with high correlation values are grouped in both slots, shown in Table 1.

3.2. Optimal PCI Self-Distributing Algorithm. The procedure of ADPCI is illustrated step by step in Figure 4 and cell identities sets are defined to facilitate the description of the algorithm in Table 2. The algorithm consists of three stages. In Stage I, the assigning order of eNBs will be given according to the well-known minimum spanning tree (MST) algorithm. In Stages II and III, PCI distributing and reuse methods are discussed in detail.

In the first stage, the network deployment is mapped into an abstract graph to reflect the actual network environment based on the location of nodes. The network deployment is mapped into a fully connected and undirected graph $G = (V, E)$, where vertices represent the network nodes, and edges represent the connection between vertices. Moreover, each edge is weighted with signal transmission propagation loss value. Specifically, the created graph $G = (V, E)$ is defined as follows.

(i) Define a set of vertexes $V = \{v_i\}$, where the element v_i represents the ith eNB.

(ii) Define a set of edges $E = \{e_{i,j}\}$, where $e_{i,j}$ is the edge between v_i, v_j.

Table 1: Groups of SSS sequences with high correlation.

Slot 0							Slot 10						
0	30	59	87	114	140	165	0						
1	31	60	88	115	141	166	1	30					
2	32	61	89	116	142	167	2	31	59				
3	33	62	90	117	143		3	32	60	87			
4	34	63	91	118	144		4	33	61	88	114		
5	35	64	92	119	145		5	34	62	89	115	140	
6	36	65	93	120	146		6	35	63	90	116	141	165
7	37	66	94	121	147		7	36	64	91	117	142	166
8	38	67	95	122	148		8	37	65	92	118	143	167
9	39	68	96	123	149		9	38	66	93	119	144	
10	40	69	97	124	150		10	39	67	94	120	145	
11	41	70	98	125	151		11	40	68	95	121	146	
12	42	71	99	126	152		12	41	69	96	122	147	
13	43	72	100	127	153		13	42	70	97	123	148	
14	44	73	101	128	154		14	43	71	98	124	149	
15	45	74	102	129	155		15	44	72	99	125	150	
16	46	75	103	130	156		16	45	73	100	126	151	
17	47	76	104	131	157		17	46	74	101	127	152	
18	48	77	105	132	158		18	47	75	102	128	153	
19	49	78	106	133	159		19	48	76	103	129	154	
20	50	79	107	134	160		20	49	77	104	130	155	
21	51	80	108	135	161		21	50	78	105	131	156	
22	52	81	109	136	162		22	51	79	106	132	157	
23	53	82	110	137	163		23	52	80	107	133	158	
24	54	83	111	138	164		24	53	81	108	134	159	
25	55	84	112	139			25	54	82	109	135	160	
26	56	85	113				26	55	83	110	136	161	
27	57	86					27	56	84	111	137	162	
28	58						28	57	85	112	138	163	
29							29	58	86	113	139	164	

Table 2: PCI sets for ADPCI algorithm.

Name	Meaning
PsetA	A set of all PCI range from 0 to N
PsetU	A set of all used PCI
$CRset(n)$	A Set of IDs that combine both sets where $index = n$ affiliated to in slot 0 and 10
PsetK	PCI set corresponding to the ordered list K
PsetL	eNB set corresponding to the ordered list L
PsetN	PCI set of direct neighbor
PsetNN	PCI set of neighbor of direct neighbors

(iii) Define an $m \times m$ matrix of weight values W:

$$W = \begin{pmatrix} w_{11} & \cdots & w_{1m} \\ \vdots & \ddots & \vdots \\ w_{m1} & \cdots & w_{mm} \end{pmatrix}, \quad (2)$$

where $w_{i,j}$ is the weight value of edge $e_{i,j}$. The weight values can be defined by different principles to

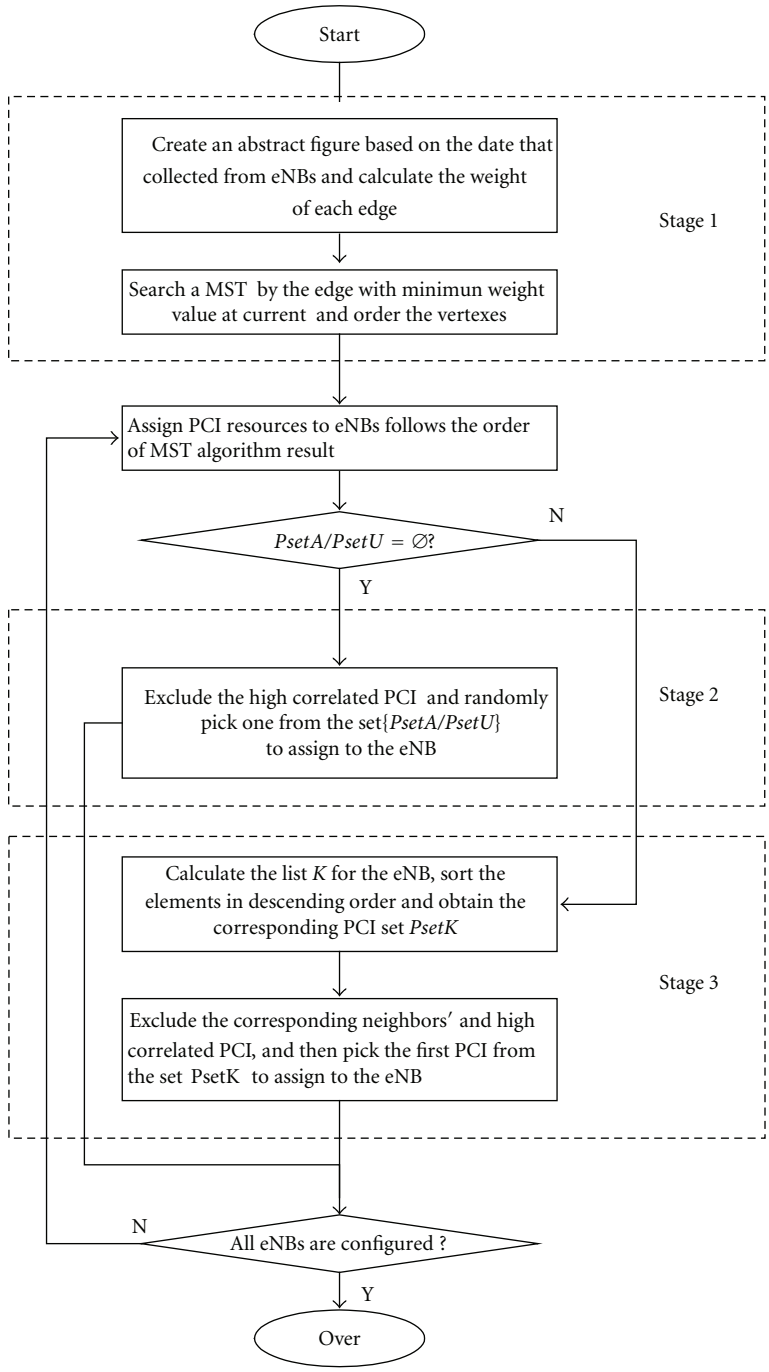


Figure 4: The Flow chart of the proposed ADPCI.

meet the specific requirements. Here we use different experiment transmission models to calculate the path loss values based on the practical network environment, $w_{i,j}$ is defined as calculated linear propagation loss value. The calculation is based on different experiment transmission models.

The MST algorithm (see the Appendix) is applied and extended to search the minimum propagation loss values. Figure 5 shows the procedure of adding edges with minimum weights and the growth of the MST. Independent trees are connected by edge with minimum weight till a spanning tree is formed. The algorithm result returns the assigning order of eNBs, allowing assignment to avoid reuse interference by using different identities.

In Stage II, PCIs are assigned to eNBs one by one following the order. Suppose that there are N IDs for the macro-eNBs, thus the top N eNBs in the MST result can acquire ID without causing reuse interference; as long as the high correlated IDs are separated.

Let m be the neighbors' identities of the ith eNB. Different PCIs are randomly picked from the unused set

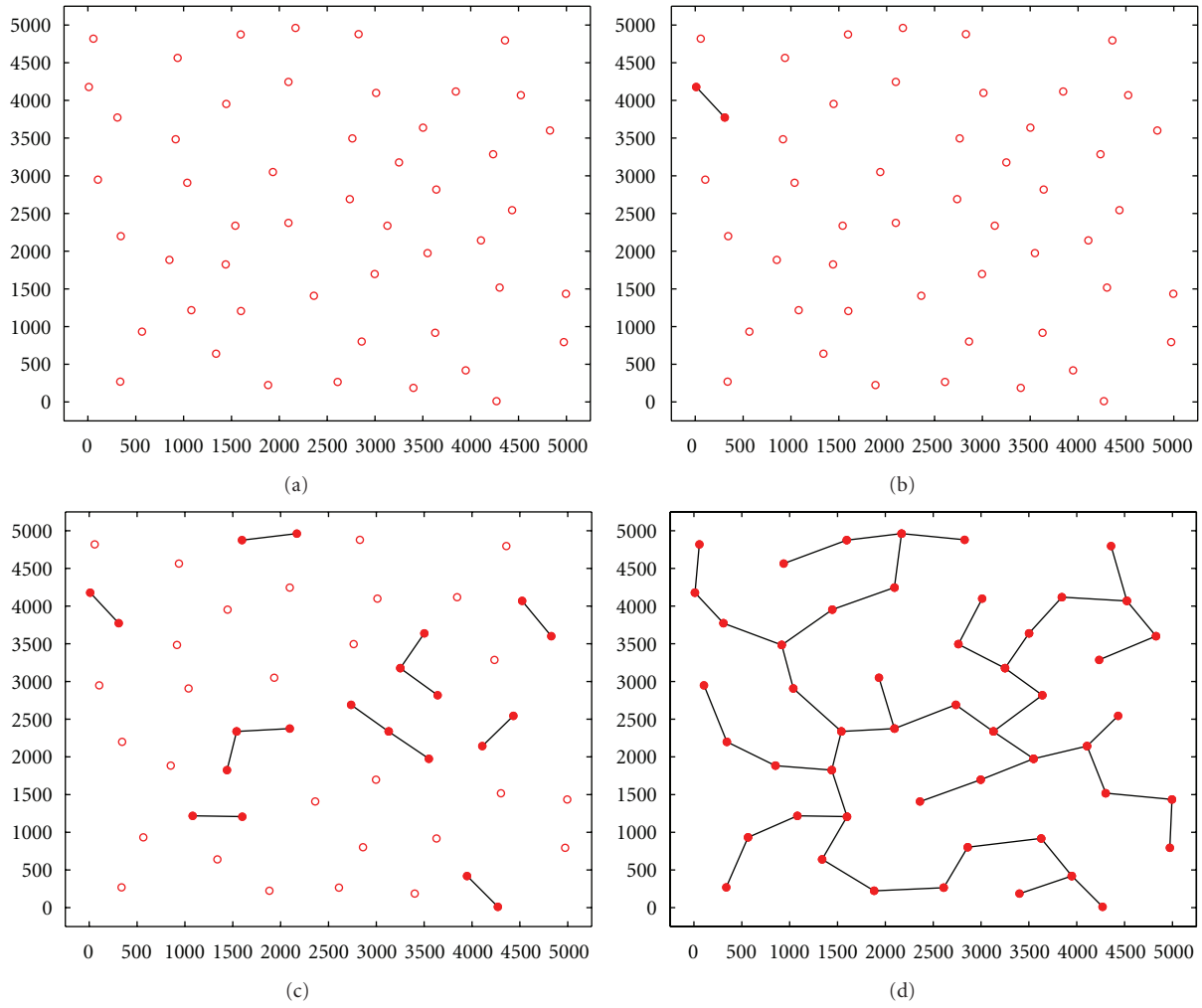


Figure 5: MST searching procedure.

$PsetA \setminus PsetU \setminus \bigcup_m CRset(m)$ and added to the used set once they have been used. In this way, the highly correlated PCIs from neighbor cells are excluded from the assigning sets to guarantee an interference-free assignment.

However, when all unused PCIs have been assigned, PCI reuse is inevitable. It is more accurate to estimate the reuse effect before actually using it. Let $k_i(n)$ be the reuse impact factor (RIF) for the ith eNB, it can be calculated and expressed by cumulative propagation loss values as follows:

$$k_i(n) = \sum_{j=1}^{d} w_{i,j}, \tag{3}$$

where $w_{i,j}$ is the linear propagation loss value from interfered eNB_j to reference eNB_i. The parameter d is the number of eNB that assigned with identical value $index = n$ at the current state. In this way, the RIFs of all IDs can be expressed, a column list K is composed as follows:

$$K = \begin{pmatrix} k_i(1) \\ k_i(2) \\ \vdots \\ k_i(N) \end{pmatrix}, \tag{4}$$

where $k_i(n)$ with different PCI indexes are calculated and included in the list, the length of K equals to the number of available PCI for macro cell. The RIF quantifies the interferences of each PCI and helps to select the optimal PCI, shown in Figure 6. Next we sort elements in list K in descending order. The set $PsetK$ is the PCI set corresponding to the ordered list K. The results return an ordered list of elements from high to low. The higher $k_i(n)$, the less serious eNB_i can be interfered. The return result may contain IDs used by neighbors and highly correlated resources, which need to be fixed to avoid

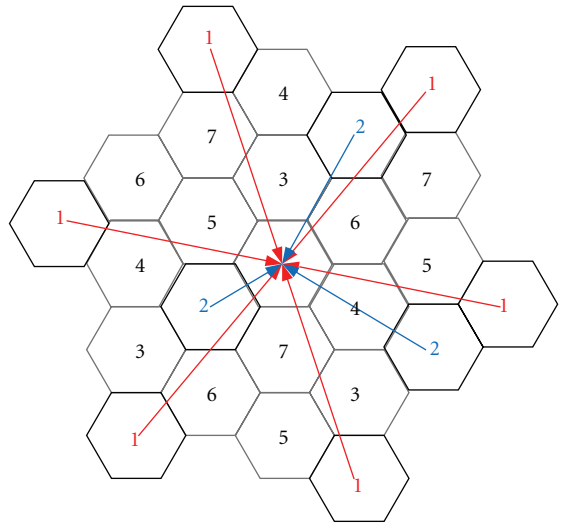


Figure 6: Calculation of RIFs ($k_i(1)$ and $k_i(2)$) for different IDs.

collision confusion. The main steps are presented as follows:

(i) If $PsetK_i \setminus \{PsetN_i \cup PsetNN_i\} \setminus \bigcup_m CRset(m)$ is not empty, pick the first element from it and assign it to the eNB.

(ii) If $PsetK_i \setminus \{PsetN_i \cup PsetNN_i\} \setminus \bigcup_m CRset(m) = \varnothing$, use the first element from set as $PsetK_i \setminus \{PsetN_i \cup PsetNN_i\}$ suboptimal solution.

(iii) Otherwise, pick the first element from set $PsetK_i$ and assign it to eNB_i.

The calculation of RIF helps eNB to evaluate the potential multiplexing interference and to make the best decision of assigning at current state. Meanwhile, removal of neighbors' PCIs can be done to avoid collision and confusion.

3.3. Suboptimal Algorithm for PCI Self-Distributing. The ADPCI provides an optimal algorithm for self-distributing, however, the high-performance causes high computational complexity. In this section, a suboptimal solution with relatively low computational complexity is given.

Define a concept of use frequency factor (UFF) for each PCI as it indicates the identity reuse times, let $U = \{u_1 u_2 \cdots u_N\}$ be the set that contains all UFFs for every PCI. In each looping, the PCI corresponding to $\min U$ is preselected, then estimate the reuse interference for every eNBs. The RIF of PCI $index = n$ for every unassigned eNBs can be calculated by (3), and a list L is defined as follows:

$$L = \begin{pmatrix} k_1(n) \\ k_2(n) \\ \vdots \\ k_i(n) \end{pmatrix}, \tag{5}$$

where $k_i(n)$ is the RIF value of eNB_i assigning with PCI $index = n$. In this case, the ID $index = n$ is the constraints and the eNBs i is the variable. Sort list L in descending order; the elements from the head of the list have a higher cumulative propagation loss value and trend to be influenced by multiplexing interferences less than that in the tail. The result may also need to exclude IDs used by neighbors and the highly correlated IDs. The set $PsetL$ is the eNB set corresponding to the ordered list L. The main steps for selecting the appropriate eNB are presented as follows.

(1) Select a PCI $index = n$ with minimum UFF when there is unused identity.

a. Pick an eNB if it satisfies $\{PsetN_i \cup PsetNN_i\} \cap CRset(n) = \varnothing$ and assigned with ID $index = n$. Otherwise, pick other eNBs.

b. If there is no eNB that can satisfy $\{PsetN_i \cup PsetNN_i\} \cap CRset(n) = \varnothing$, randomly pick an eNB to be assigned with ID $index = n$.

(2) Select a PCI $index = n$ with minimum UFF when all $u_n > 0$, calculate the RIF for every unassigned eNBs by (3) and sort list L in descending order.

a. Pick the first element from top to bottom of the set $PsetL$ that satisfies the PCI n and its highly correlated IDs are not in this eNB_i's neighbor list, $\{PsetN_i \cup PsetNN_i\} \cap CRset(n) = \varnothing$

b. If no eNB can meet the requirement above, pick the suboptimal selection that satisfies $\{PsetN_i \cup PsetNN_i\} \cap \{n\} = \varnothing$.

c. Otherwise, pick the first element from set $PsetL$ for assignment.

The major difference between SADPCI and ADPCI is that PCI is preselected before estimating the interfering influence for different eNBs. The calculation for eNBs' assigning order is unnecessary and the high computation of searching MST could be omitted; moreover, during the

PCI reuse phase, the calculation for all PCI resources' RIF needs high computational complexity; however, only a few numbers of RIF for individual PCI has to be calculated in SADPCI.

4. Evaluation and Analysis

In this section, the proposed ADPCI and SADPCI are evaluated and analyzed through the performance of users' carrier to interference ratio (CIR). We consider the simulations using a densely deployed scenario where a macrodeployment with 50 eNBs and 1000 users involved, and these eNBs and UEs are randomly distributed in the area. In order to illustrate the influence of transmission propagation models, two different signal transmission models are used to reflect the environmental changes that impacted the result of PCI distribution. The propagation loss calculation is based on the COST231 Hata urban propagation model [12]:

$$\begin{aligned} P_{L,\text{urban}}(d)\text{dB} &= 46.3 + 33.9\log_{10}(f_c) - 13.82\log_{10}(h_t) - a(h_r) \\ &\quad + \left(44.9 - 6.55\log_{10}(h_t)\right)\log_{10}(d) + C_M, \end{aligned} \tag{6}$$

where $a(h_r)$ is a correction factor for the receiver height based on the size of the coverage area; C_M is 0 dB for medium-sized cities and suburbs and is 3 dB for metropolitan areas; the carrier frequency f_c is 2 GHz; the height of transmitter h_t and receiver h_r are set to 50 m. Another model is the dual-slope model, a special case of the piecewise linear model [13]:

$$P_r(d)\text{dB} = \begin{cases} P_t + K - 10\gamma_1\log_{10}\left(\dfrac{d}{d_0}\right) & d_0 \le d \le d_c, \\ P_t + K - 10\gamma_1\log_{10}\left(\dfrac{d_c}{d_0}\right) - 10\gamma_2\log_{10}\left(\dfrac{d}{d_c}\right) & d > d_c, \end{cases} \tag{7}$$

where P_t is the transmit power; K is a constant path-loss factor; γ_1 is the path-loss exponent above some reference distance d_0 and up to some critical distance d_c, after which power falls off with path-loss exponent γ_2.

Furthermore, the available identity number N has been reduced to only 6 in order to make the solution more compelling. To compare with this extreme situation, 30 IDs are used in simulation as well. The system parameters in simulation are presented in Table 3.

To demonstrate the proposal superiority, different schemes are used in simulation for comparison. A randomly distributing scheme (RDPCI) has been introduced to assign eNB randomly; it represents the worst situation of assignment without any optimization. Besides, a hypergraph coloring PCI assigning scheme (HCPCI) uses the neighbor's relationship degree N to indicate the degree of neighbors' relationship, which is regarded to be the reuse distance. ENBs with $N \le 7$ are too close to reuse PCI, only when $N \ge 7$, PCI are allow to be reuse. The R-PCIS offers the lower bound of assigning property while the HCPCIS reduces reuse interference more effectively.

Table 3: Simulation parameters.

Parameters	Assumption
Cellular layout	Randomly distributed, 50 eNBs
User layout	Randomly distributed, 100 users/eNB
Cell radius	288 m
BS Transmit Power	46 dBm
Carrier frequency	2 GHz
Number of PCI	6 OR 30
Macro propagation models	Model no. 1 Piecewise Linear or Model no. 2 COST231 Hata

Figure 7 shows the performance comparisons among the four schemes when the number of PCI N varies from 6 to 30. The simulation results show that there are huge performance gaps between ADPCI and RDPCI. The gaps are becoming larger when the number of PCIs are richer, for example, the average CIR (CDF = 0.5) between ADPCI and RDPCI varies from 4 dB to 9 dB when PCI number changes from 6 to 30. Because of the PCI deficiency, the multiplexing interference increased inevitably with the densely distribution of PCI. The proposed ADPCI always performs better than the other schemes for achieving higher users' CIR due to the optimized reuse distance in PCI configuration; the curve of SADPCI is close to ADPCI and the performance is approximate to ADPCI. The mainly gaps between ADPCI and SADPCI is the gain of applying MST algorithm because the reuse interference can be avoided by unused PCI. With the unused PCIs increase from 6 to 30, the gap is becoming larger as the MST effect is more significantly.

Figure 8 shows the CDF curves of CIR where different propagation loss models are used. The choice of models should be based on the practical environment; the selection also affects the calculation of weight, thus influences the algorithm results. Although varied models have shown mixed results where the COST231 Hata urban propagation model performs a sharper slope of CIR curve and Broken-Line model has a smooth one, the performance of different schemes are obvious. However, the results suggest that the performances of ADPCI and SADPCI over the others schemes are similar to that of previous simulation.

5. Conclusions

The ADPCI is presented to improve the performance of PCI management by greedy search to achieve the global optimum. The mechanism benefits from two aspects: utilizing the grouped PCI resources to avoid multiplex interference; decreasing reuse interference through estimating the reuse influence of different IDs. The SADPCI has a similar property as ADPCI but needs less computational complexity. The performance of the ADPCI and SADPCI are evaluated and compared to traditional scheme in initial rollout macro

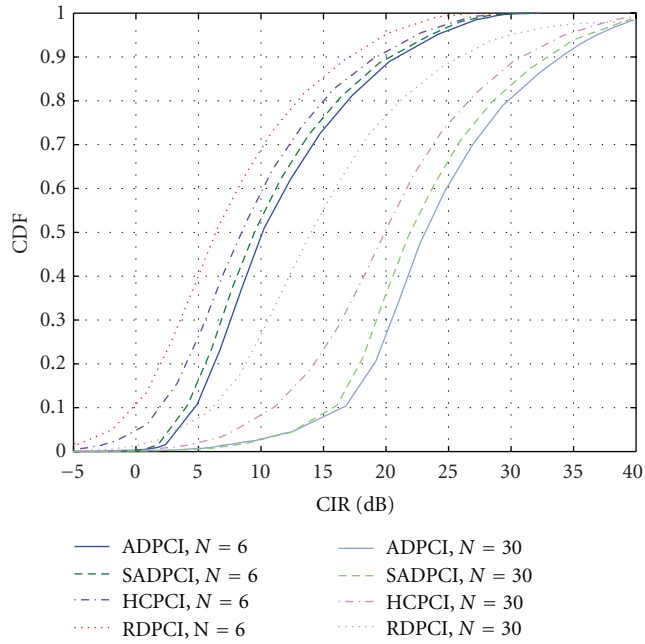


Figure 7: Users' CIR distribution when PCI number changes.

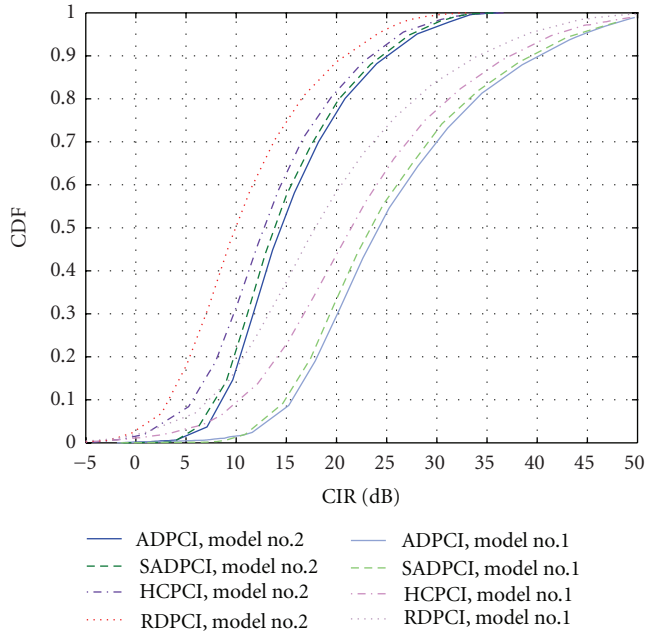


Figure 8: Users' CIR distribution when different models are applied.

scenario. As the results shown, the proposed ADPCI achieves higher users' CIR than the other schemes in condition of applying different PCI numbers and different signal transmission models.

Appendix

A. Minimum Spanning Tree (MST) Algorithm

The main MST algorithm is stated in the following steps.

(1) Create a forest F (a set of trees), where every vertex in the graph is an individual separate tree.

(2) Create a set E_{mst} that initially contains no edge in the graph.

(3) While F is not yet a spanning tree, operate as follows:

a. add an edge with minimum weight to the set E_{mst} from all valid edges;

b. if that edge connects two different trees, add it to the forest and combine two trees into a single one;

c. otherwise, discard that edge from E_{mst} and keep looping until the graph has only one component and forms a minimum spanning tree.

At the termination of the algorithm, the forest has only one component and forms a minimum spanning tree of the graph.

Acknowledgments

This work was supported by the State Major Science and Technology Special Projects (Grant no. 2011ZX03003-002-01), the Fok Ying Tong Education Foundation Application Research Projects (Grant no. 122005), and the Program for New Century Excellent Talents in University.

References

[1] M. Peng and W. Wang, "Technologies and standards for TD-SCDMA evolutions to IMT-advanced," *IEEE Communications Magazine*, vol. 47, no. 12, pp. 50–58, 2009.

[2] M. Peng, Y. Liu, D. Wei, W. Wang, and H. H. Chen, "Hierarchical cooperative relay based heterogeneous networks," *IEEE Wireless Communications*, vol. 18, no. 3, pp. 48–56, 2011.

[3] H. Akhavan et al., "Next Generation Mobile Networks—Beyond HSPA EVDO—Whitepaper Tech," Tech. Rep., NGMN Ltd, 2006, http://www.ngmn.org/.

[4] Y. Zhou, J. Wang, and M. Sawahashi, "Downlink transmission of broadband OFCDM systems—part I: hybrid detection," *IEEE Transactions on Communications*, vol. 53, no. 4, pp. 718–729, 2005.

[5] H. Zhu and J. Wang, "Chunk-based resource allocation in OFDMA systems—part I: chunk allocation," *IEEE Transactions on Communications*, vol. 57, no. 9, pp. 2734–2744, 2009.

[6] 3GPP TR 36.902, V9.3.1, "Self-configuring and self-optimizing network (SON) use cases and solutions," December 2011.

[7] R3-080376, "SON Use Case: Cell Physical ID Automated Configuration," 3GPP RAN3 #59, 2008.

[8] R3-080812, "Configuration of Physical Cell Identity Use Case," 3GPP RAN3 #59bis, 2008.

[9] T. Bandh, G. Carle, and H. Sanneck, "Graph coloring based physical-cell-ID assignment for LTE networks," in *Proceedings of the ACM International Wireless Communications and Mobile Computing Conference (IWCMC '09)*, pp. 116–120, ACM, April 2009.

[10] H. Xu, X. W. Zhou, and Y. Li, "Model of hypergraph colouring for self-configuration in LTE networks," in *Proceedings of the 4th International Conference on Information Management, Innovation Management and Industrial Engineering (ICIII '11)*, vol. 1, pp. 393–396, 2011.

[11] 3GPP TR 36.211, V10.2.0, "Physical Channels and Modulation," June 2011.

[12] European Cooperative in the Field of Science and Technical Research EURO-COST231, "Urban transmission loss models for mobile radio in the 900 and 1800 MHz bands," rev. 2, The Hague, September 1991.

[13] E. McCune and K. Feher, "Closed-form propagation model combining one or more propagation constant segments," in *Proceedings of the 47th IEEE Vehicular Technology Conference*, pp. 1108–1112, May 1997.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 721957, 13 pages
doi:10.1155/2012/721957

Research Article

Self-Organizing Energy-Saving Management Mechanism Based on Pilot Power Adjustment in Cellular Networks

Yu Peng, Li Wenjing, and Qiu Xuesong

State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to Yu Peng, yupeng@bupt.edu.cn

Received 3 May 2012; Revised 30 August 2012; Accepted 13 September 2012

Academic Editor: Zhiguo Ding

Little literature concentrates on effective and self-organizing regional energy-saving schemes in cellular networks with dynamic traffic. In this paper, we propose a regional self-organizing energy-saving management mechanism through pilot power adjustment in cellular networks. The mechanism analyzes autonomic processes for saving energy firstly. Then, it proposes regional BS TP (Trigonal-Pair) compensation method to determine BS modes for saving energy. Next, it adopts intelligent coverage compensation algorithm to adjust pilot power of compensated BSs, so as to avoid coverage hole and ensure enough capacity. Integrated evaluation model is still proposed in the mechanism to validate efficiency of proposed algorithms. Simulations show that about 17% of regional energy consumption can be saved. In comparison to other ES methods, our ES mechanism can obtain better energy-saving outcomes and keep service quality and coverage above acceptable levels.

1. Introduction

Saving energy for cellular networks (including GSM, 3G, and LTE) contributes to reduction of greenhouse effect and curtailment of operating expense, which make sense of strong practical significance [1]. Varieties of services supplied by wireless networks require higher speed and bandwidth. And accordingly wireless access points should meet high-density deployment demands as well. In cellular networks, these points (BTS, NodeB, eNodeB, etc.) are key elements of energy consumption, which consume around 80% to 90% of telecommunication networks [2]. For providing continuous services, sizes of cellular networks are designed in accordance with peak traffic. However, during night period load of several access points is fairly low, so much resource and operation costs are wasted [3]. Therefore, saving energy through sleeping several BSs (Base Stations) is feasible. Researches for improvements of wireless cellular networks mainly include two aspects, which are hardware and software separately. Former is demand for more energy efficient BS hardware modules such as power amplifiers. Latter requires enhancement for autonomic management functions. These functions are able to sleep several BSs, and meanwhile control neighbor BSs to compensate their coverage and capacity through signaling [2, 4].

As a software level method, Energy-Saving (ES) management of cellular networks has been standardized. Definition of 3GPP is followed. Management centre switches off several BS under low traffic period, and meanwhile adjusts wireless parameters (such as pilot power and down tilt) of active BSs, so as to avoid coverage hole and guarantee users' QoS (Quality of Service) demand [5]. To avoid frequent adjustment for network parameters, ES management is categorized to self-optimization use case of SON (Self-Organizing Network). In this case, when triggering conditions is satisfied, network is able to autonomic execute energy-saving actions above without artificial interference. Thereby regional energy saving is achieved [4]. And these definitions have been considered as basic instruction for software level ES method.

This paper proposes a self-organizing pilot-power adjusting mechanism (SPAM) to save energy for cellular networks. It can obtain a better ES effect comparing to other methods and guarantee regional coverage and service quality as

well. SPAM adopts centre management manner. It proposes processes of autonomic ES management and key algorithms for coverage compensation. It still considers regional interference and introduces a more integrated energy model.

Moreover, in order to provide a clear description of SPAM, in this paper, states for network and modes for BSs are specified. States of cellular networks mainly contain Normal State (NS) and Sleeping State (SS). Under NS all the BS states still keep on Normal Mode (NM). When ES triggering conditions are satisfied, networks will transfer to SS. Under SS, BS modes may be NS, Sleeping Mode (SM), and Compensation Mode (CM) determined by regional TP compensation method.

