Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2015, Article ID 690482, 11 pages
http://dx.doi.org/10.1155/2015/690482

Research Article

TR-SDTN: Trust Based Efficient and Scalable Routing in Hostile Social DTNs

Zhenjing Zhang,[1] Maode Ma,[2] and Zhigang Jin[3]

[1]*School of Computer Science and Technology, Tianjin University, Tianjin 300072, China*
[2]*School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798*
[3]*School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China*

Correspondence should be addressed to Maode Ma; emdma@ntu.edu.sg

Received 29 July 2014; Revised 1 November 2014; Accepted 1 November 2014

Academic Editor: Muhammad Khurram Khan

A delay tolerant network (DTN) presents a new communication model, which uses a store-carry-forward approach to deliver messages due to the nature of the network. With the development of the mobile internet, the research on DTNs becomes a hot topic. One critical issue in the DTNs with social characteristics is that there could exist selfish and malicious nodes which block important data and as a result intentionally disturb the communication. Although there are many research works which address the issue and adopt various trust models to encourage nodes to forward information, most of them have not considered the social characteristics of the mobile nodes and they have not handled the selfish behaviors and attacks in a systematic way. In view of these defects, we propose a novel solution to address the selfish and security issues in social DTNs. Based on nodes' social characteristics, a dynamic trust model is proposed to prevent bad-mouthing and ballot stuffing attacks and the Shannon entropy function is introduced to avoid blackhole and greyhole attacks. The simulation results show that our proposed scheme can significantly improve the performance of social DTNs. And the simulation is proved to be consistent with the results from theoretical analysis.

1. Introduction

With the development of mobile internet, wireless communication via mobile devices over a typical delay tolerant network (DTN) becomes a hot research topic. DTNs hold a new network architecture which takes use of a store-carry-forward communication model to forward messages because there is no persistent end-to-end path from the source to the destination. A lot of research works have been done on the efficient routing and transmission of messages in a DTN environment.

Pioneer studies on the routing in DTNs have mainly focused on the actions for the next hop transmission with consideration of historical information. Typical protocols include epidemic [1], prioritized epidemic (PREP) [2], probabilistic routing protocol using history of encounters and transitivity (PROPHET) [3], and some other improvements and variations proposed. As an example, the PROPHET is a routing mechanism based on historical information, which is the history of encounters and transitivity to select next hop for the message forwarding. And the transmit predictability $P_{(a,b)} \in (0, 1]$ has been adopted as the probability metrics. The main characteristic of PROPHET is that node i will forward a message to node j if node j has a higher predicted probability to the destination of the message than that of node i.

Although the protocols with history information work well to be able to achieve high message delivery ratio, messages are delivered with high latency. To reduce the delay, other properties such as mobility models and the relationship among mobile nodes have been considered to make routing decisions. In some particular DTN environments, clustering with hierarchical structures has been proposed to reduce the end-to-end delay. In [4], a hierarchical forwarding mechanism has been proposed to group the nodes according to their encounter frequency. Initially, each node is considered as a cluster consisting of a single node. And the operation which combines the two best clusters to form a new cluster, determined by a distance function, is repeated until the

cluster including all nodes is finally formed. Similar works can be found in [5, 6]. By these protocols, the mobile nodes with frequent contact and the mobile nodes with less contact will be considered differently. By these mechanisms, the efficiency of message transmission can be improved. However, the destination of the messages may not be able to receive all the messages due to the unreliability of the wireless communication channels.

By the above mentioned routing schemes, good performance may not be obtained in some DTNs which have the characteristics of social networks (social DTNs). Existing studies show that, in such scenarios, there are some active nodes which can transmit messages to their destinations with less hops. For example, in a campus scenario, students in the same group communicate with each other frequently while students in different groups have less contact. But a group leader has more contacts among different groups. Some new schemes have been proposed to address routing issue in such DTNs. In [7], according to the small world theory, a routing algorithm has been proposed to combine the similarity and centrality, where the similarity refers to the number of the same neighbors of two nodes while the ratio of the number of the shortest paths including a node over the number of all the shortest paths is defined as the centrality of the node. Additionally, in [8] in accordance with the two important characteristics of a social network, community and centrality, the author has proposed a forwarding algorithm, by which messages will be constantly forwarded to the nodes with the higher centrality because these nodes will have higher probability to meet the destination node. More social characteristics such as the social distance defined in [9], the asynchronous centrality defined in [10], the node's social relation in [11], and the impact of strangers in [12] can be employed to make the forwarding decisions.

However, in social DTNs, there could exist some selfish or malicious nodes. They are unwilling to forward messages from strangers or impair communications by launching attacks. In [13], a dynamic trust management model has been used to select next carrier to avoid messages being forwarded to malicious nodes. By analyzing real traces, in [14], a secure friend discovery scheme has been proposed by identifying potential attacks. In [15], based on the social relation of nodes, a protocol for avoiding the internet threats caused by social selfishness has been proposed. In [16], the authors have pointed out that nodes may be selfish and may not be cooperative for packet forwarding and current research mainly focuses on encouraging nodes to participate in packet forwarding. Similarly, in [17], based on node's own trust relationships, a new scheme against social selfishness has been proposed. What is more, in [18, 19], different mechanisms have been proposed to handle the selfish behaviors of the mobile nodes in the networks. And in [20], a literature survey has presented a summary on the recent routing schemes which take advantages of the positive social characteristics or overcome the negative social characteristics. In addition, there are some research solutions against malicious attacks such as the solution against the blackhole attacks in [21]. From those recent research works, we can find that most of them have not made full use of the nodes' positive social characteristics to overcome the selfish behaviors and prevent malicious attacks at the same time. They have only considered one aspect of the selfish behaviors or attacks. In general, considering the characteristics of DTNs and the properties of social networks, to avoid various threats from the selfishness of the mobile nodes and the attacks from the malicious nodes in the social DTN environments, we propose a novel routing scheme, named as trust routing in hostile social DTNs (TR-SDTN), which divides mobile nodes into different clusters according to the contact frequency. Since the nodes communicate with each other frequently in a cluster, it is assumed that nodes in the same cluster are friends. Selfish nodes and malicious nodes are only considered existing in different clusters. The main contribution of our work is that a direct trust model, an indirect trust model, and the Shannon entropy function are used to prevent the threats from selfish nodes and attacks from malicious nodes.

The remainder of this paper is organized as follows. In Section 2, the system model is introduced. The proposed routing mechanism is presented in Section 3. And in Section 4, the performance of the solution is evaluated by a mathematical model and simulation experiments. Finally, the conclusion of the paper is summarized in Section 5.

