# Epidemics on Tree-Based Communities of Wireless Sensor Networks

**Qiao Li[a], Zhendong Niu[1b]**

*School of Computer Science & Technology, Beijing Institute of Technology, Beijing, 100081, China*
*E-mail: [a]liqq007@bit.edu.cn, [b]zniu@bit.edu.cn*

**Baihai Zhang[c], Lingguo Cui[2]**

*School of Automation, Beijing Institute of Technology, Beijing, 100081, China*
*E-mail: [c]smczhang@bit.edu.cn*

**Bin Wu**

*Science and Technology on Space Physics Laboratory, Beijing, 100076, China*
*E-mail: wubin_005@163.com*

Many efficient deployments of wireless sensor networks based on the community rise into view recently. The energy efficiency is considered as the key design problem of sensor nodes, which lacks sophisticated immune mechanisms for virus attacks. The epidemic on tree-based communities of wireless sensor networks is studied in this paper. Due to random links in the community, the sensor virus extends drastically on the network. Random links add the average number of neighbors to the relevant population and accelerate the virus propagation. The Cayley tree is proposed to analyze the underlying tree-based architecture of the network. The mathematical analysis of the virus prevalence on the abstract tree-based communities is presented in this paper. The analysis and evaluation shows that the number of infected nodes increases exponentially with the prevalence time as the infection spreads. The larger the infection probability is, the higher the speed of the prevalence will be. The research results can further our understanding of epidemics on the tree-based communities of wireless sensor networks.

## 1. Introduction

The wireless sensor networks are deployed recently in variable applications including environmental surveillance [1], agriculture management[2] and security monitoring [3], etc. Different applications of heterogeneous nodes deployed in variable networks make these networks appear to the scientific community, the industry area and the military application. Many efficient deployments based on the community which is one of the common properties of the network rise into view recently [4-6]. Qualitatively, the community is defined as a subset of nodes within the graph such that connections between the nodes are denser than connections with the rest of the network [7]. In a sense, the node is easier to be attacked by the virus compared with the cyber system [8]. For the constraint of cost and resource, it is difficult for the sensor to protect its safety because it is not a complicated hardware structure or software system. In order to prevent the outbreak of the virus prevalence, the immunization strategy is usually chosen for combating the sensor virus [9].

We focus on the community structures and observe the dynamic characteristics of the virus prevalence on the assumption that the infection probability is above the infection probability threshold in this paper. The contributions include the following aspects. The Cayley tree is proposed to depict the underlying tree-based architecture of the communities. We consider three epidemiological processes on the network and find that the number of the infected nodes increases exponentially with respect to the prevalence time. The effect of infection probability on the epidemiological process is analyzed. As a network of ripples quivers momentarily across the surface of the still pool, the epidemic extends drastically on the network.

## 2. Network Model and Preliminaries

We consider the following two problems which motivate our study:

*Problem* **1.** The efficient deployments result in variable network architectures. How should we abstract the main structure characteristics to describe the hierarchical architectures?

*Problem* **2.** The random process is not absolutely suitable for the analysis of the virus prevalence on hierarchical architectures [10]. How should we describe the prevalence?

Symmetric nodes are deployed in the detection area with similar properties (e.g., distance of communication, architecture of software, etc.). The network $G$ can be split into subgraphs $G_0$, $G_1$, $G_2$ … $G_m$ according to the deployment of the communities. $m$ is the number of the subgraphs. $N_0$, $N_1$, $N_2$ … $N_m$ is the number of nodes belonging to each subgraph with $N = N_0+N_1+N_2+ …+ N_m$. Given the network, $G$ is split into $m$ subgraphs with the same size and each subgraph includes $N_0=N/m$ nodes. The largest number of edges in the subgraph is $N_0(N_0-1)/2$ and the largest number of edges outside of the subgraph is $N_0(N-N_0)$.

Fig. 1 shows our network model describing the underlying network architecture and the deployment of sensors within the community. There are several communities with dense internal links denoted by the dashed circles, but between the communities there is only a lower density of external links. In our research, the underlying architecture of tree-based communities is abstracted as the Cayley tree. Random links (denoted by dotted lines) are added to the underlying architecture, and the community model has a hybrid structure with regular bonds and random bonds. The Cayley tree is a regular structure network, in which there is no loops [11].