The remainder of this paper is organized as followed. Section 2 introduces several related work of ES methods and algorithms. Processes of SPAM management are proposed in Section 3. In Section 4, regional Trigonal-Pair (TP) compensation method is introduced to resolve BS selection problem of SPAM. This method includes compensation solution for single BS and regional BS selection scheme for ES. To resolve compensation problem of SPAM, intelligent optimization algorithm for pilot-power adjustment is described in Section 5. Integrated validation model for ES method in SPAM is analyzed in Section 6. Simulation and analysis are given in Section 7. Conclusions and future work are introduced in Section 8.

2. Related Work

Currently, little literature concentrates on effective integrated regional energy-saving schemes. ES management of cellular networks faces the following problems.

Firstly, standardizations have not proposed effective solutions for saving energy in cellular networks. ES management use case has been defined, and management functions, requirements, and assessment criteria are specified as well [4, 5]. However, no processes are proposed to execute ES actions. Merely parameters which can be adjusted to achieve ES gains are referred. These standards still does not refer implementation methods or algorithms for regional energy saving. Energy consumption evaluation metrics which are required to compute energy efficiency of ES methods have been analyzed in [6], but there are not standardized as well.

Secondly, several ES schemes are just from user perspective and are not suitable for dynamic traffic. Besides, these schemes always take on high computation complexities. Simplified centralized and distributed ES algorithms are introduced in [7, 8], but only load reallocation strategy is considered. ES problems from user perspective are formulated as NP hard problems and given effective resolution algorithms in [9, 10]. But regional coverage blind ratio and service quality under ES duration have not been evaluated. Besides, how to adjust wireless parameters in order to compensate coverage and capacity is also not introduced. Schemes in the literatures are still suitable for ES at a single time point, not the entire time period.

Thirdly, current regional ES schemes are defective. As regional ES schemes mainly resolve ES problems from traffic perspective, they can be formulated much simpler and suitable for dynamic traffic distribution. Though few regional energy-saving schemes have proposed maximized mathematical models and regional compensation methods [11–14], but interference is unfortunately neglected. Besides, the BS sites are all based on theoretical deployment. Irregular BS sites under practical scenarios are not considered. Still, BS energy consumption under different modes is very simple and not accurate.

Fourthly, saving energy by deployment of smaller sizes of BSs/cells (for instance, pico-BS/cell and femto-BS/cell) is not suitable for operation network. Though adoption of sleep mode for macrocell can save energy to a certain degree [15, 16], extra BS construction may bring out additional radio interference in the area. Still, extra costs of the new BSs are not economical.

Fifthly, current energy consumption model is not integrated. Static and dynamic components of BS power in wireless networks have been analyzed [17]. But no precise relationship between traffic and dynamic power part is concluded. Radius variation and its effect to transmit power are analyzed as well [18]. But only propagation model is considered. So evaluation of energy-saving efficiency may be not accurate.

Aiming to resolve the above problems, SPAM gives process of autonomic ES management mechanism, defines several evaluation metrics to validate effect of ES methods, proposes regional ES method for dynamic traffic suitable for operational networks, and analyzes power consumption model of BS.

3. Processes of Self-Organizing ES Management Mechanism

SPAM proposes self-organizing processes for ES management. The management centre executes ES mechanism according to these processes. These processes include five self-organizing stages (self-monitoring, self-analysis, self-planning, self-execution, and self-evaluation), which is shown in Figure 1. Each step is controlled by a centralized management centre.

Detailed description is described as follows.

(1) Monitoring regional traffic and power variation. Theoretic capacity of cellular networks depends on number of CE (Channel Element) resource. Thus, occupied CE number is adopted as an important evaluation metric for regional resource usage. Still, regional traffic is computed through the model in [12] when arrival rates and average service time for each service is known.

(2) When network is under NS, if ES triggering conditions are satisfied, which mean regional traffic is lower than ES triggering threshold and duration is longer than buffering time, ES processes are activated and transferred to (4) or else turned back to (1).

(3) When network is under SS, if recovery conditions are satisfied, which mean regional traffic is higher than

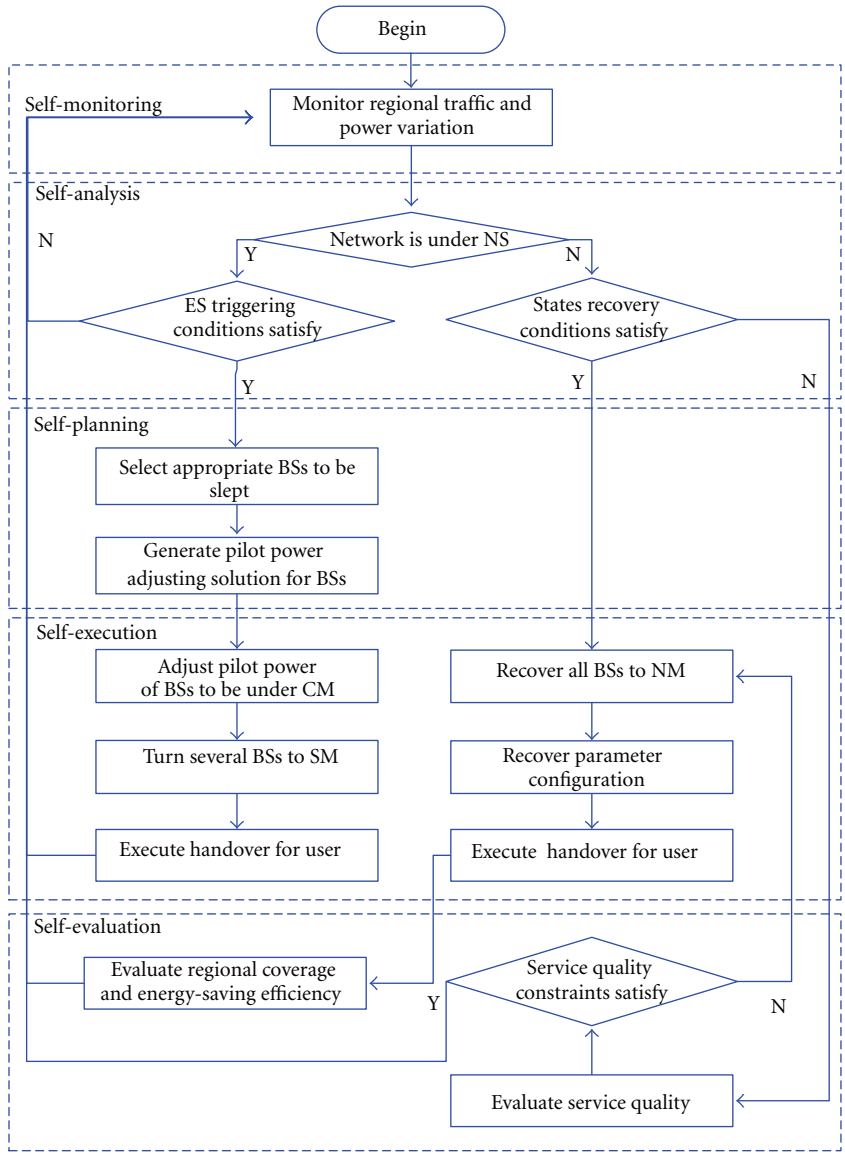


Figure 1: Processes of SPAM.

ES recovery threshold and duration is longer than buffering time, recovery processes are activated and transferred to (10)otherwise turned to (9).

(4) According to deployment of BSs, regional TP compensation method determines different BS model based on practicable local compensation.

(5) For BSs which will be to CM, intelligent coverage optimization algorithm is adopted to obtain adjusting values of pilot power. The purpose is guaranteeing regional coverage range and quality and maximizing saving energy of BSs at the triggering time point. Steps (4) and (5) are two key issues of SPAM.

(6) Set pilot power of BSs which will be turned to CM according to the adjusting values in (5).

(7) Set modes of BSs which will be turned to SM to sleep mode with low power to keep them controllable.

(8) Execute handover for users serviced by BSs under SM to appropriate BSs and CM. Network goes to SS and return to (1).

(9) When network is under SS, evaluate service quality (service block probability is taken as an important indicator here), if service quality constraints are satisfied, go back to (1) or else go to (10).

(10) Turn on all the BSs to NM, and retrieval parameters values of all the BSs to original ones;

(11) Execute handover for user whose service is downgraded to appropriate BSs. Network goes to NS.

(12) Evaluate regional coverage and energy saving ratio and go back to (1).

In analysis stage, setting of buffering time aims at avoiding frequent control and handover caused by rapid traffic variations. Traffic information for quantities of BS should be obtained and necessary handovers are required for ES management. Moreover, each self-organizing stage should be executed according to the conditions of the whole network. Hence, centralized management function can be deployed in current network management centre. Here, we use OAM (Operation Administration and Maintenance) systems to fulfill these functions. In self-execution stage, OAM systems control BSs and other important network nodes to complete transfer of BS modes and handover for users. Cooperations among BSs are essential. Meanwhile, in order to rapidly recover network states, OAM systems must record parameter values of controlled BS under NM.

4. Regional TP Compensation Method

Ideally coverage of each BS is circular and each cell is hexagonal. One BS may serve for one or multiple cells. Referring to compensation coverage method through adjustment radiuses of neighbor BSs in [12] and considering symmetry of BS and cells, a theoretical TP under regular BS deployment is proposed in [14]. In order to improve current work, regional TP compensation method proposes local TP compensation scheme, and TP-based regional BS selection scheme suitable for irregular BS deployment. Objective of this method is to determine BS modes when network is under SS.

4.1. Local TP Compensation Scheme. From local perspective, coverage of single BS can be compensated by neighbor BSs. An ideal TP example for BS with steerable antenna (often deployed in urban area) can be seen in Figure 2. BS pair (BS_1, BS_2, BS_3) is called a TP of BS_0. TPs of each BS make up of a set *TP*.

For practical situation with asymmetric deployments and limited resource, more general TP compensation scheme is shown in Figure 3. Assume that radius of BS_0 is r_0. And traffic is uniformly distributed in the ES area. Assume that BS set which storage BSs to be composed of TP is denoted as NB. Initially we set NB = Φ.

Besides, TPs of BS_0 are selected as follows.

(1) Find corresponding six BSs with strongest signal strength from neighbor list of BS_0 and then put these BSs to set NB.

(2) Select three nonadjacent BSs from NB to make up of a triangle. If every interior angle θ_i satisfies constraints that $45^\circ \le \theta_i \le 75^\circ$, and load of each BS is under individual ES triggering threshold, then select BSs constitute a TP for BS_0, such as TP (BS_2, BS_4, BS_6) in Figure 3. Put TP into set *TP* for BS_0.

(3) For other BSs in NB, repeat the operation in (2) and attempt to find another TP.

For BS with omnidirectional antenna, local TP compensation scheme is same as BS with steerable antenna, which will be ignored here.

For multiple choices exist for local TP compensation of single BS, it may bring about additional coverage effect for other BSs. So as to optimize regional coverage quality and energy efficiency, regional TP selection scheme is required to maximize coverage effectiveness of BS under CM.

4.2. Regional TP Selection Scheme. Assume that B represents set of BSs in ES area and S represents modes of BSs. Still assume $s_i = 0$, 1, and 2 mean NM, SM, and CM, separately. Then regional TP selection scheme is described in Algorithm 1. n_{ij}, o_{ij}, and c_{ij} represent numbers of BSs under NM, SM, and CM in tp_{ij}, separately. Let the Normal Ratio (NR), Compensating Ratio (CR), and Sleeping Ratio (SR) are ratios of different BS modes from the above algorithm. Assume that NBSs exist in the network. It is easy to compute that complexity of the algorithm is $O(N)$.

In Algorithm 1, initially each BS is under NM. And we only consider BS with TPs. Still, only TPs without BS under SM can be taken as effective TPs to compensate BS under NM. Moreover, in order to maximize compensation effect for BS under CM, for effective TPs of each BS, the one including maximal number of BS under CM will be chosen. After execution of this algorithm, we can obtain maximal number of BS under SM and minimal number of BS under CM.

However, settings for triggering and recovery conditions still should be determined in order to obtain n_{ij}, o_{ij}, and c_{ij}. Firstly, assume that regional ES triggering and recovery thresholds are equal and both denoted as δ. According to conclusions in [11–14], regional traffic varies are almost periodic. And the cycle is always a week. Assume regional peak traffic (monitored by OAM system) during a cycle is $\mathrm{Tr}_{\max}$, then we initially set as $\delta = 2\mathrm{Tr}_{\max}/3$. And after each period, $\delta' = \mathrm{Tr}_{\max} \cdot \min\{\delta/\mathrm{Tr}_{\max}, \mathrm{NR}+\mathrm{CR}\}$ will be taken as a new threshold. Thus, thresholds can be adjusted dynamically. ES triggering and recovery buffering time can be set as 10 minutes to avoid frequency handover for the network and guarantee enough time for analysis stage.

5. Intelligent Coverage Optimization Algorithm

In cellular networks, coverage is determined by pilot signal (here including BCCH in GSM, CPICH in UMTS, and PDCCH in LTE). When network is under SS, in order to compensate coverage for BS under SM, pilot power adjustments for BS under CM are required. Still, at the triggering time, it is a complex combinatorial optimization problem to adjust pilot power for BSs under CM with coverage and capacity constraints. Intelligent optimization algorithm is required. Mathematical model of regional coverage compensation is discussed in this chapter firstly. Then intelligent algorithm adopted to resolve this model is then introduced.

5.1. Mathematical Model. Assume ES triggering time point is t^*. At time t $M(t)$ mobile users and K services are supplied in the network. Each BS installed with steerable antenna and serves for three sectors. Still, only one carrier is configured in

(1) Input: B, S
(2) Output: S
(3) **while** $B \neq \Phi$ **do**
(4) *for* $\forall b_i \in B$
(5) **if** $s_i = 0$, **then**
(6) *find the set of opposite pairs* TP_i
(7) **while** $TP_i \neq \Phi$ **do**
(8) *for* $\forall tp_{ij} = (b_k, b_l, b_m) \in TP_i$ *count* n_{ij}, o_{ij} *and* c_{ij}
(9) $TP_i = TP_i - tp_{ij}$
(10) **end while**
(11) **if** *such* j^* *exists where* $tp_{ij^*} = (b_{k^*}, b_{l^*}, b_{m^*})$, $o_{ij^*} = 0$, and $j^* = \arg\max\{c_{ij}\}$, **then**
(12) $s_i = 1, s_{k^*} = 2, s_{l^*} = 2, s_{m^*} = 2$
(13) $B = B - b_{k^*} - b_{l^*} - b_{m^*}$
(14) **end if**
(15) **end if**
(16) $B = B - b_i$
(17) **end while**

Algorithm 1: Regional TP selection scheme.

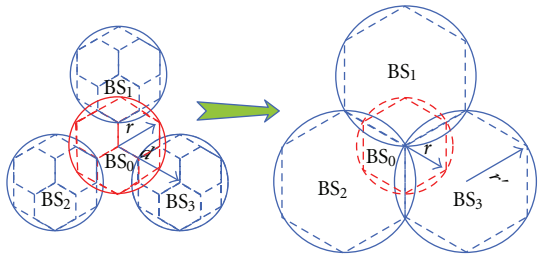


Figure 2: Ideal TP compensation for BS with steerable antenna.

each sector. And $d_{ij}(t)$ denotes distance between user i and BS j. $r_j^d(t)$ and $r_i^u(t)$ represent downlink coverage radius of BS j and uplink coverage radius of user i separately. $c_{ijk}(t)$ denotes service state between user i and BS j for service k determined by traffic model. Transmit power of user equipment i is $p_{tx}^i(t)$. $P_j(t)$, $P_{Tx}^j(t)$, and, $P_{\mathrm{BS}}^j(t)$ separately represent pilot power, transmit power, and entire power for BS j at time t. $p_{ik}(t)$ denotes minimal TCH (Traffic Channel) power of service k for user i at time t. And ν_k denotes CE number required for service k. Still, An optimal pilot power vector $\mathbf{P} = [P_1, P_2, \ldots, P_N]$ is required in order to achieve regional coverage compensation objective at time t^*.

$P_{Tx}^j(t)$ in cellular networks consists of the following parts [19]:

$$P_{Tx}^j(t) = (1+\alpha) \cdot P_j(t) + (1+\gamma) \cdot \sum_{k=1}^{K} \sum_{i=1}^{M(t)} p_{ik}(t) \cdot c_{ijk}(t), \tag{1}$$

α denote ratios of other control channels comparing to pilot power. γ is soft handover factor. In cellular networks offset between $p_{ik}(t)$ and $P_j(t)$, which is called o_k, is constant and is shown in the following:

$$p_{ik}(t) = \sum_{j=1}^{N} P_j(t) \cdot c_{ijk}(t) + o_k. \tag{2}$$

Still, components of $P_{\mathrm{BS}}^j(t)$ (in Watt) are described in (3):

$$P_{\mathrm{BS}}^j(t) = n_{\mathrm{sector}} \left[n_{Tx} \left(\frac{P_{Tx}^j(t)}{\eta} + P_{\mathrm{Trans}} \right) + P_{\mathrm{Proc}} + P_{\mathrm{Rect}} \right] + P_{\mathrm{micro}} + P_{\mathrm{airco}}, \tag{3}$$

n_{sector} represents sectors served in BS. n_{Tx} and η denote antenna number in each sector and efficiency of power amplifier. P_{Trans}, P_{Proc}, P_{Rect}, P_{micro}, and P_{airco} are power of the transceiver, the digital signal processing, the rectifier, the microwave link, and the air conditioning [17].

It is true that executing ES algorithms and schemes always put additional computation and management burden of the system, and energy consumption may increase as well. However, in our mechanism, these algorithms and schemes are mainly executed in OAM system. Number of these nodes is fairly lower than number of BSs, so their energy consumption is much lesser than BSs. Besides, as we adopt algorithms and schemes with low computation complexity, their additional energy consumption is inappreciable comparing to energy-saving gains for BSs. Considering that this additional energy is minor and hard to be quantified, we just ignore it here.

Relation f between $P_{\mathrm{BS}}^j(t)$ and $P_j(t)$ for BS j can be obtained from the above three equations. From perspective

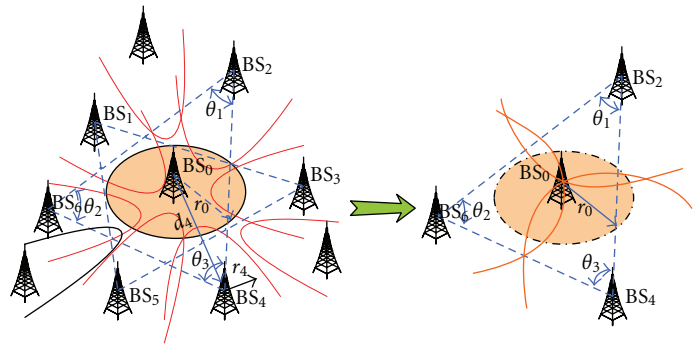


Figure 3: General TP compensation for BS with steerable antenna.

of saving energy, minimal regional consumed power should take the following objective at time t^*:

$$\min G(\mathbf{P}) = \sum_{j=1}^{N} P_{\mathrm{BS}}^{j}(t^*) = \sum_{j}^{N} f\left(P_j(t^*)\right). \tag{4}$$

It is easy to find that $G(\mathbf{P})$ is a increasing function for vector $\mathbf{P}$.

Assume that minimal acceptable signal strength of BS receiver is P_{au}. Then for user i, maximal uplink path-loss $L_u^i(t^*)$ at time t^* is shown as follows:

$$L_u^i(t^*) = p_{tx}^i(t^*) - P_{\mathrm{au}}. \tag{5}$$

Still, assume that minimal acceptable pilot power strength of edge user is P_{ad}. Then for BS j, maximal downlink path-loss $L_d^j(t^*)$ is shown as follows:

$$L_d^j(t^*) = P_j(t^*) - P_{\mathrm{ad}}. \tag{6}$$

According to propagation models decided by different scenarios, mapping relation between path-loss L and coverage radius r is shown as follows [20] $b(\cdot)$ is a universal mapping function and suitable for (5) and (6):

$$r = b(L). \tag{7}$$

From (5), (6), and (7), relations between $P_j(t^*)$ and $r_j^d(t^*)$, $p_{tx}^i(t^*)$, and $r_i^u(t^*)$ can be obtained and denoted as $g(\cdot)$.

Notice that from (1) to (7) is suitable for BS under CM and NM. For BS under SM, only a little power is consumed in order to keep it controllable, which is denoted as P_s.

From perspective of coverage, minimal regional coverage gap is required:

$$\min H(\mathbf{P}) = 1 - \frac{\left[\sum_{j=1}^{N} \pi g^2\left(P_j(t^*)\right) - \sum_{l=1}^{N}\sum_{q=1}^{N} O_{lq}\right]}{A} \times 100\%. \tag{8}$$

In (8), $H(\mathbf{P})$ denotes coverage gap ratio in ES area. A and O_{lq} represent entire ES area and overlap area between BS l and BS q. It is can be easily proved that $H(\mathbf{P})$ is a decreasing function for vector $\mathbf{P}$. Then $H(\mathbf{P})$ and $G(\mathbf{P})$ are contradictory. However, coverage requirement should be guaranteed when ES actions are executed. So $H(\mathbf{P})$ can be taken as a constraint for $G(\mathbf{P})$. Based on the above analysis, when user i is on service, values for item $c_{ijk}(t)$ are shown as follows:

$$c_{ij^*k}(t) = \begin{cases} 1, & \forall i, j, k, d_{ij^*}(t) \le r_j^d(t) \le r_i^u(t), \\ 0, & \text{else.} \end{cases} \tag{9}$$

In order to guaranteeing resource, transmit power and coverage gap constraints, following requirements should be satisfied for $G(\boldsymbol{P})$:

$$\text{s.t.} \begin{cases} H(\mathbf{P}) \le \beta \\ \forall i, k, \sum_{k=1}^{K}\sum_{j=1}^{N} c_{ijk}(t^*) = 1 \\ \forall j, P_{Tx}^{\min} \le P_{Tx}^{j}(t^*) \le P_{Tx}^{\max} \\ \forall j, \sum_{k=1}^{K}\sum_{i=1}^{M(t^*)} v_k \cdot c_{ijk}(t^*) \le \left(1 - \omega_j\right) V_j. \end{cases} \tag{10}$$

In (10), β is upper limit for coverage gap. $P_{Tx}^{\min}$ and $P_{Tx}^{\max}$ separately denote the minimal and maximal values of BS transmit power. V_j is available CE number of BS j, and ω_j denotes margin ratio in V_j for soft handover and inference.

5.2. Simulated Annealing Algorithm for Coverage Compensation. Assume that P_T is value set for $\mathbf{P}$, and then P_T is a real value space. It's easy to prove that problem described above is NP-hard. As a general random search algorithm, Simulated Annealing (SA) algorithm can effectively resolve NP-hard problems. Still, its advantages include avoidance of local optimization, independence of initial value and theoretical global optimization [21]. Though convergence time of this algorithm is a little low, but our problem is not time sensitive. So this algorithm is suitable for above model.

Based on analysis in last section, natural number coding will be adopted to represent state for $\mathbf{P}$. Noticed that SA algorithm is suitable to resolve problem without constraints,

so above model should be converted to the following unconstraint model.

$$\begin{aligned}\min z &= G(\mathbf{P}) \\ &+ \left(\frac{\sigma}{T_k}\right)\sum_{j=1}^{N}\left[\sum_{P_{Tx}^{j}(t^*)\in P_U}\left(P_{Tx}^{j}(t^*)-P_{Tx}^{\max}\right)^2\right. \\ &\left.+\sum_{P_{Tx}^{j}(t^*)\in P_L}\left(P_{Tx}^{\min}-P_{Tx}^{j}(t^*)\right)^2\right] \\ &+ \left(\frac{\chi}{T_k}\right)\sum_{j=1}^{N} sg^2\left(\sum_{k=1}^{K}\sum_{i=1}^{M(t^*)} v_k\cdot c_{ijk}(t^*)+\omega_j V_j - V_j\right) \\ &+ \left(\frac{\varsigma}{T_k}\right)\left|H(\mathbf{P})-\beta\right|^2\end{aligned} \tag{11}$$

σ, χ and ς are separately penalty factors for BS transmit power, BS capacity and coverage. T_k is temperature of SA. $P_U = \{y \mid y > P_{Tx}^{\max}\}$ denotes number set which exceeds maximal transmit power of BS. $P_L = \{y \mid y < P_{Tx}^{\min}\}$ denotes number set which is lower than minimal transmit power of BS. $sg(x)$ is a piecewise nonnegative function and shown as followed:

$$sg(x) = \begin{cases} 0, & x \le 0, \\ x, & x > 0. \end{cases} \tag{12}$$

Penalty will enlarge along with temperature decrease. Thus in the initial stage global search will be executed, and local search will be obtained in the final stage.

Procedures of SA algorithm for coverage compensation are shown in Figure 4. The detailed steps are shown followed.

Step 1. Choose an initial solution $\mathbf{P} \in P_T$, give initial temperature T_0, and terminal T_f. Set iterative metrics $k = 0$, $T_k = T_0$. Specify inner loop time $n(T_k)$ and set inner loop counter $n = 0$.

Step 2. Randomly generate a neighbor solution $\mathbf{P}' \in N(\mathbf{P}) \in P_T$. $N(\mathbf{P})$ denotes neighbor solution set of $\mathbf{P}$. $\mathbf{P}'$ can be obtained through change value of a random item in $\mathbf{P}$. Set $n = n+1$ and compute increment of objective named Δz and $\Delta z = z(\mathbf{P}') - z(\mathbf{P})$;

Step 3. If $\Delta z < 0$, let $\mathbf{P} = \mathbf{P}'$ and go to Step 4; otherwise generate $\xi = U(0,1)$, if $\exp(-\Delta z/T_k) > \xi$, then let $\mathbf{P} = \mathbf{P}'$.

Step 4. If heat balance is achieved, which means $n > n(T_k)$, go to Step 5; otherwise go to Step 2.

Step 5. Decrease T_k and let $k = k + 1$. If $T_k < T_f$, terminate the algorithm, otherwise reset $n(T_k)$ and let $n = 0$; go to Step 2.

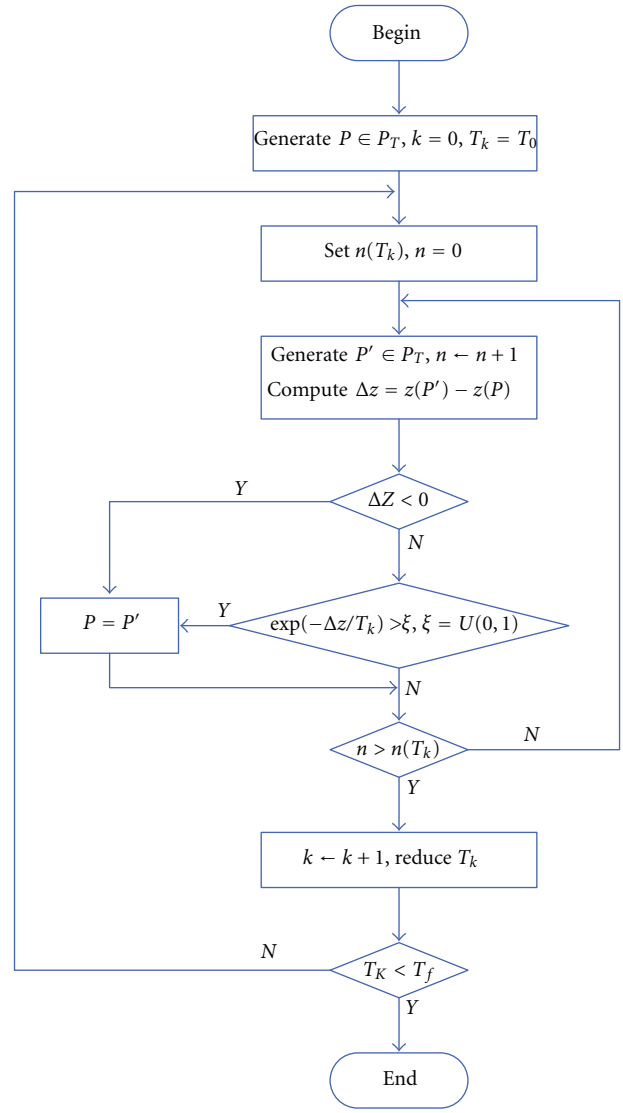


Figure 4: Procedures of SA algorithm for coverage compensation.

However, SA algorithm may accept several bad solutions, so the final solution may worse than best solution. Thus, best solutions during the computed process should be saved.

The pilot power of each BS will adjust to the value generated by this SA algorithm. Thus, regional coverage and capacity compensation are achieved.

6. Integrated Evaluation Model

Algorithms in Sections 4 and 5 mainly aim to time point t^*. And when network is under SS, BS modes and pilot values will keep on the outcomes generated by these algorithms. As shown in processes of SPAM, in order to verify ES outcome, we should evaluate the service quality during the ES interval. Moreover, regional coverage and energy efficiency still should be evaluated when network returns back to NS. As a tradeoff, ES gain is obtained through several scarifications for other properties. In this paper, in order

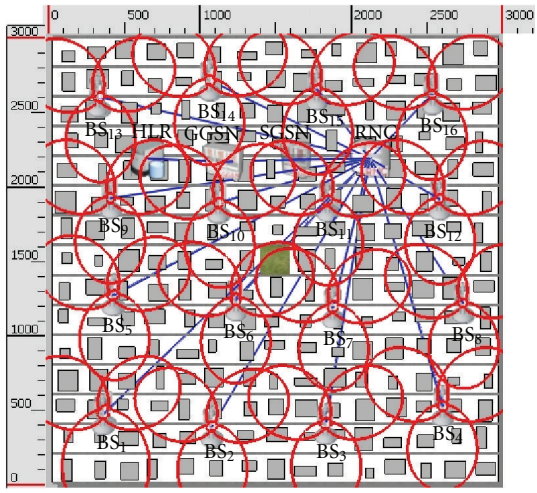


Figure 5: Top view of simulation scenario.

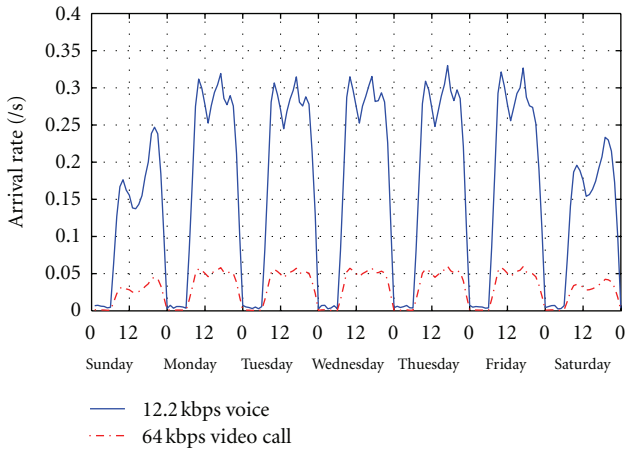


Figure 6: Average arrival rate for the two services for one cell.

to compensate coverage, capacity, and quality for BS under SM, we adjust pilot power of neighbor BS to accommodate coverage and capacity. That is, power and resource of BS under CM will be sacrificed. In order to verify the effeteness of our mechanism, we should explore the scarifications and gains from perspectives of service quality, regional coverage, and energy efficiency.

6.1. Service Quality Evaluation. In cellular networks, service blocking probability is an important indicator of service quality [12], which will be adopted here to evaluate the effect of SPAM.