2. System Model

2.1. Characteristics of DTNs and Social Networks. The communication in DTNs is a type of asynchronous communication without an end-to-end path at any moment. One of its main characteristics is that the links between mobile nodes are volatile and may break down for a long period of time at any moment, which makes the DTNs always suffer from partitioning from time to time. There are ubiquitous scenarios of the DTNs including mobile sensor networks, disaster recovery, military deployment, and interstellar communications to hold those characteristics. And there are some scenarios to have the DTNs formed by several clusters. In such scenarios, the mobile nodes in the same cluster communicate with each other frequently and several active nodes will be able to communicate with the nodes in different clusters. These active mobile nodes show same characteristics as the nodes in the social networks. Therefore, the characteristics of social networks can be borrowed to design the routing protocols in those scenarios.

A social network is one type of network formed by the corporations of people. Due to different characteristics of different people in nationality, environment, interests, gender, age, and so forth, different communities could be established by the people with same characteristics. One person is more likely to only interact with another one in the same community than a randomly selected person. On the other hand, within a community, there are some individuals who are active in communication with other people. In addition to contacting frequently with the persons in the same community, they also have more contacts with others in other communities. These facts can be further exploited to design the routing protocols for the DTNs with the social characteristics. However, the assumption that all nodes are cooperative is unrealistic because in a social network there

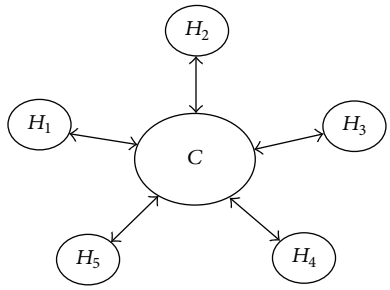


Figure 1: Mobility model.

exist some nodes which are selfish. They are unwilling to forward messages from others. And there also exist malicious nodes which can launch a variety of attacks to impair normal communications.

2.2. System Model. In this paper, the system under the study is a social DTN, in which each mobile node is in continuous movement. Some nodes move in a small range at low speeds and they communicate with each other frequently in the range. And some other nodes move frequently over a larger area at high speeds. Due to the movement of the nodes, temporary clusters can be formed from time to time and there exist some mobile nodes which can be used to transmit messages among clusters. Therefore, in our system, clustering technology can be used with different forwarding mechanisms used for the intracluster and intercluster communications. Since in intracluster, nodes communicate with each other frequently, we assume that they are cooperative with each other as friends with low probability to behave selfishly and maliciously. On the contrary, selfishness and attacks must be considered in intercluster communication. So, in this paper, our solution will target preventing the selfishness and malicious attacks appearing only in the intercluster communication.

On the mobility of a node in the system, we have the following mobility model to describe movement of the nodes in the system. To simplify the scenario, we adopt the mobility model in [6] with some modifications. In the system, there are five hot spots and one cold spot denoted by H_1–H_5 and C as shown in Figure 1. Each mobile node is assumed to have a "home" hot spot where it stays most of time. And each mobile node always goes to the cold spot when it leaves a hot spot. It is further assumed that 20% of the nodes will take more time to move among different spots. As shown in Figure 2, when a node is home, it will have a probability of P_H to stay or $1 - P_H$ to move to the cold spot in the next time slot. While at the cold spot, it will go home with probability of P_H, or have a probability of P_C to stay, or move to other hot spots with probability of $1 - P_H - P_C$ ($P_H + P_C < 1$). Finally, when the node stays in a hot spot which is not its home spot, it will have a probability of P_H to move to the cold spot or stay with probability of $1 - P_H$.

In the model, each node will have a unique ID and maintain the contact information with others by a list of parameters including node ID, contact probability, contact time, and trust value (direct trust value and indirect trust value) for other nodes (see Figure 1).

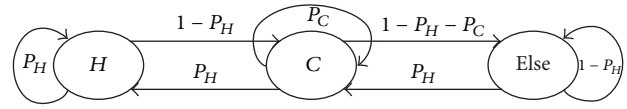


Figure 2: Change of the state of a node.

2.3. Attack Model. As mentioned in Section 2.1, in the social DTNs, there exist selfish and malicious nodes which block the important data and as a result intentionally disturb the communication over the networks. The most common attacks are blackhole attacks, greyhole attacks, bad-mouthing attacks, and ballot stuffing attacks. The blackhole attackers try to promote their importance by providing good recommendations on themselves. Subsequently, they will drop all the packets received. Instead of dropping all the packets, only partial droppings occur by the greyhole attacks which are more difficult to detect. And by the bad-mouthing attacks, a malicious node ruins the reputation of the well-behaved nodes by providing bad recommendations against those good nodes, while by the ballot stuffing attacks a malicious node provides good recommendations on the bad nodes so as to increase the opportunity of the message dropping. To some extent, these two attacks can make the blackhole and greyhole attacks more effectively. In this paper, it is assumed that malicious nodes will present the following behaviors.

(1) *Behavior 1.* A malicious node will provide good recommendations on itself and the other malicious nodes to the nodes it meets. In this way, the importance of these malicious nodes will be promoted.

(2) *Behavior 2.* To reduce the importance of the good nodes, a malicious node will provide bad recommendations on the good nodes to the nodes it meets.

(3) *Behavior 3.* A malicious node will drop packets it receives according to the dropping probability. If the dropping probability is set to 1, the node is considered as a blackhole attacker. Otherwise, if the dropping probability is between 0 and 1, the node is considered as a greyhole attacker.

(4) *Behavior 4.* In order to increase the concealment, a malicious node could receive packets and forward packets like a normal node.

3. Routing Mechanism with Security Function

The nodes are clustered with different forwarding mechanisms used for intracluster and intercluster routing. By clustering technique, each node will belong to a cluster with a cluster ID and maintain the information of the members in the cluster by a list of parameters such as the node ID. Since the nodes communicate with each other frequently in one cluster, it is assumed that the nodes in the same cluster are friends. The selfish nodes and malicious nodes are only considered possibly existing in other clusters. For intracluster

routing, the mechanism "spray and wait" is adopted and there are two copies for each message. For intercluster routing, the Shannon entropy function and a trust model have been used to select the next carrier and avoid attacks from malicious nodes. In this paper, our focus is on the intercluster routing with prevention of the selfish behaviors and malicious attacks. Since the trust value and Shannon entropy function have been used in intercluster routing, we first present the way to compute the trust value and the entropy function. And then the proposed intercluster routing scheme will be introduced.

3.1. Calculation of Trust Value. The trust value between node i and node j at time t is denoted by $T^{ij}(t)$, which is computed as follows:

$$T^{ij}(t) = \sum_{X}^{\text{all}} w^x * T_X^{ij}(t), \tag{1}$$

where X represents a trust property with the two attributes of selfishness and centrality, $T_X^{ij}(t)$ is node i's trust value in the trust property X toward node j, and w^x is the weight of the trust property X.