Every node *i* of the Cayley tree is connected with $k_i=z+1$ child nodes except that the degree of the leaf node on the boundary is 1 with *k*=1. Starting from the root node, whose generation *g* is 0, the tree-based topology is constructed and *z*+1 child nodes are linked with the root node. This process is repeated, and every child node is given *z* children. The number of layers (generations) of the tree topology increases to a fixed value *g*. The Cayley tree grows either in width with the parameter *z* or in depth with the generation *g*. The number of nodes is $n(g)=(z+1)z^{g-1}$ in generation *g* (*g*>0) in the Cayley tree. The total number of nodes of the network is calculated as $N(g)=1+\sum_{g'=1}^{g} n(g')$ . In this tree abstract, the total number of links can be easily calculated by $M(g)=N(g)-1=(z+1)(z^g-1)/(z-1)$ .

The network of hybrid architecture can be described traditionally with a random graph or a regular graph. The virus prevalence on the hybrid network is usually analyzed in terms of the mean-field theory in the hypothetical scene that all nodes of the network are moving randomly or flooding messages under no rules [12-14]. The prevalence on a regular network is usually deemed as a standard percolation problem [15]. The proposed network model in this paper depicts the hierarchical tree-based communities and indicates hybrid structure characteristics of wireless sensor networks. Our research focuses on the actual topologies of tree-based communities and analyzes the epidemiological processes on wireless sensor networks. The epidemic propagation is a susceptible-infectious (SI) process if there is no immune mechanisms
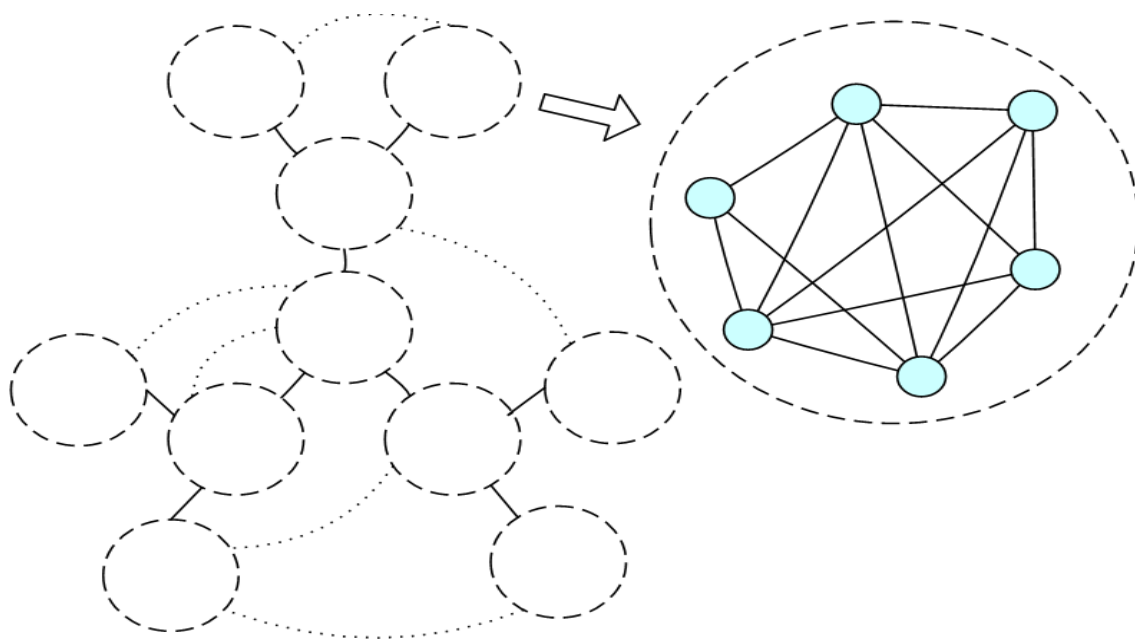


[16].