Assume that traditional traffic model based on Markovian processes with multiclass $M/M/N/0$ queue is adopted as referring to [13], and required bandwidth of service k is b_i. For each cell of BS j, users generate service k according to a poisson process with rate $\lambda_{jk}(t)$ (including incoming handovers) and no queuing is possible. Mean service time for service k call is $1/\mu_k$, the mean time spent by the user in one cell is $1/\mu_h$. Then state space of one cell can be denoted by the numbers of active users in each service, $n_k(t)$ for service k at time t, to a vector $\mathbf{q}(t) = (n_1(t), n_2(t), \ldots, n_K(t))$. Then possible states for the cell is given by $Q(t)$ as follows:

$$Q(t) = \left\{ \mathbf{q}(\mathrm{t}) = (n_1(t), n_2(t), \ldots, n_K(t)) \mid \sum_{k=1}^{K} b_i n_i(t) \le B_U \right\}. \tag{13}$$

B_U is the bandwidth that can be used in the cell. Then steady-state probabilities for this model are as follows.

$$\begin{aligned} \pi(\mathbf{q}(t)) &= \pi(n_1(t), n_2(t), \ldots, n_K(t)) \\ &= \frac{\prod_{k=1}^{K} \left(\rho_k^{n_k}(t)/n_k!\right)}{\sum_{q(t)\in Q(t)} \prod_{k=1}^{K} \left(\rho_k^{n_k}(t)/n_k!\right)}, \end{aligned} \tag{14}$$

where $\rho_k = \lambda_{jk}(t)/(\mu_k + \mu_h)$ is the traffic load of service k. Then blocking probability for service k in the cell, denoted as $p_b^k(t)$ is shown below:

$$P_b^k(t) = \sum_{\mathrm{q}(t)\in Q_i(t)} \pi(\mathbf{q}(t)). \tag{15}$$

And $Q_i(t)$ is shown as follows:

$$Q_i(t) = \left\{ q(t) \mid B_U - b_i < \sum_{i=1}^{K} b_i n_i(t) \le B_U \right\}. \tag{16}$$

Moreover, average number of active user for service k at time t is

$$E[U_i] = \sum_{\mathrm{q}(t)\in Q(t)} n_i \pi(\mathbf{q}(t)). \tag{17}$$

And $c_{ijk}(t)$ can be obtained from (17) at time t.

6.2. Regional Coverage Evaluation. Algorithm in Section 5 gives a theoretical compensation for regional coverage. In order to verify effect to coverage quality by SPAM, regional coverage should be evaluated after ES recovery. And power signal strength and corresponding E_c/I_o of pilot channel are most direct indicators. Assume that at time t received RSCP (Received Signal Code Power) received by user i from BS j which supply service is $\rho_{ij}(t)$, and corresponding pilot power E_c/I_o is $\kappa_{ij}(t)$. Wireless propagation parameters set affect these two variables is Ω, then $\rho_{ij}(t)$ and $\kappa_{ij}(t)$ can be represented by (18) shown as follows [22]:

$$\begin{aligned} \rho_{ij}(t) &= \phi_i\left(P_j(t); \Omega\right), \\ \kappa_{ij}(t) &= \varphi_i\left(P_j(t), P_{Tx}^j(t); \Omega\right). \end{aligned} \tag{18}$$

So as to guarantee regional coverage, when network is under SS probability distributions of $\rho_{ij}(t)$ and $\kappa_{ij}(t)$ should satisfy several requirements, as

$$\text{s.t.} \begin{cases} F_{\rho_{ij}(t)}\left(\rho_{ij^*}(t) \ge \rho_{\min}\right) \ge \nu, \\ F_{\kappa_{ij}(t)}\left(\kappa_{ij^*}(t) \ge \kappa_{\min}\right) \ge \upsilon, \end{cases} \qquad j^* = \arg\max\left\{\rho_{ij}(t)\right\} \tag{19}$$

$\rho_{\min}$ and $\kappa_{\min}$ denote the lower limits of $\rho_{ij}(t)$ and $\kappa_{ij}(t)$ when service is supplied. $F_x(C)$ represents cumulative probability function for x when condition C is satisfied.

6.3. Regional Energy Efficiency Evaluation. Energy efficiency denotes regional energy saving ratio of ES algorithms. It is the most important indicator to evaluate ES outcome. Assume that period of traffic variation is T_P. In one period, NM intervals and SM intervals of network are alternate. If NS durations are set as $[t_{2i}, t_{2i+1}]$, $i = 0, 1, \ldots, W$, then SM intervals are $[t_{2j+1}, t_{2j+2}]$, $j = 0, 1, 2, \ldots, W-1$. And we can set $t_0 = 0$ and $t_{2W+1} = T_P$. For the network, energy consumption during SS denoting as P_{ES} is shown below:

$$P_{\mathrm{ES}} = \sum_{j=1}^{N\cdot \mathrm{NR}} \sum_{i=0}^{W-1} \int_{t_{2i+1}}^{t_{2i+2}} P_{\mathrm{BS-NS}}^{j}(t)dt + \sum_{j=1}^{N\cdot \mathrm{CR}} \sum_{i=0}^{W-1} \int_{t_{2i+1}}^{t_{2i+2}} P_{\mathrm{BS-ES}}^{j}(t)dt + \sum_{j=1}^{N\cdot \mathrm{SR}} \sum_{i=0}^{W-1} \int_{t_{2i+1}}^{t_{2i+2}} P_S dt, \tag{20}$$

$P_{\mathrm{BS-NS}}^{j}(t)$ and $P_{\mathrm{BS-ES}}^{j}(t)$ are power of BS under NS and CS, which satisfy relation from (1) to (7). When network is under SS, energy-saving ratio during SS intervals denoting as E_s is the following:

$$E_s = \frac{\left[\sum_{j=1}^{N} \sum_{i=0}^{W-1} \int_{t_{2i+1}}^{t_{2i+2}} P_{\mathrm{BS-NS}}^{j}(t)dt - P_{\mathrm{ES}}\right]}{\sum_{j=1}^{N} \sum_{i=0}^{W-1} \int_{t_{2i+1}}^{t_{2i+2}} P_{\mathrm{BS-NS}}^{j}(t)dt}. \tag{21}$$

Assume energy-saving ratio on the whole period is E_e, it is shown as follows:

$$E_e = \frac{\left[\sum_{j=1}^{N} \sum_{i=0}^{W-1} \int_{t_{2i+1}}^{t_{2i+2}} P_{\mathrm{BS}}^{j}(t)dt - P_{\mathrm{ES}}\right]}{\sum_{j=1}^{N} \int_{0}^{T_P} P_{\mathrm{BS}}^{j}(t)dt}. \tag{22}$$

From above the analysis, we can get a more accurate energy consumption evaluation model.

7. Simulations and Discussions

This section introduces simulation and analysis of SPAM, and integrated evaluation model is still validated. Moreover, SPAM is compared to another two ES methods in the references to evaluate its efficiency.

7.1. Scenario Description. SPAM will be simulated under an urban region of WCDMA in Qualnet. Top view of simulation scenario is shown in Figure 5. Region size is 3 km × 3 km, and street width is 20 m. A square shaped garden with size of 200 m × 200 m is located in centre of this region. Heights of all the buildings are between 20 m and 40 m. 16 homogeneous BSs (here is NodeB) are deployed on top of buildings and distances among these BSs are between 600 m and 700 m. Each BS contains three sectors and are all managed by the same OAM system. ID of each cell is automatically generated by Qualnet.

For these BSs, carrier frequency is 2.13 GHz. width of horizontal half-power antenna beam is 65 degree, and width of vertical half-power antenna beam is 10 degree. Initial tilt, antenna gain, and radiation efficiency of antenna are 8 degree, 15 dBi, and 0.8, separately. Besides, available bandwidth of each cell is 2 Mbps. Assume CE number and margin ratio of every BS are same as V and ω. Values of different power parts in (3) refer to [17] for UMTS BS.

Table 1: Link budget parameters in our scenario.

Parameters	Uplink	Downlink	Unit
Total transmitter power	21	43	dBm
TX antenna gain	0	17.9	dBi
TX cable loss	0	2	dB
TX body loss	2	0	dB
Transmitter EIRP	19.0/23.5	40.9/45.9	dBm
RX antenna gain	17.9	0	dBi
Receiver noise power	−103.2	−100.2	dBm
Processing gain	25.0/17.8	25.0/17.8	dBi
Required E_b/N_o	5.1/1.7	7.2/3.8	dB
Interference margin	3	6	dB
Required signal power	−120/−116	−112/−108	dBm
RX cable loss	2	0	dB
RX body loss	0	2	dB
Diversity gain	3	0	dB
Fast fading margin	3	0	dB
Soft handover gain	2	1	dB
Shadow fading margin	7.5	7.5	dB
Indoor penetration loss	4	4	dB

In the area, AMR 12.2 kbps voice service and CS 64 kbps video call are provided. COST231-HATA propagation model is adopted to estimate path loss. Initial link budget parameters set Ω can be found in Table 1 for voice/video call service. Shadowing fading model is lognormal distributed with average value of 4 dB, and fast fading model is Rician distributed.

Practical arrival rates in one cell obtained from an urban area in Beijing is adopted and shown in Figure 6 for one week. From the arrival rates variation we can find that cycle is one week. And variation during weekend is different from weekday. $1/\mu_k$ for them are 3 minutes and 5 minutes, and $1/\mu_h$ for them are 10 minutes and 15 minutes. From service quality evaluation model, we can obtain that peak active user for voice and video call in one cell are 30 and 10.

In our simulation, other important parameter configurations for intelligent coverage optimization algorithm and integrated evaluation model are shown in Table 2.

Besides, in SA algorithm, Change step of each P_i is 1 dBm. At the beginning, each P_j is set as 33 dBm. Cool function is set $T_{k+1} = 0.8T_k$.

For transmit power, maximal transmit power of BS under NM is 43.01 dBm. Assume that when network is under SS, power of BSs under SM is 5% of maximal one. In order to maintain management function, pilot power of BS under CM should keep on the optimal values computed by SPAM.

7.2. Analysis of Simulation Result. A near-sinusoidal traffic variation is obtained in ES region for each hour can be seen in Figure 7. It is easy to find that traffic variations from Monday

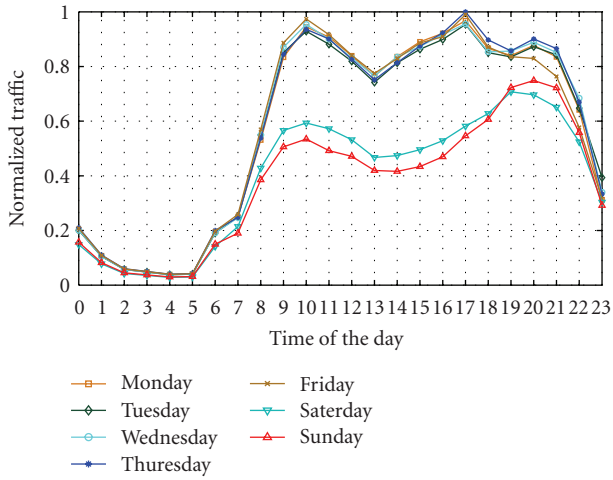


Figure 7: Normalized traffic variations during one week.

to Friday (work time) are similar. Still, traffic variations from Saturday to Sunday (weekend) are alike as well and lower than work time. Moreover, traffic during 23:00 to 7:00 (called as night time) of each day is fairly slight.

Through simulation, SPAM will be executed each time when ES triggering conditions are satisfied. For example, in Monday when SPAM is executed, the best value is obtained when iteration time is 143, and values of $H(\mathbf{P})$ and $G(\mathbf{P})$ are 1.998% and 20.039 KW separately. Moreover, all users can be accepted and none BS is overloaded.

More detailed result of SPAM during Monday is shown in Table 3. NodeB numbering is similar with numbering in Figure 5. Modes of each NodeB and pilot power value can be observed. During this period, values of NR, SR, and CR are 0.375, 0.25, and 0.375 separately. Energy saving time is 10.15 hour with $E_s = 39.78\%$ and $E_e = 14.21\%$ for Monday. In fact SR will increase with a larger ES region. So this mechanism is more suitable for large area with denser BS deployments.

SPAM will then be evaluated during one week to assess its efficiency comparing to two representative methods. One is centralized algorithm (we call it as CA) in [8] and other one is switched off scheme (we call it as SoS) in [12]. CA is an algorithm from user perspective, and SoS is an algorithm from regional traffic perspective. We choose them for the following reasons: (1) they are all centralized algorithms; (2) they are both suitable for ES of UMTS; (3) the evaluated metrics are similar. So comparisons among them will be more convincing.

Firstly, evaluation of service quality for the three methods with cell having highest load can be seen in Figure 8. In current cellular networks, target of blocking probability should below 0.01. From Figure 8(a), we can find that blocking probability of video call is higher than voice, and values for Sunday and Saturday are fairly low for the arrival rates are slow. Figures 8(b), 8(c), and 8(d) show that when network is under SS, blocking probability is higher than NS, which means that ES methods will degrade QoS. However, all the methods will guarantee the blocking probability below target value. Still, SoS adopts blocking probability as a triggering condition so its value is highest of all. And SPAM just cause a little degradation (about 0.001) to the service

Table 2: Configuration of several important parameters.

Parameter	Value
α	1
γ	0.3
o_1	−10 dB
o_2	−6 dB
n_{sector}	3
n_{Tx}	1
η	0.2
P_{au}	−112 dBm
P_{ad}	−95 dBm
$P_{Tx}^{\min}$	40.00 dBm
T_0	10
σ	1
ς	10000
$P_{Tx}^{\max}$	47.78 dBm
β	2%
ν_1	1
ν_2	4
V	384
ω	0.25
$\rho_{\min}$	−95 dBm
$\kappa_{\min}$	−12 dB
ν	97%
υ	95%
T_f	0.001
χ	1
$n(T_k)$	5

Table 3: Result of SPAM for Monday.

Node B	P_j (dBm)	Mode
# 1	33	NM
# 2	33	NM
# 3	36	CM
# 4	33	NM
# 5	36	CM
# 6	—	SM
# 7	—	SM
# 8	35	CM
# 9	33	NM
# 10	—	SM
# 11	35	CM
# 12	—	SM
# 13	36	CM
# 14	33	NM
# 15	33	NM
# 16	35	CM

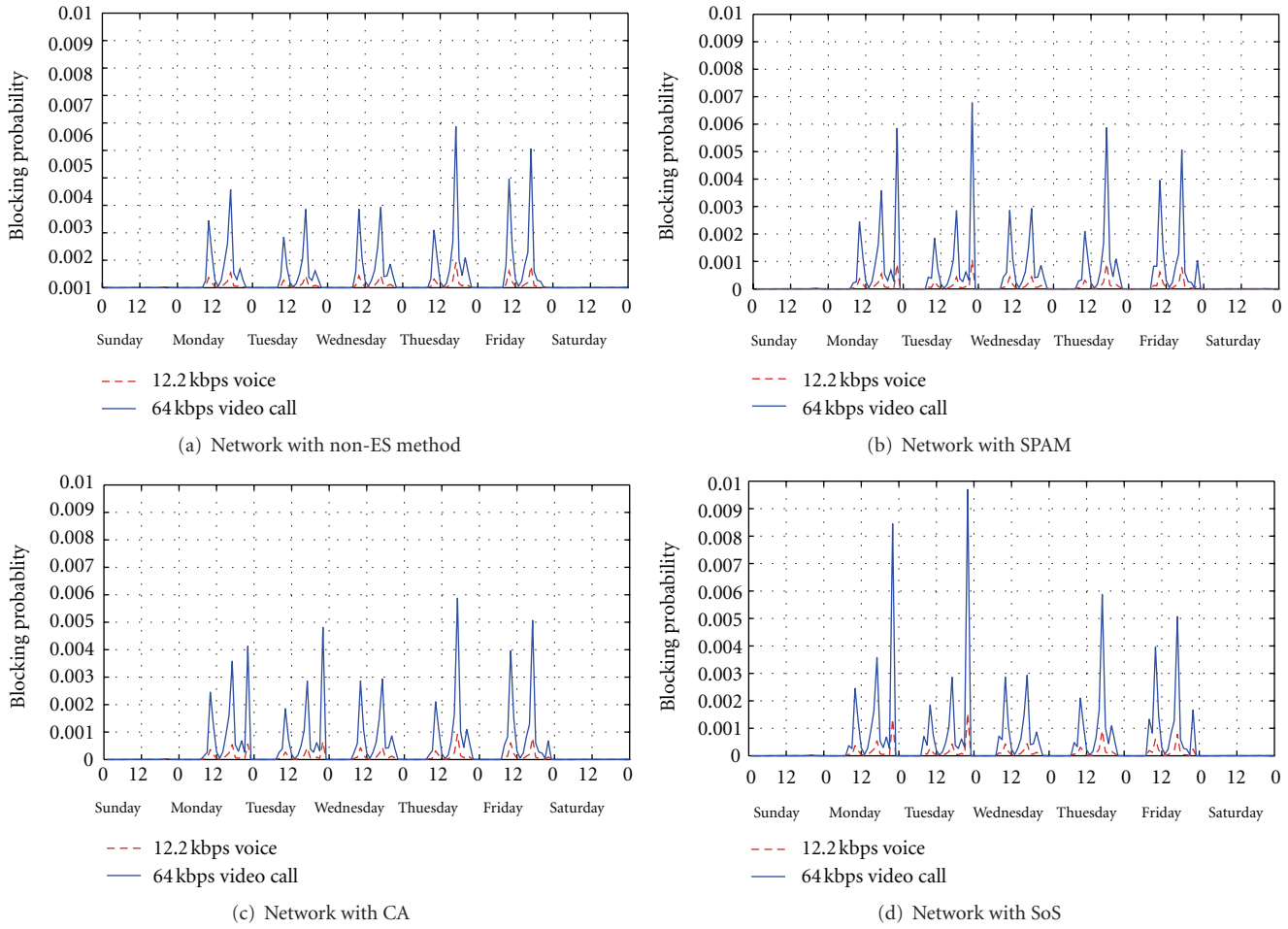


Figure 8: Blocking probability comparisons for network without and with ES method.

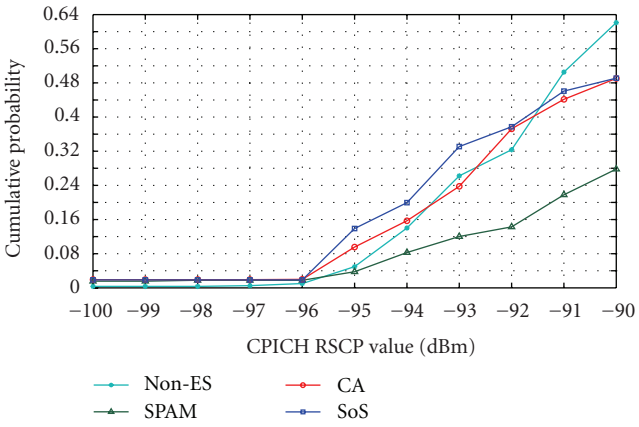


Figure 9: Comparison of CPICH RSCP for network without and with ES methods.

quality. CA takes on best performance here. And all of three ES methods can guarantee service quality.

Then regional coverage will be evaluated. Cumulative probabilities of CPICH RSCP and corresponding E_c/I_o can be seen in Figures 9 and 10. Figure 9 shows that for network with non-ES method, with SPAM, with CA, and with SoS, $F(\cdot)$ for RSCP (more than or equal to -95 dBm) are 99.48%, 98.23%, 98.02%, and 98.13%, which are above target value.

Corresponding to RSCP value, Figure 10 shows that for network with non-ES method, with SPAM, with CA, and with SoS, $F(\cdot)$ for E_c/I_o (equal to or more than -12 dB) are 100%, 97.32%, 97.74%, and 98.23%, which are above target value as well. So from coverage perspective, each method still keeps on acceptable levels.

Still, ES methods may affect signal strength a little. For the three ES methods, SPAM possesses strongest signal distribution, so the weak coverage problem caused by BS under SM can be resolved. However, due to adjustments to pilot power, interference of SPAM may be higher than CA, and SoS, that is tradeoff among ES and coverage.

Regional energy efficiency will be evaluated at last, as shown in Figure 11. Regional ES time, E_s, and E_e are compared for SPAM, CA and SoS. We can find that SPAM can obtain longest ES time during one week (about 103 hours) and takes on best energy efficient during ES time (saving about 41% of entire energy consumption) and the

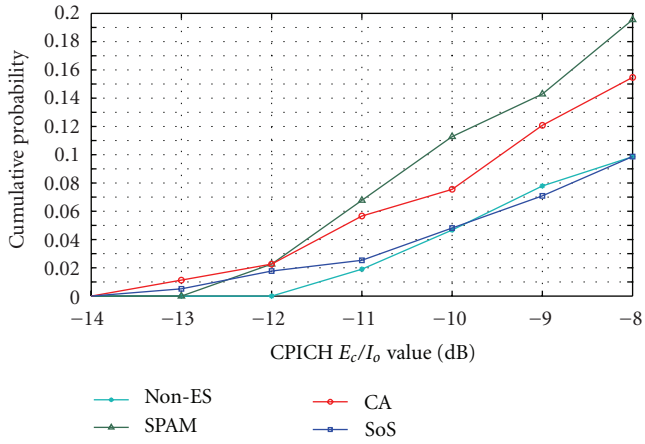


Figure 10: Comparison of CPICH E_c/I_o for network without and with ES methods.

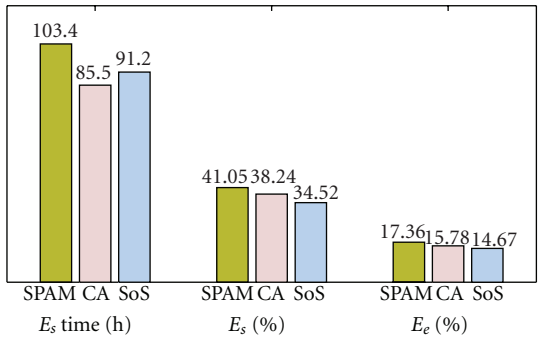


Figure 11: Comparisons of energy saving ratios for three ES methods.

whole time period (saving about 17.36% of entire energy consumption).

The above analysis shows that as a self-organizing ES method, SPAM can autonomic monitor regional traffic variations and execute corresponding actions when ES triggering or recovery conditions are satisfied. Comparing to other ES methods in cellular networks, SPAM can save most energy consumption and meanwhile guarantee service quality and better regional coverage as well.

8. Conclusions and Future Work

Aiming at design of effective self-organizing regional energy-saving scheme, SPAM is proposed for cellular networks in this paper. It describes self-organizing processes of saving energy, proposes a more general BS selection scheme, an intelligent coverage optimization algorithm, and integrated evaluation model. Simulation on Qualnet shows that SPAM can save at least 17% regional energy consumption and meanwhile guarantee regional coverage, capacity, and service quality, which take on more efficient comparing to other ES methods.

However, as a centralized management mechanism, SPAM is suitable for current cellular networks. Due to that distributed management manner is more suitable for large and complex networks with heterogeneous cell deployments (femtocell, picocell, etc.), SPAM will be extended for self-organizing saving energy under multiple frequencies, multiple cell styles, and multiple service categories. Other network parameter (such as down tilt) adjustment for saving energy are still under current research.

Acknowledgments

This paper is supported by the Funds for Creative Research Groups of China (61121061), National S&T Major Projects (2011ZX03003-002-01), National Natural Science Foundation of China (61271187), National Key Technology R&D Program (2012BAH06B02), and NCET-10-0240 and Chinese Universities Scientific Fund (BUPT2012RC0608).

References

[1] C. Lubritto, A. Petraglia, C. Vetromile et al., "Energy and environmental aspects of mobile communication systems," *Energy*, vol. 36, no. 2, pp. 1109–1114, 2011.

[2] J. T. Louhi and IEEE, "Energy efficiency of modern cellular base stations," in *Proceedings of the International Telecommunication Energy Conference (INTELEC '07)*, pp. 475–476, October 2007.

[3] M. A. Marsan and M. Meo, "Energy efficient wireless Internet access with cooperative cellular networks," *Computer Networks*, vol. 55, no. 2, pp. 386–398, 2011.

[4] E. Mino, E. Torrecilla, L. M. Del Apio, and I. Berberana, "SON use case study 'energy saving' for LTE eNBs," *IEEE Latin America Transactions*, vol. 8, no. 2, pp. 184–189, 2010.

[5] GPP, "TS 32. 551 Energy Saving Management (ESM): concepts and requirements," Release 10, 2010.

[6] Z. Hasan, H. Boostanimehr, and V. K. Bhargava, "Green cellular networks: a survey, some research issues and challenges," *IEEE Communications Surveys and Tutorials*, vol. 13, no. 4, pp. 524–540, 2011.

[7] E. Oh and B. Krishnamachari, "Energy savings through dynamic base station switching in cellular wireless access networks," in *Proceedings of the 53rd IEEE Global Communications Conference (GLOBECOM '10)*, pp. 1–5, December 2010.

[8] Z. Niu, Y. Wu, J. Gong, and Z. Yang, "Cell zooming for cost-efficient green cellular networks," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 74–79, 2010.

[9] K. Dufková, M. Bjelica, B. Moon, L. Kencl, and J. Y. Le Boudec, "Energy savings for cellular network with evaluation of impact on data traffic performance," in *Proceedings of the European Wireless Conference (EW '10)*, pp. 916–923, April 2010.

[10] J. Lorincz, A. Capone, and D. Begušić, "Optimized network management for energy savings of wireless access networks," *Computer Networks*, vol. 55, no. 3, pp. 514–540, 2011.

[11] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Optimal energy savings in cellular access networks," in *Proceedings of the IEEE International Conference on Communications Workshops (ICC '09)*, pp. 1–5, June 2009.

[12] L. Chiaraviglio, D. Ciullo, M. Meo, and M. A. Marsan, "Energy-efficient management of UMTS access networks," in *Proceedings of the 21st International Teletraffic Congress, ITC 21: Traffic and Performance Issues in Networks of the Future*, September 2009.

[13] L. Chiaraviglio, D. Ciullo, M. Meo, M. A. Marsan, and I. Torino, "Energy-aware UMTS access networks," in *Proceedings of the 11th International Symposium on Wireless Personal Multimedia Communications (WPMC '08)*, pp. 1–5, 2008.

[14] M. F. Hossain, K. S. Munasinghe, and A. Jamalipour, "An eco-inspired energy efficient access network architecture for next generation cellular systems," in *Proceedings of the 2011 IEEE Wireless Communications and Networking Conference (WCNC' 11)*, pp. 992–997, March 2011.

[15] F. Richter, A. J. Fehske, and G. P. Fettweis, "Energy efficiency aspects of base station deployment strategies for cellular networks," in *Proceedings of the IEEE 70th Vehicular Technology Conference Fall, (VTC '09)*, September 2009.

[16] B. Badic, T. O'Farrell, P. Loskot, and J. He, "Energy efficient radio access architectures for green radio: Large versus small cell size deployment," in *Proceedings of the IEEE 70th Vehicular Technology Conference Fall, (VTC '09)*, September 2009.

[17] M. Deruyck, E. Tanghe, W. Joseph, and L. Martens, "Modelling and optimization of power consumption in wireless access networks," *Computer Communications*, vol. 34, no. 17, pp. 2036–2046, 2011.

[18] S. Bhaumik, G. Narlikar, S. Chattopadhyay, and S. Kanugovi, "Breathe to stay cool: adjusting cell sizes to reduce energy consumption," in *Proceedings of the 1st ACM SIGCOMM Workshop on Green Networking*, pp. 41–46, New Delhi, India, August 2010.

[19] J. Laiho, A. Wacker, and T. Novosad, *Radio Network Planning and Optimisation for UMTS*, Wiley Online Library, 2002.

[20] K. Hiltunen and N. Binucci, "WCDMA downlink coverage: interference margin for users located at the cell coverage border," in *Proceedings of the 55th Vehicular Technology Conference (VTC '02)*, pp. 270–274, May 2002.

[21] E. Meshkova, J. Riihijärvi, A. Achtzehn, and P. Mähönen, "Exploring simulated annealing and graphical models for optimization in cognitive wireless networks," in *Proceedings of the IEEE Global Telecommunications Conference, (GLOBECOM '09)*, December 2009.

[22] Y. K. Song, D. Kim, and J. Zander, "Pilot power adjustment for saving transmit power in pilot channel assisted DS-CDMA mobile systems," *IEEE Transactions on Wireless Communications*, vol. 9, no. 2, pp. 488–493, 2010.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 438090, 15 pages
doi:10.1155/2012/438090

Research Article

Cooperative Group Localization Based on Factor Graph for Next-Generation Networks

Xuefei Zhang, Qimei Cui, Yulong Shi, and Xiaofeng Tao

Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing 100876, China

Correspondence should be addressed to Xuefei Zhang, zhangxuefeikate@gmail.com

Received 3 May 2012; Revised 11 August 2012; Accepted 25 August 2012

Academic Editor: Yonghui Li

In ill-conditioned communication environment, multiple target localization is of important practical significance. The cooperative group localization (CGL) model was firstly put forward, which has verified the effectiveness of localization performance gain and simultaneous multiple target localization in ill conditions. However, there exist two inherent difficulties: the strict demand for CGL topology and the high complexity. By the rational use of information to relax restrictions on topology and by dividing the complex problem into some simple local ones, the factor graph (FG) together with the sum-product algorithm is a perfect candidate for the problems above. In order to solve the two problems, we propose the weighted FG-based CGL (WFG-CGL) algorithm which incorporates the optimal weights based on the information reliability. In order to further reduce the complexity, we propose the low-complexity FG-based CGL (LCFG-CGL) algorithm. The Cramer-Rao lower bound (CRLB) of the localization error in CGL is first derived. Theoretical analysis and numerical results indicate that the proposed algorithms not only perform better in relaxing CGL topology requirement, but also enjoy high localization accuracy under low complexity in comparison with the existing CGL algorithm.

1. Introduction

Wireless localization technologies, which are designated to estimate the position of a mobile terminal (MT), have drawn a lot of attention over the past few decades. Federal Communication Commission (FCC) has mandated the cellular network operators to estimate the position of emergency caller with the error of less than 100 meters in the enhanced 911 (E-911) [1]. In recent years, there are increasing demands for commercial applications to adopt the localization information within their system design, such as the navigation system, the health care system, the wireless sensor network (WSN) [2–4], and the intelligent transportation system (ITS) [5, 6]. Moreover, more and more people rely on the instant localization, which may make the communication services more user-friendly. With the emergent interests in the localization-based services (LBSs), the wireless localization with enhanced accuracy becomes necessary for the applications under different circumstances.

The research of the localization techniques is an important part of wireless localization fields [7]. A variety of wireless localization techniques have been investigated, such as basic localization algorithms and their corresponding Cramer-Rao lower bound (CRLB) analysis [8–10]. CRLB provides a useful means for the analysis of the limits of localization accuracy. CRLB results for localization, based on different types of internode measurements, can be found in [2, 3, 11–15]: the localization error is shown to depend on several elements, in particular on the measurement reliability (which may be impaired by multipath, non-line of sight, synchronization errors, etc.) and the topology of the network.

Most of current academic achievements in localization field focus on the enhancement localization techniques in the well condition. The well condition means that mobile terminal (MT) to be localized has the sufficient signal resources. For example, if one MT can be accurately located by the measurement parameters of time of arrival (TOA), time difference of arrival (TDOA), angle of arrival (AOA), or received signal strength (RSS), the MT connects well with at least three base stations (BSs) and line-of-sight (LOS) propagation paths existing between BSs and MT [16].

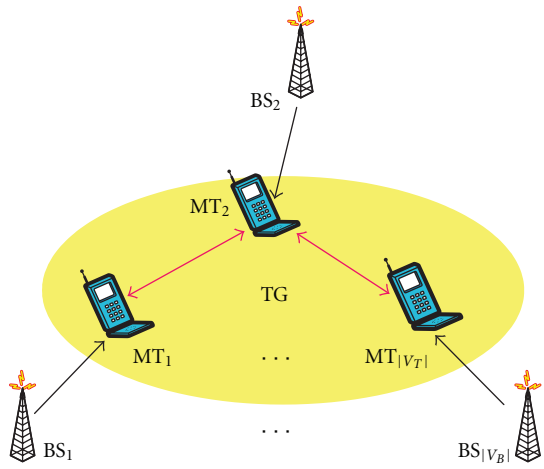


Figure 1: CGL topology diagram.

Nevertheless, the scenarios with sufficient signal resources do not always happen in real circumstances. Sometimes, the parameter measurements may be corrupted by non-line-of-sight (NLOS) error, or MT can not simultaneously connect with three BSs [9, 16]. These ill conditions with insufficient signal resources can make the accurate localization unrealistic.

The cooperative group localization (CGL), which can achieve the multiple target localization simultaneously, is proposed as an approach to solving the above ill-conditioned localization problem in [17, 18]. The existing research indicates that the CGL scheme not only has higher localization accuracy over the traditional methods in ill conditions, but also can effectively enhance the robustness of localization. However, there also exists two inherent difficulties: (1) the strict demand for CGL topology and (2) the high complexity. According to the rigid graph theory [19], the CGL topology must be global rigidity so as to satisfy the unique solvability. Furthermore, CGL has to face the problem of high complexity, especially with the increasing number of cooperative MTs.

