On Anonymous Routing in Delay Tolerant Networks

Kazuya Sakai, *Member, IEEE*, Min-Te Sun, *Member, IEEE*, Wei-Shinn Ku, *Senior Member, IEEE*, and Jie Wu, *Fellow, IEEE*

Abstract—Due to instability of links in the network, the model of delay tolerant networks (DTNs) is often adopted in many emerging mobile applications. To organizations and individuals using these mobile applications, security and privacy are considered to be two of the most significant concerns. This research seeks to address anonymous communications in DTNs. While many different anonymous routing protocols have been proposed for ad hoc networks, to the best of our knowledge, only variants of onion-based routing have been tailored for DTNs. Since each type of anonymous routing protocol has its pros and cons, there is no single anonymous routing protocol for DTNs that can adapt to the different levels of security requirements. In this paper, we first design a set of anonymous routing protocols for DTNs, called anonymous Epidemic and zone-based anonymous routing, based on the original anonymous routing protocols for ad hoc networks. Then, we propose a framework of anonymous routing (FAR) for DTNs, which subsumes all the aforementioned protocols. By tuning its parameters, the proposed FAR is able to outperform onion-based, anonymous Epidemic, and zone-based routing. In addition, numerical analyses for the traceable rate, node anonymity, and path anonymity models are built. Extensive simulations using randomly generated graphs as well as real traces demonstrate that the proposed framework for DTNs successfully achieves its design goals.

Index Terms—Delay tolerant networks, DTNs, anonymous routing

1 Introduction

In many emerging wireless applications, including people/pocket-switched networks, vehicular networks, and battlefield communications, the links among nodes are naturally intermittent. As a result, the traditional ad hoc network model, which emphasizes the stability of links for packet routing, is no longer viable for these applications. To address this issue, the model of delay tolerant networks (DTNs) has been proposed, in which each link is replaced by the probability of contact events. The DTN model is especially suitable for networks with a high level of node mobility. However, very little work has been done on the security, privacy, or their relationship with network performance in DTNs, which are of significant concern in these applications. For instance, one of the communicating parties in a battlefield is most likely to be a gateway to the infrastructure or a command operator. The identities and locations of such nodes should not be disclosed to the adversaries. Motivated by these observations, we are interested in anonymous wireless communications that prevent adversaries from violating mobile users' privacy, e.g., deriving users' identities, locations, and routing paths, by traffic analyses.

A great deal of effort has been invested in designing anonymous routing protocols for the internet [1], [2] and mobile ad hoc networks [3], [4], [5], [6]. The message that is protected by a number of encrypted layers, a so-called *onion* [7], is widely used to preserve the privacy of end hosts as well as routing paths. In onion-based routing, onion routers serve as proxies, and any given intermediate node will never know where the source and sink of the message are located. In mobile ad hoc networks, the location-based deanonymization attack [8] may reveal the physical location of nodes. To this end, the zone-based anonymous routing is proposed in [6] where the source and the last proxies perform restricted flooding, so as to make sure that the source and destination nodes are not identifiable within the flooding zone.

In the DTN research community, a few anonymous routing protocols, which use the idea of onion groups [8], [9], [10] and the threshold [11], have been proposed in order to improve the degree of privacy, such as the traceable rate, node anonymity, path anonymity. However, the following research challenges that particularly arise in anonymous routing in DTNs are yet to be addressed.

First, it is known that the use of a number of onions results in lower traceable rate. As a consequence, onion-based protocols [8], [9], [10] experience slow packet delivery. Second, the anonymity set of the source and destination nodes can be deduced, should the first and last onion relay be compromised. Third, although the zone-based approach improves

- *K. Sakai is with the Department of Electrical Engineering and Computer Science, Tokyo Metropolitan University, 6-6 Asahigaoka, Hino, Tokyo 191-0065, Japan. E-mail: ksakai@tmu.ac.jp.*
- *M.-T. Sun is with the Department of Computer Science and Information Engineering, National Central University, Taoyuan 320, Taiwan. E-mail: msun@csie.ncu.edu.tw.*
- *W.-S. Ku is with the Department of Computer Science and Software Engineering, Auburn University, Auburn, AL 36849. E-mail: weishinn@auburn.edu.*
- *J. Wu is with the Department of Computer and Information Science, Temple University, 1925 N. 12th St., Philadelphia, PA 19122. E-mail: jiewu@temple.edu.*

Manuscript received 30 Oct. 2017; revised 8 Nov. 2018; accepted 29 Nov. 2018. Date of publication 12 Dec. 2018; date of current version 31 Oct. 2019. (Corresponding author: Min-Te Sun.)
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TMC.2018.2886349

node anonymity, neither Epidemic-like nor zone-based protocol for DTNs has been proposed so far. One reason for this is the difficulty in defining a zone in DTNs where the network graph is constructed from the past contact history, rather than from physical locations of nodes. At last, to the best of our knowledge, there is no work that balances the pros and cons of these different approaches. It is interesting to design an anonymous routing framework that subsumes all the aforementioned protocols and optimizes the anonymous DTN routing based on a number of metrics, e.g., delivery rate, anonymity, delay, and forwarding cost, by tunable parameters.

To address the above challenges, we propose the framework of anonymous routing for DTNs. The contributions of this paper are as follows.

- We first design a set of anonymous DTN protocols, including Anonymous Epidemic (AE), Restricted Epidemic Routing (RER), and Zone-Based Anonymous Routing (ZBAR), based on anonymous routing protocols originally proposed for mobile ad hoc networks. The key difference from the existing solutions is the definition of "zone," where senders and receivers stay anonymous. The proposed RER guarantees that a message reaches at least one of the nodes in the next onion group, with a certain probability specified by the threshold. In addition, RER can be used as a subroutine of ZBAR.
- We next propose a framework of anonymous routing (FAR) for DTNs that subsumes all the Epidemic, zone-based, and onion-based routing protocols with tunable parameters. In FAR, a message travels along a set of onion groups with router-by-router encryption, and every communication between two consecutive onion routers on the routing path is performed by either Epidemic routing or spray-and-wait forwarding with a time constraint. By doing this, FAR enjoys the advantages of these baseline protocols, and DTN users can balance the performance, privacy, and cost based on their preferences.
- We then quantitatively analyze the privacy metrics provided by FAR. To be specific, the closed form solutions used to estimate the traceable rate, source/destination anonymity, and path anonymity are provided. The proposed mathematical models help DTN users to select appropriate routing parameters that meet their security and privacy requirements.
- Finally, we conduct extensive simulations using a well-known real trace, CRAWDAD dataset cambridge/haggle [12], as well as random graphs to demonstrate the performance and degree of privacy of the proposed scheme. Furthermore, the simulation results are compared with analytical results, and the comparisons show that our analyses provide very close approximations.

The rest of this paper is organized as follows. Section 2 presents the AE, RER, and ZBAR protocols specifically revised for DTNs. These protocols will serve as the building blocks for the proposed FAR, which is introduced in Section 3. The mathematical analysis of the proposed FAR is presented in Section 4. The performance of the proposed scheme is evaluated by both computer simulations and real trace dataset in Section 5. The discussions on how to select parameters is provided in Section 6. Section 7 reviews related works and Section 8 concludes this paper.

2 Protocol Design

In this section, we first design a set of protocols for DTNs based on anonymous broadcast and the zone-based protocols, which are originally designed for mobile ad hoc networks. These revised protocols, as well as the onion-based protocols, will serve as the building blocks for the proposed FAR protocol introduced in Section 3.

2.1 Notations and Assumptions

A DTN is represented by an undirected graph which is constructed from contact histories among nodes. Let v_i be a node i, and two nodes, say v_i and v_j, are connected in a graph if v_i and v_j have at least one contact in the past. The weight of a link between v_i and v_j is given by $\lambda_{i,j}$, where $1/\lambda_{i,j}$ is the inter-meeting time between two nodes v_i and v_j. In [10], [13], the inter-contact time between nodes in a DTN is assumed to be exponential distribution. We adopt this assumption in this paper for the protocol design and analysis. However, we will relax this assumption in the performance section by using the real trace dataset and use this dataset to access the performance of our derived protocol in the real-world DTN scenarios. The probability density function that v_i meets v_j at time t is obtained by $\lambda_{i,j}e^{-\lambda_{i,j}t}$. In addition, the probability that v_i meets v_j within T (where $T > 0$) is computed by

$$P_{i,j}(T) = \int_0^T \lambda_{i,j}e^{-\lambda_{i,j}t}dt = 1 - e^{-\lambda_{i,j}T}. \quad (1)$$

In onion-based routing, a message, denoted by m, travels a set of onions in the specified order by which each layer of an onion is to be peeled off. We denote R_i as the set of nodes for the ith onion group by which m travels. For convenience, the jth node in R_i is labeled by $r_{i,j}$, and the size of R_i is G_i. In addition, the average group size is denoted by G.