Node i updates its trust value toward node j in the trust property X when encountering a node as follows, assuming the current time is t and Δt is the encounter interval:

$$\begin{aligned} T_X^{ij}(t+\Delta t) = {} & \beta T^{ij}_{\text{direct},X}(t+\Delta t) \\ & + (1-\beta)\, T^{ij}_{\text{indirect},X}(t+\Delta t). \end{aligned} \tag{2}$$

When node i encounters the node k, the trust value will be updated. There are two cases which need to be considered. For example, node i updates its trust value towards node j: $k = j$ and $k \neq j$.

(i) When $k = j$, there are two trust properties used, selfishness and centrality.

$T^{ij}_{\text{direct, Selfishness}}(t+\Delta t)$: in social DTNs, some nodes are selfish which are not willing to forward messages from a stranger. However, friends will be cooperative toward each other even if they are selfish. By proposed protocol, each node will have a friend list, denoted by F_i for node i. When node i receives a message from node j, node i will check whether node j is in its friend list. If node j belongs to F_i, then $T^{ij}_{\text{direct, Selfishness}}(t+\Delta t)$ will be valued to 1. Otherwise, it is assumed that nodes with similar attributes with values computed from a list $(a_1, a_2, a_3, a_4, a_5)$ will be easier to become friends. And in each attribute there are five selection froms:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{15} \\ a_{21} & a_{22} & \cdots & a_{25} \\ \vdots & \vdots & & \vdots \\ a_{51} & a_{52} & \cdots & a_{55} \end{bmatrix}. \tag{3}$$

If node i and node j have m same attributes, then $T^{ij}_{\text{direct, Selfishness}}(t+\Delta t) = 0.2m$. To avoid forged information from malicious nodes, an encounter history will be stored. If the information of a node is changed frequently, it will be marked as a malicious node.

$T^{ij}_{\text{direct, Centrality}}(t+\Delta t)$: in social DTNs, nodes with high centrality will have higher probability to deliver messages. By the proposed trust protocol, nodes with higher centrality will have higher trust value towards trust property of centrality. The centrality of node i, denoted by CT_i, is defined as the number of different clusters that the node i meets per unit time. The trust value on the centrality can be computed as

$$T^{ij}_{\text{direct, Centrality}}(t+\Delta t) = \frac{CT_j}{CT_i + CT_j}. \tag{4}$$

In this case, these is no indirect trust; node i updates $T^{ij}_{\text{indirect},X}(t+\Delta t)$ as follows:

$$T^{ij}_{\text{indirect},X}(t+\Delta t) = e^{-\lambda_d \Delta t} * T^{ij}_{\text{indirect},X}(t). \tag{5}$$

In this way, the indirect trust value will be delayed if there is not new information from other nodes.

(ii) When $k \neq j$, to calculate the indirect trust value, when node i meets node k, node i uses its one-hop neighbors including node k as recommenders to update $T^{ij}_{\text{indirect},X}(t+\Delta t)$ and to prevent bad-mouthing and ballot stuffing attacks, for any neighbor node a, $T^{ia}(t) \geq T^{\text{rec}}$, where T^{rec} is a threshold to indicate whether a node is trustworthy. The indirect value can be calculated as follows:

$$\begin{aligned} & T^{ij}_{\text{indirect},X}(t+\Delta t) \\ & = \begin{cases} e^{-\lambda_d \Delta t} * T^{ij}_{\text{indirect},X}(t) & |R_i| = 0 \\ \dfrac{\sum_{k \epsilon R_i} \left(T_X^{ik}(t) * T_X^{kj}(t) \right)}{\sum_{k \epsilon R_i} T_X^{ik}(t)} & |R_i| \neq 0, \end{cases} \end{aligned} \tag{6}$$

where R_i is a set including node i's trustworthy neighbors and $|R_i|$ is the number of node i's trustworthy neighbors.

Since there is no direct trust, the direct trust will be delayed:

$$T^{ij}_{\text{direct},X}(t+\Delta t) = e^{-\lambda_d \Delta t} * T^{ij}_{\text{direct},X}(t). \tag{7}$$

3.2. Calculation of the Entropy Function. The entropy function is defined as

$$H(p) = -p \log_2 p - (1-p) \log_2 (1-p), \tag{8}$$

where p is the probability of a node forwarding a packet and $H_b(p)$ is the binary entropy function related to probability p.

According to formula (8), we define our trust value for node i as follows:

$$T^i_{\text{Entropy-Function}} = \begin{cases} 1 - H(p) & 0.5 \leq p \leq 1 \\ H(p) - 1 & 0 \leq p < 0.5. \end{cases} \tag{9}$$

In this way, $T^i_{\text{Entropy-Function}}$ is an increasing function with probability p. If a malicious node drops packets, it will have lower probability p and then have lower $T^i_{\text{Entropy-Function}}$.

The probability p is quantified based on the observations of a nodes' forwarding behavior. Assume that the distribution to observe this process follows the binomial distribution. Then, the probability of a node forwarding k packets out of n packets can be computed as follows:

$$p(t_n = k \mid \tau) = \binom{n}{k} \tau^k (1-\tau)^{n-k}, \tag{10}$$

where τ is the probability of performing forwarding, t_n represents the number packets forwarded.

According to Bayesian approach, we can get

$$p(t_{n+1} = k+1 \mid t_n = k) = \frac{p(t_{n+1} = k+1, t_n = k)}{p(t_n = k)}, \tag{11}$$

where

$$\begin{aligned} p(t_n = k) &= \int_0^1 p(t_n = k \mid \tau)\, d\tau, \\ p(t_{n+1} = k+1, t_n = k) &= \int_0^1 \tau * p(t_n = k \mid \tau)\, d\tau. \end{aligned} \tag{12}$$

Then we can compute

$$\begin{aligned} &p(t_{n+1} = k+1 \mid t_n = k) \\ &\quad = \frac{\int_0^1 \tau * p(t_n = k \mid \tau)\, d\tau}{\int_0^1 p(t_n = k \mid \tau)\, d\tau} = \frac{k+1}{n+2}. \end{aligned} \tag{13}$$

Therefore, probability p can be expressed as $(k+1)/(n+2)$.

3.3. Routing Mechanism. As shown in Algorithm 1, when node i encounters node j, the $T^j_{\text{Entropy-Function}}$ will be computed first. A threshold $T^{\text{threshold}}_{\text{Entropy-Function}}$ is defined to identify whether node j is a malicious node. If $T^j_{\text{Entropy-Function}} \geq T^{\text{threshold}}_{\text{Entropy-Function}}$, node j is admitted as a trustworthy node. And then compute the trust value $T^{ij}(t + \Delta t)$. There is another threshold of $T^{\text{threshold}}$. If $T^{ij}(t + \Delta t) \geq T^{\text{threshold}}$, the messages will be forwarded from node i to node j. Using the Shannon entropy function, the blackhole and greyhole attacks can be prevented since in these attacks more packets are dropped than the normal nodes and then the value $T^j_{\text{Entropy-Function}}$ computed will be smaller than the threshold $T^{\text{threshold}}_{\text{Entropy-Function}}$. By using the direct and indirect trust values, selfish behaviors, bad-mouthing, and ballot stuffing attacks can also be avoided because messages will not be forwarded to a node with lower trust value than the threshold. More details will be introduced in the following.