In [20], factor graph (FG) theory is proposed, which can simplify a complex problem into multiple simple sub problems. The optimal or the near-optimal solution can be obtained by solving the sub problems, respectively. According to its feasibility and simplicity, it has been introduced into cellular or sensor network based on measurement parameters TOA, TDOA, AOA, and RSS [21–25] and reaches the good tradeoff between the localization accuracy and complexity. In [26], Wymeersch et al. analyzes the feasibility of FG theory applied into the cooperative localization.

In this paper, we combine FG theory with CGL. The FG-CGL graph can adapt its constitutions and architecture based on the different measurement information, considered as self-organizing. First, a weighted FG-based CGL algorithm (WFG-CGL) is proposed to effectively solve the above problems of strict demand for CGL topology and high complexity in the ill condition. It enhances the accuracy by adopting the different information reliability of BSs and MTs into its formulation. Then, a low-complexity FG-based CGL (LCFG-CGL) algorithm is proposed to further diminish the complexity. Moreover, CRLB analysis is extended to CGL scenarios. Numerical simulation results verify that compared to the existing CGL method, the proposed algorithms can relax CGL topology requirement and provide the high localization accuracy under low complexity. And the important factors for the FG-based CGL algorithms are analyzed based on the simulation results.

The paper is organized as follows. In Section 2, CGL model is introduced. In Section 3, the principles of factor graph and sum-product algorithm are presented. Section 4 proposes WFG-CGL algorithm and LCFG-CGL algorithm. Section 5 completes CRLB analysis of CGL. Section 6 evaluates the performance and the complexity of the proposed algorithms by numerical simulation. Finally, the conclusion is included in Section 7.

2. Mathematical Model of CGL

For the description convenience, take the scenario shown in Figure 1 for example, which presents the principle of CGL in the following sections. In CGL, the conception of the terminal group (TG) is firstly put forward. TG refers to a group of MTs supporting peer-to-peer communication mode in the next-generation communication networks. To represent the most general situation, all MTs are supposed to be the localization unknown nodes, and the relative distances between some certain MT pairs are known. For the next-generation communication networks, the advantages of the cooperation between MTs can be applied widely. As shown in Figure 1, each MT in TG connects with less than three BSs, which can be included in the ill conditions mentioned above.

The CGL graph $G = (V, E)$ can be divided into three sub-graphs: the first one is BS subgraph $G_B = (V_B, E_B)$, where V_B represents the BSs, and E_B represents the existing edges between any two vertices of V_B; the second one is TG subgraph $G_T = (V_T, E_T)$, where V_T includes all the MTs in TG, and E_T represents the distances determined by direct communication among the MTs; the third is cooperative sub-graph $G_C = (V_C, E_C)$, where V_C includes the BSs or the cooperative MTs in TG, and E_C represents the edges existing between G_B and G_T. The operator $|\bullet|$ of the sets denotes the number of the elements in a set.

Given the CGL graph, the aim of CGL is to simultaneously solve the positions of all the MTs in TG. However, not all topologies of CGL have unique solvability. The unique solvability means that the positions of MTs in a TG can be located without ambiguity. It can be investigated by the rigid graph theory [19].

3. Factor Graph and Sum-Product Algorithm

3.1. Factor Graph. A factor graph [20] is a bipartite graph that illustrates how a complicated global function with many variables is simplified into the product of several simple local functions. Each local function is a function with few variables. A factor graph has the variable node for each variable, the function node for each local function, and the

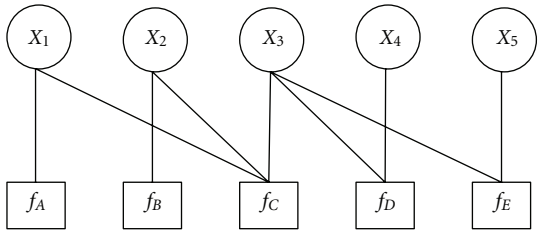


Figure 2: An example of factor graph.

edge connecting variable node to factor node if and only if the variable is an argument of the function. For example, $f(x_1, x_2, x_3, x_4, x_5)$ is a real-valued function with five variables, which is divided into the product of five local functions f_A, f_B, f_C, f_D and f_E, as is shown in Figure 2, as follows:

$$f(x_1, x_2, x_3, x_4, x_5) = f_A(x_1) \cdot f_B(x_2) \cdot f_C(x_1, x_2, x_3) \cdot f_D(x_3, x_4) \cdot f_E(x_3, x_5). \quad (1)$$

3.2. Sum-Product Algorithm. Sum-product algorithm [20] is one of the important algorithms in graphical model, which provides the convenient solution to marginal probability. The algorithm works by the passing messages along the edges between the nodes. There are two types of messages.

With two rules above, information is passed among neighboring nodes. Each variable node modifies its own value according to the received message. After a few times, the information converges. The variable can be estimated according to the product of overall message from its neighboring nodes as follows:

$$\mathrm{SI}(x) = \prod_{h \in n(x)} \mathrm{SI}(h, x). \quad (2)$$

4. FG-Based CGL Algorithms

In the following analysis, the positions of several MTs are estimated simultaneously based on TOA measurements. In order to reduce the complexity of the two-dimensional (2D) problem, the problem is divided into two one-dimensional (1D) problems [21], x-coordinate group and y-coordinate group, as illustrated in Figure 3. In Figure 3, x_i and y_i represent the coordinates for the ith MT; $\Delta x_{i,j}$ and $\Delta y_{i,j}$ represent the relative distances between the ith MT and the jth BS; $\Delta x^{\beta}_{i,j}$ and $\Delta y^{\beta}_{i,j}$ represent the relative distances between the ith MT and the jth BS after the weight adaptation; $\Delta xx_{i,t}$ and $\Delta yy_{i,t}$ represent relative distances between the ith MT and the tth cooperative MT; $\Delta xx^{\beta}_{i,t}$ and $\Delta yy^{\beta}_{i,t}$ represent relative distances between the ith MT and the tth cooperative MT after the weight adaptation; $\hat{d}_{i,j}$ and $\hat{l}_{i,t}$ are the measurement distance from the jth BS to the ith MT, and the distance from the tth cooperative MT to the ith MT; $d_{i,j}$ and $l_{i,t}$ are produced by the Gaussian distribution with the means $\hat{d}_{i,j}$ and $\hat{l}_{i,t}$, and with their corresponding variances. The meanings of function nodes are described below. The messages are processed and passed between variable nodes and function nodes.

4.1. The Functions of All Nodes in FG-Based CGL

(1) Variable x_i and y_i (in the kth iteration). The message from variable node x_i^k to function node $A^k_{i,j}$ or $M^k_{i,t}$ is a Gaussian probability density function (PDF) of x_i^k and can be expressed as

$$\mathrm{SI}\left(x_i^k, A^k_{i,j}\right) = \prod_{u=1, u\neq j}^{|V_{B,i}|} \mathrm{SI}\left(A^k_{i,u}, x_i^k\right) \prod_{v=1}^{|V_{T,i}|} \mathrm{SI}\left(M^k_{i,v}, x_i^k\right),$$
$$\mathrm{SI}\left(x_i^k, M^k_{i,t}\right) = \prod_{u=1}^{|V_{B,i}|} \mathrm{SI}\left(A^k_{i,u}, x_i^k\right) \prod_{v=1, v\neq t}^{|V_{T,i}|} \mathrm{SI}\left(M^k_{i,v}, x_i^k\right), \quad (3)$$

where $i = 1, 2, \ldots, |V_T|$, $j = 1, 2, \ldots, |V_{B,i}|$, $t = 1, 2, \ldots, |V_{T,i}|$, $t \neq i$. $|V_T|$ is the number of whole MTs to be located simultaneously. $|V_{B,i}|$ and $|V_{T,i}|$ represent the number of BSs and other cooperative MTs connecting to the ith MT, respectively. $\mathrm{SI}(A^k_{i,u}, x_i^k)$ and $\mathrm{SI}(M^k_{i,v}, x_i^k)$ are Gaussian PDF of x_i^k. Note that the product of any U Gaussian pdf is also Gaussian and can be derived as [20]

$$\prod_{u=1}^{U} N(x, m_u, \sigma_u^2) \propto N(x, m_\Lambda, \sigma_\Lambda^2), \quad (4)$$

where $N(x, m, \sigma^2) \propto \exp[-(x-m)^2/2\sigma^2]$, $\sigma_\Lambda^2 = 1/(\sum_{u=1}^{U} 1/\sigma_u^2)$, and $m_\Lambda = \sigma_\Lambda^2 \sum_{u=1}^{U} (m_u/\sigma_u^2)$. Similarly, variable node y_i can be calculated.

(2) Variable Nodes $\Delta x_{i,j}$, $\Delta y_{i,j}$, $\Delta xx_{i,t}$, $\Delta yy_{i,t}$, $\Delta x^{\beta}_{i,j}$, $\Delta y^{\beta}_{i,j}$, $\Delta xx^{\beta}_{i,t}$, $\Delta yy^{\beta}_{i,t}$, $d_{i,j}$, and $l_{i,t}$. Variable node $\Delta x_{i,j}$ and $\Delta xx_{i,t}$ can transmit the receiving message directly as follows:

$$\mathrm{SI}\left(\Delta x_{i,j}^k, R_{i,j}^k\right) = \mathrm{SI}\left(C_{i,j}^k, \Delta x_{i,j}^k\right), \quad (5)$$
$$\mathrm{SI}\left(\Delta x_{i,j}^k, C_{i,j}^k\right) = \mathrm{SI}\left(R_{i,j}^k, \Delta x_{i,j}^k\right), \quad (6)$$
$$\mathrm{SI}\left(\Delta xx_{i,t}^k, RR_{i,t}^k\right) = \mathrm{SI}\left(CC_{i,t}^k, \Delta xx_{i,t}^k\right), \quad (7)$$
$$\mathrm{SI}\left(\Delta xx_{i,t}^k, CC_{i,t}^k\right) = \mathrm{SI}\left(RR_{i,t}^k, \Delta xx_{i,t}^k\right). \quad (8)$$

Similarly, variable node $\Delta y_{i,j}$, $\Delta yy_{i,t}$, $\Delta x^{\beta}_{i,j}$, $\Delta xx^{\beta}_{i,t}$, $\Delta y^{\beta}_{i,j}$, $\Delta yy^{\beta}_{i,t}$, $d_{i,j}$, and $l_{i,t}$ all play the role as a message transfer between the corresponding two function nodes.

(3) Function Nodes $A_{i,j}$, $B_{i,j}$, $M_{i,t}$, and $N_{i,t}$. The function of these nodes is to convert the relative position information into the absolute position information, and vice versa

$$\Delta x^{\beta}_{i,j} = X_{i,j} - x_i,$$
$$\Delta y^{\beta}_{i,j} = Y_{i,j} - y_i, \quad (9)$$

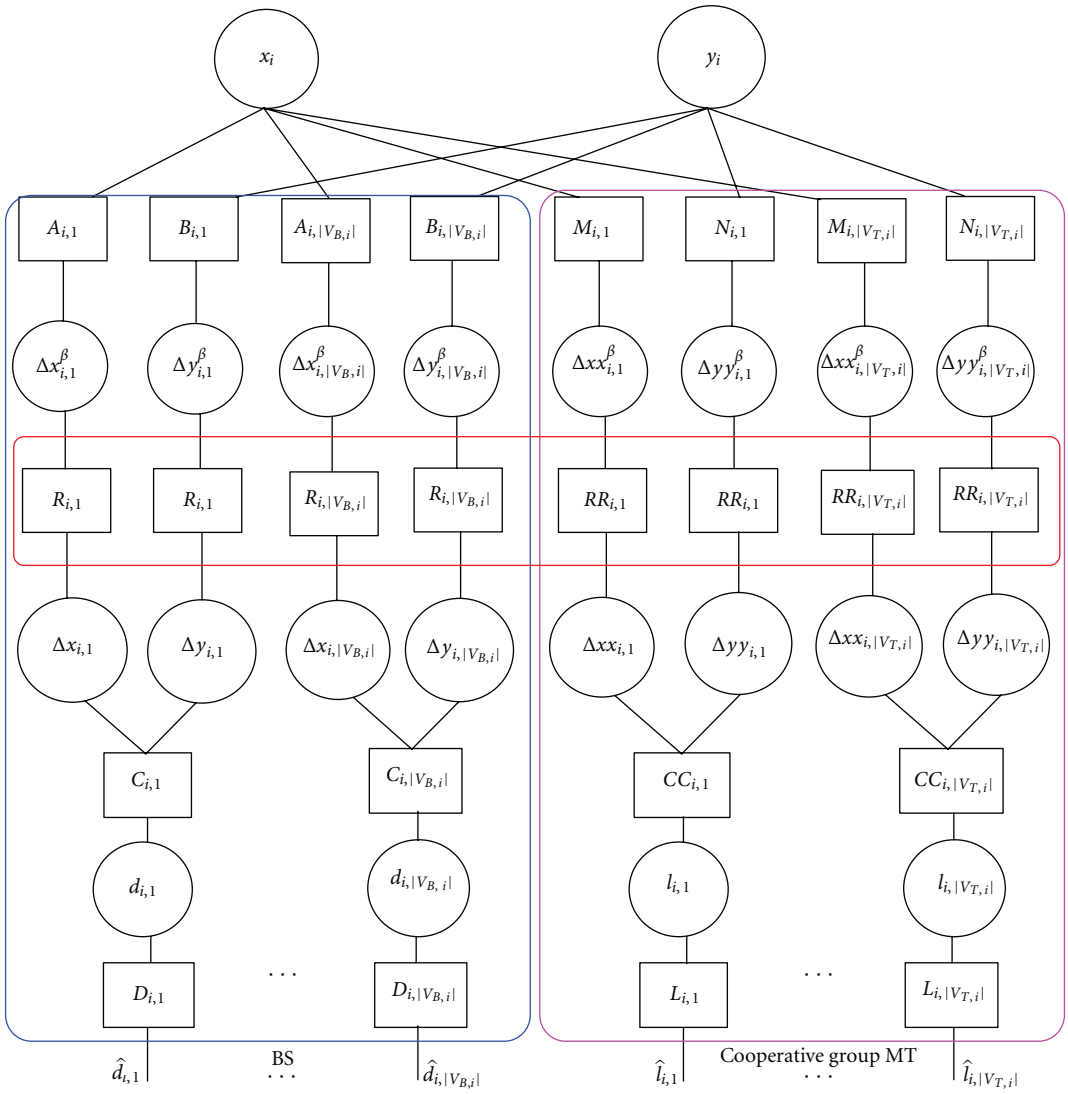


Figure 3: FG in CGL (take one of the MT as an example).

where $X_{i,j}$ and $Y_{i,j}$ represent the position of the jth BS connecting to the ith MT.

$$\begin{aligned} \Delta xx_{i,t}^{\beta} &= x_t - x_i, \\ \Delta yy_{i,t}^{\beta} &= y_t - y_i, \end{aligned} \quad t = 1, 2, \ldots, |V_{T,i}|, \tag{10}$$

$$\mathrm{SI}\left(A_{i,j}^{k}, \Delta x_{i,j}^{\beta,k}\right) = N\left(\Delta x_{i,j}^{\beta,k}, X_{i,j} - x_i^k, \sigma_{x_i^k}^2\right), \tag{11}$$

$$\mathrm{SI}\left(A_{i,j}^{k}, x_i^k\right) = N\left(x_i^k, X_{i,j} - \Delta x_{i,j}^{\beta,k}, \sigma_{\Delta x_{i,j}^{\beta,k}}^2\right), \tag{12}$$

where $\sigma_{x_i^k}^2$ and $\sigma_{\Delta x_{i,j}^{\beta,k}}^2$ are the variance of the Gaussian soft information from x_i^k and $\Delta x_{i,j}^{\beta,k}$. Similarly, the message from $B_{i,j}$, $M_{i,t}$, and $N_{i,t}$ can be calculated.

(4) Function Nodes $C_{i,j}$ and $CC_{i,t}$. The function of these nodes is to merge the separate information from x-coordinate group and the y-coordinate group and then compare it with real measurement data as follows:

$$\begin{aligned} \left(\Delta x_{i,j}\right)^2 + \left(\Delta y_{i,j}\right)^2 &= d_{i,j}^2, \\ \left(\Delta xx_{i,t}\right)^2 + \left(\Delta yy_{i,t}\right)^2 &= l_{i,t}^2. \end{aligned} \tag{13}$$

$\Delta y_{i,j}$ updates itself according to the messages from $\Delta x_{i,j}$ and $d_{i,j}$; $\Delta x_{i,j}$ updates itself according to the messages from $\Delta y_{i,j}$ and $d_{i,j}$.

$$\mathrm{SI}\left(C_{i,j}^{k}, \Delta y_{i,j}^{k+1}\right) = N\left(\Delta y_{i,j}^{k+1}, \pm\sqrt{\left(\hat{d}_{i,j}^{k}\right)^2 - \left(\Delta x_{i,j}^{k}\right)^2}, \frac{\left(\Delta x_{i,j}^{k}\right)^2\left(\sigma_{\Delta x_{i,j}}^{k}\right)^2 + \left(\hat{d}_{i,j}^{k}\right)^2\left(\sigma_{d_{i,j}}^{k}\right)^2}{\left(\hat{d}_{i,j}^{k}\right)^2 - \left(\Delta x_{i,j}^{k}\right)^2}\right), \tag{14}$$

$$\mathrm{SI}\left(C_{i,j}^{k}, \Delta x_{i,j}^{k+1}\right) = N\left(\Delta x_{i,j}^{k+1}, \pm\sqrt{\left(\hat{d}_{i,j}^{k}\right)^2 - \left(\Delta y_{i,j}^{k}\right)^2}, \frac{\left(\Delta y_{i,j}^{k}\right)^2\left(\sigma_{\Delta y_{i,j}}^{k}\right)^2 + \left(\hat{d}_{i,j}^{k}\right)^2\left(\sigma_{d_{i,j}}^{k}\right)^2}{\left(\hat{d}_{i,j}^{k}\right)^2 - \left(\Delta y_{i,j}^{k}\right)^2}\right), \tag{15}$$

where $\hat{d}_{i,j}$ and $\hat{l}_{i,t}$ are the measurement information between BSs and MTs, and between MTs. Similarly, $\mathrm{SI}(CC_{i,t}^{k}, \Delta x x_{i,t}^{k+1})$ and $\mathrm{SI}(CC_{i,t}^{k}, \Delta y y_{i,t}^{k+1})$ can be calculated.

(5) Function Nodes $R_{i,j}$ and $RR_{i,t}$. The function of these nodes is to perform the weighting process based on information reliability as follows:

$$\mathrm{SI}\left(\Delta x_{i,j}^{k}, R_{i,j}^{k}\right) = N\left(\Delta x_{i,j}^{k}, \pm\sqrt{\left(\hat{d}_{i,j}^{k}\right)^2 - \left(\Delta y_{i,j}^{k}\right)^2}, \frac{\left(\Delta y_{i,j}^{k}\right)^2\left(\sigma_{\Delta y_{i,j}}^{k}\right)^2 + \left(\hat{d}_{i,j}^{k}\right)^2\left(\sigma_{d_{i,j}}^{k}\right)^2}{\left(\hat{d}_{i,j}^{k}\right)^2 - \left(\Delta y_{i,j}^{k}\right)^2}\right), \tag{16}$$

$$\mathrm{SI}\left(\Delta y_{i,j}^{k}, R_{i,j}^{k}\right) = N\left(\Delta y_{i,j}^{k}, \pm\sqrt{\left(\hat{d}_{i,j}^{k}\right)^2 - \left(\Delta x_{i,j}^{k}\right)^2}, \frac{\left(\Delta x_{i,j}^{k}\right)^2\left(\sigma_{\Delta x_{i,j}}^{k}\right)^2 + \left(\hat{d}_{i,j}^{k}\right)^2\left(\sigma_{d_{i,j}}^{k}\right)^2}{\left(\hat{d}_{i,j}^{k}\right)^2 - \left(\Delta x_{i,j}^{k}\right)^2}\right), \tag{17}$$

$$\mathrm{SI}\left(R_{i,j}^{k}, \Delta x_{i,j}^{\beta,k}\right) = N\left(\Delta x_{i,j}^{k}, \pm\sqrt{\left(\hat{d}_{i,j}^{k}\right)^2 - \left(\Delta y_{i,j}^{k}\right)^2}, \beta_i \cdot \frac{\left(\Delta y_{i,j}^{k}\right)^2\left(\sigma_{\Delta y_{i,j}}^{k}\right)^2 + \left(\hat{d}_{i,j}^{k}\right)^2\left(\sigma_{d_{i,j}}^{k}\right)^2}{\left(\hat{d}_{i,j}^{k}\right)^2 - \left(\Delta y_{i,j}^{k}\right)^2}\right), \tag{18}$$

$$\mathrm{SI}\left(R_{i,j}^{k}, \Delta y_{i,j}^{\beta,k}\right) = N\left(\Delta y_{i,j}^{k}, \pm\sqrt{\left(\hat{d}_{i,j}^{k}\right)^2 - \left(\Delta x_{i,j}^{k}\right)^2}, \beta_i \cdot \frac{\left(\Delta x_{i,j}^{k}\right)^2\left(\sigma_{\Delta x_{i,j}}^{k}\right)^2 + \left(\hat{d}_{i,j}^{k}\right)^2\left(\sigma_{d_{i,j}}^{k}\right)^2}{\left(\hat{d}_{i,j}^{k}\right)^2 - \left(\Delta x_{i,j}^{k}\right)^2}\right), \tag{19}$$

$$\mathrm{SI}\left(\Delta x_{i,j}^{\beta,k}, R_{i,j}^{k}\right) = \mathrm{SI}\left(A_{i,j}^{k}, \Delta x_{i,j}^{\beta,k}\right), \tag{20}$$

$$\mathrm{SI}\left(\Delta y_{i,j}^{\beta,k}, R_{i,j}^{k}\right) = \mathrm{SI}\left(B_{i,j}^{k}, \Delta y_{i,j}^{\beta,k}\right), \tag{21}$$

$$\mathrm{SI}\left(R_{i,j}^{k}, \Delta x_{i,j}^{k}\right) = \mathrm{SI}\left(\Delta x_{i,j}^{\beta,k}, R_{i,j}^{k}\right), \tag{22}$$

$$\mathrm{SI}\left(R_{i,j}^{k}, \Delta y_{i,j}^{k}\right) = \mathrm{SI}\left(\Delta y_{i,j}^{\beta,k}, R_{i,j}^{k}\right), \tag{23}$$

where β_i is the weighting coefficient and $\beta_i = 1$ (without weighing) is for the information from BS. Similarly, $\mathrm{SI}(\Delta x x_{i,t}^{k}, RR_{i,t}^{k})$, $\mathrm{SI}(\Delta y y_{i,t}^{k}, RR_{i,t}^{k})$, $\mathrm{SI}(RR_{i,t}^{k}, \Delta x x_{i,t}^{\beta,k})$, $\mathrm{SI}(RR_{i,t}^{k}, \Delta y y_{i,t}^{\beta,k})$, $\mathrm{SI}(RR_{i,t}^{k}, \Delta x x_{i,t}^{k})$, and $\mathrm{SI}(RR_{i,t}^{k}, \Delta y y_{i,t}^{k})$ can be calculated.

(6) Function Nodes $D_{i,j}$ and $L_{i,t}$. The measurement information $\hat{d}_{i,j}$ and $\hat{l}_{i,t}$ enter FG through the function node $D_{i,j}$ and $L_{i,t}$ as follows:

$$\mathrm{SI}\left(D_{i,j}^{k}, d_{i,j}^{k}\right) = N\left(d_{i,j}, \hat{d}_{i,j}, \sigma_{d_{i,j}}^{2}\right), \quad \mathrm{SI}\left(L_{i,t}^{k}, l_{i,t}^{k}\right) = N\left(l_{i,t}, \hat{l}_{i,t}, \sigma_{l_{i,t}}^{2}\right). \tag{24}$$

After K iterations, information in variable nodes x_i and y_i converges. x_i and y_i can be estimated.

$$\mathrm{SI}\left(x_i^{k}\right) = \prod_{u=1}^{|V_{B,i}|}\mathrm{SI}\left(A_{i,u}^{k}, x_i^{k}\right)\prod_{v=1}^{|V_{T,i}|}\mathrm{SI}\left(M_{i,v}^{k}, x_i^{k}\right), \quad \mathrm{SI}\left(y_i^{k}\right) = \prod_{u=1}^{|V_{B,i}|}\mathrm{SI}\left(B_{i,u}^{k}, y_i^{k}\right)\prod_{v=1}^{|V_{T,i}|}\mathrm{SI}\left(N_{i,v}^{k}, y_i^{k}\right). \tag{25}$$

4.2. The Characteristics of FG-Based CGL. From Figure 3, it can be concluded that the passing messages from BSs or cooperative MTs only merge in variable nodes x_i and y_i, while the passing message in other variable nodes is just related with the information in their local branches. The estimation for x_i (or y_i) is $m_{\Lambda,i} = \sigma_{\Lambda,i}^2 \sum_{u=1}^{U}(m_{i,u}/\sigma_{i,u}^2)$, where $\sigma_{\Lambda,i}^2 = 1/(\sum_{u=1}^{U} 1/\sigma_{i,u}^2)$. $m_{i,u}$ represents the mean of the information in the uth branch (no matter the BS branch or the MT branch) connecting to the ith MT, and, $\sigma_{i,u}^2$ represents the variance of the information in the uth branch connecting to the ith MT. Moreover, it is supposed that the variances of the information from BSs are the same (denoted as $C = 1/\sigma_{i,j}^2$), and the variances of the information from cooperative MTs are the same as well (denoted as $D = 1/\sigma_{i,t}^{'2}$). Therefore, $m_{\Lambda,i} = \sigma_{\Lambda,i}^2 \sum_{u=1}^{U}(m_{i,u}/\sigma_{i,u}^2)$ can be written as $m_{\Lambda,i} = (C\sum_{j=1}^{|V_{B,i}|} m_{i,j} + D\sum_{t=1}^{|V_{T,i}|} m'_{i,t})/(C|V_{B,i}| + D|V_{T,i}|)$, where $|V_{B,i}|$ denotes the number of BSs connecting to the ith MT, and $|V_{T,i}|$ denotes the number of cooperative MTs connecting to the ith MT. The estimated localization error is

$$\Delta_i = m_{\Lambda,i} - \overline{m}_i = \frac{C\sum_{j=1}^{|V_{B,i}|} m_{i,j} + D\sum_{t=1}^{|V_{T,i}|} m'_{i,t}}{C|V_{B,i}| + D|V_{T,i}|} - \overline{m}_i, \tag{26}$$

where $\overline{m}_i$ is the real value.

According to the form of (26), the localization error does not decrease linearly with the growing number of BSs or the number of cooperative MTs, and it is affected by the value of C/D. When the number of BSs and MTs is definite, the localization error can reduce with the increasing

value of C/D. Moreover, the influence of $|V_{B,i}|$ and $|V_{T,i}|$ on localization accuracy weakens with the increasing value of C/D. Therefore, the strict demand for topology is relieved which embodies the robustness of the algorithms.

4.3. The Formulation of WFG-CGL Algorithm. Obviously, the information provided by the cooperative MT is less reliable than that of BS. Generally, the variances information represents the reliability. The information reliability lowers with the larger variance. If the variance information is multiplied by a weighted coefficient based on its reliability before performing the position estimation, the higher localization accuracy can be obtained. Considering the influence from the weighted coefficient, a *weights optimization method* is proposed, which can calculate the optimal weights for multiple MTs, respectively.

Denote the localization error as $\mathbf{e}$ and

$$\mathbf{e} = \begin{bmatrix} (x_1 - \overline{x}_1)^2 + \left(y_1 - \overline{y}_1\right)^2 \\ \cdots \\ (x_{|V_T|} - \overline{x}_{|V_T|})^2 + \left(y_{|V_T|} - \overline{y}_{|V_T|}\right)^2 \end{bmatrix}, \tag{27}$$

where x_i and y_i are the functions of β_i; $\overline{x}_i$ and $\overline{y}_i$ are the real position for the ith MT.

Take one element of the matrix in (27) *as an example to calculate the corresponding weight.*

Denote $\mu_{x,i}$ and $\sigma^2_{x,i}$ represent the mean and variance of the Gaussian variable x_i, $\mu_{y,i}$, and $\sigma^2_{y,i}$ represent the mean and variance of the Gaussian variable y_i; $\overline{x}_i$ and $\overline{y}_i$ are the constants.

$$\varepsilon_{x,i} = x_i - \overline{x}_i \sim N\left(x_i - \overline{x}_i, \mu_{e,x,i}, \sigma^2_{x,i}\right), \tag{28}$$

$$\varepsilon_{y,i} = y_i - \overline{y}_i \sim N\left(y_i - \overline{y}_i, \mu_{e,y,i}, \sigma^2_{y,i}\right), \tag{29}$$

$$\mu_{e,x,i} = E[x_i - \overline{x}_i] = \mu_{x,i} - \overline{x}_i, \tag{30}$$

$$\mu_{e,y,i} = E\left[y_i - \overline{y}_i\right] = \mu_{y,i} - \overline{y}_i. \tag{31}$$

Therefore, $\varepsilon^2_{x,i}$ and $\varepsilon^2_{y,i}$ follow the noncentral Chi-squared distribution. $E(\varepsilon^2_{x,i}) = n_x\sigma^2_{x,i} + s^2_x$, where n_x is the freedom degrees of the ith MT in x-coordinate and $s^2_x = \sum_{i=1}^{n_x} \mu^2_{e,x,i}$. Due to the fact that $\varepsilon^2_{x,i}$ is independent of $\varepsilon^2_{y,i}$ as follows:

$$\begin{aligned} E\left(\varepsilon^2_{x,i} + \varepsilon^2_{y,i}\right) &= E\left(\varepsilon^2_{x,i}\right) + E\left(\varepsilon^2_{y,i}\right) \\ &= \sigma^2_{x,i} + s^2_x + \sigma^2_{y,i} + s^2_y \\ &= \sigma^2_{x,i} + \mu^2_{e,x,i} + \sigma^2_{y,i} + \mu^2_{e,y,i}, \end{aligned} \tag{32}$$

where $\sigma^2_{x,i}$, $\mu^2_{e,x,i}$, $\sigma^2_{y,i}$, and $\mu^2_{e,y,i}$ are the functions of β_i.

The optimal weight adaptation achieves minimum localization error for multiple MTs, respectively,

$$\min_{\beta_i} \arg E\left(\varepsilon^2_{x,i} + \varepsilon^2_{y,i}\right), \tag{33}$$

$$\frac{\partial E\left(\varepsilon^2_{x,i} + \varepsilon^2_{y,i}\right)}{\partial \beta_i} = 0, \tag{34}$$

$$\frac{\partial \sigma^2_{x,i}}{\partial \beta_i} + 2\mu_{e,x,i}\frac{\partial \mu_{e,x,i}}{\partial \beta_i} + \frac{\partial \sigma^2_{y,i}}{\partial \beta_i} + 2\mu_{e,y,i}\frac{\partial \mu_{e,y,i}}{\partial \beta_i} = 0, \tag{35}$$

where $\sigma^2_{x,i}$ and $\mu_{e,x,i}$ are based on (15) and (30).

$$\frac{\partial \sigma^2_{x,i}}{\partial \beta_i} = \frac{\sum_{j=1}^{|V_{T,i}|}\left(1/\sigma^2_{\Delta xx,i,j}\right)}{\left(\sum_{j=1}^{|V_{B,i}|}(1/\sigma^2_{\Delta x,i,j}) + \sum_{j=1}^{|V_{T,i}|}(1/\sigma^2_{\Delta xx,i,j})\right)^2}, \tag{36}$$

$$\begin{aligned} \frac{\partial \mu_{e,x,i}}{\partial \beta_i} = \frac{\partial \sigma^2_{x,i}}{\partial \beta_i}&\left(\sum_{j=1}^{|V_{B,i}|}\frac{\mu_{\Delta x,j}}{\sigma^2_{\Delta x,i,j}} + \sum_{j=1}^{|V_{T,i}|}\frac{\mu_{\Delta xx,j}}{\sigma^2_{\Delta xx,i,j}}\right) \\ &- \sigma^2_{x,i}\cdot\sum_{j=1}^{|V_{T,i}|}\frac{\mu_{\Delta xx,j}}{\sigma^2_{\Delta xx,i,j}}, \end{aligned} \tag{37}$$

where only $\mu_{\Delta xx,j}$ and $\sigma^2_{\Delta xx,i,j}$ are the functions of β_i. Numerical methods can be utilized to obtain β_i estimation.