For cryptographic operations, PK_i and SK_i are defined as the public and private keys of node v_i. In addition, GK_i represents the group key of onion group R_i. The encryption and decryption functions are denoted by $Enc(.)$ and $Dec(.)$. The initialization of public and private keys is the same as the existing solution [9], [10], [14]. That is, each node obtains the keys from a key distribution server when it has an access to the server. In addition, the onion groups are assumed to be fixed before deployment, and each node in the same group shares GK_i.

The notations used in this paper are summarized in Table 1.

2.2 The Attack Model

The attack model in the Internet-based anonymous communications are categorized into either strong or weak model. The adversary is said to be *strong* if she can monitor all the traffic in the network, and *weak* otherwise. On the other hand, in the wireless networks, the adversary must be in the proximity of a node to monitor the traffic, and therefore, monitoring all the traffic is infeasible. In this sense, the adversaries are weak in the setting of DTNs. Similar to the existing anonymous routing for ad hoc networks [5], [6] and DTNs [10], the weak adversary

TABLE 1
Definition of Notations

Symbols	Definition
n	The number of nodes in a network
v_i	Node i
$1/\lambda_{i,j}$	The inter-contact time between v_i and v_j
m, σ	A message and an encrypted message
$Enc(.)/Dec(.)$	Encryption/decryption functions
L	The number of copies
K	The number of onion routers that a message travels
η	The number of hops between two nodes
R_i	A set of onion routers for the ith hop
G_i	The size of onion group R_i
G	The avg. number of nodes in an onion group
$r_{i,j}$	The jth node in R_i
T, t_i	The end-to-end and the zone i's deadlines
τ	The threshold to determine t_i
c	The number of compromised nodes
ϕ	A set of nodes
ϕ'	A set of suspicious nodes
$H(\phi')$	The entropy of a system with given ϕ'
H_{max}	The maximal entropy of a system
$D(\phi')$	The anonymity with given ϕ'

model is applied to this paper. In this model, the adversaries can obtain information from only compromised nodes, and thus, eavesdropping all the traffic is physically not possible.

While the network model in DTNs differs from that of ad hoc networks, the similar security threats such as eavesdropping and traffic analysis are possible in DTNs. For example, an adversary clandestinely stalks a legitimate mobile user to monitor whom the user meets and eavesdrops on wireless channels. Another possible attack is that an adversary blackmails a user to obtain the network log, which contains the information about from/to which node she receives/sends a message.

In this paper, we abstract the aforementioned threats by the compromise attack, where some nodes in a network are marked as being compromised and the message transmissions/receptions are monitored. Then, an adversary reasons possible routing paths and identifies source/destination based on the information disclosed from compromised nodes. Let $\{v_s, r_1, r_2, \ldots, r_K, v_d\}$ be a path with $K+1$ hops and the link between two relays be $r_k \rightarrow r_{k+1}$. Then, we define the two security attacks as follows.

Attack 1 (The Path Tracing). *An adversary tries to discover links $v_s \rightarrow r_1$, $r_k \rightarrow r_{k+1}$ for $1 \leq k \leq K-1$, and $r_K \rightarrow v_d$ which constitutes a path as much as possible. Should r_k be compromised, an adversary will be able to find the next relay r_{k+1} by stalking r_k.*

As a privacy metric against Attack 1, the traceable rate [5] can be applied, which is a weighted metric indicating what portion of a path is disclosed to adversaries when some nodes are compromised. Let η be the number of hops between the source and destination, C_{seg} be the number of compromised segments, and $c_{seg,i}$ be the length of the ith compromised segments. Then, the traceable rate, denoted as P_{trace}, is defined

$$P_{trace} = \frac{1}{\eta^2} \sum_{i=1}^{C_{seg}} (c_{seg,i})^2. \tag{2}$$

For example, let $v_1 \rightarrow v_2 \rightarrow v_3 \rightarrow v_4 \rightarrow v_5$ be a routing path where the number of hops is four, i.e., $\eta = 4$. Assume that the link between nodes v_i and v_{i+1} is disclosed to an adversary when v_i is compromised. For instance, when three nodes, v_1, v_3, and v_4, are compromised, the traceable rate will be $\frac{1^2+2^2}{4^2} = \frac{5}{16}$. If three consecutive nodes, v_1, v_2, and v_3, are compromised, the traceable rate will be $\frac{3^2}{4^2} = \frac{9}{16}$. As indicated by these cases, the traceable rate is weighted in the sense that the greater the length of the consecutive compromised segments, the greater the portion of the path that is traceable.

Attack 2 (The Node Deanonymizing). *An adversary tries to identify v_s and v_d. Should the first onion router r_1 or the last onion router r_K be compromised, the adversary may narrow the anonymity set to which v_s or v_d belongs.*

Anonymity is the state of not being identifiable among an anonymity set. Anonymity is generally modeled as an entropy-based metric [15]. Let ϕ be all the possible elements, and p be the probability that a given element is original. The elements could be nodes and routing paths in our context. The entropy of the system is given

$$H(\phi) = -\sum_{\forall i \in \phi} p_i \log_2(p_i). \tag{3}$$

When $p_i = p_j$ for all $i, j \in \phi$ $(i \neq j)$, the set of elements is anonymous. For example, assume that 10 nodes exist in an anonymous zone, and one of them is the receiver of a message. If a broadcast scheme is an anonymous protocol, then the receiver is not identifiable among the 10 nodes. In other words, any node in the set has the same probability of being the receiver.

Let ϕ' be a set of suspicious elements in the system (in this case, ϕ' is a set of nodes), $H(\phi')$ be the entropy of the system, and H_{max} be the maximal entropy that the system can achieve. Then, the degree of anonymity is defined as $D(\phi') = H(\phi')/H_{max}$. Computing p_i in Equation (3) is application-dependent. The definitions of anonymity for source and destination nodes are modeled in Section 4.2.

Attack 3 (The Path Deanonymizing). *An adversary tries to discover the set of onion groups that the copies of a message shall travel. Should r_k be compromised, an adversary will conclude that any copy of a message shall travel one of the onion routers in group R_k at the kth hop.*

The degree of privacy protection against Attack 3 can be quantified by the path anonymity. Similar to the node anonymity, the entropy of the system is defined by $H(\phi)$, but the instance of ϕ is all of the possible paths between two nodes. In addition, the path anonymity is obtained by $D(\phi') = H(\phi')/\ H_{max}$, where the elements in ϕ' are suspicious paths. In anonymous routing in DTNs, the number of hops between the source and destination nodes is defined by $\eta = K+1$, where K is the number of intermediate onion relays. When no node is compromised, any of the nodes can be an element of the original path and $\phi = \phi'$. Hence, the size of ϕ' equals to $\frac{n!}{(n-\eta)!}$, and each element in ϕ' has the equal probability of being the actual path that a message travels. When some nodes are compromised, the adversary can confine the anonymity set, i.e., $\phi' < \phi$. How to compute $H(\phi')$ is again application-dependent, and we will quantify the path anonymity for the proposed framework in Section 4.4.

2.3 Discussion on Anonymity

The node and path anonymity introduced in the previous sections is essentially computed by the ratio between the concrete entropy of a node/path and the maximal entropy. For any instance of anonymity models, the anonymity as well as the maximal anonymity range from 0 to 1. Both of them are decreasing functions with respect to the number of compromised nodes. In addition, anonymity decreases faster than maximal anonymity, when the number of compromised nodes increases.

An important aspect of normalized anonymity is that it tells us whether or not compromising some elements in an anonymity set affects the privacy of the others, and this differs from the probability-based metric.

For example, consider that an anonymity set with 10 nodes contains one source node. When no node is compromised, all the 10 nodes can be the source node with the equal probability of $1/10$, and thus, the anonymity of the source node is 1. Assume that one of the nodes, which is not the source node, is compromised. If this does not affect the other nodes, the anonymity set contains 9 nodes, and the source node can be identified with a probability of $1/9$ by the probability-based metric. In the entropy-based metric that we introduced, the achievable maximal entropy of the system also changes, since the size of an anonymity set is at most 9. As a result, the anonymity of the source node still equals to 1, since compromising one node does not divide the anonymity set, i.e., compromising one node does not affect the privacy of the others. When compromising one node affects the others, the anonymity set will contain a smaller number of nodes, resulting in a smaller node anonymity.

In the routing problem, some piece of information about the source/destination node and the path may, unfortunately, be leaked. The entropy-based anonymity metric is useful in measuring how strong an anonymous protocol is when some nodes are compromised. This is why we apply the entropy-based metric in this paper.