As shown in Figure 3, each node will have a list of its attributes, centricity, and forwarding probability and a list of the information from other nodes. And $a_{11}, a_{12}, \ldots, a_{15}$ is used to represent the value of the attribute $a_1, a_{21}, a_{22}, \ldots, a_{25}$ is used to represent the value of the attribute a_2, and it is similar for the attributes a_3, a_4, a_5. In Figure 3, node j is a malicious node to launch a greyhole attack, which drops part of the received messages, and node l launches a blackhole attack to drop all the received messages.

In step one, when messages are required to be forwarded, node s should select a relay node according to the value including direct trust, indirect trust, and entropy function. In the figure, node s meets node i and node j and updates the information of trust value and entropy function. Since node j has dropped some of the messages it received, the value of entropy function will be lower which is smaller than the threshold of 0.5, and messages will not be forwarded to node j. Similarly, messages will not be forwarded from node i to node l in step two because all messages received by node l have been dropped. Therefore, by the proposed routing mechanism, the greyhole and blackhole attacks can be prevented.

Moreover, the proposed scheme can prevent the bad-mouthing attacks, by which a malicious node ruins the reputation of well-behaved nodes by providing bad recommendations against good nodes. By the proposed routing mechanism, the recommendation value is the average value of the recommendations from the one-hop neighbor nodes. And each recommendation from the neighbor will have a valid range. The recommendation out of range will not be considered. In this way, the recommendation value from one malicious node will not make much effect to the average value. This method is also efficient for the ballot stuffing attacks. Therefore, the bad-mouthing and ballot stuffing attacks can be avoided by the routing mechanism.

4. Performance Analysis

4.1. Theoretical Analysis for Delivery Probability. As shown in the system model, the mobility model is a Markov process, where nodes will change from one state to another due to movement. For any node k, let S_n be the nth state, and $P(i, j, t)$ represents in the period time t; node k changes from state i to state j. We can get $P(i, j, t)$ as

$$P(i, j, t) = P(S_{n+1} = j, \Delta t \leq t \mid S_n = i). \tag{14}$$

In the Markov process, $P(i, j)$ indicates the probability of node k from state i to state j. And $P(i, j)$ can be represented as

$$P(i, j) = P(S_{n+1} = j \mid S_n = i). \tag{15}$$

$G(i, j, t)$ can be used to represent that the time from state i to state j is less than t. And $G(i, j, t)$ is shown as

$$\begin{aligned} G(i, j, t) &= P(\Delta t \leq t \mid S_{n+1} = j, S_n = i) \\ &= \sum_{1 \leq u \leq t} P(\Delta t = u \mid S_{n+1} = j, S_n = i). \end{aligned} \tag{16}$$

Message Forwarding Phase (for any node i and j)

(1) WHILE (node i meets node j) THEN
(2) //compute the Entropy Function $T^{j}_{\text{Entropy-Function}}$ according to formula (9)
(3) IF ($T^{j}_{\text{Entropy-Function}} < T^{\text{threshold}}_{\text{Entropy-Function}}$) THEN
(4) //Node j is marked as malicious node, do nothing
(5) ELSE
(6) FOR (any node k && $k \neq j$)
(7) //Compute indirect value $T^{ik}_{\text{indirect},X}(t + \Delta t)$ where node k is recommended by node j and one-hop neighbors of node i according to formula (6).
(8) ENDFOR
(9) //Compute the direct trust value $T^{ij}_{\text{direct},X}(t + \Delta t)$ with two attributes (selfishness, centrality).
(10) //Compute $T^{ij}_{X}(t + \Delta t)$ by combining $T^{ij}_{\text{direct},X}(t + \Delta t)$ and $T^{ij}_{\text{indirect},X}(t + \Delta t)$ according to formula (2).
(11) //Compute $T^{ij}(t + \Delta t)$ by combining all attributes (selfishness, centrality) according to formula (1).
(12) IF ($T^{ij}(t + \triangle t) \geq T^{\text{threshold}}$)
(13) //Node j is marked as a trustworthy node and messages will be forwarded to node j.
(14) ELSE
(15) //do nothing
(16) ENDIF
(17) ENDIF
(18) ENDWHILE

Algorithm 1: Message forwarding phase.

And then $P(i, j, t)$ can be computed using $P(i, j)$ and $G(i, j, t)$. Consider

$$\begin{aligned} P(i, j, t) &= P(S_{n+1} = j, \Delta t \leq t \mid S_n = i) \\ &= P(\Delta t \leq t \mid S_{n+1} = j, S_n = i) \\ &\quad \times P(S_{n+1} = j \mid S_n = i) \\ &= G(i, j, t) \times P(i, j). \end{aligned} \tag{17}$$

As $P(i, j, t)$ is the probability of node k from state i to state j directly in the period time t, however, node k may enter another state and then enter the state j. So we use $Q(i, j, t)$ represents this situation. Consider

$$\begin{aligned} &Q(i, j, t) \\ &= \sum_{r=1}^{m} \sum_{u=1}^{t} (P(i, r, u) - P(i, r, u-1)) \times Q(r, j, t-u) \\ &\qquad j \neq i, \end{aligned} \tag{18}$$

where m is the number of states that node k may enter. In other words, $Q(i, j, t)$ is, in fact, the probability of node k from the source cluster to the destination cluster, denoted by $P_{c-c}(\Delta t \leq t)$. Since in the same cluster, nodes communicate with each other frequently, it is assumed that the delivery ratio in one cluster is 1. So we can approximately compute the delivery ratio as

$$P_{\text{delivery}}(\Delta t \leq t_N) = \sum_{0 \leq i < T_N} P_{c-c}(\Delta t \leq i). \tag{19}$$

Since there is a limited capacity for the cache at each node, when enough messages are generated, the cache could be overflowed, and the messages could be dropped. To compute the delivery probability, the service efficiency of the network δ should be known. Let ρ to be the ratio of customer service per unit time; we can obtain

$$\rho = \frac{1}{T_N} \sum_{i=1}^{T_N} P_{\text{delivery}}(\Delta t \leq i). \tag{20}$$

Let λ be the arrival rate of the messages and B_{Max} be the maximum cache available. Since there are malicious nodes in the network, by the trust routing scheme, these nodes will be detected and will not be selected as next carrier. It can also be modeled that the cache becomes smaller. Therefore, we can get the new maximum cache $B'_{\text{Max}} = \omega B_{\text{Max}}$, where ω is the percentage of the malicious nodes in the whole collection of the nodes. We can compute δ as