According to the above analysis, the procedures of WFG-CGL algorithm are described as follows.

Step 1 (initialization). All variable nodes need deterministic initial values, except $d_{i,j}$ and $l_{i,t}$. Initial values can be made at random or by some specific algorithms, not detailed here. Set $k = 0$.

Step 2 (entrance of measurement information in the kth iteration). The measurement information $\hat{d}^k_{i,j}$ and $\hat{l}^k_{i,t}$ from TOA measurement enter FG through function nodes $D^k_{i,j}$ and $L^k_{i,t}$, according to (24).

Step 3 (uplink calculation in the kth iteration). Function node $C_{i,j}$ updates $\Delta y_{i,j}$ and $\Delta x_{i,j}$ according to (14), and (15). Similarly, function node $CC_{i,t}$ updates $\Delta xx_{i,t}$ and $\Delta yy_{i,t}$. β_i is calculated in (35), and function nodes $R_{i,j}$ and $RR_{i,t}$ update $\Delta x^\beta_{i,j}$, $\Delta y^\beta_{i,j}$, $\Delta xx^\beta_{i,t}$, and $\Delta yy^\beta_{i.t}$ according to (18) and other similar formulae. Function nodes $A_{i,j}$, $B_{i,j}$, $M_{i,t}$, and $N_{i,t}$ update the messages for x_i and y_i according to (12) and other similar formulae.

Step 4 (downlink calculation in the kth iteration). Variable nodes x_i transmit message to $A_{i,j}$ and $M_{i,t}$ according to (3). Similarly, y_i transmits message to $B_{i,j}$ and $N_{i,t}$. Function node $A_{i,j}$ updates $\Delta x^\beta_{i,j}$ according to (11). Similarly, function nodes $B_{i,j}$, $M_{i,t}$, and $N_{i,t}$ update $\Delta y^\beta_{i,j}$, $\Delta xx^\beta_{i,t}$, and $\Delta yy^\beta_{i,t}$. Function node $R_{i,j}$ updates $\Delta x_{i,j}$ and $\Delta y_{i,j}$ according to (22) and (23). Similarly, function node $RR_{i,t}$ updates $\Delta xx_{i,t}$ and $\Delta yy_{i,t}$.

Step 5. In the kth iteration, Variable nodes x_i and y_i are updated and estimate the position of the ith MT according to (25).

Step 6. Targeting for the mth MT who has not updated the position in the kth iteration, we make $i = m$, return to Step 2. Else, when all of the cooperative MTs in a TG have completed the position updating in the kth iteration, we make $k \leftarrow k+1$.

Step 7. If $k < K$, return to Step 2. Else, $x_i = x_i^k$ and $y_i = y_i^k$, the algorithm stops.

4.4. The Formulation of LCFG-CGL Algorithm. In WFG-CGL algorithm, the solution to the weights increases the calculation complexity. In order to further diminish the complexity, the weights for all MTs in TG are not differentiated, leading to the fact that all weights are considered as 1. Therefore, we propose LCFG-CGL algorithm to solve the problem, and the procedures of *LCFG-CGL algorithm* are described as follows.

Step 1 (initialization). All variable nodes need deterministic initial values, except $d_{i,j}$ and $l_{i,t}$. Initial values can be made at random or by some specific algorithms, not detailed here. Set $k = 0$.

Step 2 (entrance of measurement information in the kth iteration). The measurement information $\hat{d}_{i,j}^k$ and $\hat{l}_{i,t}^k$ from TOA measurement enter FG through function nodes $D_{i,j}^k$ and $L_{i,t}^k$, according to (24).

Step 3 (uplink calculation in the kth iteration). Function node $C_{i,j}$ updates $\Delta y_{i,j}$ and $\Delta x_{i,j}$ according to (14) and (15). Similarly, function node $CC_{i,t}$ updates $\Delta xx_{i,t}$ and $\Delta yy_{i,t}$. All β_i are set as 1, and function nodes $R_{i,j}$, and $RR_{i,t}$ work as transfers to $\Delta x_{i,j}^{\beta}$, $\Delta y_{i,j}^{\beta}$, $\Delta xx_{i,t}^{\beta}$, and $\Delta yy_{i,t}^{\beta}$. Function nodes $A_{i,j}$, $B_{i,j}$, $M_{i,t}$, and $N_{i,t}$ update the messages for x_i and y_i according to (12) and other similar formulae.

Step 4 (downlink calculation in the kth iteration). Variable nodes x_i transmits message to $A_{i,j}$ and $M_{i,t}$ according to (3). Similarly, y_i transmits message to $B_{i,j}$ and $N_{i,t}$. Function node $A_{i,j}$ updates $\Delta x_{i,j}^{\beta}$ according to (11). Similarly, function nodes $B_{i,j}$, $M_{i,t}$, and $N_{i,t}$ update $\Delta y_{i,j}^{\beta}$, $\Delta xx_{i,t}^{\beta}$, and $\Delta yy_{i,t}^{\beta}$; Function node $R_{i,j}$ updates $\Delta x_{i,j}$ and $\Delta y_{i,j}$ according to (22)and (23). Similarly, function node $RR_{i,t}$ updates $\Delta xx_{i,t}$ and $\Delta yy_{i,t}$.

Step 5. In the kth iteration, variable nodes x_i and y_i are updated and estimate the position of the ith MT according to (25).

Step 6. Targeting for the mth MT who has not updated the position in the kth iteration, we make $i = m$, return to Step 2. Else, when all of the cooperative MTs in a TG have completed the position updating in the kth iteration, we make $k \leftarrow k+1$.

Step 7. If $k < K$, return to Step 2. Else, $x_i = x_i^k$ and $y_i = y_i^k$, the algorithm stops.

5. CRLB Analysis

The CRLB is commonly used as a performance benchmark of an estimator because it gives the lowest possible variance [27]. Although the CRLB analysis is a classical localization metric, the CRLB analysis targeting CGL has not been performed in detail. Here, we make the CRLB analysis of CGL firstly.

Considering a vector of all nodes parameters $\gamma = [\gamma_1, \gamma_2 \ldots, \gamma_{|V|}]$ in CGL, nodes $1, 2, \ldots, |V_T|$ represent the blindfolded MTs, and nodes $|V_T|+1, |V_T|+2, \ldots, |V_T|+|V_B|$ represent the BSs. The unknown MT parameter vector is $\theta = [\theta_1, \theta_2, \ldots, \theta_{|V_T|}]$, where $\theta_i = \gamma_i$ for $i = 1, 2, \ldots, |V_T|$. Note that the BS parameter vector $\{\gamma_i,\ i = |V_T| + 1, M + 2, \ldots, |V_T| + |V_B|\}$ is known. Nodes i and j make pair-wise observations $Z_{i,j}$ with density $p_{Z|\gamma}(Z_{i,j} \mid \gamma_i, \gamma_j)$. Let $H(i) = j$: node j makes pair-wise observations with node i. By convention, a node cannot make pair-wise observation with itself, so that $i \notin H(i)$. By symmetry, if $j \in H(i)$, then $i \in H(j)$.

It is assumed by reciprocity that $Z_{i,j} = Z_{j,i}$. Thus, it is sufficient to consider only the lower triangle of the observation matrix $Z = ((Z_{i,j}))_{i,j}$ when formulating the joint likelihood function. We assume that $\{Z_{i,j}\}$ are statistically independent for $j < i$. This assumption can be somewhat oversimplified but necessary for analysis. The log of the joint conditional PDF is

$$l(Z \mid \gamma) = \sum_{i=1}^{|V|} \sum_{j \in H(i), j<i} l_{i,j}, \qquad l_{i,j} = \ln p_{Z|\gamma}\left(Z_{i,j} \mid \gamma_i, \gamma_j\right). \tag{38}$$

The CRLB on the covariance matrix of any unbiased estimator $\hat{\theta}$ is $\mathrm{cov}(\hat{\theta}) \geq F_\theta^{-1}$, where the Fisher information matrix (FIM) F_θ is defined as

$$F_\theta = -E\nabla_\theta\left(\nabla_\theta l(Z \mid \gamma)\right)^T = \begin{bmatrix} f_{1,1} & \cdots & f_{1,|V|} \\ \vdots & \ddots & \vdots \\ f_{|V|,1} & \cdots & f_{|V|,|V|} \end{bmatrix}, \tag{39}$$

where the diagonal elements $f_{k,k}$ for $k = 1, 2, \ldots, |V|$ of F_θ can be simplified to a single sum over $H(k)$ since there are $H(k)$ term in (38) that depend on $\theta_k = \gamma_k$. The off-diagonal elements can be further reduced: when $k \neq l$ for $k = 1, 2, \ldots, |V_T|$; $l = 1, 2, \ldots, |V_T|$, there is at most one summand in (38) that is a function of both k and l.

$$f_{k,l} = \begin{cases} -\sum\limits_{j \notin H(k)} E\left[\dfrac{\partial^2}{\partial {\theta_k}^2} l_{k,j}\right], & k = l, \\ -I_{H(k)}(l) E\left[\dfrac{\partial^2}{\partial \theta_k \theta_l} l_{k,l}\right], & k \neq l, \end{cases} \tag{40}$$

where $I_{H(k)}(l)$ is an indicator function: 1 if $l \in H(k)$ or 0 otherwise.

For a 2D system, $\gamma = [\gamma_1, \gamma_2 \ldots, \gamma_{|V|}]$, where $\gamma_i = [x_i, y_i]^T$. The relative localization problem corresponds to the estimation of blindfolded device coordinates $\theta = [\theta_x, \theta_y]$, where $\theta_x = [x_1, x_2, \ldots, x_{|V_T|}]$, $\theta_y = [y_1, y_2, \ldots, y_{|V_T|}]$, and the known BSs coordinates are $[x_{|V_T|+1}, \ldots, x_{|V_T|+|V_B|}, y_{|V_T|+1}, \ldots, y_{|V_T|+|V_B|}]$. In the cellular network, $T_{i,j}$, $i = 1, \ldots, |V_T|$; $j = |V_T| + 1, \ldots, |V_T| + |V_B|$ is the measured TOA between BSs and MTs. $T_{i,i'}$, $i = 1, \ldots, |V_T|$;

$i' = 1,\ldots,|V_T|$; $i \neq i'$ is the measured TOA between MTs. Assuming that both $T_{i,j}$ and $T_{i,i'}$ are independent Gaussian distributed, which is denoted as

$$T_{i,j} \sim N\left(\frac{d_{i,j}}{c}, \delta^2_{\mathrm{MB}}\right), \tag{41}$$

where $i = 1,\ldots,|V_T|$; $j = |V_T|+1,\ldots,|V_T|+|V_B|$.

$$T_{i,i'} \sim N\left(\frac{d_{i,i'}}{c}, \delta^2_{\mathrm{MT}}\right), \tag{42}$$

where $i = 1,\ldots,|V_T|$; $i' = 1,\ldots,|V_T|$; $i \neq i'$.

$$d_{i,j} = d\left(\gamma_i, \gamma_j\right) = \left\|\gamma_i - \gamma_j\right\|, \tag{43}$$

$$d_{i,i'} = d(\gamma_i, \gamma_{i'}) = \|\gamma_i - \gamma_{i'}\|, \tag{44}$$

$$f_{\mathrm{MB}}\left(\frac{T_{i,j}}{\gamma}\right) = \frac{1}{\sqrt{2\pi\delta^2_{\mathrm{MB}}}} \mathrm{e}^{-(T_{i,j}-d_{i,j}/c)^2/2\delta^2_{\mathrm{MB}}}, \tag{45}$$

$$f_{\mathrm{MT}}\left(\frac{T_{i,i'}}{\gamma}\right) = \frac{1}{\sqrt{2\pi\delta^2_{\mathrm{MT}}}} \mathrm{e}^{-(T_{i,i'}-d_{i,i'}/c)^2/2\delta^2_{\mathrm{MT}}}, \tag{46}$$

where c is the speed of light propagation, δ^2_{MB} and δ^2_{MT} denote the variance of TOA measurements between BSs and MTs, and between MTs, respectively.

For simplification, the FIM will have a similar form to (39) if partitioned into blocks

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_{xx} & \mathbf{F}_{xy} \\ \mathbf{F}_{yx} & \mathbf{F}_{yy} \end{bmatrix}, \tag{47}$$

where $\mathbf{F}_{xx}$ is given by (39) using only the x parameter vector $\theta = \theta_x$, and $\mathbf{F}_{yy}$ is given by (39) using only $\theta = \theta_y$. The off-diagonal blocks $\mathbf{F}_{xy}$ and $\mathbf{F}_{yx}$ are similarly derived. According to this cooperative structure, the elements of the submatrics of (47) are derived as follows:

$$\begin{aligned} l\left(\frac{X}{\gamma}\right) = &\sum_{i=1}^{|V_T|}\sum_{j=M+1}^{|V_T|+|V_B|} \ln\left\{f_{\mathrm{MB}}\left(\frac{T_{i,j}}{\gamma}\right)\right\} \\ &+ \sum_{i=1}^{|V_T|}\sum_{\substack{i'=1 \\ i'\neq i}}^{|V_T|} \ln\left\{f_{\mathrm{MT}}\left(\frac{T_{i,i'}}{\gamma}\right)\right\}. \end{aligned} \tag{48}$$

And the FIM is for $k = l$

$$\begin{aligned} [\mathbf{F}_{xx}]_{k,l} = &\frac{1}{c^2\delta^2_{\mathrm{MB}}} \sum_{j\in H(k)} \frac{\left(x_k - x_j\right)^2}{\left\|\gamma_k - \gamma_j\right\|^2} \\ &+ \frac{1}{c^2\delta^2_{\mathrm{MT}}} \sum_{i\in H(k)} \frac{(x_k - x_i)^2}{\|\gamma_k - \gamma_i\|^2}, \end{aligned} \tag{49}$$

$$\begin{aligned} \left[\mathbf{F}_{yy}\right]_{k,l} = &\frac{1}{c^2\delta^2_{\mathrm{MB}}} \sum_{j\in H(k)} \frac{\left(y_k - y_j\right)^2}{\left\|\gamma_k - \gamma_j\right\|^2} \\ &+ \frac{1}{c^2\delta^2_{\mathrm{MT}}} \sum_{i\in H(k)} \frac{(y_k - y_i)^2}{\|\gamma_k - \gamma_i\|^2}, \end{aligned} \tag{50}$$

$$\begin{aligned} \left[\mathbf{F}_{xy}\right]_{k,l} = &\frac{1}{c^2\delta^2_{\mathrm{MB}}} \sum_{j\in H(k)} \frac{\left(x_k - x_j\right)\left(y_k - y_j\right)}{\left\|\gamma_k - \gamma_j\right\|^2} \\ &+ \frac{1}{c^2\delta^2_{\mathrm{MT}}} \sum_{i\in H(k)} \frac{(x_k - x_i)(y_k - y_i)}{\|\gamma_k - \gamma_i\|^2}, \end{aligned} \tag{51}$$

for $k \neq l$,

$$\begin{aligned} [\mathbf{F}_{xx}]_{k,l} = &\frac{-I_{H(k)}(l)}{c^2\left(J_{\mathrm{MB}}(k,l)\delta^2_{\mathrm{MB}} + J_{\mathrm{MT}}(k,l)\delta^2_{\mathrm{MT}}\right)} \\ &\cdot \frac{(x_k - x_l)^2}{\|\gamma_k - \gamma_l\|^2}, \end{aligned} \tag{52}$$

$$\begin{aligned} \left[\mathbf{F}_{yy}\right]_{k,l} = &\frac{-I_{H(k)}(l)}{c^2\left(J_{\mathrm{MB}}(k,l)\delta^2_{\mathrm{MB}} + J_{\mathrm{MT}}(k,l)\delta^2_{\mathrm{MT}}\right)} \\ &\cdot \frac{(y_k - y_l)^2}{\|\gamma_k - \gamma_l\|^2}, \end{aligned} \tag{53}$$

$$\begin{aligned} \left[\mathbf{F}_{xy}\right]_{k,l} = &\frac{-I_{H(k)}(l)}{c^2\left(J_{\mathrm{MB}}(k,l)\delta^2_{\mathrm{MB}} + J_{\mathrm{MT}}(k,l)\delta^2_{\mathrm{MT}}\right)} \\ &\cdot \frac{(x_k - x_l)(y_k - y_l)}{\|\gamma_k - \gamma_l\|^2}, \end{aligned} \tag{54}$$

where $I_{H(k)}(l)$ is 1 when node k and node l can make pairwise measurements, else $I_{H(k)}(l)$ is zero. J_{MB} is 0 when $k,l = 1,2,\ldots,|V_T|$ or $k,l = |V_T|+1,\ldots,|V_T|+|V_B|$, else J_{MB} is 1. J_{MT} is 1 when $k,l = 1,2,\ldots,|V_T|$, else J_{MT} is 0.

The formula derivation of FIM has been completed in (47)~(52), (53) and (54). Then, the trace $\hat{\theta}_{\mathrm{CRLB}}$ of the covariance for the ith blind MT satisfies

$$\begin{aligned} \hat{\theta}_{\mathrm{CRLB},i} = \mathrm{tr}\left(\mathrm{cov}_{\hat{\theta}}(\hat{x}_i, \hat{y}_i)\right) \geq &\left(\left[\mathbf{F}_{xx} - \mathbf{F}_{xy}\mathbf{F}_{yy}^{-1}\mathbf{F}_{xy}^T\right]^{-1}\right)_{i,i} \\ &+ \left(\left[\mathbf{F}_{yy} - \mathbf{F}_{xy}\mathbf{F}_{xx}^{-1}\mathbf{F}_{xy}^T\right]^{-1}\right)_{i,i}, \end{aligned} \tag{55}$$

$$\hat{\theta}_{\mathrm{CRLB}} = \sum_{i=1}^{|V_T|} \frac{\hat{\theta}_{\mathrm{CRLB},i}}{|V_T|}. \tag{56}$$

6. Simulation Results

In this section, numerical simulations include two parts. Part I is the research on important factors for the proposed algorithms. Part II is performed to show the effectiveness of the proposed WFG-CGL and LCFG-CGL algorithms under two typical topologies for localization research, including global rigidity which has the unique solvability shown in

Table 1: Parameters configuration.

The number of BSs ($\|V_B\|$)	3
The number of MTs in a TG ($\|V_T\|$)	3
Cell radius (m)	3000
TG radius (m)	750

Figure 4(a), and rigidity which has no unique solvability and converges in the finite positions shown in Figure 4(b). The topologies with three MTs are taken as examples due to the fact that the simple and specific topologies are easy to follow. The comparison of localization accuracy and complexity between the proposed algorithms and the existing Taylor-series-based CGL algorithm [18] is performed. Moreover, CRLB analysis is conducted.

Any topology scenario is constructed by the random birth of MTs in the efficient range of BSs. The average error is evaluated over 10,000 independent trials. While in the different scenarios with the same number of MTs, the positions of MTs and BSs are the same. First, we consider the edges existing between BSs and MTs or between MTs LOS paths. The distance measurements derived from TOA between MTs and BSs are $d = d_0 + n$, the distance measurements between MTs are $l = l_0 + q$, where d_0 denotes the real distance between BS and MT, l_0 denotes the real distance between MTs, and n and q denote the corresponding measurement error and follow the Gaussian distribution with zero mean value. Moreover, in urban environment, we consider NLOS propagation case. The NLOS error was modeled as the exponential distribution random variable in the literatures [28, 29]. The probability density function (PDF) of NLOS error k follows

$$p(k) = \begin{cases} \frac{1}{K_{\text{rms}}} e^{-k/K_{\text{rms}}} & k > 0 \\ 0 & \text{otherwise} \end{cases} \quad (57)$$

where K_{rms} denotes the mean of k. Referring to the literature [29], K_{rms} is set as 100 m.

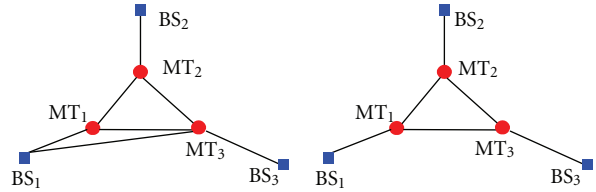
The basic parameter configuration is listed in Table 1. The localization accuracy is analyzed by means of root mean square error (RMSE) between the real MTs' position $p_{\text{MT}}(i)$ and the estimated location $p'_{\text{MT}}(i)$, that is,

$$E_{\text{RMSE}} = \frac{\sum_{i=1}^{|V_T|} \|p_{\text{MT}}(i) - p'_{\text{MT}}(i)\|_2}{|V_T|}. \quad (58)$$

6.1. The Important Factors for FG-Based CGL

(1) The Iteration Times. for the FG-based localization, the iteration times is predefined and has strong influence on the localization accuracy. Apparently, more iteration times can make the localization more precise. However, excessive iterations will lead to the increasing complexity. Therefore, the reasonable value of K should be investigated.

The intuitive results can be obtained from Figures 5 and 6 which shows the localization error in terms of the iteration times from $K = 1$ to $K = 15$ with the topologies of



(a) CGL topology(global rigidity) (b) CGL topology (rigidity)

Figure 4: CGL topology.

Table 2: Statistics analysis for difference σ_q^2.

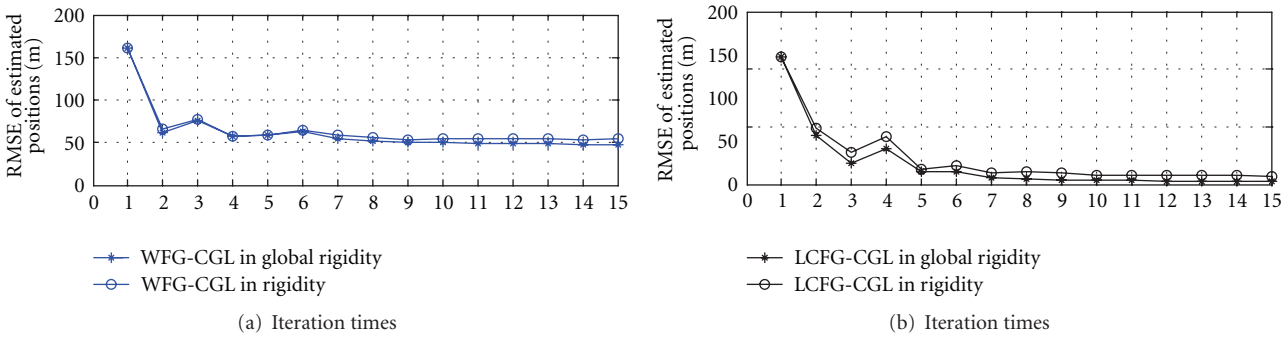
		G: (μ, σ^2)	
		Global rigidity	Rigidity
LCFG	LOS	(0.134, 0.143)	(0.200, 0.162)
	NLOS	(0.131, 0.002)	(0.198, 0.002)
WFG	LOS	(0.219, 0.057)	(0.175, 0.101)
	NLOS	(0.159, 0.002)	(0.167, 0.001)

Figures 4(a) and 4(b) in LOS and NLOS scenarios. It can be observed from Figures 5 and 6 that no matter LOS or NLOS scenario, the performance has no significant improvement with the growth of K after $K = 7$ in both the rigidity and the global rigidity. Considering the high accuracy and low complexity simultaneously, $K = 7$ is adopted in the following simulations.

(2) The Influence of σ_n and σ_q. Figures 7 and 8 indicate the localization error of LOS and NLOS paths between MTs and BSs, respectively. As shown in Figure 7(a) and Figure 8(a) where $\sigma_q^2 = 10\,\text{m}^2$ in LOS and NLOS scenarios, and Figure 7(b) and Figure 8(b), where $\sigma_q^2 = 100\,\text{m}^2$ in LOS and NLOS scenarios, they indicate that the standard deviation between MT and BS σ_n has stronger effects on the localization accuracy than the standard deviation between MTs σ_q for the proposed algorithms in both LOS and NLOS scenarios. The main reason is attributive to the fact that BS can provide the information of higher accuracy, compared to the cooperative MT. However, the information in NLOS scenario is more inaccurate than that in LOS scenario due to the fact the influence of σ_n in NLOS is slighter than that in LOS. This can be illustrated in Table 2.

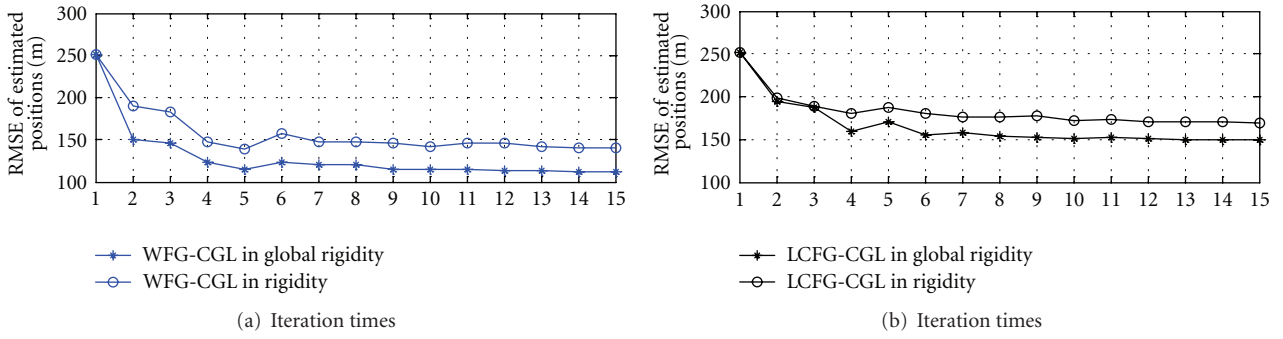
Denote $G = (\text{RMSE}_{100} - \text{RMSE}_{10})/\text{RMSE}_{100}$, where RMSE_{100} represents the RMSE in $\sigma_q^2 = 100\,\text{m}^2$, and RMSE_{10} represents the RMSE in $\sigma_q^2 = 10\,\text{m}^2$. μ_G and σ_G^2 represent the mean and variance of G.

The statistic data can be obtained from Table 2 which shows μ_G and σ_G^2 in terms of σ_n from $\sigma_n = 5\,\text{m}$ to $\sigma_n = 45\,\text{m}$ for the two proposed algorithms in different scenarios. No matter the LCFG-CGL algorithm or the WFG-CGL algorithm, no matter the global rigidity topology or the rigidity topology, and no matter the LOS or NLOS, the values of μ_G and σ_G^2 are all small which verifies the little performance difference between $\sigma_q^2 = 10\,\text{m}^2$ and $\sigma_q^2 = 100\,\text{m}^2$. Thereinto,



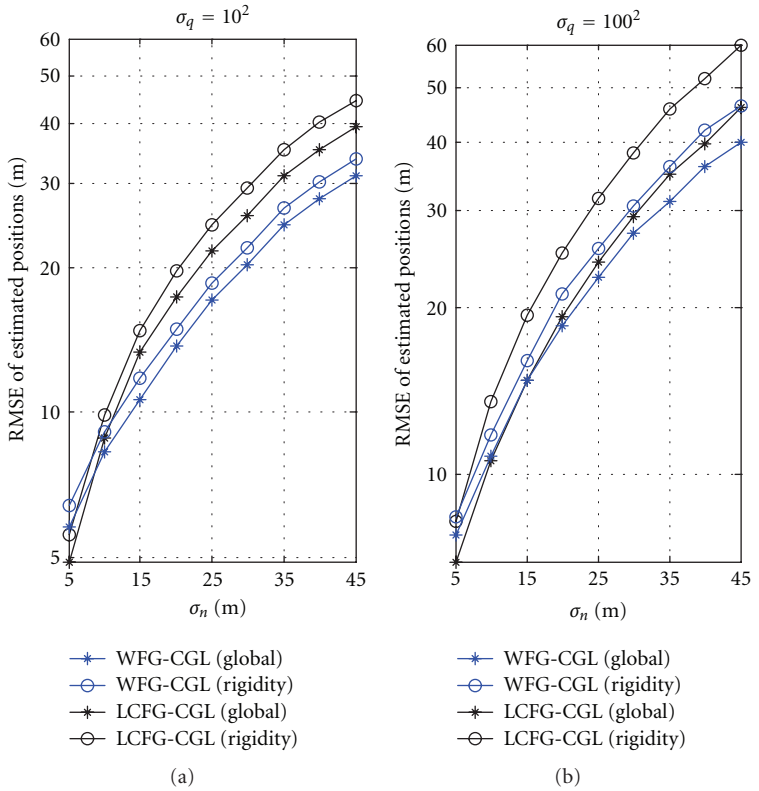
(a) Iteration times (b) Iteration times

Figure 5: The influence of iteration times in LOS.



(a) Iteration times (b) Iteration times

Figure 6: The influence of iteration times in NLOS.



(a) (b)

Figure 7: RMSE of difference σ_q^2 in LOS.

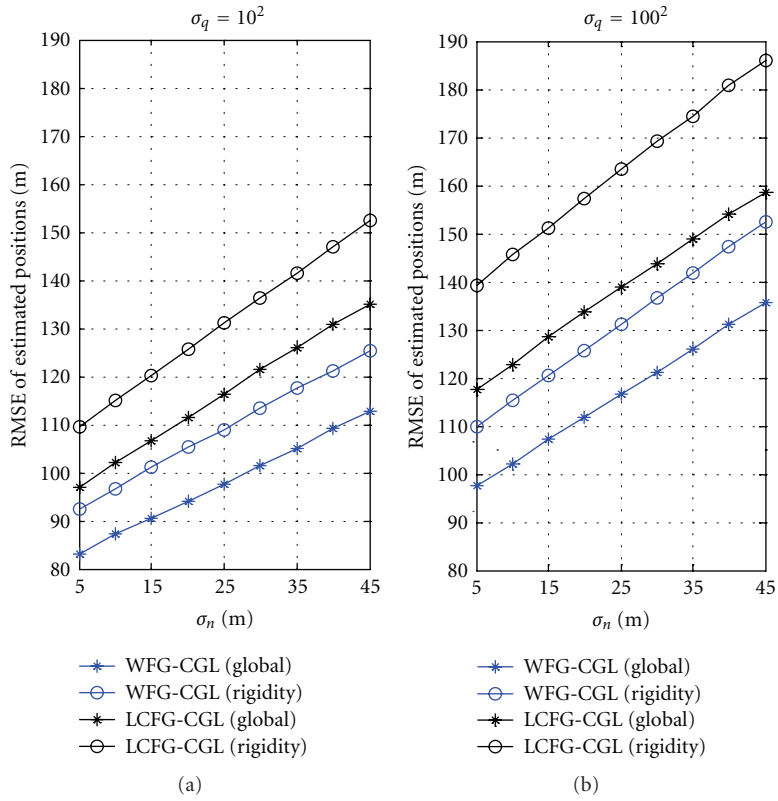


Figure 8: RMSE of difference σ_q^2 in NLOS.

we just take $\sigma_q^2 = 10\,\text{m}^2$ as an example in the following simulations.

(3) The Influence of $|V_B|$. According to the analysis in Section 4.2 and the simulation results in Figure 9, the localization error decreases with the growing number of the cooperative BSs, while the decline rate of the localization error becomes smaller with the growing number of the cooperative BSs. It is assumed that the variance of the measurement error between MTs σ_q^2 is definite. Figure 9 illustrates that the value of C/D reduces with the rising of σ_n^2, and the influence of $|V_B|$ on localization accuracy weakens with the increasing value of C/D, no matter in LOS and NLOS scenarios. The simulation results consist with the analysis in Section 4.3, which can verify the robustness of our algorithms for the topology of CGL graph. It also clarifies the localization accuracy cannot be reduced through increasing the amount of measurement information (the number of BSs) when the measurement information is inaccurate. Therefore, the localization accuracy can be improved, only through optimizing the algorithm. Fortunately, our proposed algorithms can achieve this.