2.4 Anonymous Epidemic Routing

Each node is assumed to have its private key and the public key of the other nodes. Let v_s be the node who wishes to deliver message m to destination node v_d. The message header includes the message ID and the end-to-end deadline, denoted by $m.id$ and $m.T$, respectively. Note that $m.id$ can be either a unique sequence number or a random number, and m cannot be deduced from $m.id$. For simplicity, we denote the header of m by $m.hdr := (m.id, m.T)$. First, v_s encrypts m by v_d's public key, say PK_d. Let σ be the ciphertext computed by $Enc(PK_d, ID(v_d)||m)$, where $ID(v_d)$ is the ID of node v_d and $||$ is a concatenation. In other words, only when v_d successfully decrypts σ, can v_d confirm that it is the corresponding destination node. The information about the source node is not included in the header, as to preserve source anonymity. Such information should be stored in m so that only v_d can tell where the message comes from.

Afterwards, a pair of $m.hdr$ and σ is sent based on Epidemic routing. Consider that node v_i has encrypted message $(m.id, \sigma)$ and meets another node v_j. Nodes v_i and v_j check whether this is the first time for v_j to receive σ by exchanging $m.id$. If v_j has received m previously, no action will be taken. If it is the first time for v_j to see $m.id$, v_i forwards $(m.hdr, \sigma)$ to v_j. If the receiver, v_j, is the destination, it successfully decrypts σ by computing $Dec(SK_d, \sigma)$. Otherwise, v_j continues the Epidemic process. For message m, the end-to-end deadline is initialized by parameter T, and m is discarded if the deadline has passed.

In the case of mobile ad hoc networks, v_d will also broadcast m to pretend that it is not the destination against the location-based deanonymization attack. However, in DTNs, a network is constructed by contact events, and thus, such an attack is not of concern. The pseudo code of anonymous Epidemic routing is described in Algorithm 1.

Algorithm 1. AE(v_s, v_d, m, T)

```
/* v_s does the following */
v_s sets v_d, m.id, and m.T ← T.
v_s sets m.hdr ← (m.id, m.T)
v_s computes σ ← Enc(PK_d, ID(v_d)||m).
/* v_i does the following at a contact with v_j */
v_i and v_j establish a secure link.
if v_j has not seen m.id then
    v_i sends (m.hdr, σ) to v_j.
/* When v_j is v_d, it does the following */
if v_j = v_d then
    v_d obtains m by ID(v_d)||m ← Dec(SK_d, σ), return SUCCESS.
/* Error handling */
if m is not delivered within T then
    v_i discards m, and returns FAIL.
```

2.5 Restricted Epidemic Routing Mode

For the proposed protocol to use anonymous Epidemic routing as a subroutine, we extend Algorithm 1 in the previous section as anonymous restricted Epidemic routing. Specifically, not only source and destination nodes, but also any two relay nodes like onion routers, for example, can use anonymous Epidemic routing. One example is to apply Epidemic as a variant of partial flooding, which is used in the zone-based routing.

The first extension is the introduction of a *zone*, where Epidemic routing is performed. Note that the zone in anonymous Epidemic routing is the entire contact graph. In the zone-based protocol for ad hoc networks, an anonymous zone is defined by euclidean distance or topological distance, i.e., the number of hops. However, the network representation of a DTN does not indicate the physical location of nodes, and so euclidean distance cannot be applied. For topological distance, a small value of time to live (TTL), say two or three hops, is normally used as an anonymous zone. The small TTL value will, unfortunately, make a protocol susceptible to the topology-based deanonymization attack. Therefore, in order to anonymously control the area of Epidemic zone, the zone deadline, which is denoted by t, is used. Here, the value of t is much smaller than the end-to-end deadline T, but is large enough for a message to reach the expected receiver within the deadline with high probability.

Let v_i be the node with message m, and v_j be the expected receiver. We define τ as the probability that v_j receives m within the zone deadline, t. Here, τ is a system parameter required by v_i, and t is dynamically computed from a given τ. Let $P_{i,j}(t)$ be the probability that v_i and v_j have a contact

within t. If we set τ to be $P_{i,j}(t)$ in Equation (1), i.e., the probability that v_i has a contact with v_j within t, the appropriate zone deadline t can then be computed as shown

$$t = -\frac{ln(1-\tau)}{\lambda}. \quad (4)$$

In the case of anycast-like forwarding, i.e., a message transmission from node v_i to any node r in R_k, we may set λ to be $\sum_{\forall r \in R_k} \lambda_{i,r}$.

The second extension is the introduction of a group, where any node in the next group can serve as a relay. Let $v_i \in R_k$ be the node which wishes to relay message m to any node $r_{k,j} \in R_{k+1}$. We define $GID(R_{k+1})$ as a group ID of R_{k+1}.

The message header of RER includes the message ID, the end-to-end deadline, and the zone deadline, which is denoted by $m.hdr := (m.id, m.T, m.t)$. The ciphertext σ_k is defined using the group key with the corresponding group ID being concatenated, i.e., $Enc(GK_{k+1}, GID(R_{k+1})||\sigma_{k+1})$, where GK_{k+1} is the group key of R_{k+1}.

The RER works as follows. At every contact between v_i and v_j, v_j checks if it has seen $m.id$ before. If so, they do nothing. Otherwise, v_i sends σ_k to v_j. Then, Epidemic routing is repeated until the zone deadline, $m.t$, has expired. If $v_i \in R_k$, it identifies itself as a next-relay by the group ID. Using the corresponding group key of R_{k+1}, v_j peels off a layer of the encrypted message, i.e., $GID(R_{k+1})||\sigma_{k+1} \leftarrow Dec(GK_{k+1}, \sigma_k)$. Only a member of the corresponding group RK_{k+1} can obtain σ_{k+1}. If either a zone or end-to-end deadline has passed, m is discarded.

The pseudocode of RER is presented in Algorithm 2.

Algorithm 2. RER(v_i, R_{k+1}, σ_k, τ, T)

```
1: /* v_i ∈ R_k does the following */
2: v_i sets m.t_{k+1} from τ.
3: v_i sets m.hdr ← (m.id, m.T, m.t_{k+1})
4: /* v_i does the following at a contact with v_j */
5: v_i and v_j establish a secure link.
6: if v_j has not seen m.id then
7:    v_i sends (m.hdr, σ_k) to v_j.
8:    if v_j ∈ R_{k+1} then
9:       v_j computes GID(R_{k+1})||σ_{k+1} ← Dec(SK_{GK_{k+1}}, σ_k).
10:      return SUCCESS;
11: /* Error handling */
12: if m is delivered within neither m.t_k nor m.T. then
13:    v_i discards σ_k from its buffer.
```

2.6 Zone-Based Anonymous DTN Routing

A zone-based anonymous DTN routing can be constructed from Epidemic and spray-and-wait protocol, each of which is replaced with partial flooding and unicast routing (e.g., geographical routing). That is, Algorithm 2 is used for message transmission from the source to its proxy and from the destination proxy to the destination. Between the proxies, source/binary spray-and-wait is used.

An anonymous spray-and-wait forwarding between two proxies is basically the same as the one used between two intermediate relays in onion-based routing. Based on these ideas, we construct a zone-based anonymous DTN routing, as follows.

The message header of ZBAR includes the message ID, the end-to-end deadline, the zone deadline, and the mode, denoted by $m.hdr := (m.id, m.T, m.t, mode)$. The source node v_s selects the source and destination proxies, say r_s and r_d, respectively. Then, $\sigma_d \leftarrow Enc(PK_d, ID(v_d)||m)$, $\sigma.r_d \leftarrow Enc(PK_{r_d}, ID(r_d)||\sigma_d)$, and $\sigma.r_s \leftarrow Enc(PK_{r_s}, ID(r_s)||\sigma_{r_d})$ are computed. The encryption structure is the same as that of an onion, where v_d can decrypt the encrypted data after r_s and r_d peel off the outer layers. In addition, $m.id$ and $m.t$ are calculated. An encrypted message is composed of ($m.hdr$, $\sigma.r_s$). The value of $mode$ could be either the restricted epidemic RE or spray-and-wait SW forwarding mode.

From v_s to r_s, restricted Epidemic routing is performed. A receiving node first attempts to decrypt σ_{r_s}. If it fails, the node is not the proxy, and the Epidemic process continues as long as $m.t$ has not expired. Otherwise, r_s decrypts the outmost layer of the onion, and it switches the mode of the message to the spray-and-wait forwarding mode. From r_s to r_d, a message ($m.hdr$, $\sigma.r_d$) is forwarded by anonymous spray-and-wait with single-copy forwarding. When the destination proxy, r_d, receives the message, the corresponding layer of $\sigma.r_d$ is decrypted, and $m.t$ is computed. Then, the restricted Epidemic routing for the message ($m.hdr$, $\sigma.d$) is again performed. The destination identifies itself by successfully decrypting σ_d using SK_d. The pseudo code of ZBAR is presented in Algorithm 3.

