

Dynamics in small worlds of tree topologies of wireless sensor networks

Qiao Li^{1,*}, Baihai Zhang¹, Zhun Fan², and Athanasios V. Vasilakos³

1. School of Automation, Beijing Institute of Technology, Beijing 100081, P. R. China;

2. Department of Mechanical Engineering, Technical University of Denmark, Kongens Lyngby 2800, Denmark;

3. Department of Computer and Telecommunications Engineering, University of Western Macedonia, Kozani GR 50100, Greece

Abstract: Tree topologies, which construct spatial graphs with large characteristic path lengths and small clustering coefficients, are ubiquitous in deployments of wireless sensor networks. Small worlds are investigated in tree-based networks. Due to link additions, characteristic path lengths reduce rapidly and clustering coefficients increase greatly. A tree abstract, Cayley tree, is considered for the study of the navigation algorithm, which runs automatically in the small worlds of tree-based networks. In the further study, epidemics in the small worlds of tree-based wireless sensor networks on the large scale are studied, and the percolation threshold is calculated, at which the outbreak of the epidemic takes place. Compared with Cayley tree, there is a smaller percolation threshold suffering from the epidemic.

Keywords: epidemic, navigation algorithm, small world, tree topology.

DOI: 10.1109/JSEE.2012.00040

1. Introduction

In recent years, it is seen that the deployments of wireless sensor networks (WSNs) exist in a variety of applications including habitat and environmental monitoring [1], precision agriculture, security surveillance [2], etc. More and more efficient ways of sensor deployments rise into view. The tree topology is a kind of architecture used frequently, which is ubiquitous in the deployment of wireless sensor nodes. Some routing protocols, topology control algorithms and aggregation schedules of WSNs are helpful to construct tree topologies.

Distributed quad-tree (DQT) is an in-network tree framework, which achieves the distance sensitivity and resiliency for event-based querying, as well as greatly reduces the cost of complex range querying [3]. The hybrid address assignment (HAA) scheme uses a tree address structure to make the proposed scheme less susceptible to

physical distribution of WSN devices [4]. Semantic/spatial correlation-aware tree (SCT) is a simple, scalable and distributed tree structure that addresses the practical challenges in the context of aggregation in WSNs [5]. With this structure, the total cost of the aggregation tree can be minimized. The localized energy-efficient multicast algorithm (LEMA) uses a function to locally estimate the energy-efficient paths to multiple destinations [6]. It is able to deal with the inherent errors of WSNs. Several tree topology protocols based on minimal spanning tree (MST) [7] have attracted much attention recently. The base-station controlled dynamic clustering protocol (BCDCP) introduces an MST to connect cluster-heads and adopts iterative cluster splitting algorithms to choose cluster-heads or form clusters [8]. It distributes energy dissipation evenly among all sensor nodes to improve the network lifetime and average energy saving. The cluster-based and tree-based power efficient data collection and aggregation (CTPEDCA) protocol is based on the clustering and MST routing strategy for cluster heads, which uses the MST to improve the transmission routing mechanism between cluster heads so that only one cluster head communicates directly with the far-away base station in each round [9]. Most of the tree topology protocols are multi-hop protocols, which are famous for energy saving in data gathering and transferring.

In the work of tree-based networks, random link additions among nodes take place inevitably for the use of omnidirectional antennae. Obstacles, adjustments of radio energy, joins of new members and errors of the location precision all incur link additions. Our researches and contributions focus on studying small world phenomena [10] existing in tree-based WSNs. It is complex to consider dynamics in the small worlds. A tree abstract, Cayley tree, is considered, and the navigation algorithm, which runs automatically for searching short paths in the small worlds of tree-based networks, is studied. Due to shortcuts, the epidemic propagation becomes much drastic in the network. If the distribution of infected nodes is random, the prob-

Manuscript received March 10, 2011.

*Corresponding author.

This work was supported by the National Natural Science Foundation of China (61104086) and the National Defense Advanced Research Project of China (40405020401).

lem when an epidemic takes place becomes equivalent to a standard percolation problem [11]. In this paper, the percolation threshold, at which the outbreak of the epidemic takes place in the small worlds of tree-based networks, is calculated.

The rest of the paper is organized as follows. Section 2 outlines the related work. Our small-world model and basic ideas are described in Section 3. The navigation algorithm which runs automatically in the proposed small world is studied in Section 4. The epidemic threshold and mathematical calculation are given in Section 5. Numerical simulations and analyses are presented in Section 6. The paper concludes in Section 7.

2. Related work

Small world phenomena were first investigated in sociology that individuals are often linked by a short chain of acquaintances. M. Stanley and his group conducted a series of mail delivery experiments and found that an average of “six degrees of separation” exists between senders and receivers [10]. J. W. Duncan and H. S. Steven proposed an alternative model for small world phenomena by using the graph theory [12]. Recent researches show that small world phenomena are ubiquitous in nature, society and technology. Small worlds were also observed in wireless networks [13]. Random link additions among nodes take place inevitably for the use of omnidirectional antennae of sensors. The α -model [12], which is used in the category of sociology, cannot be directly used to analyze the small world characteristics of WSNs. In our research, link additions are conducted on tree topologies to observe small world phenomena.

The varieties of the length characteristic and the clustering characteristic affect message transmissions in small worlds. K. Jon proposed an infinite family of network models in [14]. A decentralized algorithm, which achieved a rapid delivery time, was proposed. M. R. Roberson and B. A. Daniel studied the navigation by the greedy algorithm on fractal small world networks with random long-range connections taken from a power-law distribution [15]. S. A. Hill and D. Braha presented a dynamic preferential attachment mechanism to explain real-world dynamic networks and qualitatively reproduce these dynamic centrality phenomena [16]. However, the above models are not suitable for small worlds of tree-based WSNs. Our work focuses on the navigation algorithm, which runs automatically in small worlds of tree-based networks.