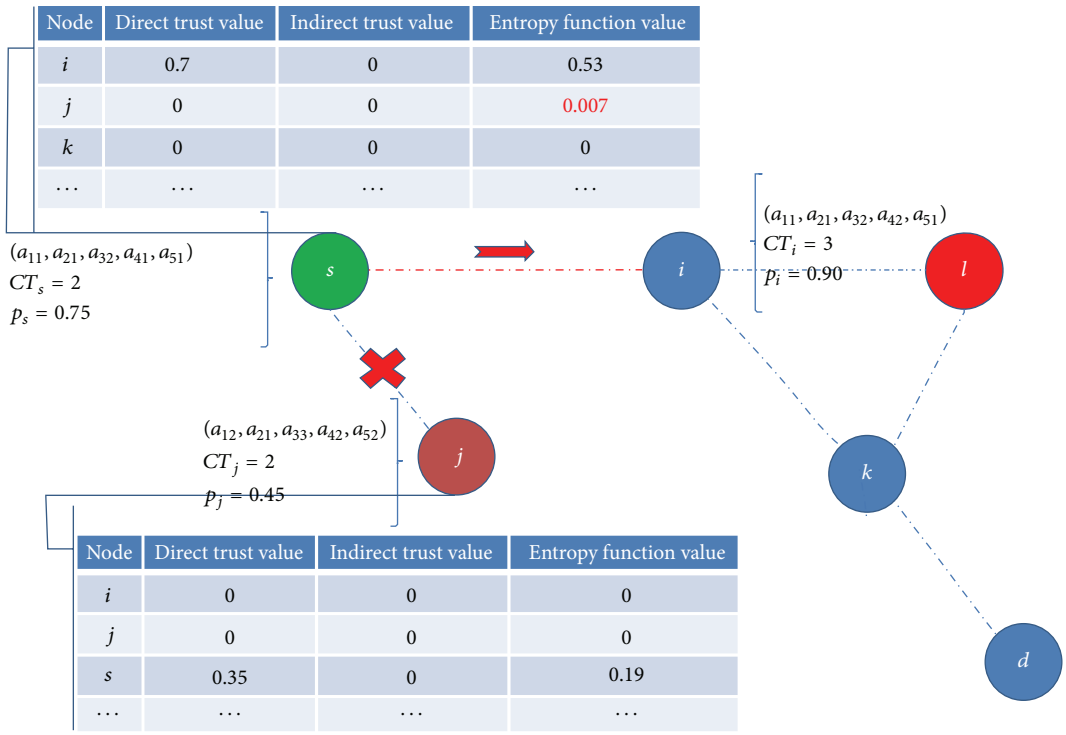
$$\delta = \frac{(\lambda * T_N * \rho + B'_{\text{Max}})}{(\lambda * T_N)}. \tag{21}$$

Since we have obtained the delivery probability without considering the cache in formula (19), and according to the service efficiency, the delivery probability with the limited cache $P'_{\text{inter-cluster}}(\Delta t \leq T_N)$ can be approximated as

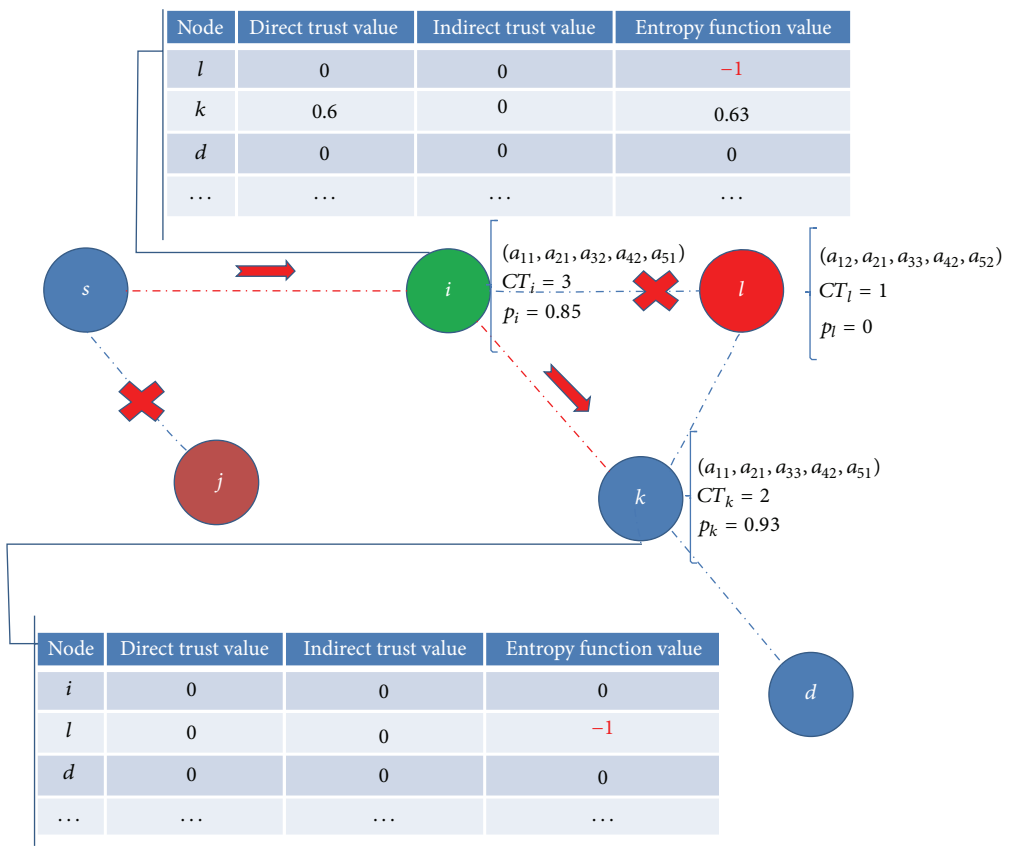
$$P'_{\text{delivery}}(\Delta t \leq T_N) \approx \delta * P_{\text{delivery}}(\Delta t \leq t_N). \tag{22}$$

What is more, the service efficiency of the network δ can be approximately used as the forwarding probability to select proper $T^{\text{threshold}}_{\text{Entropy-Function}}$.

4.2. Analysis for Average End-to-End Delay. As we all know, if the delivery ratio is close to 1, it indicates that most of the messages have been delivered.



(a) Step one



(b) Step two

Figure 3: Routing procedure.