Algorithm 3. ZBAR(v_s, v_d, m, T)

```
1: /* v_s does the following */
2: v_s selects two proxies, r_s and r_d.
3: v_s computes σ_d ← Enc(PK_d, ID(v_d)||m), σ.r_d ← Enc
   (PK_{r_d}, ID(r_d)||σ_d), and σ.r_s ← Enc(PK_{r_s}, ID(r_s)||σ_{r_d}).
4: v_i sets m.t_{k+1} from τ.
5: v_s sets m.hdr ← (m.id, m.T, m.t, RE).
6: v_s executes Algorithm 2 RER(v_s, {r_d}, m, m.t).
7: /* r_s meets node v_i. */
8: v_i and v_j establish a secure link.
9: if v_i identifies itself as r_d then
10:    v_i computes ID(r_d)||σ_d ← Dec(SK_{r_d}, σ.r_s).
11:    v_i sets m.t and m.mode ← RE.
12:    v_i executes Algorithm 2 RER(r_d, {v_d}, m, T).
13: if m.t expires then
14:    v_i removes m from its buffer.
15: /* v_d does the following */
16: v_d obtains m by ID(v_d)||m ← Dec(SK_d, σ), return
    SUCCESS.
17: /* Error handling */
18: if m is not delivered in T then
19:    v_i discards m, and returns FAIL.
```

3 Framework of Anonymous Routing

3.1 Motivation and Basic Idea

We first point out two problems regarding the existing anonymous routing with onion-based [9], [16] and threshold-based [11] schemes for DTNs. The first issue is that the source (or the destination) node is anonymous only within its onion group. Hence, the identity of a source or destination node will be revealed if the first or the last onion router is compromised. The second issue is that an intermediate onion router knows the previous and subsequent onion routers. These problems significantly reduce the node/path anonymity and the path untraceability.

To alleviate the first problem, we have proposed the ZBAR protocol based on zone-based routing [6] in Section 2. However, the second issue still remains unresolved with the zone-based approach. To preserve anonymity, an intermediate onion router should not know the exact previous and next forwarding nodes. In addition, the first and last onion routers should not know they are located at the edge of an onion path.

To achieve these desirable properties, we propose a Framework for Anonymous Routing for DTNs that subsumes all the anonymous routing protocols. That is, the source node sets up a set of onion routers, and then all nodes on the path forward a message with the restricted Epidemic routing. Note that the proposed FAR does not just combine different anonymous routing protocols, but creates a framework that subsumes all the protocols. In other words, FAR serves as either an anonymous Epidemic, ZBAR, or onion-based protocol, when its parameters are set differently. By adjusting the parameters appropriately, FAR enjoys the advantages of all these anonymous routing protocols.

3.2 The Protocol Overview

In this section, we describe the high-level overview of the proposed FAR. Let v_s be the source node which wishes to deliver message m to destination v_d. The routing parameters, $\{K, L, G, F\}$, are selected by v_s, where K is the number of onion relays that m shall travel, L is the number of copies, G is the size of the onion group, and $F = \{f_1, f_2, \ldots, f_K\}$ is a set of forwarding modes. A forwarding mode can be either restricted Epidemic RE or source spray-and-wait SW. Similar to RER, the message header includes the message ID, the end-to-end deadline, and the zone deadline. The forwarding mode is protected by the layered encryption.

After initializing the routing parameters, v_s randomly selects a set of K onion groups ($K \geq 0$), along which m travels and creates an onion. When $K = 0$, no intermediate onion is used as AE does. How to forward m from one node to another differs, depending on the forwarding mode utilized. In the RE mode, a node, say v_i, with m sends a copy to all the nodes contacted by v_i within the zone deadline. In the SW mode, a node with m sends a copy to any node in the next onion group as long as the tickets (the number of copies allowed to duplicate) are available. The forwarding mode for the ith hop is determined by f_i. When a node, say r_j, in the next onion group R_{i+1} receives m, the outer layer of the onion is peeled off by the corresponding group key. At this time, $m.id$ is randomly generated to improve the privacy against path tracing and path deanonymizing attacks. Then, the forwarding process continues based on the forwarding mode specified in f_{i+1}. This process is repeated until the destination v_d receives m.

3.3 Framework of Anonymous Routing

To initialize the anonymous network system, an approach for onion group routing, proposed in [9], can be used. The nodes in a network are divided into $\lceil n/G \rceil$ groups, where G is the average number of nodes in a group. For simplicity, we assume n to be divisible by G. Nodes in the same group are assumed to be able to encrypt/decrypt the corresponding layer of an onion by common secret or public/private keys.

The pseudo code of FAR is provided in Algorithm 4. As inputs, the system parameters $\{K, L, G, F\}$ and the end-to-end deadline, T, are selected by v_s. Lines 1 to 6 represent the initialization phase. The source node, v_s, randomly selects a set of onion groups by which m travels. First, v_s obtains σ_0 by computing $Enc(PK_d, ID(v_d)||m)$ with v_d's public key. Then, an encrypted onion is created by applying a set of group keys associated with R_i, i.e., $\sigma_i \leftarrow Enc(GK_{R_i}, GID(R_{i+1})||f_{i+1}||\sigma_{i+1})$ for $1 \leq i \leq K$. Note that the forwarding mode is encrypted to prevent adversaries from distinguishing the first/last onion relay from the intermediate onion relays. Finally, v_s sets the timer, denoted as $\sigma.t$, by Equation (1).

Algorithm 4. FAR(v_s, v_d, m, K, L, G, F, T, τ)

```
1:  /* v_s does the following */
2:  v_s selects K onion groups.
3:  v_s computes σ_{K+1} ← Enc(PK_d, ID(v_d)||m).
4:  for i from K to 1 do
5:     v_s computes σ_i ← Enc(GK_{R_{i+i}}, GID(R_{i+1})||f_{i+1}||σ_{i+1}).
6:  v_s computes σ_1.t_1 from τ.
7:  v_s executes RER(v_s, R_1, σ_1, T).
8:  /* On receiving σ_k from v_j ∈ R_{k-1}, v_i ∈ R_k does the
    following */
9:  if v_i ∈ R_k receives σ_k from v_i ∈ R_{k-1} then
10:    if v_i identifies itself as v_d then
11:       v_d obtains m by ID(v_d)||m ← Dec(SK_d, σ_k).
12:       returns SUCCESS.
13:    else
14:       v_i sets σ_k.t_k.
15:       if σ_k.f_k is RE then
16:          /* Restricted Epidemic mode */
17:          v_i executes RER(v_i, R_{k+1}, σ_k, T).
18:       else if σ_k.f_k is SW then
19:          /* Anonymous spray-and-wait mode */
20:          v_i forwards σ_k when it has a contact r ∈ R_{k+1} if r
             has not seen σ_k.
21: /* Error handling */
22: if m is not delivered in T then
23:    v_i discards m, and returns FAIL.
```

The forwarding process at the kth Epidemic zone is shown from Lines 8 to 20. For each zone, RER or spray-and-wait forwarding is executed until m reaches v_d. During the RE forwarding mode, σ is discarded if the zone deadline $\sigma.t$ expires. When the destination node, v_d, receives σ, it applies its private key to obtain the original message, m. If the destination does not obtain m by the deadline T, the routing process fails.

FAR subsumes Epidemic, zone-based, and onion-based anonymous routing protocols. The parameters ($K = 0, null, null, S = \{RE\}$) indicate an AE protocol, in which Epidemic is performed by hiding the source and destination nodes. In the case of (K, L, G, $\{f_1 = SW, f_2 = SW, \ldots, f_K = SW\}$), the protocol is reduced to onion-based routing. In addition, depending on G and L, the protocol can be onion ($G = 1$) or onion group ($G \geq 2$) routing with single/multi copies ($L = 1$ or $L \geq 2$). The configuration of ($K = 2, L = 1$, G, $\{f_1 = RE$, $f_2 = SW, \ldots, f_{K-1} = SW, f_K = RE\}$) serves as the ZBAR protocol.

4 Security Analyses

In this section, analytical models are built for traceable rate and node/path anonymity of the proposed FAR under

Attacks 1 and 2, respectively. Our analysis provides the closed form solutions to different metrics, by which DTN users select the system parameters that meet their security and privacy requirements. Note that the analyses of AE (Algorithm 1) and ZBAR (Algorithm 3) are trivial and thus omitted.