Compared with regular computer systems, it is even easier for sensors to be compromised by virus attacks. Sensor nodes do not have the complicated hardware architecture or operating system to protect program safety due

to the cost and resource constraints. Nodes in the same network are homogeneous in both hardware and software. Y. Yang et al. studied the worm propagation in the WSNs and considered the propagation as a random process in a random network [17]. P. De et al. investigated the potential disastrous threat of node compromise spreading in the WSNs. They focused on the possible epidemic breakout based on the random network [18]. Considering routing protocols, topology control algorithms and aggregation schedules, random graphs cannot completely indicate the structure characteristics of WSNs. For the existence of small world phenomena in wireless networks, we analyze dynamics and epidemics in the small world of tree-based WSNs. Some researches were conducted in dynamics of epidemic propagations on small world networks. M. E. J. Newman and J. W. Duncan proposed a simple model for the disease propagation on an infinite small world in which the communication of the disease takes place with 100% efficiency [19]. R. Pastor-Satorras and A. Vespignani studied a dynamical model for the spread of epidemics in complex networks by analytical methods and large-scale simulations [20]. They made use of the susceptible-infected-susceptible (SIS) model [21] and studied analytically the prevalence and the persistence of infected individuals on complex networks. M. M. Telo da Gama and A. Nunes studied the effect of the network structure on immunization models for life diseases and found that in addition to the reduction of the effective transmission rate, through the screening of infectives, spatial correlations might strongly enhance the stochastic fluctuations [22]. E. S. Thomas et al. studied the relative effects of vaccinations and avoidance of the infected individuals in a susceptible-infected-recovered (SIR) epidemic model on a dynamic small world network [23].

3. Network model and basic idea

We consider a network composed of N nodes distributed in the network. The following properties of the WSN are supposed:

- (i) nodes are not moveable after they are randomly distributed in the network,
- (ii) all nodes are symmetric and have similar characteristics (e.g., range of radio coverage, energy of batteries, etc.),
- (iii) there is only one root, from which only one tree is constructed in the sensor network.

An abstract small-world model for tree-based networks is needed for the study of dynamics. We choose Cayley tree to construct the underlying lattice of tree-based networks. Random link additions are conducted on it. Cayley tree is such an abstract with one root node and K children

nodes for each parent node, $K \geq 2$. After the root node is selected out randomly, it broadcasts its location information to other nodes in the field. K neighbor nodes, which are nearest to the root node, are connected to the root node. Then the K nodes broadcast their location information to other nodes which are not connected and each is connected with $K - 1$ nearest neighbor nodes. This process continues until all nodes are included in the tree. Random link additions are conducted on the tree model. The small-world model neglects the actual distance information between the nodes. If two nodes are connected with each other, the logical distance between them is 1. It simplifies extremely the problem and represents a spatial graph model. Fig. 1 shows the small world of a tree-based network and our proposed model. The small world of an SCT [5] is shown in Fig. 1(a). Circles denote nodes. The black line denotes the link between two nodes. Dashed lines denote link additions of the small world. The actual distance information is neglected, and only logical links of the SCT are described. Cayley tree with link additions is shown in Fig. 1(b). In this figure, each node has three neighbors on the underlying topology with $K = 3$, and dashed lines denote link additions.

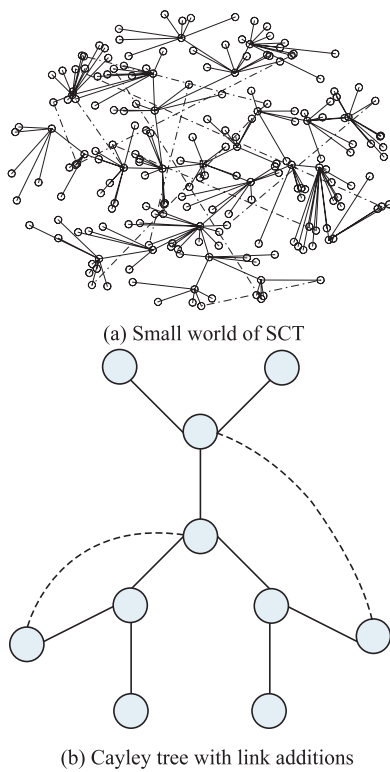


Fig. 1 Small worlds of a tree-based network and our proposed model

4. Navigation algorithm

When small world phenomena occur, the length character-

istic of the network varies greatly. The choice of transmission paths is of vital importance for the normal transmission of messages in small worlds. The navigation algorithm, which runs automatically in the small worlds of tree topologies, is studied.

In tree-based networks, the existence of small world phenomena does not mean that the nodes, which are not as smart as humans, can find short paths automatically. According to the theories of the random graph and regular graph, no mechanism tells us how to construct such short paths although they really exist in the small world. Small world networks are rich in local connections, with a few long-distance connections (shortcuts). Because the logical distance between two nodes is defined to be 1 if they are connected directly with each other, the logical distance between every two nodes is clear and unique. Every node has K local connections and a few long-distance connections.

The message holder u in a given step has the following knowledge:

- (i) the set of local contacts among all nodes;
- (ii) the location information, on the tree-based network, of the target t ;
- (iii) the locations and long-distance contacts of all nodes that have come in contact with the message.

The expected delivery time of a decentralized algorithm is the expected number of steps taken by the algorithm to deliver messages. We suppose that the i th directed edge from u has endpoint v with probability proportional to $[(K - 1)^{d(u,v)}]^{-r}$ with $K \geq 2$, r is the clustering exponent and $d(u, v)$ is the lattice distance between two nodes u and v . The quantity is divided by the appropriate normalizing constant $\sum_v [(K - 1)^{d(u,v)}]^{-r}$ to get a probability distribution.