6.2. The Localization Accuracy Analysis. The simulation results are shown in Figures 10 and 11, where the average localization errors versus σ_n of the range measurement errors are compared in LOS and NLOS scenarios. The comparison is made among the Taylor-series-based CGL algorithm, the LCFG-CGL algorithm, the WFG-CGL algorithm and CRLB in the same scenario, where the Taylor-series-based CGL algorithm is analyzed in [18]. As expected, no matter the global rigidity or the rigidity, (1) the average localization errors increase when the range measurement error increases, (2) the proposed algorithms are superior to the traditional Taylor series CGL algorithm, (3) the WFG-CGL algorithm has higher localization accuracy than LCFG-CGL algorithm primarily due to its exploitation of the information reliability as the weights, (4) in NLOS scenario, LCFG-CGL algorithm and WFG-CGL algorithm perform far well than the traditional Taylor-series CGL algorithm, and (5) especially, the WFG-CGL algorithm is approximate to CRLB results in both global rigidity and rigidity regarding the LOS scenario.

Denote $G_{tp} = (\text{RMSE}_R - \text{RMSE}_G)/\text{RMSE}_R$, where RMSE_R represents the RMSE in rigidity topology, and RMSE_G represents the RMSE in global rigidity topology. $\mu_{G_{tp}}$ and $\sigma_{G_{tp}}^2$ represent the mean and variance of G_{tp}.

Remarkably, Taylor-series-based CGL algorithm performs better in global rigidity topology than in rigidity topology. It is obvious that the global rigidity topology can provide more localization information. However, the two proposed algorithms in rigidity have the similar localization accuracy in global rigidity, as shown in Figures 10 and 11. Table 3 derived from Figures 10 and 11 which shows the $\mu_{G_{tp}}$ and $\sigma_{G_{tp}}^2$ in terms of σ_n from $\sigma_n = 5\,\text{m}$ to $\sigma_n = 45\,\text{m}$ in different scenarios and different algorithms. It denotes that the two proposed algorithms in rigidity have

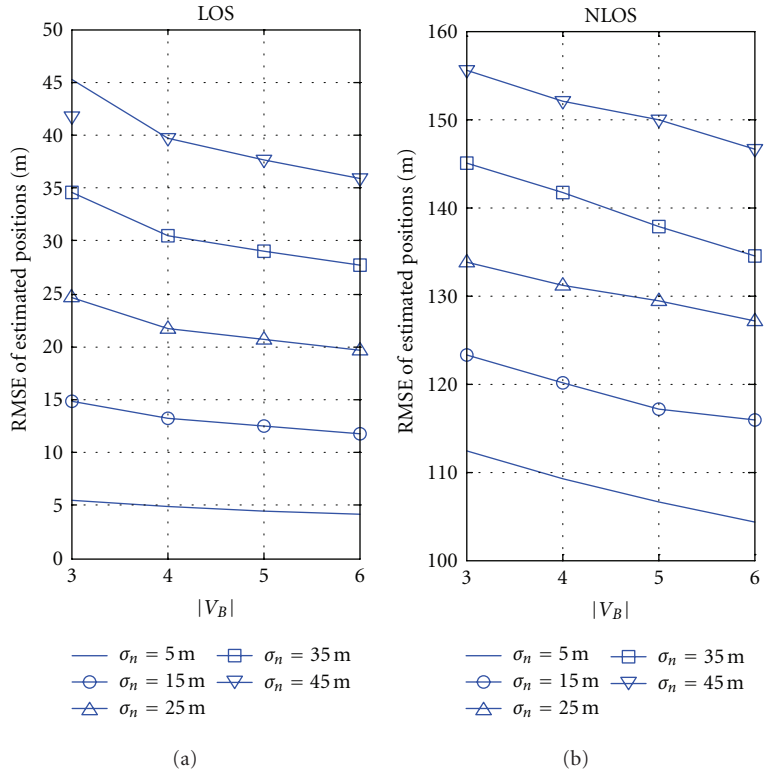


Figure 9: RMSE of difference $|V_B|$.

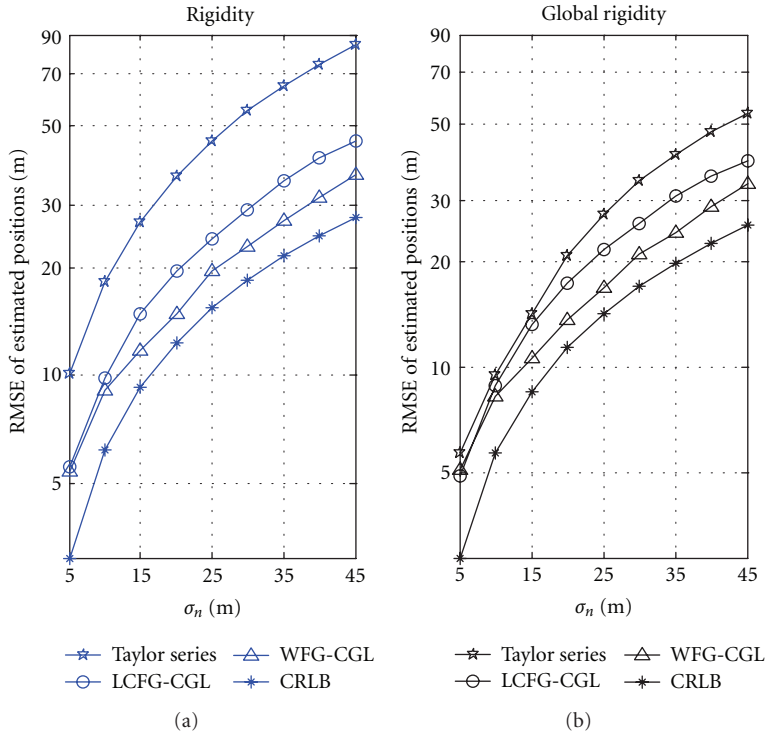


Figure 10: RMSE of different topologies in LOS.

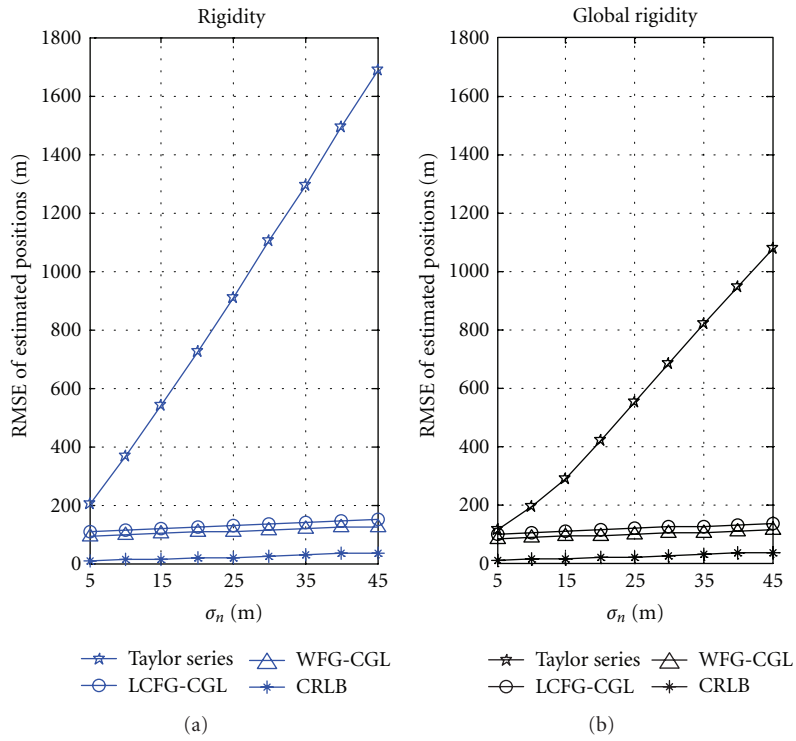


Figure 11: RMSE of different topologies in NLOS.

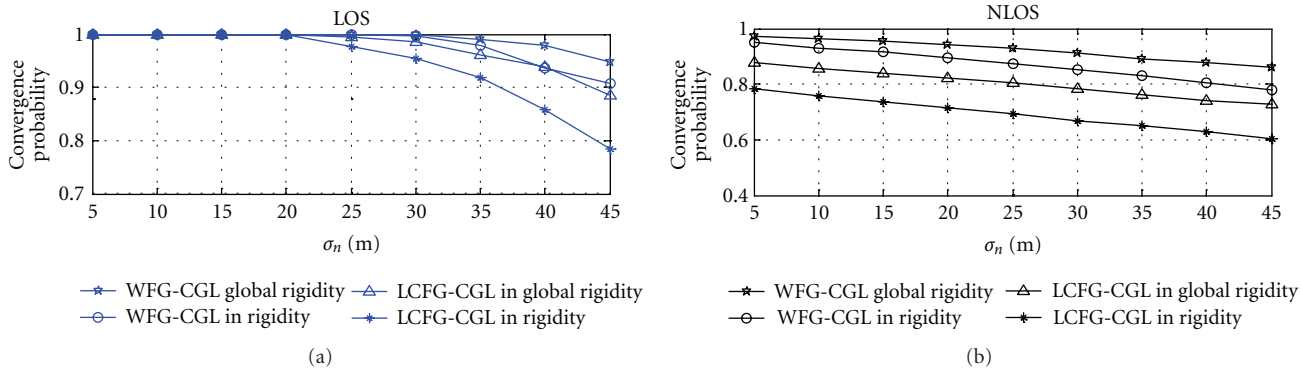


Figure 12: Convergence probability of FG-based CGL.

the similar localization accuracy in global rigidity, which demonstrates the strong robustness of the proposed algorithms for different CGL topologies.

The localization accuracy depends on how to utilize the available information. Generally, the variances can represent the practical environment effectively. The proposed algorithms make full use of the variances in nearly each node of FG, which is considered as a great help for performance upgrade in rigidity topology, even with less information than global rigidity. Furthermore, the power of the proposed algorithms lies in the ability to distributively process the problem and to provide the optimum or near-optimum solutions, which makes it a good approach to solving distributed problem. On the other hand, CGL is a distributed problem, that is, measurement parameters are only locally related. Therefore, the excellent combination between FG and CGL can relax the CGL topology requirement under ensuring the localization accuracy to some extent.

6.3. Complexity Analysis. This section analyzes the computation complexity of the proposed algorithms and Taylor-series-based algorithm, respectively.

(1) Complexity Analysis of Taylor-Series-Based CGL Algorithm. In Taylor-series-based CGL algorithm [18], the most complex step is the solution to pseudoinverse of large scale matrix. In each iteration, the computation complexity is $O(N^3)$, where N is the number of edges in CGL topology. Moreover, the iteration number is not definite, because it is determined by the tolerable localization error and

Table 3: Statistics analysis for different topologies.

		$\mu_{G_{tp}}$	$\sigma^2_{G_{tp}}$
Taylor series	LOS	0.41	0.0021
	NLOS	0.409	0.002
LCFG	LOS	0.12	0.0002
	NLOS	0.11	0
WFG	LOS	0.09	0.0002
	NLOS	0.10	0
CRLB	LOS	0.08	0
	NLOS	0.002	0

the distance measurement error. Especially when the TG includes more cooperative MTs, the complexity of CGL-based Taylorseries localization algorithm increases sharply.

(2) Complexity Analysis of WFG-CGL Algorithm and LCFG-CGL Algorithm. In the two proposed FG-based CGL algorithms, the amount of messages passed between nodes is limited. And the iteration times is pre-defined, commonly $K = 7$. Because after $K = 7$, the performance has no significant improvement with the growth of K. The performance of the proposed algorithms approaches high accuracy after 7 iterations, which shows that the convergence speeds of the two proposed algorithms are reasonably high.

Although the solution of the weights in WFG-CGL algorithm results in the increasing calculation complexity to some extent, they can be calculated distributively, which means that it just solves multiple equations, respectively, instead of the solution to pseudo-inverse of large scale matrix. The computation complexity of weights solution is $O(N)$, which is less than the Taylor-series-based CGL algorithm.

On the other hand, the convergence probabilities of the two proposed algorithms are shown in Figure 12. The proposed algorithms have very high convergence probabilities in both LOS and NLOS scenarios. Even when the standard derivation is big or the topology is not a unique solvability, the convergence probabilities of the two proposed algorithms are still acceptable. Those all lead to the low computational complexity and transmission loads while ensuring the localization accuracy.

From the above analysis and simulation results, it can be concluded that (1) both of the proposed algorithms can relax the CGL topology requirement and provide high localization accuracy. (2) WFG-CGL algorithm possesses higher accuracy than LCFG-CGL algorithm while LCFG-CGL algorithm possesses lower complexity than WFG-CGL algorithm. Although the complexity of WFG-CGL algorithm increases to some extent, it still performs with lower complexity than the traditional Taylor-series-based CGL algorithm.

7. Conclusion

In this paper, the WFG-CGL algorithm and the LCFG-CGL algorithm are proposed in ill conditions. Especially for WFG-CGL algorithm, the weights based on the information reliability are incorporated into the algorithm in order to obtain higher localization accuracy. The corresponding analysis and simulations show that the proposed algorithms can relax the demand for CGL topology and provide high localization accuracy under low complexity in comparison with the existing CGL algorithm. Moreover, the important factors and CRLB for FG-based localization are analyzed.

Acknowledgments

This work is supported by the National Nature Science Foundation of China Project (Grant 61001119) and Fundamental Research Funds for the Central Universities (G470712).

References

[1] F. C. Commission, Revision of the Commissions Rules to Insure Compatibility with Enhanced 911 Emergency Calling Systems, 1996.

[2] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero III, R. L. Moses, and N. S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 54–69, 2005.

[3] S. Gezici, Z. Tian, G. B. Giannakis et al., "Localization via ultra-wideband radios: a look at positioning aspects of future sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 70–84, 2005.

[4] S. Hara, D. Zhao, K. Yanagihara et al., "Propagation characteristics of IEEE 802.15.4 radio signal and their application for location estimation," in *Proceedings of the IEEE 61st Vehicular Technology Conference (VTC '05)*, pp. 97–101, June 2005.

[5] S. Feng and C. L. Law, "Assisted GPS and its impact on navigation in intelligent transportation systems," in *Proceedings of the IEEE International Conference on Intelligent Transportation Systems*, pp. 926–931, June 2002.

[6] P. Farradyne, *Vehicle Infrastructure Integration (VII) Architecture and Functional Requirements*, draft, version 1.0, 2005.

[7] M. Shafl, W. Aamer, and K. Butterworth, "Next generation wireless networks," in *Proceedings of the The IEEE International Symposium on Circuits and Systems (ISCAS '01)*, pp. 721–729, Sydney, Australia, May 2001.

[8] Y. Zhao, "Standardization of mobile phone positioning for 3G systems," *IEEE Communications Magazine*, vol. 40, no. 7, pp. 108–116, 2002.

[9] A. H. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location: challenges faced in developing techniques for accurate wireless location information," *IEEE Signal Processing Magazine*, vol. 22, no. 4, pp. 24–40, 2005.

[10] J. Chaffee and J. Abel, "GDOP and the Cramer-Rao bound," in *Proceedings of the IEEE Position Location and Navigation Symposium*, pp. 663–668, Las Vegas, Nev, USA, April 1994.

[11] E. G. Larsson, "Cramér-rao bound analysis of distributed positioning in sensor networks," *IEEE Signal Processing Letters*, vol. 11, no. 3, pp. 334–337, 2004.

[12] Y. Shen, H. Wymeersch, and M. Z. Win, "Fundamental limits of wideband cooperative localization via fisher information," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC '07)*, pp. 3951–3955, March 2007.

[13] Y. Qi, H. Kobayashi, and H. Suda, "On time-of-arrival positioning in a multipath environment," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 5, pp. 1516–1526, 2006.

[14] D. B. Jourdan, D. Dardari, and M. Z. Win, "Position error bound and localization accuracy outage in dense cluttered environments," in *Proceedings of the IEEE International Conference on Ultra-Wideband (ICUWB '06)*, pp. 519–524, September 2006.

[15] J. N. Ash and R. L. Moses, "On the relative and absolute positioning errors in self-localization systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5668–5679, 2008.

[16] Y. Zhao, "Standardization of mobile phone positioning for 3G systems," *IEEE Communications Magazine*, vol. 40, no. 7, pp. 108–116, 2002.

[17] Y. Zhang, Q. Cui, and X. Tao, "Cooperative group localization for 4G wireless networks," in *Proceedings of the IEEE 70th Vehicular Technology Conference Fall (VTC '09)*, Anchorage, Alaska, September 2009.

[18] Q. Cui and X. Zhang, "CGL-based modified taylor-series localization algorithms for future wireless networks," in *Proceedings of the IEEE International Conference on Communication Technology (ICCT '11)*, Jinan, China, September 2011.

[19] J. Aspnes, T. Eren, D. K. Goldenberg et al., "A theory of network localization," *IEEE Transactions on Mobile Computing*, vol. 5, no. 12, pp. 1663–1678, 2006.

[20] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.

[21] J. C. Chen, Y. C. Wang, C. S. Maa, and J. T. Chen, "Network-side mobile position location using factor graphs," *IEEE Transactions on Wireless Communications*, vol. 5, no. 10, pp. 2696–2704, 2006.

[22] C. Mensing and S. Plass, "TDoA positioning based on factor graphs," in *Proceedings of the IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '06)*, pp. 1–5, September 2006.

[23] B. Omidali and S. A. A. B. Shirazi, "Performance improvement of AOA positioning using a two-step plan based on factor graphs and the Gauss-Newton method," in *Proceedings of the 14th International CSI Computer Conference (CSICC '09)*, pp. 305–309, October 2009.

[24] C. T. Huang, C. H. Wu, Y. N. Lee, and J. T. Chen, "A novel indoor RSS-based position location algorithm using factor graphs," *IEEE Transactions on Wireless Communications*, vol. 8, no. 6, pp. 3050–3058, 2009.

[25] M. Üney and M. Çetin, "Target localization in acoustic sensor networks using factor graphs," in *Proceedings of the IEEE 16th Signal Processing, Communication and Applications Conference (SIU '08)*, pp. 1–4, April 2008.

[26] H. Wymeersch, J. Lien, and M. Z. Win, "Cooperative localization in wireless networks," *Proceedings of the IEEE*, vol. 97, no. 2, pp. 427–450, 2009.

[27] N. Patwari, A. O. Hero, M. Perkins, N. S. Correal, and R. J. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2137–2148, 2003.

[28] K. Yu, I. Sharp, and Y. J. Guo, *Non-Line-of-Sight Identification, Ground-Based Wireless Positioning*, John Wiley & Sons, Chichester, UK, 2009.

[29] P. C. Chen, "A non-line-of-sight error mitigation algorithm in location estimation," in *Proceedings of the IEEE Wireless Communications and Networking Conference*, pp. 316–320, New Orleans, La, USA, September 1999.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 146516, 8 pages
doi:10.1155/2012/146516

Research Article

A Novel Hybrid Self-Organizing Clustering Routing Algorithm

Peng Zhu[1] and Fei Jia[2]

[1] *Department of Information Management, School of Economics and Management, Nanjing University of Science and Technology, 200 Xiao Ling Wei Street, Nanjing 210094, China*
[2] *Division of Education Affairs, Nanjing Forest Police College, Nanjing 210046, China*

Correspondence should be addressed to Peng Zhu, pzhu@ieee.org

Received 9 May 2012; Revised 17 July 2012; Accepted 30 July 2012

Academic Editor: Mugen Peng

This paper discusses the distributed routing algorithm of wireless self-organized network and puts forward a new type of hybrid self-organizing clustering routing protocol by combining energy sense and maximum connectedness. This routing protocol adopts a mechanism which is mainly based on cluster routing with the theory of "on demand-based plane routing discovery mechanism" becoming secondary. The routing discovery and routing maintenance algorithm are also described in this paper. As the emulation experiment result shows, there is no significant difference between the routing protocol and the demand-based routing protocol; however, when there are quite a number of nodes in the network, the performance can be greatly enhanced compared with the demand-based routing protocol. The routing protocol discussed in this paper has better extendability, lower routing overhead, and better data transmission rate; thus it can extend network lifetime and enhance the network performance.

1. Introduction

In most cases, the on-demand routing protocol works effectively; however, as the network scale expands and the node number increases, the number of the routing control messages increases dramatically; as a result, the network congestion is caused and the network performance index decreases [1]. At present, the clustering algorithm is recognized as an effective self-organized method which can improve network extendability, realize smaller route and control overhead, reduce number of nodes that share same channels, and also reduce the collision rate [2]. In addition, the cluster algorithm can help implementation of the functions such as routing selection and resource management. As a commonly used clustering algorithm, the maximum link degree clustering algorithm (MAXD) [3], its node firstly derives number of the neighboring nodes via the interaction control message and then broadcasts its connectedness to the neighboring nodes. Among this node and all other neighboring ones, the node with maximum connectedness to the neighbor one should be selected as the cluster head. If the maximum connectedness between the nodes and their neighborhoods are the same, the node with smallest ID should be selected as the cluster head. The one-hop neighboring node of the cluster head becomes the member node in the cluster. The above-said procedure can be repeated until all nodes join one certain cluster. This cluster algorithm has the advantage of fewer number of clusters, that is to say, there are no many average hops between the source node and the destination one. By using the algorithm, the packet delivery delay can be reduced, but the overlapping between the created clusters is still high, and the network lifetime can be affected because the network loading balance is not taken into consideration. In such a background, we put forward a new self-organized clustering routing protocol based on link degree and energy aware. In this paper, the proposed clustering routing protocol is denoted as NSCR.

The basic principle of the NSCR is described as follows. On one hand, by referring to the theory of "on demand-based plane routing discovery mechanism" at the initial stage, the routing protocol later evolves as a mechanism which is mainly based on cluster routing; thus the theory of "on demand-based plane routing discovery mechanism" becomes secondary. On the other hand, when choosing the cluster head node, the system considers node connectedness and remaining energy. At the same time, each cluster head

node chooses one candidate cluster head node (means the node which will become the cluster head once the node in the cluster head is not the head node) in its one-hop neighbor domain; when energy of the cluster head node is smaller than a certain value, the candidate node works as the new cluster head node instead, so that the network load balance can be achieved and network lifetime can be extended.

The rest of the paper is organized as follows. In Section 2, we present related works. Section 3 describes the proposed algorithm. Experimental results are reported in Section 4 and conclusions are drawn in Section 5.

2. Related Work

Grouping sensor nodes into clusters has been widely pursued by the research community in order to achieve the network scalability objective. The objective of clustering is mainly to generate stable clusters in environments with sensor nodes. In addition to supporting network scalability, clustering has numerous advantages [4].

"Self-organizing" is defined as the process where a structure or pattern appears in a system without intervention by external directing influences. It organizes through direct interaction in a peer-to-peer method [5]. Several self-organizing clustering protocols were studied as follows. Tournus et al. [6] propose a routing to self-organization, in which thin films obtained by deposition of size-selected CoPt clusters on graphite surface. The preformed clusters can easily diffuse on the surface and gather to form "islands" or "bunches" of clusters. By changing the cluster size, very different morphologies can be obtained, going from large-ramified islands to bunches of noncontacting clusters having the size of the initially deposited particles. Ahmed et al. [7] propose that energy efficiency and enhanced backbone capacity are obtained by exploiting the geometric orientation of cooperative nodes in wireless sensor network. The cooperative communication in wireless sensor networks gives people leverage to get the inherent advantages of its random node's locations and the direction of the data flow. Depending on the channel conditions and the transmission distance, the number of cooperative nodes is selected, that participate in an energy-efficient transmission/reception. Hasan and Jue [8] associate survivability and energy efficiency with the clustering of WSNs and show that such a proactive scheme can actually increase the lifetime. They present an easy-to-implement method named DED (distributed, energy-efficient, and dual-homed clustering) which provides robustness for WSNs without relying on the redundancy of dedicated sensors, that is, without depending on node density. DED uses the already gathered information during the clustering process to determine backup routes from sources to observers, thus incurring low message overhead. It does not make any assumptions about network dimension, node capacity, or location awareness and terminates in a constant number of iterations. Ahmed et al. [9] presents an energy-efficient selection of cooperative nodes with respect to their geographical location and the number of nodes participating in cooperative communications in wireless sensor networks. The cooperative communication in wireless sensor networks gives people leverage to get the inherent advantages of its random node's locations and the direction of the data flow. Depending on the channel conditions and the transmission distance, the number of cooperative nodes is selected, that participate in an energy-efficient transmission/reception. Simulation results show that increasing the cooperative receive diversity decreases the energy consumption per bit in cooperative communications. It has also been shown that the network backbone capacity can be increased by controlled displacement of antennas at base station at the expense of energy per bit. Sun and Gu [10] propose and evaluate an energy-efficient clustering scheme based on LEACH (low energy adoptive clustering hierarchy), that is, LEACH-Energy Distance (LEACH-ED). In LEACH-ED, cluster heads are elected 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 another self-organized protocol that is based on LEACH. AbdelSalam and Olariu [11] propose to construct what they call a network skeleton that is constructed immediately after network deployment and provides a topology that makes the network more tractable. The skeleton provides sensors with coarse localization information that enables them to associate their sensory data with the geographic location in which the data was measured. Moreover, it promotes a geographic routing scheme that simplifies data communication across the network through skeleton sensors. Younis et al. [12] 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. Sangjoon [13] introduces a clustering strategy and self-organizing scheme for cluster-based wireless sensor networks, while maintaining the merits of a clustering approach. This scheme is a clustering method to configure cluster by diffusing an interest from a sink node. When a sink node diffuses an interest, every node decides which node is elected as a cluster head or intermediary node by sending and receiving messages.

In summary, the routing protocol in the above-mentioned studies often consider connectivity or energy awareness singly, and their research results always have some defects. In this regard, combined with flat-based routing mechanism and clustering arithmetic, this paper presents an on-demand self-maintenance clustering routing protocol based on connectivity and energy awareness which cited a mechanism with clustering routing key point and with flat-based on-demand routing supplementary point. This protocol keeps the advantages of flat-based on-demand routing, improves the scalability, and enhances the network performance.

3. Algorithm Description

Extensive research has been conducted in the area of clustering routing algorithm in dynamic network, including MANETs and mobile sensor networks; a great deal of algorithms and protocols have been proposed. Most of the

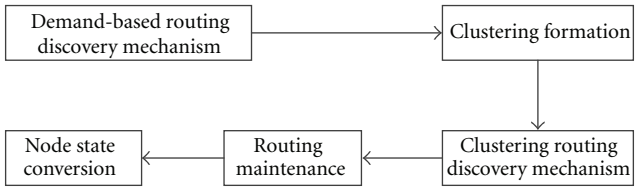


Figure 1: Algorithm Description.

research focused on reducing routing overhead and increasing latency of the network. There are, without doubt, some open issues that are still worth to investigate. In our work, we assume the following content as our clustering routing protocol.

The ad hoc network, which consists of multiple free moving nodes, can be extracted as an undirected graph of $G = (V, E)$. Among the parameters, V means network node set and E means bidirectional link set between nodes. For any symmetrical link $(i, j) \in E$ and $(j, i) \in E$. For any node $i, j \in V$ which has the unique ID, the node can provide remaining energy information at any time. If the distance between node i and node j is smaller than the transmission radius R, then the transmission between the two nodes can be achieved successful; the two nodes are called one-hop neighbor to each other.

The self-organized clustering routing protocol (NSCR) proposed by this paper starts performing route query when there is data which needs to be transmitted; in such a case, the nodes form a cluster. In addition, the system considers node connectedness and remaining energy of cluster head election, and at the same time it sets the candidate cluster head. The proposed algorithm consists of five parts as shown in Figure 1.

3.1. Demand-Based Routing Discovery Mechanism. When the source node (s) wants to communicate with the destination node (d), it initiates the routing discovery procedure. At the beginning, all nodes are in pending state, meaning that it is not in the cluster.

The source node (s) firstly checks its own routing table, to see whether there is routing information to the destination node (d). If no, it creates RREQ [14] and then broadcasts the RREQ to all neighboring nodes. RREQ = (Type, Source_addr, Dest_Addr, Dest_squence, RList, ToL). Among the parameters, Type means packet type, Source_addr means source address of the packet, Dest_Addr means destination address of the packet, and Dest_squence means packet serial number. Please note that (Source_addr, Dest_squence) can uniquely identify the RREQ, RList records the route information, and ToL means RREQ's value of lifetime.

After the intermediate node (m) receives one RREQ, it performs the following procedure.

Step 1. Substract 1 from ToL. If ToL value is 0, the intermediate node (m) shall discard the RREQ because the value means that RREQ's value of life is 0.

Step 2. Check whether the RREQ has been received according to RREQ's (Source_addr, Dest_squence). If yes, the intermediate node (m) discards the RREQ; otherwise, it checks whether the Rlist of RREQ carries the node address; if yes, it discards the RREQ.

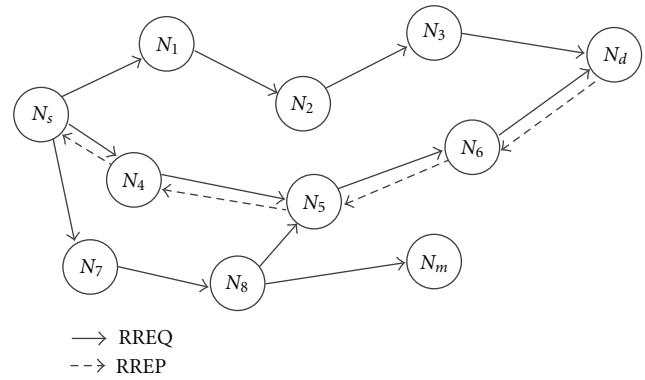


Figure 2: Demand-based routing discovery mechanism.

Step 3. If the intermediate node (m) which carries the route information to the destination node returns the RREP [15] to the source node (s) after it receives the RREQ; otherwise, the intermediate node (m) continues broadcasting the RREQ. At the same time, the node (m) records the neighbor address carried in the RREQ and creates the reverse path.

Step 4. When there are a number of roads between the intermediate node (m) and the destination node (d) can be selected, the algorithm selects the optimal next hop node (k) from the alternative nodes to transmit information according to the probability P^d_{mk}. The trust value between the alternative next hop node (k) and the intermediate node (m) should be greater than a threshold value

$$P^d_{mk} = \frac{\left(D^d_{mk}\right)^{\alpha}}{\sum_{j \in N^d_s} \left(D^d_{mj}\right)^{\alpha}}. \quad (1)$$

Among the parameters, D^d_{mk} means the distance of the path (m, k) on which the current node (m) reaches the destination node (d) through node (k). The distance of same paths to reach different destination nodes is mutative. N^d_s means the nodes set in which the neighbor nodes of node (m) can reach the destination node (d) and α means the parameter of wayfinding control information, $\alpha \gg 1$.

Step 5. After the destination node (d) receives the RREQ from the source node (s), it triggers the routing response procedure, which sends the RREP to the source node (s) in the created reverse path.

Figure 2 shows the demand-based routing discovery mechanism.

(i) Node Weight Calculation Method. In the reverse path where the destination node (d) sends the RREP to the source node (s), the passed node v and its neighboring node weight in the RREP can be obtained via calculation.

Definition 1. Weight (W_v) of node (v) is associated with its connectedness and the current remaining energy. The connectedness can be obtained from node (v), and the weight can be derived by combining the current remaining energy. Weight W_v of node (v) is defined as follows:

$$W_v = \alpha \frac{C_v}{C_{\max}} + (1 - \alpha) \frac{E_{\text{re}}}{E_{\max}}, \quad 0 \le \alpha \le 1. \tag{2}$$

Among the parameters, C_v means connectedness, $C_{\max}$ means maximum connectedness in one-hop neighbor domain, E_{re} means current remaining energy, and $E_{\max}$ means maximum initial energy. In case that the node energy is sufficient, when considering reducing the transmission delay, a big value of α should be taken; when considering achieving the energy balance, a small value of α should be taken, so that the network lifetime can be extended.

(ii) Generating Clustering Head Node. Node v broadcasts its weight to its neighboring nodes, and at the same time it updates neighboring node information list based on broadcast information of the neighboring nodes. After this procedure is finished, node v obtains weight and ID of its one-hop neighboring nodes. Then, node (v) compares the weights of this node with all other one-hop neighboring nodes and selects the node with maximum weights as the cluster head. If the maximum weight of the nodes is same, the node with smallest ID shall be selected as the cluster head. In the meanwhile, the node (as new cluster head) sends cluster head broadcast message to all its one-hop neighboring nodes; upon receiving the broadcast message from the cluster head, the neighboring nodes become member of the cluster and add the cluster head address into the head list.