4.1 Approximation of Traceable Rate

The traceable rate is computed by Equation (2) against the path tracing attack defined in Attack 1. The proposed FAR employs anonymous Epidemic forwarding, and the path can be revealed only by the reverse order from the destination. Thus, the number of compromised segments C_{seg} in Equation (2) equals either 0 or 1. Let X be the random variable that represents the length of the compromised segments $c_{seg,1}$, then $E[X]$ can be computed by the geometric distribution with the limited number of trials. The probability of a node being compromised is c/n. Denoting $p = 1 - c/n$ and $q = c/n$, $E[X]$ can be obtained as follows:

$$E[X] = \sum_{i=1}^{\eta} iq^{i-1}p + \eta q^{\eta} \tag{5}$$

$$= qE[X] + \sum_{i=1}^{\eta} q^{i-1}p + \eta q^{\eta}. \tag{6}$$

By defining $\epsilon_1 = \sum_{i=1}^{\eta} q^{i-1}p$ and $\epsilon_2 = \eta q^{\eta}$, we will have

$$E[X] = \frac{n(\epsilon_1 + \epsilon_2)}{n - c}. \tag{7}$$

Since the traceable rate is weighted, we need to compute $E[X^2]$, which can be obtained as follows:

$$E[X^2] = \sum_{i=1}^{\eta} i^2 q^{i-1}p + \eta^2 q^{\eta} \tag{8}$$

$$= qE[X^2] + 2qE[X] + \sum_{i=1}^{\eta} q^{i-1}p + \eta q^{\eta} \tag{9}$$

$$= \frac{n(n+c)(\epsilon_1 + \epsilon_2)}{(n-c)^2}. \tag{10}$$

Since $(c_{seg,1})^2 = E[X^2]$, the traceable rate is computed by $\frac{1}{\eta^2}E[X^2]$, and therefore, we derive

$$P_{trace} = \frac{1}{\eta^2}\left\{\frac{n(n+c)(\epsilon_1 + \epsilon_2)}{(n-c)^2}\right\}. \tag{11}$$

The number of hops, η, increases in proportion to the value of the number of onion routers, K. This is because all the messages must travel at least one onion relay in a particular onion group, in the predefined order. Thus, in a high-level view, we can consider that one hop from an onion router to the next onion router is a link. For a DTN user to find an appropriate routing parameter K, we may simply set η to be $K + 1$.

4.2 Source and Destination Anonymity

Quantifying anonymity is application-dependent, and thus, we model source and destination anonymity as follows. In FAR, the anonymity of source and destination nodes are computed in the same way, and only two parameters, the number of nodes n and the number of compromised nodes c, are related to this metric. In the case of a node not being compromised, the node is identified among the non-compromised nodes with the probability of $1/(n-c)$. Thus, the maximal entropy of a node is defined as

$$H_{node,max} = -\sum_{\forall \text{nodes in } \phi} \frac{1}{n-c}\log_2\left(\frac{1}{n-c}\right). \tag{12}$$

If a node is compromised, it is identified with 100 percent probability. In other words, the anonymity set to which the node belongs is of size one. Otherwise, it is still anonymous among the set with size $(n-c)$. Let ϕ' be a set of suspicious nodes. The entropy of a node, denoted by $H_{node}(\phi')$, is obtained by

$$H_{node}(\phi') = \begin{cases} -\sum_{\forall \text{nodes in } \phi'} \frac{1}{1}\log_2\left(\frac{1}{1}\right) = 0 & \text{if compromised} \\ -\sum_{\forall \text{nodes in } \phi'} \frac{1}{n-c}\log_2\left(\frac{1}{n-c}\right) & \text{otherwise.} \end{cases} \tag{13}$$

The node anonymity of FAR is formally defined as follows.

Definition 1 (Node Anonymity). *For given suspicious node set ϕ' and the maximal entropy of a system $H_{node,max}$, the node anonymity is defined as*

$$D_{node}(\phi') = \frac{H_{node}(\phi')}{H_{node,max}}. \tag{14}$$

Here, $H_{node,max}$ and $H_{node}(\phi')$ are computed by Equations (12) and (13), respectively.

In short, the anonymity of the source/destination node equals to 1, if it is not compromised and 0 otherwise.

4.3 Approximation of Source and Destination Anonymity

We will formulate the closed form solution to approximate the node anonymity from the system parameters. For given n and c, the expected entropy of a node, denoted by $\tilde{H}_{node}(\phi')$, is formulated by

$$\tilde{H}_{node}(\phi') = -\sum_{\forall \text{nodes in } \phi'} \left(1 - \frac{c}{n}\right) \cdot \frac{1}{n-c}\log_2\left(\frac{1}{n-c}\right). \tag{15}$$

Here, $|\phi'| = n - c$. Therefore, we will derive the approximate solution to the node anonymity as follow:

$$\tilde{D}_{node}(\phi') = \frac{H_{node}(\phi')}{H_{node,max}} = 1 - \frac{c}{n}. \tag{16}$$

4.4 Path Anonymity

Path anonymity is the state of not being identifiable among a set of candidate paths. Epidemic-based and onion-based with L-copy forwarding generally return multiple paths between end hosts. If some of the nodes on a path are compromised, the path anonymity of the other paths may decrease. Thus, this metric differs from the traceable rate in the sense that a set of paths, along which a message travels, affects one another. For example, consider that there are two paths along which

two of the same onion groups travel, and one of the onion routers, say $r_{k,1} \in R_k$, in a path is compromised. Even if none of the nodes on the other path is compromised, an adversary knows that a packet travels with one of the nodes in R_k at the kth hop on the second path. Thus, the size of the anonymity set of possible paths decreases.

The path anonymity for DTNs is proposed in [10], and the same definition can be applied to the proposed FAR. As we discussed in Section 2.2, there are $\frac{n!}{(n-\eta)!}$ possible η-hop paths between two nodes, and each element in the all possible paths set ϕ has the equal probability of being the actual path that a message travels. Thus, the maximal entropy for the path anonymity, denoted by $H_{path,max}$, is formulated

$$H_{path,max} = -\sum_{\forall \text{paths in } \phi} \frac{(n-\eta)!}{n!} \log_2 \left(\frac{(n-\eta)!}{n!} \right). \quad (17)$$

Path anonymity cannot be computed, until the actual paths along which a message and its copies travel are identified. In FAR with the RE mode, either source or destination node must be compromised for an adversary to confine the anonymity set of the possible paths between them. If at least one of the nodes in the kth onion group as well as all the intermediate onion groups from the source (or destination) node are compromised, the kth onion relay is anonymous within G nodes, instead of $n-k$. In addition, if the onion relay at the kth hop is compromised, the kth node is identified. To derive the entropy of the system, the following four cases are considered.

1) *In the case that no node on the paths is compromised:* The entropy of the system will equal to the maximal entropy.
2) *In the case that the source node is compromised, but not the destination node:* Let c_1 be the number of consecutive onion groups in which at least one of the members is compromised from the source node, but not the relay node itself. In addition, we define c_2 as the number of compromised relay nodes, which serve as the intermediate nodes. Here, $c_1 + c_2 < c$ and $c_1 + c_2 < \eta$ hold. Then, the anonymity set size of a path will be $G^{c_1} \cdot \frac{n}{(n-\eta+c_1+c_2)!}$.
3) *In the case that the destination node is compromised, but not the source node:* Let c_3 and c_4 be the number of consecutive onion groups in which at least one of the members is compromised from the destination node, but not the relay node itself, and be the number of compromised relay nodes, which serve as the intermediate nodes, respectively. Using c_3 and c_4, the anonymity set size is computed by exactly the same way as Case 2.
4) *In the case that both the source and destination nodes are compromised:* The anonymity set size is computed by combining Cases 2 and 3. Then, using $(c_1 + c_3)$ and $(c_2 + c_4)$, the same way as Case 2 is applied.

Note that when $c_1 = c_2 = c_3 = c_4 = 0$, Cases 2, 3, and 4 are reduced to Case 1 as its definition. The entropy of the system, denoted by $H_{path}(\phi')$, is obtained

$$H_{path} = -\sum_{\forall \text{paths in } \phi'} \frac{1}{|\phi'|} log_2 \left(\frac{1}{|\phi'|} \right). \quad (18)$$

Here, $|\phi'|$ is computed based on the above four cases as follow:

$$|\phi'| = \begin{cases} 1 & \text{if Case 1 holds} \\ G^{c_1} \cdot \frac{n}{(n-\eta+c_1+c_2)!} & \text{if Case 2 holds} \\ G^{c_3} \cdot \frac{n}{(n-\eta+c_3+c_4)!} & \text{if Case 3 holds} \\ G^{c_1+c_3} \cdot \frac{n}{(n-\eta+c_1+c_2+c_3+c_4)!} & \text{otherwise.} \end{cases} \quad (19)$$

The path anonymity of FAR is formally defined as follows.

Definition 2 (Path Anonymity). *For given suspicious node set ϕ' and the maximal entropy of a system $H_{path,max}$, the path anonymity is defined as*

$$D_{path}(\phi') = \frac{H_{path}(\phi')}{H_{path,max}}. \quad (20)$$

Here, $H_{path,max}$ and $H_{path}(\phi')$ are computed by Equations (17) and (18), respectively.