The normalization term can be obtained by

$$\begin{aligned} \sum_{v \neq u} [(K - 1)^{d(u,v)}]^{-r} &= \sum_{i=1}^{M'} K (K - 1)^{i-1} [(K - 1)^i]^{-r} = \\ &= \frac{K}{K - 1} \sum_{i=1}^{M'} (K - 1)^{(1-r)i} = \\ &= \frac{K}{K - 1} \sum_{i=1}^{M'} [(K - 1)^{1-r}]^i \end{aligned} \tag{1}$$

where M' is the depth of the tree. With $r = 1$, $(K - 1)^{1-r} = 1$, and

$$\sum_{v \neq u} [(K - 1)^{d(u,v)}]^{-r} = \frac{KM'}{K - 1}. \tag{2}$$

For $0 < r < 1$, $(K - 1)^{1-r} > 1$, and

$$\sum_{v \neq u} [(K - 1)^{d(u,v)}]^{-r} = \frac{K}{K - 1} \frac{(K - 1)^{(1-r)(M'+1)} - (K - 1)^{1-r}}{(K - 1)^{1-r} - 1}. \quad (3)$$

For $r > 1$, $0 < (K - 1)^{1-r} < 1$, and

$$\sum_{v \neq u} [(K - 1)^{d(u,v)}]^{-r} = \frac{K}{K - 1} \frac{(K - 1)^{1-r} - (K - 1)^{(1-r)(M'+1)}}{1 - (K - 1)^{1-r}}. \quad (4)$$

With $r = 1$, j shells of radius R , $j - 1 < R < j$, surround the target node t , where $j = 1, 2, \dots$. We suppose that the message holder is in shell j . The probability that the holder is connected to a long-distance link to a node in shell $j - 1$ can be calculated as follows:

$$P_1 = \sum_{i=j-1}^j \frac{K(K - 1)^{i-1} [(K - 1)^i]^{-1}}{\frac{KM'}{K - 1}} = \sum_{i=j-1}^j \frac{1}{M'} = \frac{1}{M'}. \quad (5)$$

Let X_1 denote the total number of steps spent in the process, EX_1 can be obtained through the following relation

$$EX_1 = \sum_{i=1}^{\infty} iP_1(1 - P_1)^{i-1} = P_1 \sum_{i=1}^{\infty} i(1 - P_1)^{i-1} = \frac{1}{P_1} = M'. \quad (6)$$

Since we get the largest shell with $j = M'$, the number of shells between the source and the target is of the order of M' . Thus, $EX \sim (M')^2$, where EX is the expected total number of steps required to reach the target.

For $0 < r < 1$, a circle of radius l , $0 < l < M'$, surrounds the target node t . The probability that i connects to any node in the circle can be given by

$$P_2 = \frac{[(K - 1)^{1-r} - 1] \sum_{i=1}^l [(K - 1)^{1-r}]^i}{[(K - 1)^{(1-r)(M'+1)} - (K - 1)^{1-r}]} = \frac{(K - 1)^{(1-r)(l+1)} - (K - 1)^{1-r}}{(K - 1)^{(1-r)(M'+1)} - (K - 1)^{1-r}}. \quad (7)$$

For $0 < l < M'$, the source node lies almost outside of the circle. When the source chooses a long-distance link in the

circle to transfer messages, the expected number of steps spent in the process can be calculated through the relation

$$EX_2 = \sum_{i=1}^{\infty} iP_2(1 - P_2)^{i-1} = \frac{1}{P_2} = \frac{(K - 1)^{(1-r)(M'+1)} - (K - 1)^{1-r}}{(K - 1)^{(1-r)(l+1)} - (K - 1)^{1-r}}. \quad (8)$$

Any l -step path between the source and the target must contain at least one long-distance link in the circle. If EX is the total steps for the transmission, $EX \geq EX_2 + l$. With $EX_2 + l = 1$, l_{\min} can be gotten. $EX \geq EX_2 + l_{\min}$. For $(K - 1)^{M'} \sim N$, $EX \sim N^{(1-r)}$, where N is the total number of nodes in the network and $N \rightarrow \infty$.

For $r > 1$, the probability that a node has a long-distance link longer than ε , $0 < \varepsilon < M'$, can be calculated. Then,

$$P_3 = \frac{[1 - (K - 1)^{1-r}] \sum_{i=\varepsilon}^{\infty} [(K - 1)^{1-r}]^i}{(K - 1)^{1-r} - (K - 1)^{(1-r)(M'+1)}} = \frac{(K - 1)^{(1-r)\varepsilon}}{(K - 1)^{1-r} - (K - 1)^{(1-r)(M'+1)}}. \quad (9)$$

When the message holder chooses a long-distance link longer than ε to transfer messages, the expected number of steps spent in the process can be given as follows:

$$EX_3 = \sum_{i=1}^{\infty} iP_3(1 - P_3)^{i-1} = \frac{1}{P_3} = \frac{(K - 1)^{1-r} - (K - 1)^{(1-r)(M'+1)}}{(K - 1)^{(1-r)\varepsilon}}. \quad (10)$$

The number of steps spent in the process to jump a distance larger than ε within γ steps can be calculated by $EX = \gamma EX_3$, where $\varepsilon \gamma \sim M'$ and $N \rightarrow \infty$. With $\gamma_{\min} = 1$, $\varepsilon \sim M'$. $EX \sim N^{(r-1)}$.

The results show that the most efficient navigation is achieved when the power exponent, r , for the random connection is 1 with $N \rightarrow \infty$.

With $r = 0$, the abstract tree model has the uniform distribution over long-distance contacts, which is used in the basic small-world model of J. W. Duncan and H. S. Steven, because the long-distance connections are independent of their position information in the model of J. W. Duncan and H. S. Steven.

5. Epidemics on small worlds

If we ignore the artificial interventions and consider the natural propagation of messages on tree-based networks, the WSN is vulnerable to computer viruses. The high-density deployment of wireless sensors implies that any

virus can be highly contagious. Sensor nodes are severely resource constrained, and lack sophisticated defense mechanisms to fight virus attacks. In view of the wide occurrence of small worlds in tree-based networks, it becomes an interesting issue to inspect the characteristics of networks on dynamics of the epidemic propagation.