**Figure1:** Community Model based on the Cayley Tree

## 3. Epidemiological Analysis

Due to high-density internal connections, the epidemic propagates drastically in each community of the network. After the community is attacked, the epidemic extends along the external links and the sensor virus starts to attack other communities. In the community, almost all nodes are connected with each other for strong internal connections. Supposing that one node be infected in the community, all susceptible neighbors will suffer from the virus with the same infection probability. We analyze the prevalence in the hypothetical scene that each community

could be deemed as a whole for strong internal connections and the highly contagious worm. In the analysis, the epidemic starts from certain communities. Random external links between communities are chosen uniformly. Following the description above, the infected number in the network can be given by

$$I(t) = \sum_{t'=0}^{t} hN_0 C(t')[1 + \gamma \delta_{out} I(t-t')]$$ ,

(3.1)

where $\delta_{out}$ is the outside connection probability of communities with $\delta_{out} = \dfrac{2m^2 S_{out}}{N(N-m)}$.

$S_{out}$ is the number of outside edges. $\gamma$ is the coefficient which relates the connection probability to the number of nodes in the community. If $C(t')$ is a constant $C_0$, a fixed number of communities are attacked by the virus in one time unit, and the epidemiological process is calculated as

$$I(t) = \sum_{t'=0}^{t} \frac{hNC_0}{m}[1 + \frac{2m^2 \gamma S_{out}}{N(N-m)} I(t-t')]$$ .

(3.2)

The sum is approximately calculated with an integral

$$I(t) = \int_0^t \frac{hNC_0}{m}[1 + \frac{2m^2 \gamma S_{out}}{N(N-m)} I(t-t')]dt'$$ .

(3.3)

The first-order derivative of (3.3) is obtained

$$\frac{dI(t)}{dt} = \frac{hNC_0}{m} + \frac{2mh \gamma C_0 S_{out}}{N-m} I(t)$$ .

(3.4)

Supposing there be only one infected community on the underlying architecture on the initial stage, we can get the solution

$$I(t) = (\frac{N}{m} + \frac{N(N-m)}{2m^2 \gamma S_{out}})e^{\frac{2mh \gamma C_0 S_{out}}{N-m}t} - \frac{N(N-m)}{2m^2 \gamma S_{out}}$$ .

(3.5)

We can see from (3.5) that the number of infected nodes increases exponentially with $t$ in the network. The epidemic drastically extends to all nodes of the wireless sensor network in a short time.

In consideration of the underlying architecture of the network, $C(t')$ may have more complex forms. If the number of attacked communities on the underlying architecture increases linearly with $C(t')=C_0 t'$, the number of infected nodes $I(t)$ will be rewritten as

$$I(t) = \sum_{t'=0}^{t} \frac{hNC_0}{m}t'[1 + \frac{2m^2 \gamma S_{out}}{N(N-m)} I(t-t')]$$ .

(3.6)

The sum is approximately calculated with the integral

$$I(t) = \int_0^t \frac{hNC_0}{m}t'[1 + \frac{2m^2 \gamma S_{out}}{N(N-m)} I(t-t')]dt'$$ .

(3.7)

Both sides of (3.7) are differentiated with respect to $t$, and

$$\frac{dI(t)}{dt} = \frac{hNC_0 t}{m} + \frac{2mh \gamma C_0 S_{out}}{N-m} \int_0^t I(t')dt'$$ .

(3.8)

The second-order derivative of (3.8) is given with

$$\frac{d^2 I(t)}{dt^2} = \frac{hNC_0}{m} + \frac{2mh\,\gamma\,C_0 S_{out}}{N-m} I(t) \qquad (3.9)$$

We can get the solution as follows

$$I(t) = C_1 e^{\sqrt{\frac{2mh\,\gamma\,C_0 S_{out}}{N-m}}\,t} + C_2 e^{-\sqrt{\frac{2mh\,\gamma\,C_0 S_{out}}{N-m}}\,t} - \frac{N(N-m)}{2m^2\,\gamma\,S_{out}} , \qquad (3.10)$$

where $C_1$ and $C_2$ are two constants. Supposing that only one community be infected on the underlying architecture on the initial stage, the following equation is gotten with

$$C_1 + C_2 = \frac{N}{m} + \frac{N(N-m)}{2m^2\,\gamma\,S_{out}} . \qquad (3.11)$$

For $C(t^{'}) = C_0 t^{'}$, the number of infected nodes increases exponentially with $t$ as the epidemic extends on the wireless sensor network.