4.5 Approximation of Path Anonymity

For the given system parameters (n, K, and G) and the number of compromised nodes c, we will formulate the closed form solution for the path anonymity of FAR with the RE mode. According to [10], the expected path anonymity, denoted by $\tilde{D}_{path}(\phi')$, can be defined as

$$\tilde{D}_{path}(\phi') = \frac{(\eta - c_o)(\ln(n) - 1) + c_o \ln(G)}{\eta(\ln(n) - 1)}, \quad (21)$$

where c_o is the average number of compromised onion groups on a path. Note that an onion group is compromised if at least one of the nodes in the onion group is compromised. Our model is very different from the onion-based routing [10] in how c_o is computed.

Let Y be the random variable that represents the number of compromised groups on a path. Denoting $p' = (1 - c/n)^L$ and $q' = 1 - (1 - c/n)^L$, we can derive

$$E[Y] = \sum_{i=1}^{\eta} i q'^{i-1} p' + \eta q'^{\eta-1} \frac{c}{n} \quad (22)$$

$$= \frac{n(\gamma_1 + \gamma_2)}{n - c}, \quad (23)$$

where $\gamma_1 = \sum_{i=1}^{\eta} q'^{i-1} p'$ and $\gamma_2 = \eta q'^{\eta-1} \frac{c}{n}$. For the path anonymity, we simply set η to be $K + 1$. Let $c_o = E[Y]$, and then we can obtain the path anonymity from Equation (21).

Note that the path anonymity can be computed only when at least one of message copies is delivered to its destination. Otherwise, identifying a path is impossible. Therefore, the path anonymity is independent from the zone deadline τ and the contact frequency λ, both of which affect the delivery rate and delay.

4.6 Relation Between Node and Path Anonymity

The node and path anonymity for FAR quantified above have some relations between them. In this section, we argue that FAR provides higher privacy than onion group routing (OGR) and ZBAR do in this respect.

In FAR, an adversary must identify either source or destination node to reduce the anonymity set of the paths along which a message and its copies travel as discussed. Thus, breaking the node anonymity is a necessary condition for breaking the path anonymity. However, it is not a sufficient condition. Thanks to the property of the RE mode, no node in the intermediate onion groups knows whether or not it is the first or last relay, and thus, all of the other nodes can be the source or destination. Therefore, breaking the path anonymity does not help an adversary to break the node anonymity.

In OGR and ZBAR, breaking the path anonymity reduces the node anonymity, and vice versa. To be specific, in OGR, any node in the first intermediate onion group R_1 receives a message from the source node, and thus, compromising more than two nodes in R_1, allows an adversary to uniquely identify the source node. Similarly, any node in the last intermediate onion group R_K forwards a message to the destination node. Compromising more than two nodes in R_K results in disclosing the identity of the destination node. On the other hand, identifying the source/destination node helps an adversary to confine the first/last relay node within the size G, which results in a smaller path anonymity.

In ZBAR, the forwarding mode is switched at the first and last onion groups. Hence, should the first or last relay node be compromised, the source/destination node is anonymous within an Epidemic zone. This indicates that the source/destination node is anonymous within that zone, and compromising the path anonymity reduces the node anonymity.

Therefore, from the above discussion, the proposed FAR provides a higher degree of anonymity in the sense that an adversary must identify either the source or destination node to break the path anonymity, while breaking either the node or path anonymity reduces the other in OGR and ZBAR.

5 Performance Evaluation

In this section, computer simulations are conducted to evaluate the performance of the proposed scheme. A set of the proposed protocols, AE, ZBAR, FAR, as well as the existing, onion-group routing (OGR) [10] are implemented. To the best of our knowledge, OGR in [10] is the latest and the most viable anonymous routing protocol for DTNs. Note that AE, ZBAR, and OGR are special cases of FAR. For simplicity, we refer to them as AE, ZBAR, and OGR instead of FAR with specific parameters.

5.1 Simulation Configurations

In our simulations, two scenarios are considered. One is a randomly generated contact graph for evaluating the proposed schemes in large-scale DTNs; the other is a real contact trace [12] to demonstrate that our FAR works well in realistic environments.

Random Graphs. A contact graph with 200 nodes is generated by assigning inter-contact times to each pair of two nodes. The inter-contact time is exponentially distributed with parameter $\lambda_{i,j}$ for a pair of nodes v_i and v_j ($i \neq j$). The initial value of $1/\lambda_{i,j}$ is generated by the normal distribution in which the mean and variant are set to be 360 and 720 time units, respectively. The group size is set to be 10 or 20, the number of onion routers is set to be 3, and the number of copies is set to be the same as the group size (i.e., $L = g$). The message deadline T is randomly selected in the range of 10 to 2,000 time units, and the percentage of compromised nodes is set to be $0\% \leq c/n \leq 50\%$, where c is the number of compromised nodes and n the total number of nodes. The simulation parameters are summarized in Table 2.

TABLE 2
Simulation Parameters for Random Graphs

Parameter	Value (default value)
The number of nodes	200
The inter-contact time	0 to 720 unit time
The group size	10 or 20
The number of onion routers	3
The number of copies	10 or 20
The message due	10 to 2,000 unit time
The threshold of RER	0.8 to 0.99
The % of compromised nodes	0% to 50% (10%)

The source and destination nodes are randomly selected, and each node runs an anonymous routing protocol with given parameters. If a message is delivered from source to destination within the deadline, T, the message delivery is successful. For a given percentage of compromised nodes, i.e., c/n, randomly selected nodes of such a portion are marked as compromised, and then security metrics are computed. For each set of parameters, 1,000 contact graphs are generated for the simulation.

Real Traces. CRAWDAD dataset cambridge/haggle [12] contains a set of contact trace experiments. In our simulations, Experiments 2 and 3, the so-called Cambridge and Infocom 2005 traces, are used as inputs. In these scenarios, we only consider the contacts between mobile nodes, i.e., iMotes, and omit contacts among stationary nodes and external devices. There are 12 and 41 mobile nodes in the Cambridge and Infocom 2005 traces, respectively. Each piece of contact information contains two node IDs, the time that the two nodes meet, the time that they lose a connection, the number of contact times, and the elapsed time of the last time the two nodes met. Contact events are recorded in the order of seconds. Since the contact events are traced over three to five days, there exist time periods in which there is no contact, e.g., off-business hours and night time. Thus, a source node is assumed to initiate a message transmission at any time after it has a contact with any node, which implies that message delivery starts during business hours, but not at night time.

For a given trace file, the number of nodes and inter-meeting times are calculated. The other simulation parameters, i.e., K, L, G, c, and T are set in the same way as the random graphs. For each trace file, 500 different sets of source, destination, and intermediate onion routers are randomly selected, and the average performance is computed.

5.2 Results Using Synthesize Graphs

Fig. 1 shows the cumulative distribution function (CDF) of the delivery rate with respect to the deadline. AE results in the fastest delivery, and the CDF of FAR reaches 0.95 within 70 time units. This indicates that Epidemic-based routing delivers a message much faster than does OGR. ZBAR incurs slightly longer delay than AE and FAR, since it forwards a message by the stop-and-wait between the first and last onion routers. In addition, it is intuitive that a larger group size

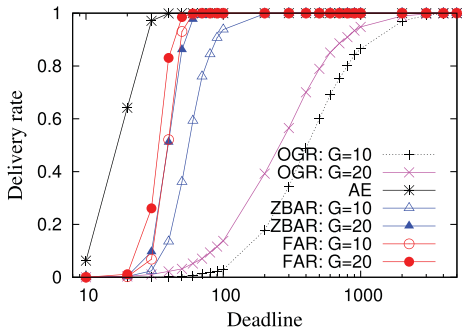


Fig. 1. The CDF of delivery rate.

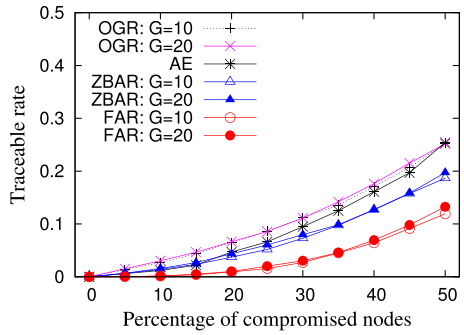


Fig. 2. The traceable rate.

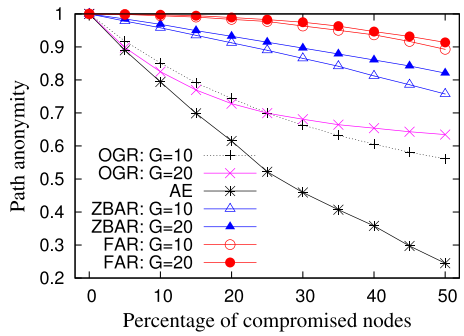


Fig. 3. The path anonymity.

leads to faster message delivery, and this can be clearly observed in this figure.