Due to shortcuts in small worlds, the epidemic propagation becomes much drastic in sensor networks. In our small-world model, we start with Cayley tree and add shortcuts between pairs of nodes chosen uniformly at random. More than one bond (link) between any two nodes, or a bond which connects a node to itself, is allowed. A certain fraction h of nodes of tree topologies is assumed to be susceptible to the disease, and the bonds between them represent the physical contacts by which a disease can spread. In this part, we assume that each node could be connected with any other node in the network. The epidemic propagation begins with a single infected node. The nodes will be occupied or not depend on whether they are susceptible to the disease. If the distribution of occupied nodes is random, the problem when an epidemic takes place becomes equivalent to a standard percolation problem on the small worlds of tree-based networks. The node is denoted by a site. The percolation probability h_c , at which the outbreak of epidemic takes place, can be calculated. In this paper, the site percolation is only considered.

Some characteristics of Cayley tree have been analyzed in [11], the percolation probability h_c can be calculated by

$$h_c = \frac{1}{K-1}. \tag{11}$$

In a large-scale network, Cayley tree reflects different characteristics with the ring understratum introduced by J. W. Duncan and H. S. Steven. In the proposed small-world model, the probability that two randomly chosen sites have a shortcut between them is calculated as

$$\theta = 1 - \left(1 - \frac{2}{N^2}\right)^{p(N-1)} \approx \frac{2p(N-1)}{N^2} \approx \frac{2p}{N} \tag{12}$$

where the parameter p is defined as the average number of shortcuts per bond on the underlying topology with $r = 0$. Occupied sites connected together will construct local clusters in the small world, and they can be connected together by shortcuts. The average number of local clusters of size i in the network can be derived by

$$X'_i = h^i (1-h)^{2+(K-2)i} N. \tag{13}$$

In order to construct a so-called ‘‘giant component’’ as in the random graph [24], we start with one particular local

cluster, and add all other local clusters to it, which can be reached by traveling along a single shortcut. Then all other local clusters are added to the new ones. This process continues until the connected cluster, the giant component, is constructed.

For calculating the percolation probability h_c , a vector \mathbf{V} is defined at each step in this process, whose component v_i is the probability that a local cluster of size i is added to the overall connected cluster. We define another vector \mathbf{V}' , whose component v'_i can be gotten in terms of the value of \mathbf{V} at the previous step. At or below the percolation threshold the component v_i is small and we can calculate the vector \mathbf{V}' using a transition matrix \mathbf{M} . The following formula reflects the relationship between \mathbf{V} and \mathbf{V}' :

$$v'_i = \sum_j M_{ij} v_j \tag{14}$$

where

$$M_{ij} = X'_i [1 - (1-\theta)^{ij}] \tag{15}$$

in which X'_i is the number of local clusters of size i as before, $[1 - (1-\theta)^{ij}]$ is the probability of a shortcut, which connects a local cluster of size i with one of size j , and there are ij possible pairs of sites by which these can be connected.

The largest eigenvalue λ of the transition matrix \mathbf{M} is considered. For $\lambda < 1$, the vector \mathbf{V} tends to 0 according to (14). The rate at which new local clusters are added falls off exponentially and the connected clusters are finite with an exponential size distribution. Conversely, for $\lambda > 1$, \mathbf{V} keeps growing until the size of the connected cluster becomes limited by the size of the whole system. The percolation threshold occurs at the point $\lambda = 1$.

It is difficult to find the largest eigenvalue of the transition matrix \mathbf{M} for finite N . If p is a constant, θ tends to 0 with $N \rightarrow \infty$. Equation (15) can be simplified by

$$M_{ij} = ij\theta X'_i. \tag{16}$$

If we set $v'_i = \lambda v_i$, (14) is rewritten as

$$\lambda v_i = i\theta X'_i \sum_j j v_j. \tag{17}$$

And v_i can be calculated by

$$v_i = C' \lambda^{-1} i\theta X'_i \tag{18}$$

where $C' = \sum_j j v_j$ is a constant. From (18) we know

$$\begin{aligned} \sum_i v_i &= C' \lambda^{-1} \theta \sum_i i X'_i \\ \sum_i i v_i &= C' \lambda^{-1} \theta \sum_i i^2 X'_i \end{aligned}$$

$$C' = C'\lambda^{-1}\theta \sum_i i^2 X'_i$$

$$\lambda = \theta \sum_i i^2 X'_i. \tag{19}$$

With $K = 2$,

$$X'_i = h^i(1-h)^2 N. \tag{20}$$

This formula is the same as that of the ring understratum proposed by J. W. Duncan and H. S. Steven. In this state, λ can be derived by

$$\lambda = \theta N h \frac{1+h}{1-h} = 2ph \frac{1+h}{1-h}. \tag{21}$$

We set $\lambda = 1$ to get the value of p at the percolation threshold h_c ,

$$p = \frac{1-h_c}{2h_c(1+h_c)}. \tag{22}$$

This special topology with $K = 2$ are found in WSNs. PEGASIS is a well-known network protocol, which constructs a chain in the network [25]. In this chain, each node has two neighbors with $K = 2$. If the end and the head of the chain are connected, it becomes a ring. F. L. Tang et al. proposed a routing algorithm named chain-cluster-based mixed (CCM) routing, which divided a WSN into a few chains and caused a longer delay for data transmission [26]. S. Pal et al. introduced a chain structure among the sensor nodes in every cluster. Cluster heads got accumulated data from this chain and sent them to the nearest base station [27].