On the tree-based communities of wireless sensor networks, the sensor virus will attack the neighbor nodes on the underlying tree-based architecture and along external links between communities. Supposing the underlying tree-based architecture be idealized as the Cayley tree and the virus may attack all neighbor nodes, $C(t^{'})$ can be simplified as $C(t^{'}) = (z+1)z^{t^{'}}$. In the tree abstract, $z+1$ is the number of child communities of each community. The number of infected nodes $I(t)$ is given by

$$I(t) = \sum_{t^{'}=0}^{t} \frac{hN(z+1)z^{t^{'}}}{m} [1 + \frac{2m^2\,\gamma\,S_{out}}{N(N-m)} I(t-t^{'})] . \qquad (3.12)$$

The sum is approximately calculated with the integral

$$I(t) = \int_0^t \frac{hN(z+1)z^{t^{'}}}{m} [1 + \frac{2m^2\,\gamma\,S_{out}}{N(N-m)} I(t-t^{'})]dt^{'} . \qquad (3.13)$$

Let $F(t) = \int_0^t z^{t^{'}} I(t-t^{'})dt^{'}$, and both sides of (3.13) are differentiated with respect to $t$,

$$F^{'}(t) = \ln z\, F(t) + I(t) . \qquad (3.14)$$

From (3.13), we get

$$I(t) = \frac{hN(z+1)}{m} \int_0^t z^{t^{'}} dt^{'} + \frac{2mh\,\gamma\,(z+1)S_{out}}{N-m} F(t) . \qquad (3.15)$$

Then, the following equation is gotten as

$$F(t) = \frac{N-m}{2mh\,\gamma\,(z+1)S_{out}} I(t) - \frac{N(N-m)}{2m^2\,\gamma\,S_{out}} \int_0^t z^{t^{'}} dt^{'} . \qquad (3.16)$$

The first-order derivative of (3.16) is generated by

$$F^{'}(t) = \frac{N-m}{2mh\,\gamma\,(z+1)S_{out}} I^{'}(t) - \frac{N(N-m)}{2m^2\,\gamma\,S_{out}} z^{t^{'}} . \qquad (3.17)$$

From (3.14), (3.16) and (3.17), the following equation is gotten as

$$\frac{N-m}{2mh\,\gamma\,(z+1)S_{out}} I^{'}(t) - \frac{N(N-m)}{2m^2\,\gamma\,S_{out}} z^{t^{'}} =$$

$$\ln z [\frac{N-m}{2mh\,\gamma\,(z+1)S_{out}} I(t) - \frac{N(N-m)}{2m^2\,\gamma\,S_{out}} \int_0^t z^{t^{'}} dt^{'}] + I(t)$$

$$(3.18)$$

The second-order derivative of (3.18) with respect to t is calculated as

$$I''(t) = [\ln z + \frac{2mh\,\gamma\,(z+1)\,S_{out}}{N-m}]\,I'(t)$$

(3.19)

The final result will be solved as

$$I(t) = C_3 + C_4 e^{[\ln z + \frac{2mh\,\gamma\,(z+1)\,S_{out}}{N-m}]t},$$

(3.20)

where $C_3$ and $C_4$ are two constants. Supposing there is one infected community on the underlying architecture on the initial stage, the following equation is gotten with

$$C_3 + C_4 = \frac{N}{m}.$$

(3.21)

If the sensor virus attacks all its neighbors on the underlying tree-based architecture and attacks nodes in other communities along external links, the number of infected nodes increases exponentially with $t$ as the epidemic extends on the wireless sensor network.

The said mathematical analysis shows that the total number of infected nodes increases exponentially when the prevalence extends on the hybrid network and three different infections occur on the underlying tree-based architecture. The existence of random bonds is similar to the addition of susceptible neighbors to the relevant nodes. The sensor virus may infect the neighbor nodes in the community and attack nodes in other communities along random external links, which increase the average number of neighbors of the node and accelerate the virus prevalence in the network. As a network of ripples quivers momentarily across the surface of the still pool, the epidemic extends drastically. The virus prevalence is a Susceptible-Infectious process when there is no immune scheme in the design of the software architecture of the sensor node.