(iii) Generating Distributed Gateway and Gateway Nodes. The member node i broadcasts the distributed gateway request message to its neighboring nodes. If the neighboring node j which receives the message finds that its cluster head node is different from that of the sending node i, then the neighboring node j shall returns a distributed gateway response message to node i; in such a case, node i becomes a distributed gateway in the cluster; later it adds its neighboring node j into its neighboring gateway list, and at the same time, the node i sends the distributed gateway message to its neighboring head node so as to allow the cluster head node to add the node i into its own distributed gateway list.

After this phase ends, the cluster is completely formed. The node is thus comprised of cluster head, gateway, distributed gateway, ordinary member, and pending node. All cluster head nodes are not directly neighbored to each other, which means that they are out of one-hop transmission range. The distance between cluster head node and cluster member node is always one hop.

3.2. Clustering Routing Discovery Mechanism. If the node becomes a cluster member, then the system shall set a cluster routing tag in the RREQ it sends and then query the routing according to the cluster routing discovery mechanism. In addition, the node which forwards RREQ determines whether to query the routing by adopting cluster routing discovery mechanism or demand-based plane routing discovery mechanism according to the fact whether the cluster routing tag is set in the received RREQ. If the node fails to find the route in the specified time, the plane routing discovery mechanism based on demand shall be used for routing discovery [16]. Figure 3 shows the clustering routing discovery mechanism.

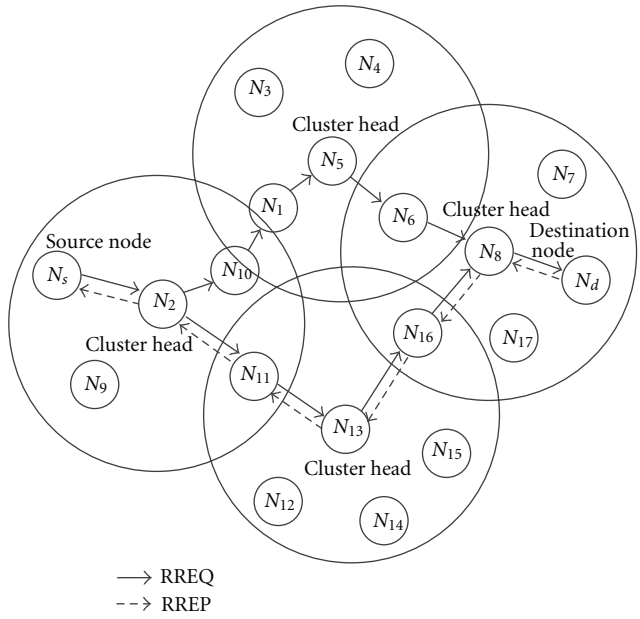


Figure 3: Clustering routing discovery mechanism.

State of any node v can be converted under certain conditions.

(i) Clustering Head Node. When the cluster head node finds its cluster member is 0, it converts its own state to pending; when remaining energy of the cluster head node is below one certain value, the candidate head node becomes the head and the previous head node becomes an ordinary member.

(ii) Clustering Ordinary Member Node. When the cluster ordinary member node finds the cluster head in its neighboring node is not the same as its own cluster head, then its state shall be converted to distributed gateway. When the cluster ordinary member node receives the broadcast message from the heads in different clusters, its state can be converted to gateway. If the cluster ordinary member node is selected as the candidate node, when remaining energy of the cluster head node is below one certain value, its state becomes cluster head.

(iii) Gateway Node. If the gateway node finds that its cluster head is not reachable, then it deletes this node from its cluster head list. If the cluster head list contains only one head node, and the head in one-hop neighbor domain is not the same as its own cluster head, then its state shall be converted to distributed gateway; otherwise, its state shall be converted to cluster ordinary member.

(iv) Distributed Gateway Node. If the distributed gateway node finds that the node in its neighbor domain does not include the node from other clusters, then its state shall be converted to cluster ordinary member. When it finds the new cluster head node, its state shall be converted to gateway.

3.3. Routing Maintenance. After the routing discovery is finished, the network link loses effectiveness and reestablishes itself because it is affected by mode movement, energy reduction, and so forth. At the same time, the network topology structure also changes; so some nodes shall join new clustering from the old one, thus causing clustering structure change. In such a situation, the routing protocol requires the node to obtain state of the neighboring node via the broadcast message sent in a periodic manner and informs the nodes affected by disconnected links to perform the maintenance mechanism such as updating routing table by sending routing error group; besides, the related cluster maintenance mechanism should be adopted to maintain stability of the cluster structure.

(i) Local Routing Repair. If the node detects that its link to one neighboring node disconnects, then all routes using this link shall also fail. In this case, the direct route repair method can be adopted for route maintenance at link disconnection node. When the intermediate node m detects the link disconnection, it firstly puts the data from source node s into cache and then sends the route request to start new routing discovery by taking this node as the source node. If the destination node d receives the request, it replies with the routing response; upon the route is repaired successfully, the link disconnection node sends the data packets to the destination one in the new routing.

(ii) Self-Maintenance of Clustering Head Node. The clustering ordinary member node v in the cluster head node's one-hop neighboring domain can be used to obtain the weight $W_{\text{candidate}}(v)$ that is related to candidate cluster head node. $W_{\text{candidate}}(v)$ is associated with three parameters: similarity between node v and its clustering head node, connectedness of node v, and remaining energy. In general, the higher the similarity of node v is, the higher the connectedness is and the greater the remaining energy is; hence, the $W_{\text{candidate}}(v)$ becomes bigger, and the chance of being selected as the candidate cluster head node is greater.

Definition 2. Weight $W_{\text{candidate}}(v)$ of node v's candidate cluster head node is defined as follows:

$$W_{\text{candidate}}(v) = \alpha S + \beta \frac{C_v}{C_{\max}} + \gamma \frac{E_{\text{re}}(v)}{E_{\max}}. \tag{3}$$

Among the parameters, S means similarity between node v and its clustering head node h, C_v means connectedness of node v, $C_{\max}$ means node v and maximum connectedness of one-hop neighboring domain, E_{re} means current remaining energy of node v, and $E_{\max}$ means maximum initial energy of the node, $0 \le \alpha, \beta, \gamma \le 1$ and $\alpha + \beta + \gamma = 1$.

Definition 3. Similarity between node v and its cluster head node h is defined as S:

$$S = \frac{|N_v| \cap |N_h|}{|N_v| \cup |N_h|}. \tag{4}$$

Among the parameters, N_v means one-hop neighbor domain of node v, $|N_v|$ means number of nodes in one-hop neighbor domain of node v, N_h means one-hop neighbor domain of the cluster head, and $|N_h|$ means number of nodes in one-hop neighbor domain of cluster head.

The clustering head node sends the candidate head request message to the node whose $W_{\text{candidate}}(v)$ is maximum in the clustering in unicast mode after comparison; the node which receives the message returns the candidate head request message to the cluster head node, which indicates that the node works as the candidate cluster head. When remaining energy of the cluster head node is below one certain value, the candidate head node becomes the new head node, replacing the previous one. The previous head node becomes an ordinary member node in the cluster.

(iii) Joining and Quitting the Clustering. When status of node v is standby, if it receives the clustering head node broadcast message, it shall send cluster join request message to the clustering head node. After node v receives the response message from the clustering head node, it saves the head node address, which indicates node v has joined the clustering. To reduce the overhead, the on demand-based method [17] is adopted to allow new nodes to join the clustering. When the clustering member nodes are not in head node's one-hop neighboring domain, they modify their statuses to pending state, which indicates that they have quit the clustering.

4. Experiment and Result Analysis

4.1. Experiment Design. By routing protocol simulation based on the Linux environment, the simulation experiment adopts NS-2 [18] as the emulation tool. NS-2 is a discrete event-driven network simulation tool; it is open source and free software and can be extended according to the needs of users. NS-2 can perform a variety of network protocols, offer a variety of data sources, achieve a variety of router queue management algorithms, bring about multicast and MAC algorithm, and provide communication model, random topology, and node mobility model generation tool. The scenario documents used in the simulation experiments are generated by the stochastic modeling tools. When the emulation begins, each node waits for a stopping duration at the initial location and then moves to a randomly selected direction with a random speed between 0 and the maximum moving speed. If the node reaches the randomly selected destination, it waits for the same period of stopping duration and then repeats the procedure described earlier until the emulation ends. All moving procedures should be recorded in the scenario documents. In the experiment, the average value is used as reference.

4.2. Parameter Settings. In the emulation scenario, we set a flat rectangular virtual environment with size as 1500 m × 1000 m and set the emulation duration as 2000 s. During the emulation, half of the nodes move randomly in the waypoint mode and the other half stay still. The maximum moving speed of all the nodes is 40 m/s, and the node stopping duration is 50 s. Besides, CBR data source is used, the data transmission rate is 10 packets per second [19], and MAC protocol is using 802.11DCF.

In this paper, two simulation scenarios are set.

(i) Simulation Scenario 1. The number of nodes is set to 100, 150, 200, 250, 300, 350, 400, 450, and 500, respectively.

(ii) Simulation Scenario 2. The number of nodes is set to 5, 10, 20, 40, 80, 120, 160, 250, and 300, respectively.

In order to analyze the performance of routing protocols, the following performance indicators of the routing protocols are assessed. First, lifetime: it refers to the time interval from the beginning of the simulation to the time when the first node in the network runs out of energy. Lifetime can be used to measure the viability of the ad hoc network. Second, routing overhead: it is used to establish and maintain the number of control packets generated by routing. Routing overhead can be used to measure routing protocol's scalability and ability to adapt to network congestion. Third, packet delivery fraction: it means the ratio of the total number of packets the destination node receives with the total number of packets the source node sends. Packet delivery fraction can be used to measure the efficiency of the routing protocol. Fourth, number of clusters: it directly reflects the structure and characteristics of the cluster network.

4.3. Result Analysis. In the emulation experiment, according to the parameters and performance indicators proposed in Section 4.2, we select representative plane on-demand distance vector (AODV) routing protocol [20] and maximum connectivity clustering algorithm (MAXD) as references, do simulation experiments to the proposed clustering routing protocol NSCR via AODV and MAXD, and select some representative simulation results be analyzed and discussed.

(i) Life Time of the Ad hoc Network. In simulation scenario 1, lifetime of the ad hoc network in NSCR clustering algorithm is greatly extended compared with the MAXD. The reason is that the NSCR clustering algorithm considers the node energy in the cluster head election process and thus adopts the candidate cluster head mechanism; by doing so, the network loading is balanced and lifetime of the ad hoc network is extended. Figure 4 shows the comparison of the lifetime of the network between NSCR and MAXD.

(ii) Routing Overhead. Figure 5 shows the routing overhead for both NSCR and AODV in simulation scenario 2. As the figure shows, when the number of the nodes in the network is only a few, routing overhead of NSCR and AODV is almost the same; however, as the number of network nodes increases, especially when the number of nodes exceeds 100,

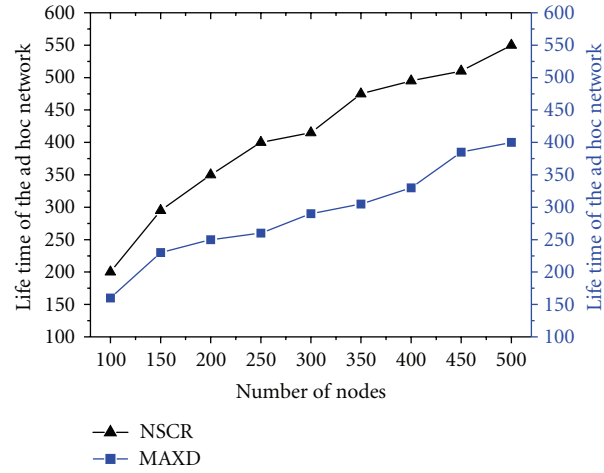


Figure 4: Life time of the Ad hoc network.

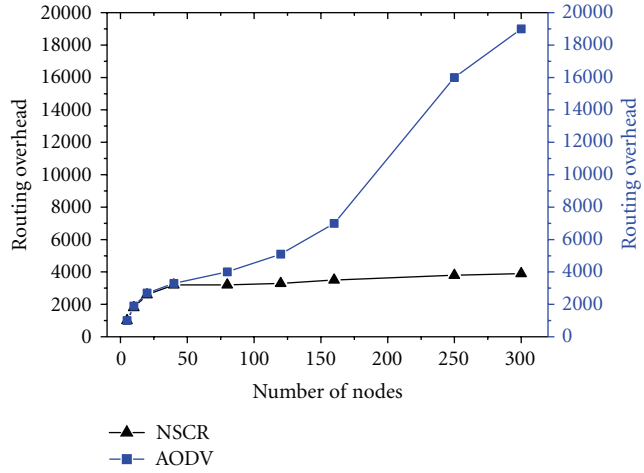


Figure 5: Routing overhead.

AODV's routing overhead increases at a relatively big pace; when the number reaches 200, AODV's routing overhead increases significantly but at the same time NSCR's routing overhead goes up in a slow pace. This phenomenon is mainly due to the NSCR routing protocol considers cluster routing mechanism, which to a certain extent improves the network scalability and reduces the network congestion. This experiment result shows that NSCR routing protocol can control network routing overhead effectively.

(iii) Packet Delivery Fraction. In simulation scenario 2, packet delivery fraction of both NSCR and AODV decreases when the number of the nodes in the ad hoc network increases, but the packet delivery fraction of NSCR is obviously higher than that of AODV when there are quite a number of nodes in the ad hoc network. The cause of this phenomenon is for the following reasons: on one hand, the NSCR reduces the routing control information and network congestion by using the cluster routing discovery mechanism; on the other hand, NSCR uses the partial

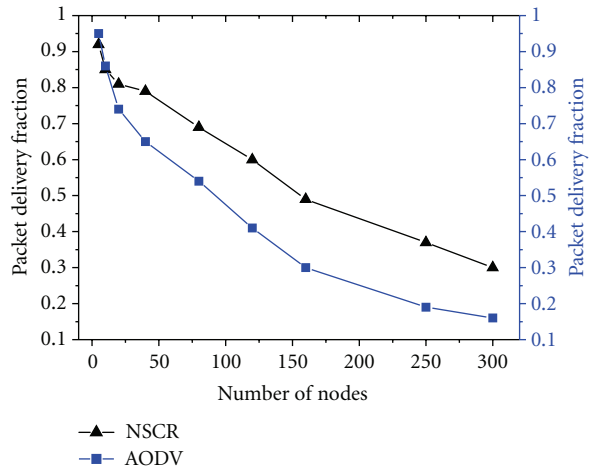


Figure 6: Packet delivery fraction.

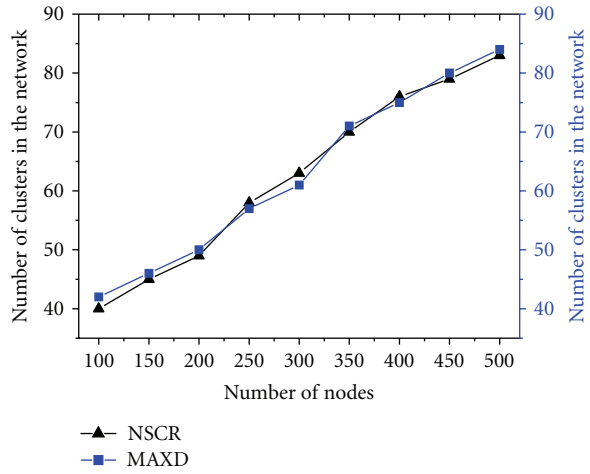


Figure 7: Number of clusters.

routing repair to implement routing maintenance, which can reduce packet loss caused by link disconnection. Figure 6 shows the comparison of packet delivery fraction of NSCR and AODV.

(iv) Number of Clusters. In simulation scenario 1, the number of clusters produced by NSCR and MAXD is similar, which indicates that the NSCR clustering algorithm maintains the advantages of clustering algorithm of MAXD. Figure 7 shows the comparison of the number of clusters of NSCR and MAXD.

5. Conclusion

This paper puts forward a demand-based self-maintenance clustering routing protocol based on energy sense and maximum connectedness cluster algorithm. By analyzing basic principle of the clustering routing protocol, describing protocol work mechanism, and conducting the emulation experiments, the paper concludes that there is no much difference between performance of the proposed clustering routing protocol and demand-based plane routing protocol when the network is in a small scale. However, when there are quite a number of network nodes, their performance can be significantly improved compared with the demand-based plane routing discovery protocol. The HSCR routing protocol has better extendability, lower routing overhead, and better data transmission rate; at the same time, compared with the maximum connectedness cluster algorithm, the HSCR has the advantage of maximum connectedness cluster algorithm; thus it can extend network lifetime and meet routing requirement of the ad hoc network.

Acknowledgment

This research was supported by Ministry of Education of the People's Republic of China, Humanities and Social Sciences Project (no. 12YJC870036).

References

[1] S. Jingfang, W. Muqing, Z. Yan, and Z. Qinjuan, "Robust on-demand routing mechanism for wireless multi-hop networks," *IET Communications*, vol. 5, no. 5, pp. 620–628, 2011.

[2] M. Peng, Y. Liu, D. Y. Wei, W. B. Wang, and H. H. Chen, "Hierarchical cooperative relay based heterogeneous networks," *IEEE Wireless Communications*, vol. 18, no. 3, pp. 48–56, 2011.

[3] A. Bari, S. Wazed, A. Jaekel, and S. Bandyopadhyay, "A genetic algorithm based approach for energy efficient routing in two-tiered sensor networks," *Ad Hoc Networks*, vol. 7, no. 4, pp. 665–676, 2009.

[4] K. Lee and H. Lee, "A self-organized and smart-adaptive clustering and routing approach for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2012, Article ID 156268, 13 pages, 2012.

[5] F. Dressler, *Self-Organization in Sensor and Actor Networks*, John Wiley & Sons, New York, NY, USA, 2007.

[6] F. Tournus, L. Bardotti, and V. Dupuis, "Size-dependent morphology of CoPt cluster films on graphite: a route to self-organization," *Journal of Applied Physics*, vol. 109, no. 11, Article ID 114309, 4 pages, 2011.

[7] I. Ahmed, M. G. Peng, and W. B. Wang, "Exploiting geometric advantages of cooperative communications for energy efficient wireless sensor networks," *International Journal of Communications, Networks and System Sciences*, vol. 1, no. 1, pp. 55–61, 2008.

[8] M. M. Hasan and J. P. Jue, "Survivable self-organization for prolonged lifetime in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2011, Article ID 257156, 11 pages, 2011.

[9] I. Ahmed, M. G. Peng, and W. B. Wang, "Energy efficient cooperative nodes selection in wireless sensor networks," in *Proceedings of the International Conference on Parallel Processing Workshops (ICPPW'07)*, pp. 273–277, Xian, China, September 2007.

[10] 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.

[11] H. S. AbdelSalam and S. Olariu, "A lightweight skeleton construction algorithm for self-organizing sensor networks," in *Proceedings of the IEEE International Conference on Communications (ICC'09)*, pp. 426–430, Dresden, Germany, June 2009.

[12] 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.

[13] J. Sangjoon, "Self-organizing clusters for routing algorithm by diffusing an interest in wireless sensor networks," in *Proceedings of the 5th ACIS International Conference on Software Engineering Research, Management, and Applications (SERA'07)*, pp. 702–707, Busan, South Korea, August 2007.

[14] C. L. Chang and T. L. Lin, "Delay-based RREQ routing scheme for wireless mesh network," *Journal of Internet Technology*, vol. 10, no. 2, pp. 103–109, 2009.

[15] R. D. Bai, M. Singhal, and Y. Luo, "Enhancing performance by salvaging route reply messages in on-demand routing protocols for MANETs," *Ad Hoc & Sensor Wireless Networks*, vol. 5, no. 3-4, pp. 161–188, 2008.

[16] X. X. Wu, G. Ding, and W. W. Zhu, "Load-based route discovery through searching range adaptation for MANET throughput improvement," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 2055–2066, 2009.

[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, article 18, 2008.

[18] J. L. Font, P. Iñigo, M. Domínguez, J. L. Sevillano, and C. Amaya, "Analysis of source code metrics from ns-2 and ns-3 network simulators," *Simulation Modelling Practice and Theory*, vol. 19, no. 5, pp. 1330–1346, 2011.

[19] C. E. A. Campbell, K. K. J. Loo, O. Gemikonakli, S. Khan, and D. Singh, "Multi-channel distributed coordinated function over single radio in wireless sensor networks," *Sensors*, vol. 11, no. 1, pp. 964–991, 2011.

[20] G. Li and H. M. Sun, "Discovering AODV-based multipath routes in wireless ad hoc networks," *Security and Communication Networks*, vol. 5, no. 4, pp. 374–383, 2012.

Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 878595, 10 pages
doi:10.1155/2012/878595

Research Article

Self-Optimization of Coverage and Capacity in LTE Networks Based on Central Control and Decentralized Fuzzy Q-Learning

Jingyu Li, Jie Zeng, Xin Su, Wei Luo, and Jing Wang

Tsinghua National Laboratory for Information Science and Technology, Tsinghua University, Beijing 100084, China

Correspondence should be addressed to Jie Zeng, zengjie@tsinghua.edu.cn

Received 4 May 2012; Accepted 18 July 2012

Academic Editor: Yiqing Zhou

To reduce capital expenditures (CAPEX) and operational expenditures (OPEX) in network operations, self-organizing network (SON) has been introduced as a key part of long-term-evolution (LTE) system. Self-optimization of coverage and capacity is one of the most important tasks in the context of SON. This paper proposes a central control mechanism that utilizes the fuzzy Q-learning algorithm in a decentralized fashion for this task. In our proposed approach, each eNB is a learning agent that tries to optimize its antenna downtilt automatically using information from its own and its neighboring cells, and the initialization and the termination of the optimization processes of all agents are in the control of the central entity. The simulation results verify that our proposed approach can achieve remarkable performance enhancement as well as fast convergence, indicating that it is able to meet different levels of demands defined by 3GPP for coverage and capacity optimization.

1. Introduction

In order to achieve higher network performance and flexibility, while reducing capital expenditures and operational expenditures, 3GPP (Third Generation Partnership Project) introduces the concept of SON in [1]. In the context of SON, LTE systems have self-organizing capabilities of automating the configuration and optimization of the wireless network by introducing functionalities of self-configuring, self-optimizing, and self-healing [2, 3].

The research and development of online algorithms are vital parts in self-optimization process due to the challenges existing in realistic networks. Usually there is no definite mapping function from the input vector and adjustment parameters to the optimization objective, and thus it is difficult to solve the optimization problem directly. Besides, the input space and the parameter space can be very large, thereby full search algorithms are computationally prohibited. Furthermore, the information about the input may be incomplete and partly incorrect, which may cause negative impact on decision making.

The simulated annealing algorithm may be one of the candidates for solving the above optimization problem. It is a probabilistic heuristic searching solution to combinatorial optimization problems by simulating the physical process of annealing where a substance is gradually cooled to reach a minimum-energy state [4]. It has been widely applied in the network planning and optimization. For example, Siomina et al. [5] adopts the simulated annealing as an automated optimization engine with the aim of minimizing CPICH power by adjusting antenna parameters. However, it achieves optimization of service coverage in the offline mode with hypothetical network settings, and it does not run in the SON entity, a realistic node with self-optimizing functionalities in the network. The idea of simulated annealing algorithms is applied to the coverage and capacity optimization task in the context of SON by Cai et al. [6]. In his study, the optimization is achieved by controlling downlink transmit power with the Annealed Gibbs Sampling method in a decentralized SON manner. His approach relies on the accurate information about the interferences the evolved Node B (eNB) will bring to its neighboring cells at its all transmit power levels. However, as for adjusting antenna parameters, such as downtilt angles, this application of simulated annealing algorithms is limited, since such information cannot be predicted due to the ill-defined influence

of changing downtilt angles on users' received signal-to-interference-and-noise ratio (SINR). Admittedly, simulated annealing algorithms can overcome these difficulties if they are utilized in a centralized way, in which the central entity runs the algorithm to search for the global optimal solution. Nevertheless, the change interval (i.e., the time needed between two successive iterations) has to be relatively large due to long operating time for collecting the global information.

In this paper, we apply Q-Learning technique in a decentralized manner for the joint optimization of coverage and capacity task. Q-Learning is a practical form of Reinforcement Learning (RL) which is an important subfield of machine learning. RL is a type of learning involving an agent learning behavior that achieves a goal by directly interacting with its uncertain environment and by properly utilizing past experience derived from previous actions [7]. And we combine the fuzzy rule with the Q-Learning to deal with the realistic problem whose input and output variables are continuous. In addition, we introduce the central control mechanism which is responsible for the initialization and the termination of the optimization process of every learning agent deployed in every eNB.

The rest of the paper is outlined as follows. In Section 2, we describe the SON concept including its basic framework, use cases, and architectures. In Section 3, we discuss about the implementation of our approach with the hybrid architecture for the coverage and capacity optimization task. In Section 4, we present the details of the learning technique applied into our approach with the decentralized multiagent manner. In Sections 5 and 6, we show the simulation setting and the corresponding results. In Section 7, we conclude the paper.

2. SON Concept

2.1. Basic Framework and Use Cases. Aiming at reducing CAPEX and OPEX, self-organizing functions have highly automatic features requiring minimum manual intervention. It consists of self-configuring, self-optimizing, and self-healing functionalities. Figure 1 depicts a basic framework for these functionalities [8].

Self-configuration provides the capabilities for newly deployed eNBs to finish the configuration with automatic installation procedures for obtaining the basic configuration information to operate the system. This process works in preoperational state, which is known as the state from when the eNB is powered up and has backbone connectivity until the RF transmitter is switched on. The use cases for self-configuration include automatic configuration of physical cell identity, neighbor-list configuration, and coverage/capacity-related parameters [2].

Self-optimization offers the benefits of the dynamic optimization in the operational state, which is known as the state where the RF interface is switched on, as shown in Figure 1. This functionality reduces the workload for site visit and analysis of network performance manually, resulting in reduction of OPEX. The main use cases for self-optimization are coverage and capacity optimization, energy savings, mobility robustness optimization, mobility load balancing optimization, RACH optimization and interference reduction [2].

The introduction of self-healing process is to solve or mitigate the faults that could be solved automatically by activating proper recovery actions in the operational state. It includes not only the automatic detection as well as localization of the failure, but also the recovery and the compensation actions. The use cases of self-healing are cell outage detection, cell-outage compensation, cell outage recovery, and return-from-cell-outage compensation [3].

2.2. SON Architecture. Regarding the SON entity allocation, there are three basic ways to implement the SON architecture, namely centralized SON, decentralized SON and hybrid SON, as shown in Figure 2 [9].

In the centralized architecture, the SON entity (i.e., the optimization algorithms) resides on the network management system (NMS) or a central SON server that manages all the eNBs. With the global information from all the eNBs, centralized architecture can readily approach the global optimum. However, the updates of parameters are rather slow, and the optimization process would cost much time because of the large operating interval to collect all the information from all the eNBs who also need to receive the measurements from their UEs (user equipments). Therefore, the applicability of many effective algorithms with slow response time is restricted, that is why the purely centralized architecture does not gain much support from the vendors.

In the decentralized architecture, each eNB has a SON entity, which means each eNB runs the optimization process and gives the decision autonomously based on its UEs' feedback information as well as other information from some other eNBs via the X2 interface. This decentralized architecture allows for ease of deployment and optimization on faster time scale. However, with such decentralized manner, there may be challenges on the stability of the network as well as the overall optimization performance.

The hybrid architecture, as its name indicates, is a combination of the centralized and the decentralized SON, which means part of a SON entity is located in the eNB, while the other is deployed in the NMS. Thus, it provides a tradeoff between the two architectures. For example, the selection of eNBs participating in the optimization, the initial parameter setting, and the objective setting is done in the central NMS, while the core optimization algorithm runs within the involved eNB.

Each architecture has its own advantages and disadvantages on implementation. Which is the most suitable approach depends on the requirement of the use case, the optimization objective, the infrastructure deployment, and other specific conditions.

3. Joint Optimization of Coverage and Capacity under Hybrid Architecture

Optimizing coverage and capacity of the network is one of the typical operational tasks of SON. The automating of

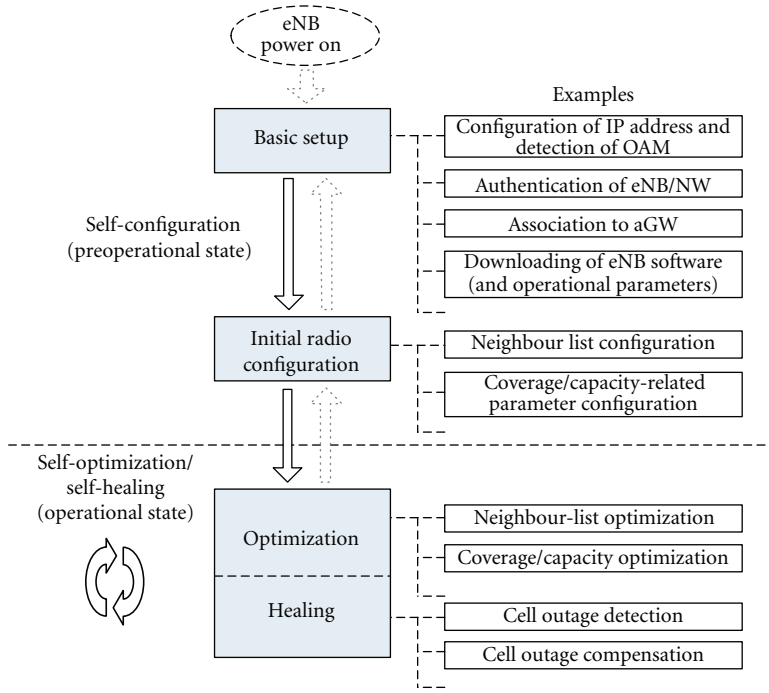


Figure 1: Basic framework for SON functionalities.

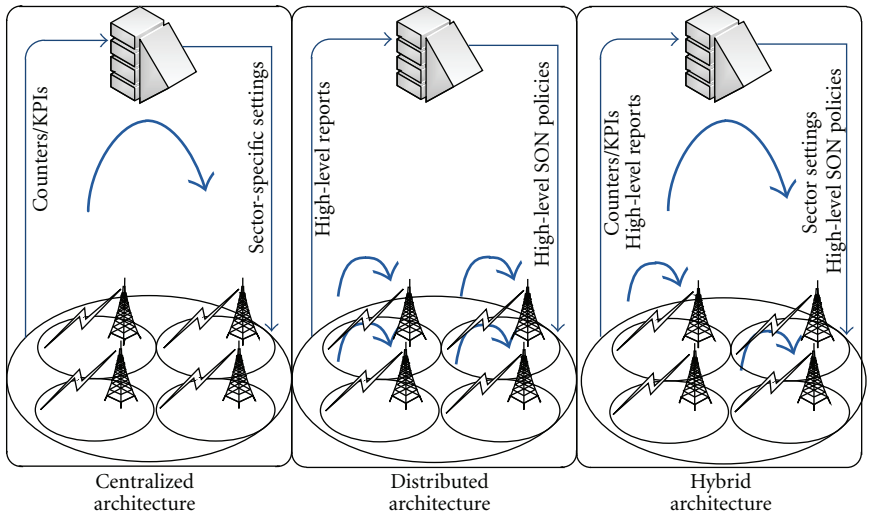


Figure 2: SON architectures.

this task can minimize human intervention in the network by discovering the coverage and capacity problems through the eNB and UE measurements and by optimizing RF parameters through intelligent algorithms automatically. Its objectives include coverage optimization, capacity optimization, and joint optimization of coverage and capacity. The challenges of joint optimization of coverage and capacity include the tradeoff between coverage and capacity, the SON architecture implementation, the selection of indicators to observe and parameters to optimize, and so forth.

3.1. The Hybrid Architecture. For this task, the coverage and/or capacity problem is usually regional, and thus, only one or two cells are usually involved in the optimization; even if more than two cells are involved, antenna-parameter adjustment causes only local impact, which means only neighbors' information is required to ensure stability and convergence. Hence, it is possible to run the algorithm in a decentralized manner to achieve the optimization. In addition, requiring only the information from its own UEs and from its neighboring cells, such manner can provide fast timescale of operation, as discussed in Section 2.2.