When the delivery ratio $P'_{\text{inter-cluster}}(\Delta t \le T_N)$ is close to 1, the average delay can be obtained as

$$D_{\text{delay}} \approx \sum_{i=1}^{T_N} P'_{\text{delivery}}(\Delta t \le i). \quad (23)$$

4.3. Complexity Analysis. It is assumed that there are M nodes in the network; according to the mechanism introduced in Section 3, the complexity analysis can be divided to four parts: the complexity of direct trust value, the complexity of indirect trust value, the complexity of entropy function, and the complexity of message forwarding. First, when the direct trust value is computed, in the worst case, any two nodes need to make finite computations, and there are maximum $M * (M-1)/2$ combinations of nodes. Therefore, the complexity of the direct trust value will be $O(M^2)$. Second, when the indirect trust value is computed, for each node, it should integrate the indirect trust value from its $M-1$ neighbors. For each node, it should compute the indirect trust value for the other maximum $M-1$ nodes. So the complexity of indirect value is $O(M^2)$, and for M nodes, the complexity of the indirect value is $O(M^3)$. When the entropy function is computed, for each node, it should make finite computations for other $M-1$ nodes. Therefore, the complexity of the entropy function for each node is $O(M)$ and for M nodes, the complexity is $O(M^2)$. Last, when messages are forwarded, there are finite comparisons for each transmission. Assume that there is no loop in the transmission of the messages, the complexity of message forwarding for each message is $O(M)$. If there are N messages required to be transmitted, the complexity will be $O(M * N)$. For the entire routing scheme, the complexity is $O(M * N + M^3)$ by combining all the four parts.

4.4. Simulation Analysis. The simulation experiments have been performed by using QualNet. Simulation parameters are shown in Table 1. As mentioned in system model, we set $P_H = 0.7$ and $P_C = 0.1$ for 80% of the nodes at each hot spot and $P_H = 0.5$ and $P_C = 0.1$ for other 20% nodes. And 100 nodes in total are distributed. The WIFI is adopted as the communication medium among all the mobile nodes and the radio range is about 250 meters. The simulation time is set to 10000 s with a 500 s warm-up period. And the size of simulation area is set to 3000 m * 3000 m. CBR traffic is used for the data generation with 0.1 arrival rate. What is more, the malicious nodes could launch different types of attacks including blackhole attacks, greyhole attacks, bad-mouthing attacks, and ballot stuffing attacks randomly.

We compare the performance of our proposed solution with that of the clustering and network coding based efficient routing in social DTNs (CCS-DTN) [22], Bayesian [13], and PROPHET [3]. The successful delivery rate, defined as the ratio of the number of messages that have been successfully received by the destination over the number of messages that the source nodes have sent, and the average end-to-end delay, defined as the average time used for a message transmitted from a source to its destination, have been measured. What is more, the simulation results have been compared with those obtained from the theory analysis.

Table 1: Simulation parameters.

Parameter	Value
Simulation time	10000 S
Size of area	3000 * 3000
Warm-up time	500 S
Number of nodes	50
Traffic	CBR
Transmission range	250 M
Message size	150 Byte
Rate of message sending	0.1
MAC protocol	802.11
$T^{\text{threshold}}_{\text{Entropy-Function}}$	0.7
$T^{\text{threshold}}$	0.65

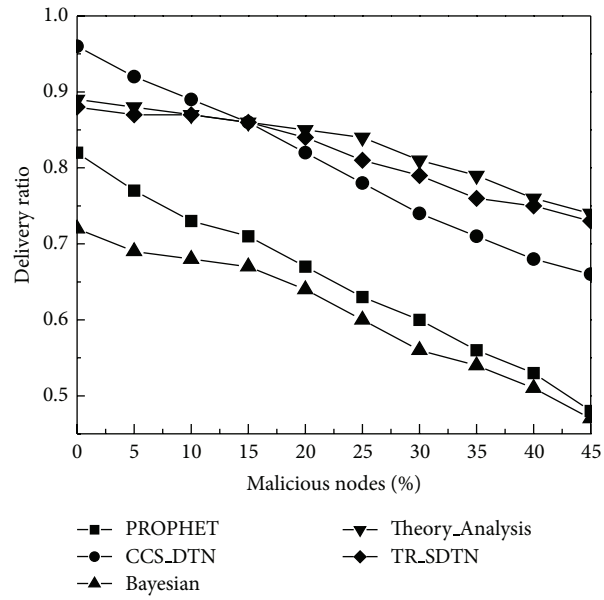


Figure 4: Impact of malicious nodes on delivery ratio.

With the change of the number of malicious nodes, the performance in terms of successful delivery ratio has been shown in Figure 4. Clearly, the successful delivery ratio has become lower with the increase of the percentage of the malicious nodes because the number of messages dropped due to the attacks from the malicious nodes has been increased. When there are small malicious nodes, the delivery ratio of CCS-DTN is higher than TR_SDTN since the few packets dropped can be recovered by the network coding technology. But with the increase of malicious nodes, in CCS-DTN, many messages are dropped by the malicious nodes which cannot be recovered by the network coding technology while in TR_SDTN, messages will not be forwarded to the malicious nodes and there are smaller messages dropped. Therefore, the delivery ratio of TR_SDTN is higher than that of CCS-DTN when the number of malicious nodes is enough. What is more, When there are less malicious nodes, the

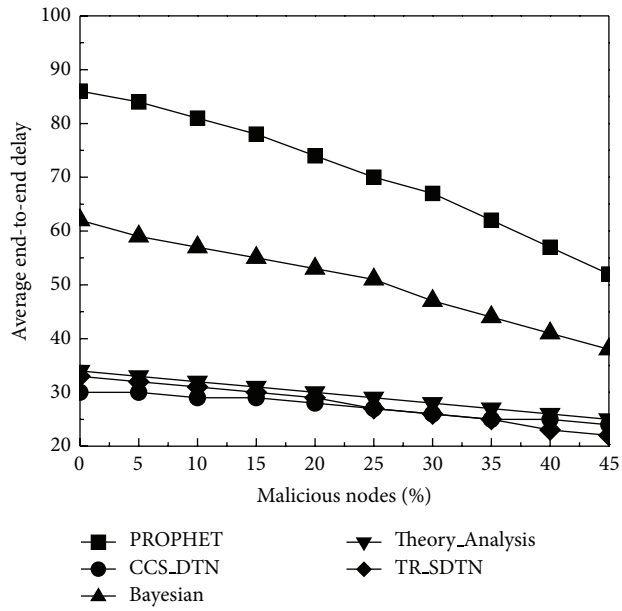


Figure 5: Impact of malicious nodes on average end-to-end delay.