## 4. Evaluation and Analysis

Mathematical evaluations and real architecture evaluations of epidemic propagations on the tree-based communities are presented in this Section.

For $C(t') = C_0$, the fixed number of communities of the network will be attacked by the virus on the proposed community model in each time unit. In the simulation, the parameters are set with $S_{out}$=2, $h$=0.8, $C_0$=3 and $\gamma$=0.8. One community of the network includes 4 nodes with $N/m = 4$. Each time unit includes 10 seconds in the simulation. In Fig. 2(a), the infection extends exponentially with respect to the prevalence time $t$ on the tree-based communities of the wireless sensor network on the basis of (3.5). If there is no immune strategy, the number of infected nodes will increase rapidly until all nodes become infectious in the network. From the figure we can see that the virus propagates drastically on the community network and about $2.8 \times 10^6$ nodes are infected at 50s.

For $C(t') = C_0 t'$, the number of communities attacked by the virus increases linearly on the community model in each time unit. In the simulation, the parameters are set with $S_{out}$=2, $h$=0.8, $C_0$=3 and $\gamma$=0.8. One community of the network includes 4 nodes with $N/m = 4$. Each time unit includes 10 seconds in the simulation. The constants $C_1$ and $C_2$ in (3.10) can be approximately calculated at $t = 0$ and $t = 10$s. Fig. 2(b) shows that the virus expands exponentially with respect to the prevalence time $t$ on the tree-based communities of the wireless sensor network on the basis of (3.10) and the random links in the communities
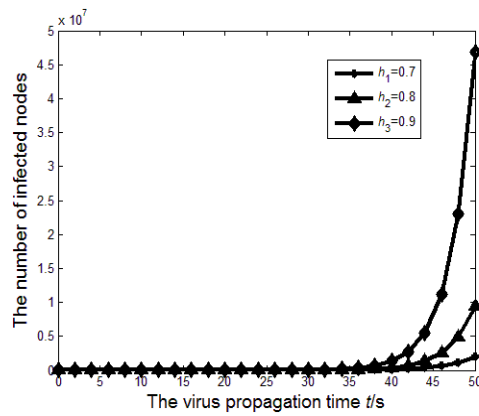
accelerate the epidemic propagation. The number of infected nodes increases rapidly after the attack from the sensor virus until all nodes of the network are infected. The virus propagates drastically on the community network, and about $9 \times 10^3$ nodes are infected at 50s.
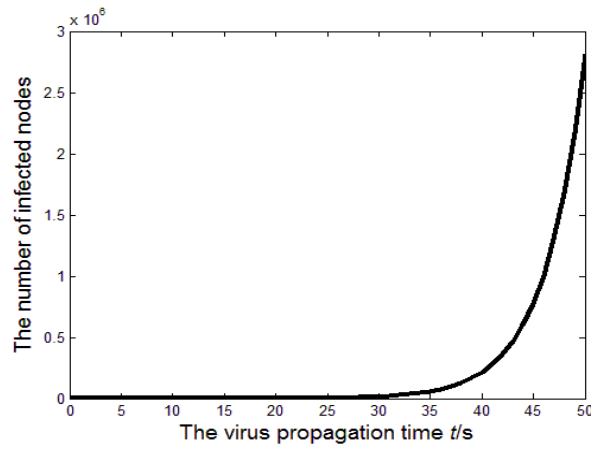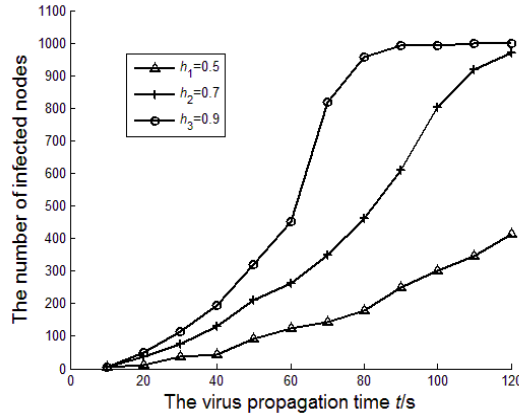