On the other hand, the capacity and coverage problem detection, the selection of eNBs, when to start/stop their optimization, and initial parameters of the algorithm need

to be chosen in a central entity that can obtain the global information.

Based on the above considerations, we use hybrid architecture in our approach, which means we implement the control functionality including optimization activation/deactivation and initial parameter setting in a central entity, and deploy the optimization algorithm in every eNB in a decentralized manner.

3.2. Indicators. For self-optimization tasks, many indicators for observation, such as UE signaling reports, pilot-strength measurements and traffic counters, are recommended in [2].

The spectral efficiency (SE), referring to the information rate that can be transmitted per unit bandwidth, is a frequently used metric for evaluating system performance. We can use the CDF of 50% tile of the spectral efficiency distribution to denote the system capacity [10]. As for the system coverage, considering that the users on the cell edge experience significantly poorer performance and greater probability of outage than the others in the cell, their performance metric can be reflected by the CDF of 5%-tile of the spectral efficiency [11]. Given that the spectral efficiency can be easily predicted by the RSRQ information that each eNB collects from its active users, RSRQ is a good choice as the indicator for the joint optimization of coverage and capacity.

To balance coverage and capacity, we define a joint performance metric (JPM) as the weighted sum of the spectral efficiency CDF of 50% tile and 5% tile by a fixed factor λ $(0 \le \lambda \le 1)$

$$\mathrm{JPM} = (1-\lambda)\mathrm{SE}_{50\%} + \lambda \mathrm{SE}_{5\%}. \tag{1}$$

In addition, to take into account the impact of interference from neighboring cells, the optimization processes should be performed jointly. Hence, we use the key performance indicator (KPI) defined in formula (2) as the optimization objective of each eNB (cell).

$$\mathrm{KPI}_i = w_i \cdot \mathrm{JPM}_i + \frac{(1-w_i)}{|N(i)|} \sum_{j \in N(i)} \mathrm{JPM}_j, \tag{2}$$

where JPM_j is the joint performance metric for cell j, w_i is the weight for cell i, and $N(i)$ is the set of its neighbors. Each cell calculates its KPI with its own and its neighboring cells' JPM by formula (2).

3.3. Outputs. The outputs, which are the parameters to adjust during the optimization process, could be the transmit power, antenna azimuth, antenna downtilt, and so forth.

In the process of self-optimization, each eNB may tend to increase transmit power in order to improve its capacity and coverage, but it would cause serious interference to its neighboring cells, thereby making little improvement for the overall system performance. In addition, antenna azimuth is one of the antenna configuration parameters which have great impacts on the cell overlap, and adjusting it can reduce cell overlap significantly, thereby improving the performance of the edge users. However, adjusting antenna azimuth has to be done mechanically, which would require costly site visits and significant time consumption. Comparatively, antenna downtilt is a better candidate for self-optimization of coverage and capacity. On one hand, signal power levels from the home cell can be improved with interference to the neighboring cells being effectively reduced by increasing the downtilt angle; on the other hand, coverage problems on the cell edge can be alleviated by decreasing the downtilt angle. Thus, we can make a tradeoff between the coverage and capacity by adjusting the downtilt angle. And fortunately, the advances in electrical downtilting enable the automated optimization task without any costly site visits by utilizing remote electrical downtilt (RET) controllers [12]. Therefore, we choose antenna downtilt as the parameter to be optimized in this paper.

3.4. Process Flow. The flow chart of the decentralized optimization approach is shown in Figure 3. After mapping RSRQ to SE, the two statistics, that is, $\mathrm{SE}_{50\%}$ and $\mathrm{SE}_{5\%}$, are both obtained. With its own and its neighbors' SE statistics, JPM is calculated by formula (1), and KPI is calculated by formula (2). The optimization algorithm tries to search for a better downtilt configuration with a certain learning technique based on the current KPI as well as some previous KPIs.

4. Learning Technique

Q-learning [7] involves an agent with the learning behavior of achieving a goal by interacting with its uncertain environment (i.e., performs an action and receives rewards from the outside) and by learning from the past experience. In the Q-learning algorithm, there is a set of states and a set of actions. Each time the agent selects one action from the action set according to a quality function which is the quality metric of a state-action pair. After performing the action, the agent moves into a next state, and obtains a reward from the environment, with which it updates its quality function. The goal of the agent is to maximize the total reward.

Fuzzy logic is usually introduced in RL algorithms as an approximate approach when realistic problems have input and output variables with a large or infinite number of possible values (e.g., continuous downtilt) [13].

In this paper, we propose the Fuzzy Q-Learning (FQL) algorithm which combines the fuzzy logic with QL algorithm in the decentralized manner where multiple agents aim at achieving their own optimal goal by interacting with the environment. And the factor considering the performance of the neighboring cells is introduced in the goal.

4.1. FQL Algorithm. Let X be the state space, and A be the action space. $r(x, a)$ denotes the reward received at the state $x \in X$ with the action $a \in A$ performed. The objective of an agent is to find an optimal policy $\pi^*(x)$ for the state x to maximize the utility function R which is defined as a long-term sum of discounted rewards.

$$R = E_\pi \left[\sum_{t=0}^{\infty} \gamma^t r(x_t, a_t) \right], \tag{3}$$

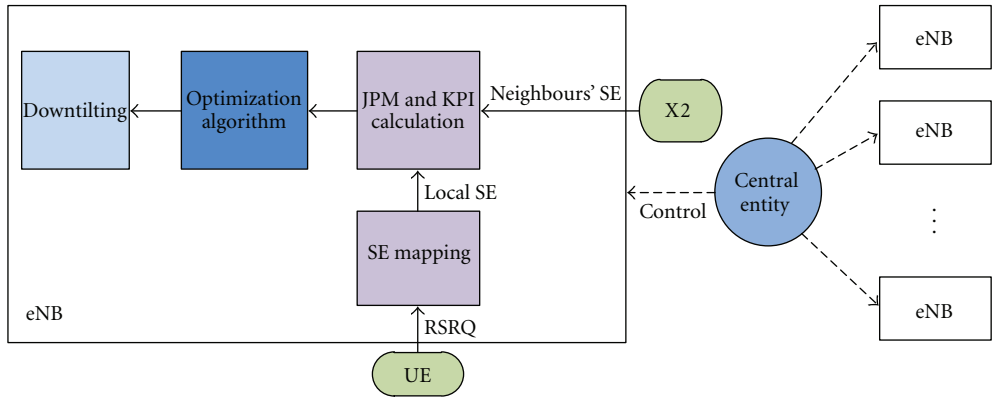


Figure 3: Process flow.

where x_t and a_t denote the state and the action of the agent at step t ($t = 0, 1, 2, \ldots$), respectively, and γ is the discount factor.

The solution of this maximization problem uses the quality function $Q^\pi(x, a)$ ($x \in X$, $a \in A$) that is defined as the expected sum of discounted rewards from the initial state x_0 under policy π as follows:

$$Q^\pi(x_0, a_0) = E_\pi\left[\sum_{t=0}^{\infty} \gamma^t r(x_t, a_t) \mid x = x_0\right]. \tag{4}$$

The low discount factor means immediate rewards are optimized, while the high one counts future rewards more strongly.

The Q-learning algorithm solves this with the Temporal Difference (TD) scheme by updating the quality function iteratively, as indicated in the following equation:

$$Q_{t+1}(x_t, a_t) = Q_t(x_t, a_t) + \xi\left(r_{t+1} + \gamma \max_{a'} Q_t(x_{t+1}, a') - Q_t(x_t, a_t)\right), \tag{5}$$

where ξ is the learning rate. With higher learning rate, the agent would learn faster, taking shorter time to achieve the optimal, but it might violate the convergence condition if the learning rate is too high.

The introduction of fuzzy inference systems (FIS) enables the possibility to treat continuous state and action spaces. An FIS is described by fuzzy rules as follows [13]:

$$\begin{aligned} &\text{if } x \text{ is } S_i \quad \text{then} \quad y = a[i,1] \quad \text{with } q[i,1],\\ &\qquad\qquad \text{or} \quad\ y = a[i,2] \quad \text{with } q[i,2],\\ &\qquad\qquad \cdots\cdots\\ &\qquad\qquad \text{or} \quad\ y = a[i,J] \quad \text{with } q[i,J], \end{aligned} \tag{6}$$

where S_i is the modal value corresponding to rule i ($i = 1, 2, \ldots, I$), and $(a[i,j])_{j=1}^{J}$ are potential actions whose quality values $(q[i,j])_{j=1}^{J}$ are initialized to zero.

For every rule i, let $k(i) \in [1 : J]$ be the subscript of the rule action chosen by an exploration/exploitation policy (EEP) using ε-greedy method as follows:

$$k(i) = \begin{cases} \underset{j\in\{1,J\}}{\text{random}}(j), & \text{with prob. } 1-\varepsilon,\\ \arg\max\limits_{j\in\{1,J\}} q(i,j), & \text{with prob. } \varepsilon, \end{cases} \tag{7}$$

where ε is the probability of taking a nongreedy action, less than but close to 1. It is a tradeoff factor between exploration and exploitation.

The inferred action $a(x)$ for input vector x and its quality are given by

$$a(x) = \sum_{i=1}^{I} a[i, k(i)]\alpha_i(x), \tag{8}$$

$$Q(x, a(x)) = \sum_{i=1}^{I} q[i, k(i)]\alpha_i(x), \tag{9}$$

where $\alpha_i(x)$, the degree of truth in the FIS for rule i, is defined by a certain membership function which must satisfy $\alpha_i(S_i) = 1$ and $\alpha_j(S_i) = 0$ for $j \neq i$.

At time step t, after performing action $a(x_t)$, the state becomes x_{t+1}. The value of this state is defined as

$$V_t(x_{t+1}) = \sum_{i=1}^{I} \max_k q[i,k]\alpha_i(x_{t+1}). \tag{10}$$

The incremental quantity ΔQ can be calculated by

$$\Delta Q = r_t + \gamma V_t(x_{t+1}) - Q(x_t, a(x_t)), \tag{11}$$

and the elementary quality $q[i,j]$ can be updated by the quantity

$$\Delta q[i,j] = \begin{cases} \xi \cdot \Delta Q \cdot \alpha_i(x_{t+1}), & \text{if } j = k[i],\\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

The detail of the FQL algorithm is presented in Table 1.

Table 1: FQL algorithm.

(1) $t = 0$, $q[i, j] = 0$ for all $i \in [1 : I]$ and $j \in [1 : J]$, and observer the state x_t.
(2) For every rule i, determine $\alpha_i(x_t)$ according to the membership function i.
(3) For every rule i, select $k[i]$ with an EEP (formula (7)).
(4) Calculate the inferred action $a(x_t)$ (formula (8)).
(5) Calculate the corresponding $Q(x, a(x))$ (formula (9)).
(6) Perform the action $a(x_t)$, receive the reward r_{t+1}, and observe the next state x_{t+1}.
(7) For every rule i, calculate $\alpha_i(x_{t+1})$.
(8) Calculate $V_t(x_{t+1})$ (formula (10)).
(9) Calculate ΔQ (formula (11)).
(10) Update $q[i, j]$: $q[i, j] \leftarrow q[i, j] + \Delta q[i, j]$ (formula (12)).
(11) $t \leftarrow t = 1$, go to (3).

4.2. Decentralized Self-Optimization Approach on the FQL. In our decentralized self-optimization approach, the multiple learning agents are the eNBs involved in the optimization process and each eNB runs its own process based on the FQL. In such a multiagent setting where each agent learns independently, we can regard the other agents as part of the environment. Although for such case the convergence to the optimal point could not be proved rigorously, this multi-agent learning approach has been shown to converge in multiple applications [14].

We adopt the current downtilt of each eNB as the state x of the FQL for each agent, which is fuzzified by 5 membership functions (one function denotes one rule, i.e., $I = 5$), as depicted in Figure 4. And the action a is the angle modification applied to the current downtilt with the following settings:

$$J = 5, \quad a[i,1] = -2, \quad a[i,2] = -1, \quad a[i,3] = 0, \quad a[i,4] = 1, \quad a[i,5] = 2 \quad (i = 1,2,3,4,5). \tag{13}$$

The change of the KPI is chosen as the reward function (i.e., $r_t = \mathrm{KPI}_t - \mathrm{KPI}_{t-1}$). In brief, by adjusting the antenna downtilt and receiving the feedback of reward, each eNB tries to maximize the KPI with the learning technique.

5. Simulation Setting

The simulation results presented in this section are derived from a dynamic LTE-based system-level simulator developed with Matlab tool in the light of [15]. The simulation scenario is based on a hexagonal network deployment with a number of eNBs and UEs randomly distributed in the network area.

In the downlink, the received signal power of each UE in its serving cell is influenced by the thermal noise and interferences from its neighboring cells. The received SINR is calculated as

$$\mathrm{SINR}_u = \frac{P_b g_{b,u}}{\sigma^2 + \sum_{i \neq b} P_i g_{i,u}}, \tag{14}$$

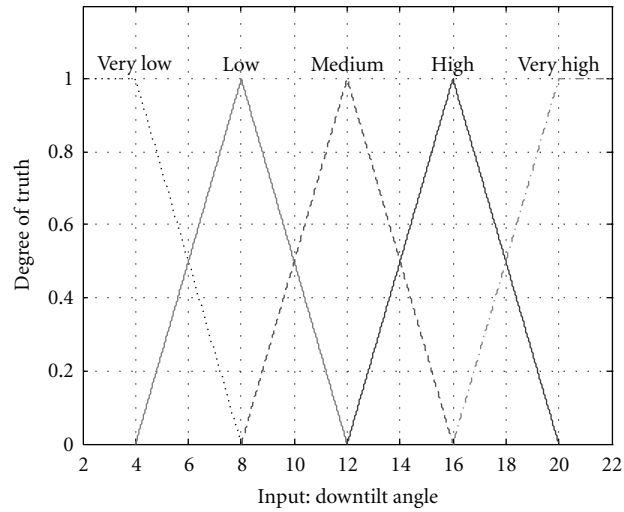


Figure 4: Membership functions.

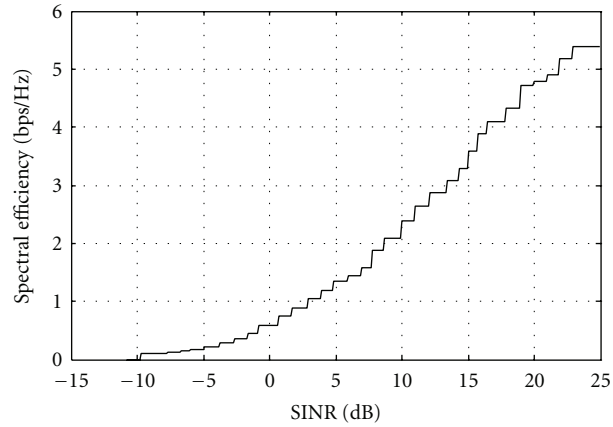


Figure 5: Mapping function from SINR to spectral efficiency.

where P_i is the transmit power (in linear unit) of the cell i (assuming the cell b is the severing cell for user$u$); $g_{i,u}$ is the path loss (in linear unit) from the cell i to the user u; σ^2 is the variance of receiver thermal noise modeled as AWGN; the spectral efficiency for each UE is mapped from the received SINR as follows:

$$\mathrm{SE}_u = \Omega(\mathrm{SINR}_u), \tag{15}$$

where Ω is a step function mapping SINR to the spectral efficiency, which is derived from LTE link-level simulation results and illustrated in Figure 5 [16].

The key parameters for scenario configuration are listed in Table 2, many of which are from 3GPP reports [17]. Three use cases defined by 3GPP [9] are considered in our simulations. In the use case shown in Figure 6(a), there are coverage holes among the cells, and the RF parameters are to be adjusted to expand the coverage area of each cell. In the use case shown in Figure 6(b), LTE systems are deployed with islands of coverage, but due to poor RF planning, the deployed network may fall short of the designed footprint. As a result, the whole coverage is to

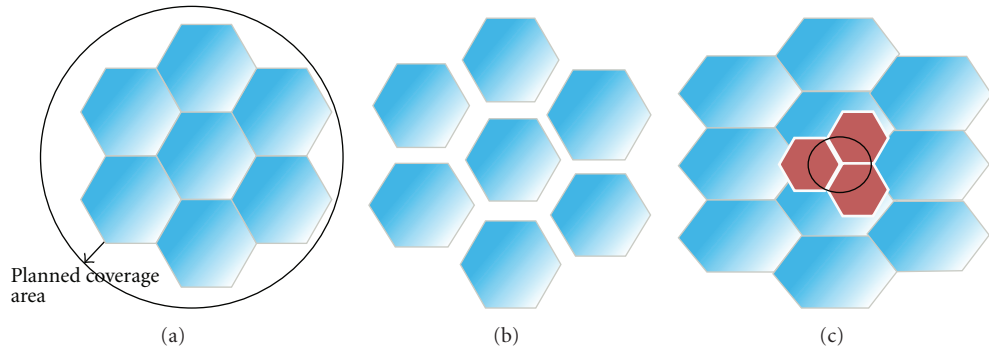


Figure 6: Three use cases: (a) isolated island; (b) coverage hole; (c) newly added site.

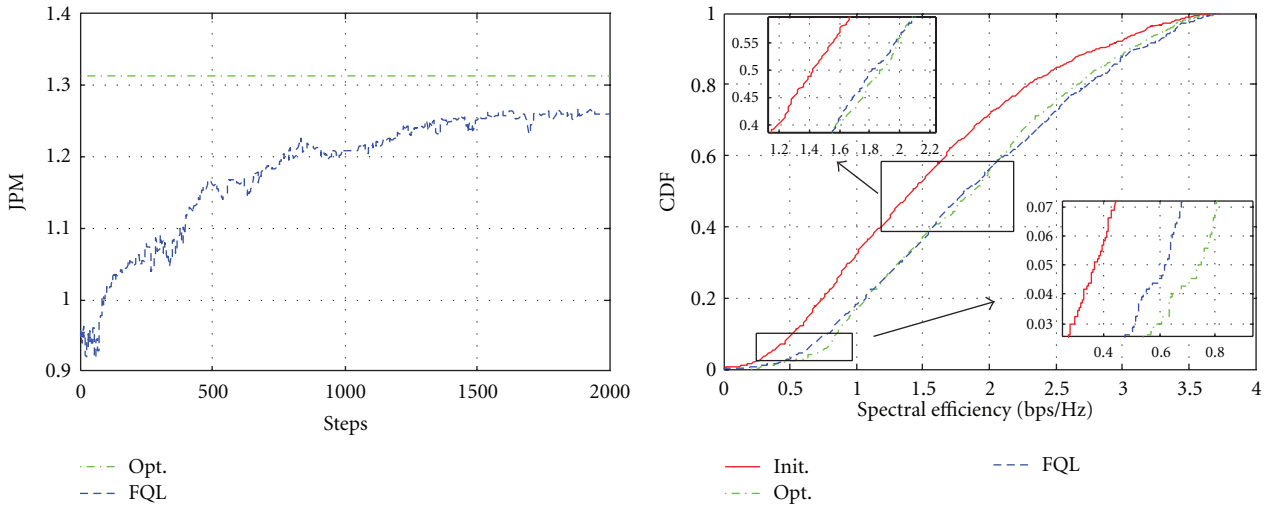


Figure 7: Performance in the isolated island case: (a) the convergency of the average JPM of the involved eNBs; (b) the CDF of the spectral efficiency.

be enlarged by adjusting the RF parameters of the eNBs. In the last case shown in Figure 6(c), a new site is added into the existing coverage. In such case, the new site and the surrounding eNBs need to automatically adjust their RF parameters to minimize interference in the area, while maintaining the service coverage.

In this paper, each site has three sectors, and we call sector as cell. Each eNB controls one cell, and it is an SON entity that runs the FQL algorithm. The central entity is in control of all the decentralized entities, deciding when to start/stop the FQL algorithms running in each SON entity, as well as the setting of the FQL algorithm parameters listed in Table 3.

Considering the total delay to be at the level of 100 ms, we set the time-step size as 200 ms in this simulation. And at each time step, only one agent is activated, in order to ensure the validity of the approximation that the other agents are considered as part of the environment. For comparison, we find the optimal solution by searching all possible downtilt angles (from 4.0° to 20.0°, with 0.5° as the step) for all eNBs in the three use cases respectively.

6. Results and Discussion

In practical application, when to stop the optimization process should be decided by the central entity who compares the average of all the JPMs reported by the involved eNBs with preset threshold, for example, the initial one minus an expected gain from the optimization process. But in order to illustrate the convergence of the FQL algorithm, we run all the simulations for 2000 steps in the three cases.

Figure 7(a) shows the convergence curves for the proposed approach and the global optimal value searched offline in the Isolated Island Case; and Figure 7(b) presents the spectral efficiency distribution before and after the optimization process in this case. Figures 8 and 9 indicate the results in the coverage hole case and the newly added site case, respectively.

From Figure 7(a), we can see that in the isolated island case, our proposed approach obtains remarkable gain (about 34%), and it is close to the global optimal value which has about 38% gain. Figure 7(b) illustrates that the system coverage (i.e., 5% tile of the spectral efficiency) and capacity

Table 2: Scenario configuration.

Parameter	Value		
Use case	Isolated island	Coverage hole	Newly added site
Cell layout	7 sites (3 sectors)	7 sites (3 sectors)	7 + 1 sites (3 sectors)
Wrap around	No	Yes	Yes
Users active per sector	40	100	40
Intersite distance	500 m	1200 m	500 m
Carrier frequency	2.0 GHz		
Minimum distance to base station	35 m		
Propagation loss	$-128.1-37.6^{*}\log10(d)$, d in km		
Shadowing standard deviation	8 dB		
Shadowing correlation distance	50 m		
Shadowing correlation	intersite: 0.5; intersector: 1.0		
Penetration loss	20 dB		
Receiver noise figure	7 dB		
Thermal noise density	−174 dBm/Hz		
Channel bandwidth	10 MHz		
BS maximum transmit power	46 dBm		
BS antenna gain (boresight)	14 dBi		
BS antenna beamwidth @3 dB	70 deg		
BS antenna front-to-back ratio	20 dB		
BS antenna height	32 m		
UE antenna gain	0 dBi		
User movement	Random walk: 3 km/h		
Traffic model	Full buffer		
Scheduling algorithm	Round robin		

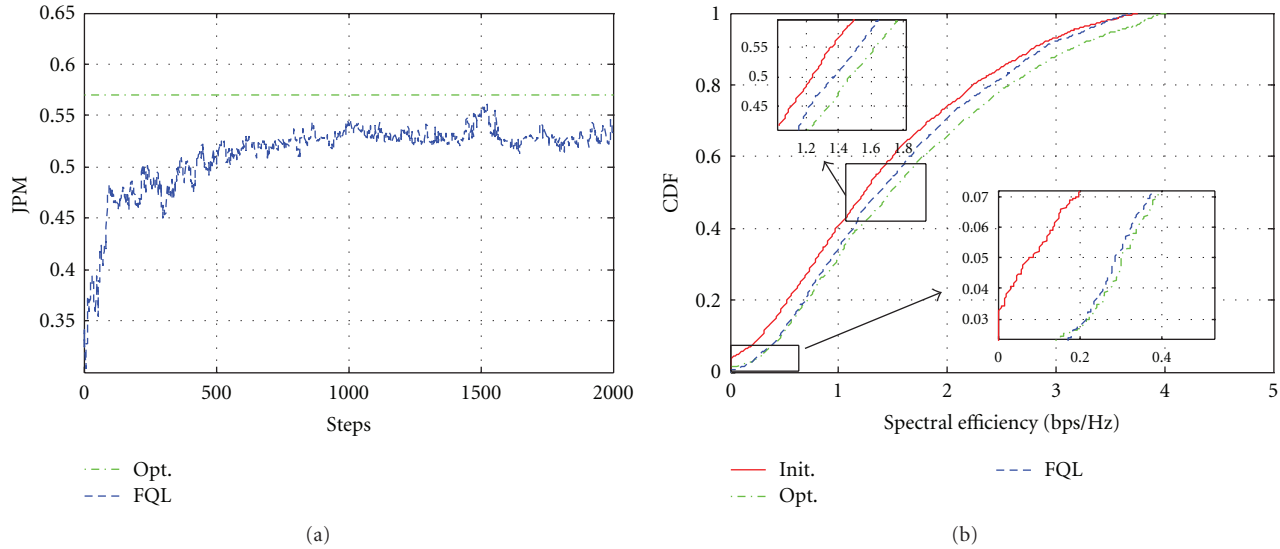


Figure 8: Performance in the coverage hole case: (a) the convergency of the average JPM of the involved eNBs; (b) the CDF of the spectral efficiency.

(i.e., 50% tile of the spectral efficiency) are improved by roughly 67% and 27% from the initial coverage and capacity performance, respectively.

In the coverage hole case as shown in Figure 8, the final JPM obtained by our proposed approach is very close to the global optimal as well, and the system coverage improvement is about 380%, which is much more remarkable than the capacity improvement. It is because in this use case, the coverage performance is rather poor at the beginning due to the large coverage hole at the cell edges, and in the optimization task the coverage is greatly emphasized with the JPM factor $\lambda = 0.8$.

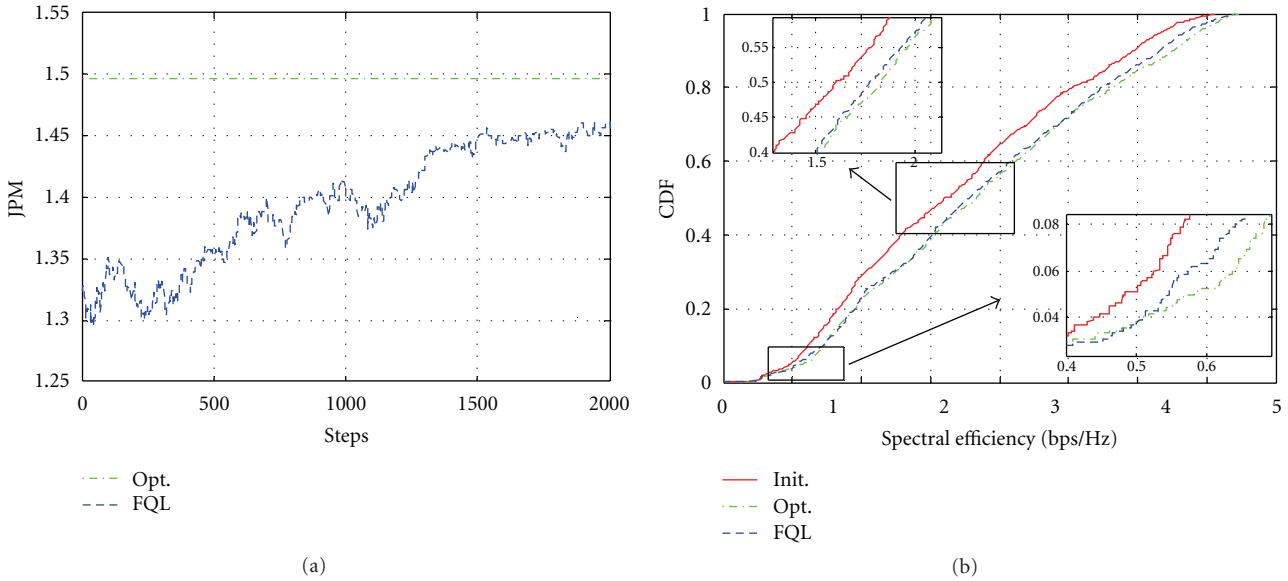


Figure 9: Performance in the newly added site case: (a) the convergency of the average JPM of the involved eNBs; (b) the CDF of the spectral efficiency.

Table 3: Algorithm parameters.

Parameter	Value		
Use case	Isolated island	Coverage hole	Newly added site
Learning rate ξ	0.5	0.6	0.4
Discount factor γ	0.7	0.7	0.5
Greedy action factor ε	0.1	0.1	0.1
The JPM factor λ	0.6	0.8	0.3
The KPI weight w	0.5	0.5	0.5
No. of cells involved	21	21	24

In the newly added site case, the coverage problem is relatively not so serious while the capacity problem may be notable due to the newly added site which would cause great impact on the central users in the cells close to the new site, so we emphasize more about the capacity (with $\lambda = 0.3$). As depicted in Figure 9(b), our approach achieves 13% capacity improvement and the similar coverage improvement.

In sum, the results in all these cases demonstrate that the proposed approach is able to achieve high-performance gain in terms of coverage and capacity, which is close to the optimal. And the tradeoff between coverage and capacity can be easily balanced by setting the factor λ. Besides, from the results we claim that 1000~1500 steps are enough for the convergence in these three cases, since there is little improvement of the average JPM after 1500 steps in Figures 7(a) and 9(a), and after 1000 steps in Figure 8(a). Yet the actual time for reaching the optimal is determined by the time demanded for the eNB to perform the adjustment, and the time for the user equipment to feedback measurement. Considering the total latency is at the level of 0.1 second, we set the time step to be 0.2 second in the simulations, and thus the total time for 1000~1500 steps is 200~300 seconds. Hence, the convergence rate of the proposed approach can meet the need of practical applications.

7. Conclusion

This paper presented an online approach for optimizing coverage and capacity autonomously in LTE networks, which is based on the central control mechanism, and the decentralized fuzzy Q-learning algorithm. All learning agents are in the control of the central entity, and try to optimize its antenna downtilt automatically using information from its own and its neighboring cells.

From the simulation results obtained in different use cases, we can draw the conclusion that our proposed approach not only achieves remarkable performance gain in terms of coverage and capacity, but also has good characters of convergence rate and stability in the multi-agent system. In addition to coverage and capacity optimization, this automatic approach with the ability to learn from the changing environment may also provide other self-optimizing capabilities for LTE self-organizing networks.

Acknowledgment

This work was supported by 973 Project (no. 2012CB 31600), Beijing Natural Science Foundation Funded Project (no. 4110001), National S&T Major Project (no. 2010ZX 03005-003 and no. 2011ZX 03003-002-01).

References

[1] 3GPP TS 32.500, "Self-organizing networks (SON); Concepts and requirements (Release 9)," V9.0.0, December 2009.

[2] 3GPP TR 36.902, "E-UTRA; Self-configuring and self-optimizing network (SON) use cases and solutions (Release 9)," V9.3.1, March 2011.
[3] 3GPP TS 32.541, "SON; Self-healing concepts and requirements (Release 10)," V2.0.0, March 2011.
[4] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
[5] I. Siomina, P. Värbrand, and Y. Di, "Automated optimization of service coverage and base station antenna configuration in UMTS networks," *IEEE Wireless Communications*, vol. 13, no. 6, pp. 16–25, 2006.
[6] T. Cai, G. P. Koudouridis, C. Qvarfordt, J. Johansson, and P. Legg, "Coverage and capacity optimization in E-UTRAN based on central coordination and distributed Gibbs sampling," in *Proceedings of the 71st Vehicular Technology Conference (VTC '10)*, pp. 1–5, May 2010.
[7] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, 1998.
[8] 3GPP TS 36.300, "E-UTRA and E-UTRAN; Overall description; Stage 2 (Release 9)," V9.9.0, December 2011.
[9] 3G Americas, "The benefit of SON in LTE: self-optimization and self-organizing networks," July 2011.
[10] ITU-Recommendation, "Requirements related to technical performance for IMT-Advanced radio interface(s)," Rec. ITU-R M.2134, November 2008.
[11] IEEE 802.16m-07/004r4, "IEEE 802.16m Evaluation Methodology Document (EMD)," November 2008.
[12] W. C. Y. Lee, *Mobile Communications Engineering*, McGraw-Hill, 1998.
[13] P. Y. Glorennec, *Reinforcement Learning: An overview*, European Symposium on Intelligent Techniques, Aachen, Germany, 2000.
[14] L. Panait and S. Luke, "Cooperative multi-agent learning: the state of the art," *Autonomous Agents and Multi-Agent Systems*, vol. 11, no. 3, pp. 387–434, 2005.
[15] ITU-Recommendation, "Guidelines for evaluation of radio transmission technologies for IMT-2000," Rec. ITU-R M.1225, 1998.
[16] 3GPP TSG-RAN WG1 R1-071987, "LTE downlink performance verification for frame structure type 2," Telephone Conference, April 2007.
[17] 3GPP TR 36.814, "E-UTRA; Further advancements for E-UTRA physical layer aspects (Release 9)," V9.0.0, March 2010.