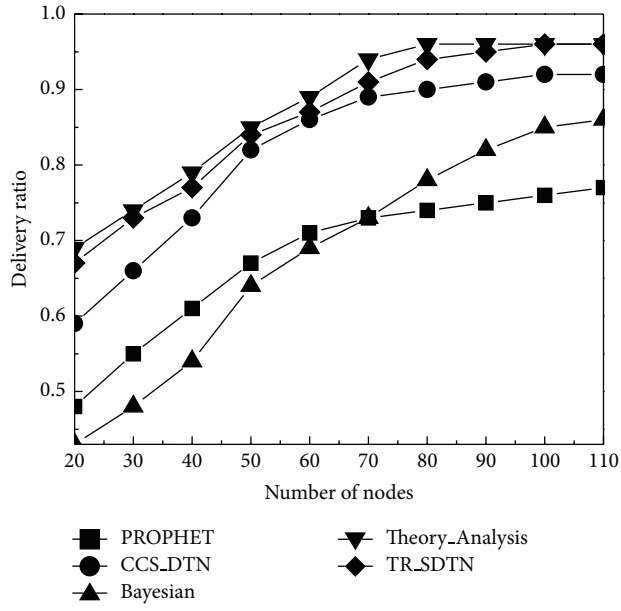


Figure 6: Impact of number of nodes on delivery ratio.

delivery ratio by the proposed TR_SDTN scheme is much higher than that by the Bayesian scheme because, by the proposed TR_SDTN scheme, the social characteristics of the nodes in the network have been efficiently used. What is more, since there is no mechanism to tackle the malicious attacks in the Bayesian scheme, the delivery ratio has declined much quicker than that by the proposed TR_SDTN scheme.

Meanwhile, as shown in Figure 5, as more malicious nodes exist in the network, more messages will be dropped. In this process, the messages to be delayed longer time will have higher probability to be dropped. Therefore, with the increase of percentage of malicious nodes, the average end-to-end delay becomes lower. Besides, compared with the Bayesian scheme and PROPHET scheme, the proposed TR_SDTN scheme has shown much lower average end-to-end delay due to the efficient use of the social characteristics of the nodes in the network. And the proposed TR_SDTN scheme can make the packet forwarded to destination much more quickly.

Compared with the CCS-DTN and the PROPHET schemes, the delivery ratio of TR_SDTN has a slower decrease since the trust mechanism can detect the malicious nodes and avoid messages which are forward to these nodes. Using this method, messages dropped are declined. Another reason for the decline of delivery probability of TR_SDTN is that, with the increase of percentage of the malicious nodes, the number of available nodes becomes smaller which leads more messages to be dropped since available cache becomes smaller. What is more, we can find that the delivery probability of CCS-DTN decreases slower than that of the PROPHET scheme because some messages dropped are recovered by the network coding mechanism.

Another measurement on the network states for the evaluation is the number of mobile nodes. In the simulation, the percentage of the malicious nodes has been set as 20% of the total number of the nodes in the network. As shown in Figure 6, the average delivery ratio increases with the increase of the number of mobile nodes. Firstly, we can find when the number of nodes is in the range of 20 to 70, the increase of delivery ratio is faster. When there are only a small number of the nodes, the messages cached in the network will be limited and the contacts among mobile nodes will be not frequent, and many messages will be dropped. With the increase of the number of the nodes, this situation has been quickly mitigated. When the number of the nodes reaches a certain level, there is enough capacity of the caches in the network and nodes have more chances to contact with each other; the delivery ratio becomes stable. Lastly, the delivery ratio by the TR_SDTN scheme is significantly greater than those of the other three schemes including the CCS-DTN, the PROPHET, and the Bayesian scheme when the number of nodes reaches a certain level because, by the proposed trust mechanism, the messages will not be forwarded to the malicious nodes and there is enough capacity of the caches to keep these messages. We can conclude from the above mentioned facts that the proposed TR_SDTN scheme can approach a much better performance than those of the other three schemes for the social DTNs. Moreover, it is found that the results obtained from the mathematical analysis and the results from the simulations have shown only a little difference. That is to say, these results have been proved by each other to show the consistency.

The next simulation is to study the impact of the trust threshold on the delivery ratio. As shown in Figure 7, when the threshold is determined, the delivery ratio will become lower with the increase of the number of malicious nodes. And it is clear that, with different trust thresholds, the delivery ratios will be different. In the figure, when the trust threshold is 0.65, the delivery ratio is the largest among the other values of the threshold. When the trust threshold is too high, although the malicious nodes have been excluded, some

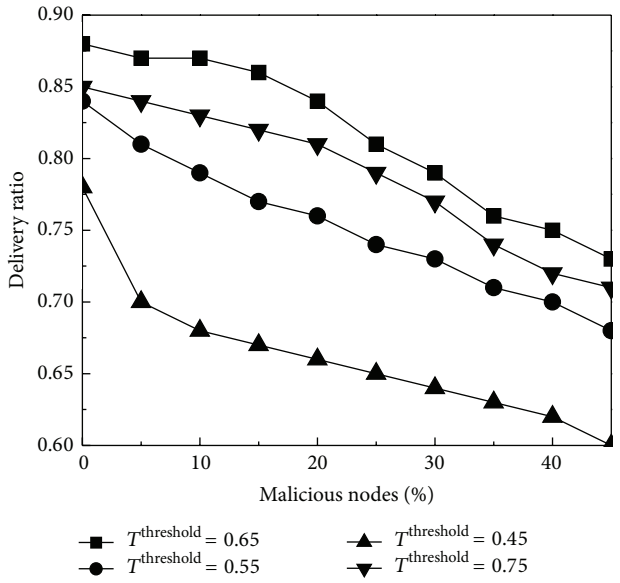


Figure 7: Impact of the trust threshold on delivery ratio.

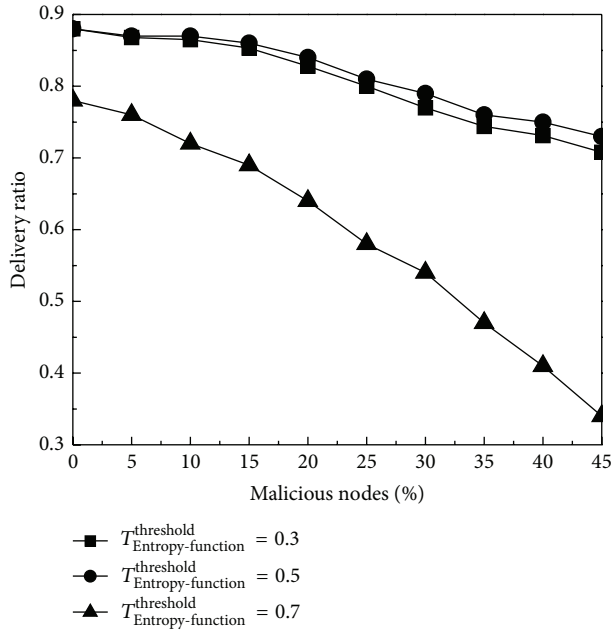


Figure 8: Impact of entropy function threshold on delivery ratio.

normal nodes will be prevented from receiving messages because their trust values are lower than the threshold. With the reduction of the trust threshold, more and more malicious nodes could be allowed to receive messages. In this way, more and more messages will be dropped by the malicious nodes, and the delivery ratio becomes lower.