Fig. 2 illustrates the traceable rate with respect to the percentage of compromised nodes. Since every path is considered independently, the group size does not affect the traceable rate. In addition, it is intuitive that the traceable rate gradually increases as the percentage of compromised nodes increases. In the proposed FAR, a routing path can be traced only by the consecutive compromised segments from the destination node, and thus, the traceable rate is much lower than that of the other protocols. From the figure, the traceable rate resulting from FAR is at most half of that by OGR. Similar to OGR, ZBAR forwards a message between intermediate onion routers by spray-and-wait forwarding. As a result, the traceable rate of ZBAR is higher than that of FAR, but smaller than that of OGR.

Fig. 3 presents the path anonymity with respect to the percentage of compromised nodes. The path anonymity by FAR is mostly 1.0 when the percentage of compromised nodes is less than or equal to 20 percent, and even when the percentage of compromised nodes is 50 percent, FAR maintains a 0.9 path anonymity. Since FAR and ZBAR always result in higher path

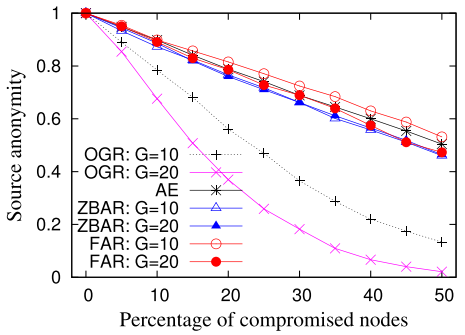


Fig. 4. The source anonymity.

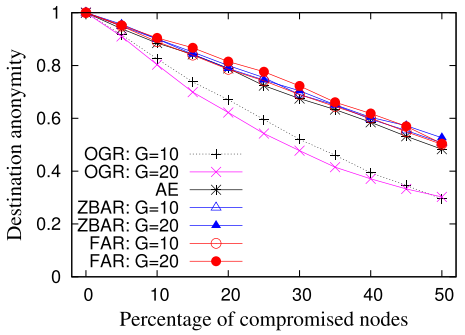


Fig. 5. The destination anonymity.

anonymity than OGR, we can conclude that Epidemic-based forwarding, except AE, preserves a higher degree of privacy. AE results in the lowest path anonymity, even though it is an Epidemic-based protocol. This is because the path is uniquely identified by adversaries, should either the source or destination node be compromised.

Figs. 4 and 5 illuminate the source and destination anonymity with respect to the percentage of compromised nodes. In OGR, the large group size results in low source and destination anonymity due to its design issue. On the contrary, the source and destination anonymity resulting from FAR is independent of the group sizes, since each of the communications between onion routers is performed by the RER (Algorithm 2). This indicates that the onion routers are indistinguishable if they are the first/last onion routers, or the intermediate ones. Hence, unless the source and destination nodes are compromised, adversaries cannot confine the anonymity set in which the source/destination is included. Similarly, AE reveals no information about the identity of source and destination nodes unless they are compromised. In ZBAR, the size of the anonymity set to which the source/destination belongs decreases if at least one of the nodes in the first/last onion groups is compromised. Therefore, ZBAR results in slightly smaller node anonymity than FAR and AE. For OGR, the destination anonymity is better than the source anonymity. This is because the destination can be ambiguous in identifying the onion group as the destination, as proposed in [9]; however, this technique cannot be applied to the source node.

Fig. 6 depicts the amount of message forwarding, introduced by anonymous protocols with respect to the size of onion groups. Note that AE does not use intermediate onion routers, and so it is independent of the group size. Apparently, Epidemic-based routing, i.e., AE, ZBAR, and FAR, incur more message overhead than OGR. FAR introduces the greatest amount of message forwarding, as it forwards a

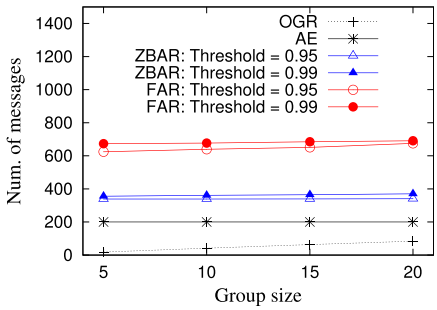


Fig. 6. The number of messages.

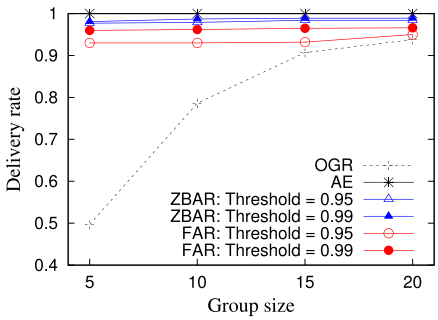


Fig. 7. The delivery rate.

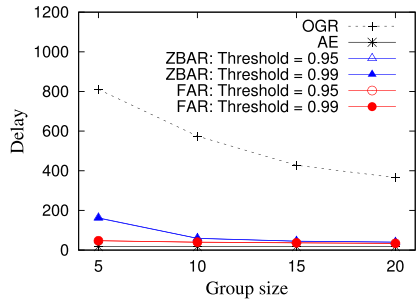


Fig. 8. The delay.

message by RER (Algorithm 2) at every communication between two onion routers. However, we claim that achieving the highest privacy in terms of the traceable rate and node/path anonymity with FAR is still worth a large amount of control overhead.

Fig. 7 shows the delivery rate with respect to the group size, where the number of message copies equals the group size, i.e., $L = G$. Since the number of message copies is bounded by the group size, this figure presents how the introduction of more message overhead increases the delivery rate. For any group size, FAR results in the higher delivery rate than ZBAR and OGR. While the delivery rate of FAR slightly increases when the group size increases, the differences are very small. This means that setting $L = G = 5$ provides sufficient forwarding opportunities. On the other hand, due to the limited forwarding opportunities, OGR results in very low delivery rate when the group size is small.

Fig. 8 illustrates the delay with respect to the group size, where the number of message copies equals the group size, i.e., $L = G$. Thanks to the property of Epidemic forwarding, the delays of FAR and ZBAR are much shorter than that of OGR. Intuitively, introducing more message overhead, the delay becomes shorter. The delay slightly decreases, when the group size increases. Similar to the delivery rate, however, significant improvement is not observed. In addition,

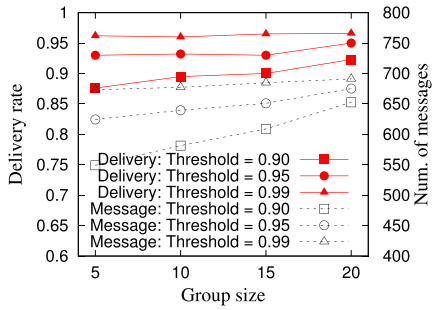


Fig. 9. The tradeoff between performance and message overhead.

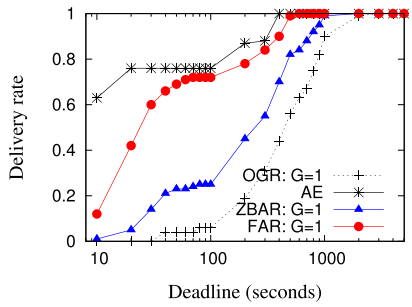


Fig. 10. The delivery rate w/ the Cambridge trace.

since the forwarding opportunities are limited in OGR, the delay of OGR is much longer than FAR and ZBAR.

Fig. 9 presents the delivery rate and the number of messages of FAR under different thresholds with respect to the group size, where the number of message copies equals the group size, i.e., $L = G$. This figure clarifies why the performance increases when the message overhead becomes large. As can be seen in the figure, the number of messages gradually increases as the group size increases. On the other hand, the delivery rate slightly increases when the group size is large. In other word, the performance increases if more message overhead is introduced. This figure indicates that having $L = G = 5$ is sufficient for higher delivery rate. However, as discussed from Figs. 2, 3, 4, and 5, adding more message overhead still plays a critical role for the privacy perspective.

5.3 Results Using Real Traces

The Cambridge trace, i.e., Experiment 2 in [12] is relatively small-scale and dense (12 mobile nodes), and thus, the number of onion routers and the group size are set to be $K = 3$ and $G = 1$, respectively. The number of copies in OGR and in the stop-and-wait mode in ZBAR are set to be $L = G$. Note that having more than one copy in OGR and ZBAR does not help message delivery when $G = 1$. On the other hand, the Infocom 2005 trace (i.e., Experiment 3 in [12]) is a medium-sized contact network with 41 mobile nodes. The number of onion routers, the group size, and the number of copies are set to be $K = 3$, $G = 5$, and $L = G$, respectively.

Figs. 10 and 11 show the delivery rate for different protocols resulting from the Cambridge and Infocom 2005 traces, respectively. In Fig. 10, the proposed FAR achieves faster delivery than ZBAR and OGR. In addition, the message delivery is mostly completed within 1,000 seconds, which is much faster than the results shown in Fig. 11. This is because the Cambridge trace is generated by the students and faculty members of the same lab group, and there is a landmark where they meet very often.