With $K = 3$, (13) can be rewritten as

$$X'_i = h^i(1-h)^{2+i} N = (1-h)^2 [h(1-h)]^i N. \tag{23}$$

λ can be calculated by

$$\lambda = \theta N(1-h)^2 \frac{g(1+g)}{(1-g)^3} \tag{24}$$

where $g = h(1-h)$. Then λ can be calculated as

$$\lambda = \theta N(1-h)^2 \frac{h(1-h)[1+h(1-h)]}{[1-h(1-h)]^3} = \theta N \frac{h(1-h)^3 [1+h(1-h)]}{[1-h(1-h)]^3}. \tag{25}$$

We set $\lambda = 1$ to get the value of p at the percolation threshold h_c ,

$$p = \frac{[1-h_c(1-h_c)]^3}{2h_c(1+h_c)^3 [1+h_c(1-h_c)]}. \tag{26}$$

For general K , the average number of local clusters of size i in the network can be given as follows:

$$X'_i = h^i(1-h)^{2+(K-2)i} N = (1-h)^2 [h(1-h)^{K-2}]^i N. \tag{27}$$

λ can be calculated by

$$\lambda = \theta N(1-h)^2 \frac{g'(1+g')}{(1-g')^3} \tag{28}$$

where $g' = h(1-h)^{K-2}$. Then λ can be derived as

$$\lambda = \theta N(1-h)^2 \frac{h(1-h)^{K-2} [1+h(1-h)^{K-2}]}{[1-h(1-h)^{K-2}]^3}. \tag{29}$$

We set $\lambda = 1$ to get the value of p at the percolation threshold h_c ,

$$p = \frac{[1-h_c(1-h_c)^{K-2}]^3}{2h_c(1-h_c)^K [1+h_c(1-h_c)^{K-2}]}. \tag{30}$$

p is a constant in our study, and the percolation threshold h_c for general K can be calculated using (30).

With the vectors and the transition matrix, the percolation threshold h_c is obtained. The largest eigenvalue λ of the transition matrix is considered, and h_c occurs at the point $\lambda = 1$. We can get the same result if we focus on the distribution of local clusters. The quantity $P(n)$ is defined as the probability that a randomly chosen site belongs to a connected cluster of n sites. Then we define

$$H(z) = \sum_{n=0}^{\infty} P(n)z^n. \tag{31}$$

With $h < h_c$, the distribution of clusters falls off exponentially with cluster size. In this state, the probability of two shortcuts connecting the same pair of clusters can be neglected. $H(z)$ satisfies the Dyson equation-like iterative condition. Equation (31) is calculated by

$$H(z) = \sum_{n=0}^{\infty} P_0(n)z^n \sum_{n=0}^{\infty} P(m/n)[H(z)]^m \tag{32}$$

where m is the number of shortcuts, which connects to other clusters. $P_0(n)$ is the probability that a randomly chosen site belonging to a local cluster of size n . We know

$$P_0(n) = \begin{cases} 1-h, & n=0 \\ nh^n(1-h)^{2+(K-2)n}, & n \geq 1 \end{cases}. \tag{33}$$

$P(m/n)$ is the probability which exactly has m shortcuts emerging from a local cluster of size n . There are $2p(N-1)$ ends of shortcuts in the network, $P(m/n)$ can be given as

$$P(m/n) = \binom{2p(N-1)}{m} \left[\frac{n}{N} \right]^m \left[1 - \frac{n}{N} \right]^{2p(N-1)-m}. \tag{34}$$

When N is large enough, (32) is rewritten as

$$\begin{aligned}
 H(z) &= \sum_{n=0}^{\infty} P_0(n)z^n [1 + (H(z) - 1)\frac{n}{N}]^{2p(N-1)} = \\
 &\sum_{n=0}^{\infty} P_0(n)[ze^{\frac{2(N-1)p[H(z)-1]}{N}}]^n \approx \\
 &\sum_{n=0}^{\infty} P_0(n)[ze^{2p[H(z)-1]}]^n. \tag{35}
 \end{aligned}$$

$H_0(z)$ is defined as

$$H_0(z) = \sum_{n=0}^{\infty} P_0(n)z^n. \tag{36}$$

Compared (35) with (36), $H(z)$ can be calculated as follows:

$$H(z) = H_0(ze^{2p[H(z)-1]}). \tag{37}$$

From (31), the mean outbreak size can be gotten by the first derivative of H , and

$$\langle n \rangle = H'(1). \tag{38}$$

From (37), $H'(z)$ is calculated as

$$\begin{aligned}
 H'(z) &= H'_0(ze^{2p[H(z)-1]})(ze^{2p[H(z)-1]})' = \\
 &e^{2p[H(z)-1]}H'_0(ze^{2p[H(z)-1]})[1 + 2pzH'(z)]. \tag{39}
 \end{aligned}$$

We know from (31) that $H(1) = 1$, and $H'(1)$ can be derived by

$$\langle n \rangle = H'(1) = \frac{H'_0(1)}{1 - 2pH'_0(1)}. \tag{40}$$

From (34) and (36), $H_0(z)$ can be obtained as

$$H_0(z) = 1 - h + \frac{h(1-h)^K z}{[1 - h(1-h)^{K-2}z]^2}. \tag{41}$$

The first derivative of $H_0(z)$ is calculated by

$$H'_0(z) = \frac{h(1-h)^K [1 + h(1-h)^{K-2}z]}{[1 - h(1-h)^{K-2}z]^3}. \tag{42}$$

With $z = 1$, $H'_0(1)$ is given as

$$H'_0(1) = \frac{h(1-h)^K [1 + h(1-h)^{K-2}]}{[1 - h(1-h)^{K-2}]^3}. \tag{43}$$

The value of $\langle n \rangle$ can be derived by

$$\begin{aligned}
 \langle n \rangle &= \frac{H'_0(1)}{1 - 2pH'_0(1)} = \\
 &\frac{h(1-h)^K [1 + h(1-h)^{K-2}]}{[1 - h(1-h)^{K-2}]^3 - 2ph(1-h)^K [1 + h(1-h)^{K-2}]}. \tag{44}
 \end{aligned}$$

The mean outbreak size diverges at the percolation threshold h_c . This threshold marks the onset of epidemic behavior and occurs at the zero of the denominator of (44). The value of p at the percolation threshold h_c can be obtained.

We have

$$p = \frac{[1 - h_c(1 - h_c)^{K-2}]^3}{2h_c(1 - h_c)^K [1 + h_c(1 - h_c)^{K-2}]} \tag{45}$$

which agrees with (30).