(a)
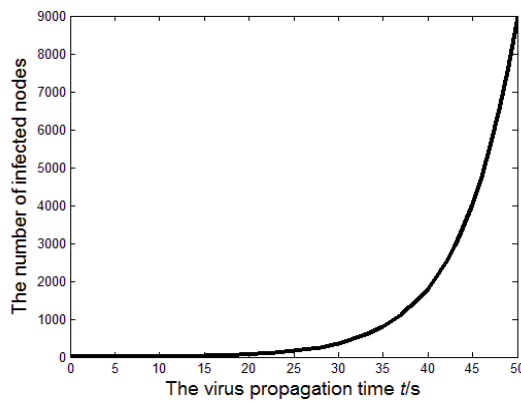


(b)



(c)

(d)

**Figure 2:** Epidemiological Propagations on the Tree-based Communities

For $C(t')=(z+1)z^{t'}$ , the virus will attack all neighbors on the community model in each time unit. The epidemic propagation and the effect of the infection probability $h$ on the epidemiological process are analyzed in the evaluation. The underlying architecture is the Cayley tree, and each community is connected with 3 child communities with $z=2$. One community of the network includes 4 nodes with $N/m = 4$. The parameters are set with $S_{out}=2$ and $\gamma=0.8$. Each time unit includes 10 seconds in the simulation. The constants $C_3$ and $C_4$ in (3.20) can be approximately calculated at $t = 0$ and $t = 10s$. In Fig. 2(c), the virus expands exponentially with the prevalence time $t$ as well on the basis of (3.20). If there is no immune strategy, the number of infected nodes increases rapidly until all nodes of the network become infectious. Fig. 2(c) also shows the time evolution of the infected number with the infection probability $h_1=0.7$, $h_2=0.8$ and $h_3=0.9$. At $t = 50s$, the infected number in the network is $0.2\times10^7$ with $h_1=0.7$, $1.0\times10^7$ with $h_2=0.8$ and $4.75\times10^7$ with $h_3=0.9$. The infection extends exponentially at different speeds when the infection probability $h$ changes. The larger the infection probability $h$ is, the higher the speed of the prevalence will be.

Experiments of the epidemic propagation on the communities of ZIGBEE are presented. ZIGBEE can construct the tree-based communities in a broad deployment area if the node is the Full Function Device (FFD). Assuming there be a small section of infectious nodes on the initial stage, the infected numbers of nodes is observed by using a large number of experiments. We assume that node $i$ be susceptible and infected with the infection probability $h$ if there are infected neighbors. The effect of the infection probability $h$ on the epidemiological process is tested. In the experiment, 1000 nodes are randomly deployed in a 1000m×1000m surveillance area. The parameters are set with $\delta_{in} = 0.9$ and $\delta_{out} = 0.3$. One community of the network includes maximum 15 nodes. In the experiments, each time unit includes 10 seconds. The figure shows the time evolution of the infected number with the infection probability $h_1=0.5$, $h_2=0.7$ and $h_3=0.9$. The infection extends exponentially at different speeds when the infection probability $h$ changes. The experiment shows that the larger the infection probability is, the higher the speed of prevalence will be. At 100s, almost all nodes are infected with $h_3=0.9$, 800 nodes are infected with $h_2=0.7$, and 300 nodes are infected with $h_1=0.5$ in Fig. 2(d). The

experiments show that the mathematical analysis reveals dynamic characteristics of the epidemic propagation on the communities.

Above experiments show the number of infected nodes increases exponentially with respect to the prevalence time although different infections occur on the underlying tree-based communities. The larger the infection probability *h* is, the higher the speed of the viral transmission will be. The number of infected nodes increases rapidly until the whole network is attacked when the network suffers from the sensor virus and there is no immune strategy.