The last simulation it to study the impact of the entropy function threshold on the delivery ratio. As shown in Figure 8, with the different threshold values, the delivery ratios will be different. When the threshold is 0.5, the delivery ratio is the largest among the other values of the thresholds. From the entropy function defined in formula (8), we can find that the major impact factor of the entropy function is the forwarding probability of each node. Since the cache is limited, some messages will be dropped at each node. There is a normal range of the values of the forwarding probability for each node. When the threshold of the entropy function is too high, the threshold of the forwarding probability will be also high. In this situation, many normal nodes would not be able to forward messages, which leads to the lower delivery ratio. With the increase of malicious nodes, the available normal nodes become fewer and fewer with more and more messages dropped. The delivery ratio declines faster. When the threshold of the entropy function is small, the threshold of the forwarding probability will be also small. In this case, the malicious nodes which launch greyhole attacks could be allowed to receive messages. At these nodes, more messages will be dropped than the normal nodes making the delivery ratio lower. So a proper threshold of the entropy function is important for the design of the routing mechanism.

5. Conclusion

In this paper, according the characteristics of the nodes and by analyzing the possible attacks in hostile social DTNs, we have proposed a novel TR_SDTN routing algorithm, by which, in the hostile social DTN environments, a direct trust model, an indirect trust model, and the Shannon entropy function have been combined to make full use of nodes' selfishness and prevent blackhole attacks, greyhole attacks, bad-mouthing attacks, and ballot stuffing attacks from the malicious nodes. The theoretical analysis and the simulation results have consistently shown that, in the hostile social DTN environments, better performance can be achieved by the proposed TR_SDTN scheme.

Conflict of Interests

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

Acknowledgments

This work is supported in part by the National Natural Science Foundation of China under Grant 61162003, the Plan of Science and Technology of Qinghai Province under Grant 2012-ZR-2989, the Plan of Science and Technology of Hainan Province under Grant ZDXM2014086, and the Natural Science Foundation of Hainan Province under Grant 614229.

References

[1] A. Vahdat and D. Becker, "Epidemic routing for partially-connected ad hoc networks," Tech. Rep. CS-200006, Duke University, 2000.

[2] R. Ramanathan, R. Hansen, P. Basu, R. Rosales-Hain, and R. Krishnan, "Prioritized epidemic routing for opportunistic networks," in *Proceedings of the 1st International MobiSys Workshop*

on Mobile Opportunistic Networking (MobiOpp '07), pp. 62–66, June 2007.

[3] A. Lindgren, A. Doria, and O. Schelen, "Probabilistic routing in intermittently connected networks," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 7, no. 3, pp. 19–20, 2003.

[4] S. Ahmed and S. S. Kanhere, "Cluster-based forwarding in delay tolerant public transport networks," in *Proceedings of the 32nd IEEE Conference on Local Computer Networks (LCN '07)*, pp. 625–632, October 2007.

[5] J. Whitbeck and V. Conan, "HYMAD: hybrid DTN-MANET routing for dense and highly dynamic wireless networks," *Computer Communications*, vol. 33, no. 13, pp. 1483–1492, 2010.

[6] H. Dang and H. Wu, "Clustering and cluster-based routing protocol for delay-tolerant mobile networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 6, pp. 1874–1881, 2010.

[7] E. M. Daly and M. Haahr, "Social network analysis for routing in disconnected delay-tolerant MANETs," in *Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '07:)*, pp. 32–40, Montréal, Canada, September 2007.

[8] P. Hui, J. Crowcroft, and E. Yoneki, "BUBBLE rap: social-based forwarding in delay tolerant networks," in *Proceedings of the 9th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '08)*, pp. 241–250, May 2008.

[9] K. Jahanbakhsh, G. C. Shoja, and V. King, "Social-greedy: a socially-based greedy routing algorithm for delay tolerant networks," in *Proceedings of the 2nd International Workshop on Mobile Opportunistic Networking (MobiOpp '10)*, pp. 159–162, February 2010.

[10] T. Hossmann, T. Spyropoulos, and F. Legendre, "Know thy neighbor: towards optimal mapping of contacts to social graphs for DTN routing," in *Proceedings of the IEEE INFOCOM*, pp. 1–9, March 2010.

[11] C. M. Kim, I. S. Kang, and Y. H. Han, "An efficient routing scheme based on social relations in delay-tolerant networks," in *Ubiquitous Information Technologies and Applications*, vol. 280 of *Lecture Notes in Electrical Engineering*, pp. 533–540, 2014.

[12] P.-Y. Yuan, H.-D. Ma, and P.-R. Duan, "Impact of strangers on opportunistic routing performance," *Journal of Computer Science and Technology*, vol. 28, no. 3, pp. 574–582, 2013.

[13] I.-R. Chen, F. Bao, M. Chang, and J.-H. Cho, "Dynamic trust management for delay tolerant networks and its application to secure routing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 5, pp. 1200–1210, 2014.

[14] W. Dong, V. Dave, L. Qiu, and Y. Zhang, "Secure friend discovery in mobile social networks," in *Proceedings of the IEEE INFOCOM*, pp. 1647–1655, April 2011.

[15] X. Guan, C. Liu, M. Chen, H. Chen, and T. Ohtsuki, "Internal threats avoiding based forwarding protocol in social selfish delay tolerant networks," in *Proceedings of the IEEE International Conference on Communications (ICC '11)*, pp. 1–6, June 2011.

[16] K. Chen and H. Y. Shen, "Multicent: a multifunctional incentive scheme adaptive to diverse performance objectives for DTN routing," in *Proceedings of the 10th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON '13)*, pp. 532–540, New Orleans, La, USA, June 2013.

[17] Y. J. Liu, L. Li, Z. S. Li, and Y. Ye, "Trust-based optimized routing scheme in Mobile Social Networks," in *Proceedings of the International Conference on Communications, Circuits and Systems (ICCCAS '13)*, pp. 87–90, November 2013.

[18] S. Yang, U. Adeel, and J. A. McCann, "Selfish mules: social profit maximization in sparse sensornets using rationally-selfish human relays," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 6, pp. 1124–1134, 2013.

[19] H. Gong, L. Yu, and X. Zhang, "Social contribution-based routing protocol for vehicular network with selfish nodes," *International Journal of Distributed Sensor Networks*, vol. 2014, Article ID 753024, 12 pages, 2014.

[20] Y. Zhu, B. Xu, X. Shi, and Y. Wang, "A survey of social-based routing in delay tolerant networks: positive and negative social effects," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 1, pp. 387–401, 2013.

[21] Y. H. Guo, S. Schildt, T. Pougel, S. Rottmann, and L. Wolf, "Mitigating blackhole attacks in a hybrid VDTN," in *Proceedings of the IEEE 15th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM '14)*, pp. 1–6, Sydney, Australia, June 2014.

[22] Z. J. Zhang, M. Ma, and Z. G. Jin, "CCS-DTN: clustering and network coding based efficient routing in social DTNs," in *Proceedings of the IEEE International Conference on (GLOBECOM '13)*, December 2013.