Now we consider $h > h_c$, and there is a giant component of connected nodes with a large number of smaller clusters whose distribution falls off exponentially with cluster size. $P(n)$ is redefined to be the probability that a randomly chosen site belongs to a cluster of size n which is not part of the giant components. The volume of the giant component is $x = 1 - H(1)$, and $P(n)$ sums not to 1 now. From (37), x is derived by

$$x = 1 - H_0(e^{-2px}). \tag{46}$$

The derivatives of both sides is calculated as

$$1 = 2pH'_0(e^{-2px})e^{-2px}. \tag{47}$$

From (42), (47) can be calculated by

$$1 = 2pe^{-2px} \frac{h(1-h)^K [1 + h(1-h)^{K-2}e^{-2px}]}{[1 - h(1-h)^{K-2}e^{-2px}]^3}. \tag{48}$$

With $x = 0$, all the values of h are suitable. For $x < 0$, it is unphysical. The threshold is gotten with $x = 0$. From (48), the value of p at the percolation threshold h_c can be derived. We have

$$p = \frac{[1 - h_c(1 - h_c)^{K-2}]^3}{2h_c(1 - h_c)^K [1 + h_c(1 - h_c)^{K-2}]}. \tag{49}$$

The result agrees with (30) and (45).

Above analyses present the mathematical method to calculate the percolation threshold in our small-world model relying on the percolation theory, although real network characteristics are complex. The percolation threshold provides a standard for our judgment to protect tree-based WSNs. If the infection probability is larger than the percolation threshold, the epidemic has an exponential increase with time, which spreads rapidly in the WSN with no suitable immunization strategies similar with the development of fads. In fact, this infection is a susceptible-infected (SI) process until all nodes are infected in the WSN.

6. Experiments and simulations

The percolation probability, at which the outbreak of the epidemic takes place, is observed. We know that the percolation threshold of Cayley tree is 1 without shortcuts [11]. If the random link addition probability p increases from 0

to 1, $h_c = \frac{\sqrt{(2p+1)^2 + 8p} - (2p+1)}{4p}$, when the tree topology transforms into a chain structure with $K = 2$. As shown in Fig. 2, the percolation threshold h_c keeps decreasing when the random link addition probability p increases. $h_c = 0.67$ with $p = 0.15$, $h_c = 0.61$ with $p = 0.2$, $h_c = 0.52$ with $p = 0.3$, and $h_c = 0.41$ with $p = 0.5$.

This result coincides with the reality. In a regular tree topology, the epidemic propagates along the regular bonds with a large percolation threshold. Due to shortcuts in small worlds, the epidemic propagation becomes much drastic in sensor networks when the random link addition probability p increases from 0. The percolation threshold decreases at the same time. In an entire random network, the percolation threshold is small and viruses easily attack the network from one side to another. For $K > 2$, it is more difficult to solve (49), but the variety is similar with that with $K = 2$. When the random link addition probability p increases, the percolation threshold h_c keeps decreasing. At $p = 1$, there exists the smallest percolation threshold.

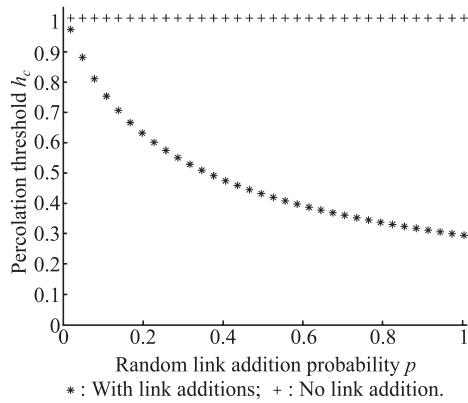


Fig. 2 Percolation threshold h_c with $K = 2$

In the following simulations, epidemiological processes are observed in Figs. 3–5 when small world phenomena occur with $h > h_c$. Fig. 3 shows the time evolution of the infected number in the epidemiological process on a small world of the SCT. There is one infected node in the initial stage. Say that the node i is susceptible, and it has k_i neighbors, of which k_{inf} are infected. Then, i will become infected with probability k_{inf}/k_i . In the simulation, some parameters are set with $p = 0.3$ and $N = 2000$. For the theoretical parameters based on [5] are used in Fig. 3, the infection extends exponentially in most time on the small world of the SCT. In the late stage of the epidemic, the exponential evolution process experiences a decline for the reduction of remaining susceptible nodes in the network. The infected number keeps increasing until all nodes are infected in the network. Fig. 4–5 show epidemiological processes

on the proposed small world of Cayley tree according to the mathematical analyses. The simulations show that the infection extends exponentially although different propagations occur on the underlying tree-based topology, which coincides with the real epidemic on the small world of the SCT.