## 5. Conclusion

The epidemic on tree-based communities of wireless sensor networks is studied in this paper and becomes more drastic due to random links in the communities. The underlying architecture of the network is abstracted as the Cayley tree. The dynamic characteristics of the virus prevalence is focused on the assumption that the infection probability is above the infection probability threshold. The evaluation and analysis shows that the number of infected nodes increases exponentially with respect to the prevalence time. The larger the infection probability is, the higher the speed of the viral transmission is. The existence of random links is similar to the addition of susceptible neighbors to the relevant nodes. The sensor virus attacks the neighbors on the underlying network architecture and attacks nodes in other communities along external links of the network. The random bonds of the network accelerate the virus prevalence. As a result, the network is easily attacked by the virus from one side to another. The research can further our understanding of epidemics on the tree-based communities of wireless sensor network.

## References

[1] E. Kampianakis, J. Kimionis, K. Tountas, C. Konstantopoulos, E. Koutroulis, A. Bletsas. *Wireless environmental sensor networking with analog scatter radio and timer principles*[J]. IEEE Sensors Journal.14 (10): 3365-3376(2014 )

[2] T. R. Rao, D. Balachander, N. Tiwari, M. V. S. N. Prasad. *Ultra-high frequency near-ground short-range propagation measurements in forest and plantation environments for wireless sensor networks*[J]. IET Wireless Sensor Systems. 3 (1): 80-84(2012)

[3] S. Roy, M. Conti, S. Setia, S. Jajodia. *Secure data aggregation in wireless sensor networks: filtering out the attacker's impact*[J]. IEEE Transactions on Information Forensics and Security. 9 (4): 681-694(2014)

[4] C. Pham. *Communication performances of IEEE 802.15.4 wireless sensor motes for data-intensive applications: a comparison of WaspMote, Arduino MEGA, TelosB,MicaZ and iMote2 for image surveillance*[J]. Journal of Network and Computer Applications. 46(): 48-59(2014)

[5] F. Wei, M. E. Haque, X. D. Lu, K. Mori. *Autonomous community construction and coordination technology to achieve real-time transmission in multiple emergencies' situation*[J]. Telecommunication Systems. 54(1): 61-78(2013)

[6] J. Y. Wu, X. Y. Shao, H. P. Zhu. *A novel clustering routing protocol with community structure detection for wireless sensor networks*[J]. Applied Mechanics and Materials. 472(): 460-465(2014)

[7] M. E. J. Newman, T. P. Peixoto. Generalized *Communities in networks*[J]. Physical Review Letters. 115(8): 088701-1-5(2015)

[8] S. S. Tang, W. Li. *An epidemic model with adaptive virus spread control for wireless sensor networks*[J]. International Journal of Security and Networks. 6(4): 201-210(2011)

[9]  H. M. Salmon, C. M. D. Farias, P. Loureiro, L. Pirmez, S. Rossetto, P. H. Rodrigues. *Intrusion detection system for wireless sensor networks using danger theory immune-inspired techniques*[J]. International Journal of Wireless Information Networks. 20(1): 39-66(2013)

[10] P. Kopietz. *Mean-field theory and the Gaussian approximation*[J]. Lecture Notes in Physics. 798(): 23-52(2010)

[11] J. P. Bagrow. *Communities and bottlenecks: trees and treelike networks have high modularity*[J]. Physical Review E. 85(): 066118-1-9(2012)

[12] C. H. Li, C. C. Tsai, S. Y. Yang. *Analysis of epidemic spreading of an SIRS model in complex heterogeneous networks*[J]. Communications in Nonlinear Science and Numerical Simulation. 19(4): 1042-1054(2014)

[13] Y. Muroya, H. Li, T. Kuniya. *Complete global analysis of an SIRS epidemic model with graded cure and incomplete recovery rates*[J]. Journal of Mathematical Analysis and Applications. 410(2): 719-732(2014)

[14] L. J. Hao, G. R. Jiang, S. Y. Liu, L. Ling. *Global dynamics of an SIRS epidemic model with saturation incidence*[J]. Biosystems. 114(1): 56-63(2013)

[15] S. Dietrich, A. Ammon. *Introduction to Percolation Theory*[M]. Burgess Science Press, Great Britain. 15-20 (1992)

[16] T. J. Norman. *The Mathematical Theory of Infectious Diseases*[M]. Hafner Press, New York. 113-126 (1975)