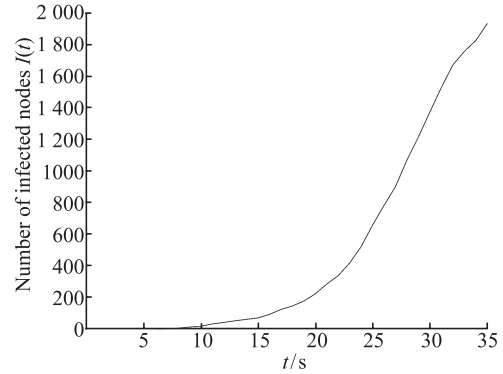


Fig. 3 Epidemic on the small world of the SCT

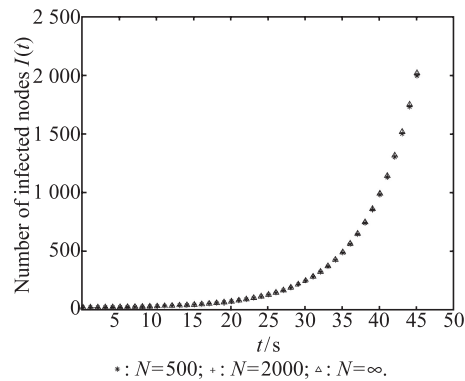


Fig. 4 Epidemic process with a certain number of infected nodes

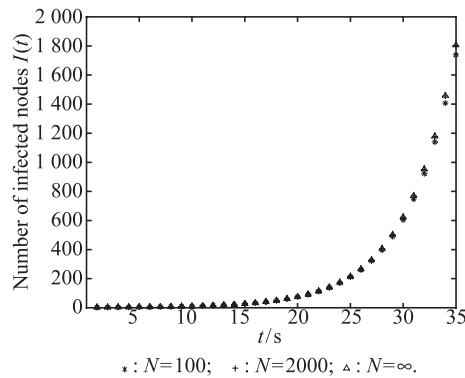


Fig. 5 Epidemic process with all infected children nodes

In the following simulation, a certain number of nodes, C_K , are infected on Cayley tree in each time unit. Certain parameters are set with $p = 0.3$, $h = 0.8$ and $C_K = 3$. Fig. 4 shows the time evolution of the infected number in the epidemiological process with $N = 500, 2000$ and ∞ . Each time unit includes 10 s.

When the infected parent node attacks almost all its children nodes in each time unit, the epidemic spreads drastically in the network. Fig. 5 shows the time evolution of

the infected number in the epidemiological process with $N = 100, 2\ 000$ and ∞ . Certain parameters are set with $p = 0.3, h = 0.8$, and $K = 3$ in the simulation. Each time unit includes 10 s.

Above simulations show that the infection extends exponentially with time t on the small world of Cayley tree. Fig. 4 and Fig. 5 describe ideal propagations of viruses, which coincide with the epidemic propagation in the small world of the SCT. The simulations reflect the basic prevalence characteristics of epidemics in the small worlds of tree-based WSNs. From the figures we can see that the prevalence is so drastic that immunizations attract much attention for protecting the networks from attacks.

7. Conclusion and future work

For small world phenomena in wireless networks, dynamics become much more complex in tree-based WSNs. An abstract tree model, Cayley tree, is considered, and the random link additions are conducted on it to construct a small-world model. A common decentralized navigation algorithm is considered, which runs automatically for searching short paths in the small worlds of tree-based WSNs. The expected delivery time of a decentralized algorithm is defined to be the expected number of steps taken by the algorithm to deliver messages over a small world network. The most efficient navigation is achieved when the power exponent, r , is 1 for the random connection. In the further study, the epidemic on the small worlds of tree topologies on large scales will be studied and the percolation probability, at which the outbreak of the epidemic takes place, will be calculated. Compared with Cayley tree, the small world has a smaller percolation threshold suffering from the epidemic. In the future, we will pay attention to the immunization of tree-based WSNs.

References

- [1] C. Alippi, R. Camplani, C. Galperti, et al. A robust, adaptive, solar-powered WSN framework for aquatic environmental monitoring. *IEEE Sensors Journal*, 2011, 11(1): 45–55.
- [2] X. Wang, S. Wang, D. W. Bi. Distributed visual-target-surveillance system in wireless sensor networks. *IEEE Trans. on Systems, Man, and Cybernetics–Part B: Cybernetics*, 2009, 39(5): 1134–1146.
- [3] M. Demirbas, X. M. Lu, P. Singla. An in-network querying framework for wireless sensor networks. *IEEE Trans. on Parallel and Distributed Systems*, 2009, 20(8): 1202–1215.
- [4] Y. C. Wong, J. T. Wang, N. H. Chang, et al. Hybrid address configuration for tree-based wireless sensor networks. *IEEE Communications Letters*, 2008, 12(6): 414–416.
- [5] Y. J. Zhu, R. Vedantham, S. J. Park, et al. A scalable correlation aware aggregation strategy for wireless sensor networks. *Information Fusion*, 2008, 9(3): 354–369.
- [6] J. A. Sanchez, P. M. Ruiz. Energy-efficient geographic multicast routing for error-prone wireless sensor networks. *Wireless Communications and Mobile Computing*, 2009, 9(3): 395–404.
- [7] H. Shen. Finding the k most vital edges with respect to minimum spanning tree. *Proc. of the IEEE National Aerospace and Electronics Conference*, 1997: 255–262.
- [8] S. D. Muruganathan, D. C. F. Ma, R. I. Bhasin, et al. A centralized energy-efficient routing protocol for wireless sensor networks. *IEEE Communication Magazine*, 2005, 43(3): S8–13.
- [9] W. Wang, B. W. Wang, Z. Liu, et al. A cluster-based and tree-based power efficient data collection and aggregation protocol for wireless sensor networks. *Information Technology Journal*, 2011, 10(3): 557–564.
- [10] M. Stanley. The small world problem. *Psychology Today*, 1967, 1(1): 60–67.
- [11] S. Dietrich, A. Ammon. *Introduction to percolation theory*. Boca Raton: CRC Press, 1992.
- [12] J. W. Duncan. *Small worlds, the dynamics of networks between order and randomness*. New Jersey: Princeton University Press, 1999.
- [13] A. Helmy. Small worlds in wireless networks. *IEEE Communications Letters*, 2003, 7: 490–492.
- [14] K. Jon. Navigation in a small world. *Nature*, 2000, 406: 845.
- [15] M. R. Roberson, B. A. Daniel. Kleinberg navigation in fractal small-world networks. *Physical Review E*, 2006, 74(1): 017101-1–3.
- [16] S. A. Hill, D. Braha. Dynamic model of time-dependent complex networks. *Physical Review E*, 2010, 8(4): 046105-1–7.
- [17] Y. Yang, S. C. Zhu, G. H. Cao. Improving sensor network immunity under worm attack: a software diversity approach. *Proc. of the 9th ACM international symposium on Mobile Ad Hoc Networking and Computing*, 2008: 149–158.
- [18] P. De, Y. Liu, S. K. Das. Modeling node compromise spread in wireless sensor networks using epidemic theory. *Proc. of the International Symposium on World of Wireless, Mobile and Multimedia Networks*, 2006: 237–243.
- [19] M. E. J. Newman, J. W. Duncan. Scaling and percolation in the small-world network model. *Physical Review E*, 1999, 60(6): 7332–7342.
- [20] R. Pastor-Satorras, A. Vespignani. Epidemic dynamics and endemic states in complex networks. *Physical Review E*, 2001, 63(6): 066117-1–8.
- [21] T. J. Norman. *The mathematical theory of infectious diseases*. New York: Hafner Press, 1975.
- [22] M. M. Telo da Gama, A. Nunes. Epidemics in small world networks. *The European Physical Journal B*, 2006, 50(112): 205–208.
- [23] E. S. Thomas, M. J. Matthew, R. M. Susan. Comparative effects of avoidance and vaccination in disease spread on a dynamic small-world network. *Physica A: Statistical Mechanics and its Applications*, 2010, 389(23): 5515–5520.
- [24] N. Alon, J. H. Spencer. *The probabilistic method*. 2nd ed. Hoboken, NJ: Wiley, 2000.
- [25] S. Lindsey, C. Raqhavendra, K. M. Sivalingam. Data gathering algorithms in sensor networks using energy metrics. *IEEE Trans. on Parallel and Distributed Systems*, 2002, 13(9): 924–935.
- [26] F. L. Tang, L. You, S. Guo, et al. A chain-cluster based routing algorithm for wireless sensor networks. *Journal of Intelligent Manufacturing*, 2010: 10.1007/s10845-010-0413-4.
- [27] S. Pal, B. Debnath, T. H. Kim. Chain based hierarchical routing protocol for wireless sensor networks. *Communications in Computer and Information Science*, 2010, 78: 482–492.

Biographies



Qiao Li was born in 1980. He is now a lecturer of the Artillery Academy of the PLA. He joined the System Engineering Laboratory of Beijing Institute of Technology to pursue the M.S. degree in 2006. And he is pursuing the Ph.D. degree in the School of Automation from Beijing Institute of Technology, Beijing, China. He partly takes on researches of the Chinese Defence Advance

Research Program of Science and Technology and the Pre-Research Foundation of Chinese Ordnance Industry Group from 2009. His research interests include theories and applications of wireless sensor networks, particularly in issues on epidemic dynamics, routing protocols, and topology control algorithms.

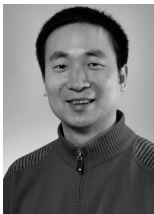
E-mail: liqq007@gmail.com



Baihai Zhang received his B.S., M.S., and Ph.D. degrees from Harbin Institute of Technology, China, in 1988, 1991 and 1994, respectively. From November 1994 to May 1997, he was a post-doctoral research fellow in Beijing Institute of Technology, China. From April 2001 to October 2001, he studied mechatronic systems modeling by bond graphs as a visiting scholar at the

Department of Mechanical Engineering, Michigan State University. From January 2006 to April 2006, he studied systems identification as a visiting scholar at the Faculty of Computing, Engineering and Mathematical Sciences (CEMS), University of the West of England. He is now a professor in the School of Automation, Beijing Institute of Technology. His research interests include theories and applications of systems engineering, evolutionary computation, and wireless sensor networks.

E-mail: smczhang@bit.edu.cn



Zhun Fan received his B.S. degree (control engineering) in 1995 and M.S. degree (control engineering) in 2000 from Huazhong University of Science and Technology, China. He received the Ph.D. (electrical and computer engineering) in 2004 from the Michigan State University. From 2004 to 2007, he was an assistant professor in the Department of Mechanical Engineering at

the Technical University of Denmark. He is currently an associate professor in the Department of Management Engineering at the Technical University of Denmark. His research interests include applying evolutionary computation and computational intelligence in design automation and optimization of MEMS, mechatronics, and robotic systems, and intelligent communication, transportation and power system.

E-mail: zhfa@man.dtu.dk



Athanasios V. Vasilakos is currently a professor at the Department of Computer and Telecommunications Engineering, University of Western Macedonia, Greece and visiting professor at the Graduate Programme of the Department of Electrical and Computer Engineering, National Technical University of Athens (NTUA), Greece. His research interests include computer networks, mobile computing, wireless communications, game theory, artificial ubtekkugebte, bioinformatics, digital arts.

He is coauthor of the books *Computational Intelligence in Telecommunications Networks*, *Ambient Intelligence*, *Wireless Networking*, *Ubiquitous Computing*, *Autonomic Communications*, *Digital Arts*, *Delay Tolerant Networks: Protocols and Applications*. He has published more than 200 publications in top international journals (i.e., IEEE/ACM Transactions on Networking, IEEE T-Information Theory, IEEE JSAC, IEEE Transactions on Communications, IEEE Transactions on Wireless Communications, IEEE Transactions on Mobile Computing, ACM Transactions on Autonomous and Adaptive Systems, IEEE Transactions on Neural Networks, IEEE Transactions on Systems, Man, and Cybernetics, IEEE T-ITB, etc) and conferences in the area of computer networks, mobile computing, wireless networking, game theory, evolutionary game theory, bioinformatics, wireless healthcare, digital and internet arts. He is the editor-in-chief of the inderscience publishers journals: International Journal of Adaptive and Autonomous Communications Systems, International Journal of Arts and Technology. He was or he is at the editorial board of several international journals: IEEE Communications Magazine, IEEE Transactions on Systems, Man and Cybernetics, IEEE Transactions on Information Theory in Biomedicine, IEEE Transactions on Wireless Communications, ACM/Springer Wireless Networks, Wireless Communications and Mobile Computing, EURASIP journal on Wireless and Communications Networks (WCN), Computer Communications, International Journal of Ad Hoc and Ubiquitous Computing, ACM Computers in Entertainment, International Journal of Mobile Communications, International Journal of Internet Protocol Technology, Cluster Computing, Security and Communications Journal, Journal of Sensors, Journal of Supercomputing, Telecommunications Journal, etc. He is chairman of the Telecommunications Task Force of the Intelligent Systems Applications Technical Committee (ISATC) of the IEEE Computational Intelligence Society (CIS).

E-mail: vasilako@ath.forthnet.gr