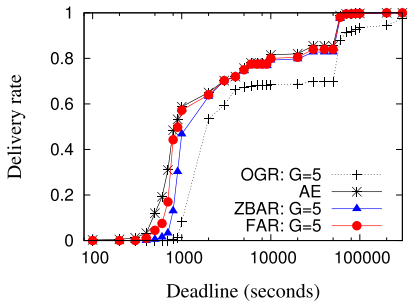


Fig. 11. The delivery rate w/ the Infocom 2005 trace.

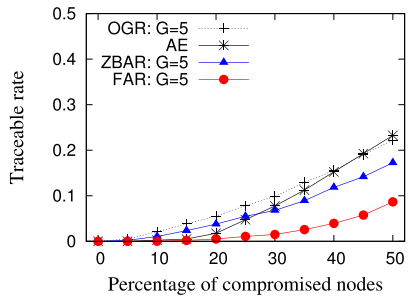


Fig. 12. The traceable rate w/ the Infocom 2005 trace.

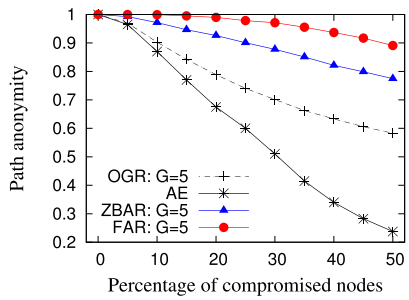


Fig. 13. The path anonymity w/ the Infocom 2005 trace.

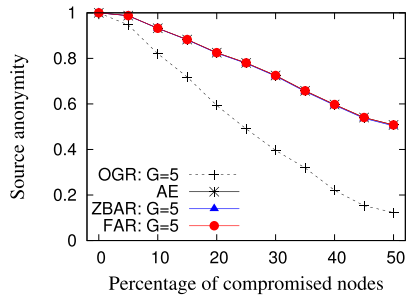


Fig. 14. The source anonymity w/ the Infocom 2005 trace.

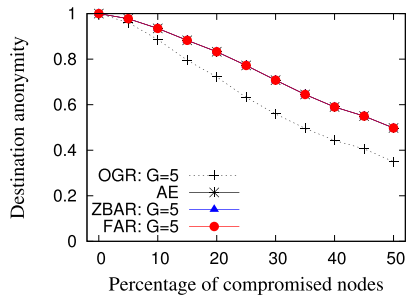


Fig. 15. The destination anonymity w/ the Infocom 2005 trace.

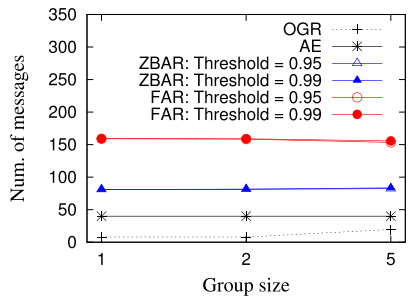


Fig. 16. The number of messages w/ the Infocom 2005 trace.

The Infocom 2005 trace contains fewer contact events than the Cambridge trace. The x-axis of Fig. 11 is scaled longer. As can be seen from the figure, the delivery rate of all the protocols increases toward 1,000 seconds, and then a stable period is observed from 5,000 to 30,000 seconds. This implies that there are no contact events during business off-hours. The delivery rate of all the protocols reaches 99 percent around 60,000 seconds (approximately 16.5 hours), and most of the message transmissions are likely to go through business off-hours. OGR always results in smaller delivery rate than the other protocols, and no significant difference between AE and FAR can be seen. ZBAR incurs a slightly longer delay than AE and FAR, as it uses onion-based forwarding between source and destination proxies.

Fig. 12 presents the traceable rate using the Infocom 2005 trace with respect to the percentage of compromised nodes. Note that the traceable rate is independent of the inter-meeting time among nodes. As can be seen in the figure, the traceable rate of FAR is at least half of AE, ZBAR, and OGR when 50 percent of the nodes are compromised.

Fig. 13 demonstrates the path anonymity for different protocols with respect to the percentage of compromised nodes. The figure shows a similar trend as that resulting from random graphs, and FAR achieves much higher path anonymity than the other protocols in a real trace. We can conclude that the proposed FAR achieves both faster delivery and higher degree of security by introducing message overhead.

Figs. 14 and 15 illustrate the source and destination anonymity resulting from the Infocom 2005 trace. The node anonymity of AE, ZBAR, and FAR linearly decreases when the percentage of compromised nodes increases. Since the contact trace is not large, i.e., the Infocom 2005 trace contains 41 mobile nodes, the difference among AE, ZBAR, and FAR is not significant. On the other hand, OGR always results in less source and destination anonymity than the other protocols.

Fig. 16 depicts the number of messages for different protocols resulting from the Infocom 2005 trace. While OGR results in the smallest message overhead, its delivery rate is not acceptable as shown in Fig. 11. FAR and ZBAR introduce more redundant message forwarding than AE and OGR do. However, we stress that they provide lower traceable rate and high node anonymity. Since the trace is a relatively small scale network containing 41 mobile nodes, the difference value of the thresholds does not affect the message overhead.

5.4 Comparisons between Simulation and Analysis

Fig. 17 shows the traceable rate resulting from simulations and analysis. Note that, according to our analysis, the traceable rate is independent of the size of onion groups, and thus simulation results with different group sizes are very close to

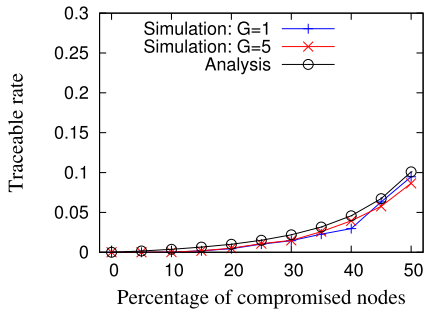


Fig. 17. The traceable rate analysis w/ the Infocom 2005 trace.

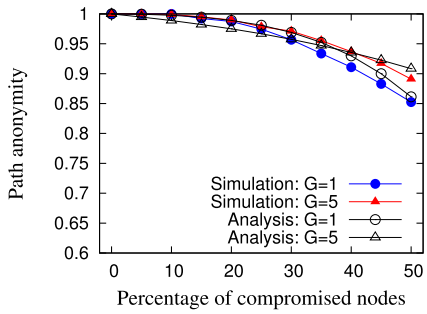


Fig. 18. The path anonymity analysis.

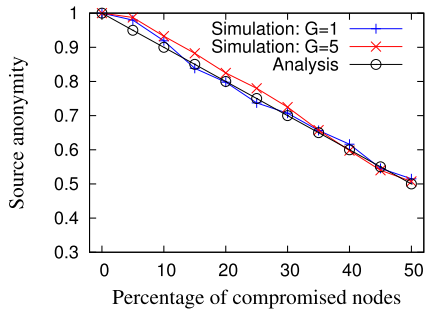


Fig. 19. The source anonymity analysis w/ the Infocom 2005 trace.

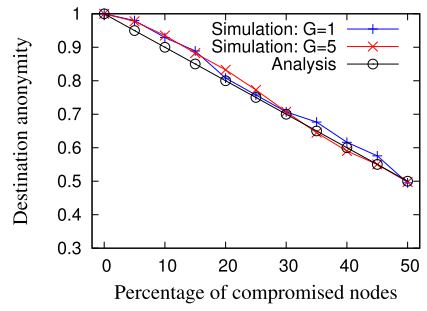


Fig. 20. The destination anonymity analysis w/ the Infocom 2005 trace.

each other. This figure demonstrates that the analytical result provides a very close approximation for the traceable rate.

Fig. 18 presents the path anonymity resulting from simulations and analysis. Even though we use simplified assumptions to derive the closed form solution in the analytical model, the difference between the simulation and analysis for both the cases of $G = 1$ and $G = 5$ is less than five percent. Thus, this comparison validates the mathematical model for the path anonymity.

Figs. 19 and 20 provide the source and destination anonymity resulting from simulations and analyses. As the proposed analysis indicates, both the source and destination anonymity decrease as the number of compromised nodes increases. In addition, a significant difference between different group sizes is not observed, since the node anonymity is independent of the size of onion groups.

6 Considerations on Parameter Selection

In this section, we discuss how to select parameters that satisfy a given security, performance, and cost requirements. The first consideration is a set of forwarding modes, SW and RE. There is no obvious advantage of using SW except smaller amount of message overhead. Thus, RE mode should be applied for achieving both the faster delivery and a higher degree of privacy as long as the message overhead is acceptable.

There are three tunable parameters: the number of message copies L, onion group size G, and the number of intermediate onion routers K. Among them, G and K are specified by the network administrator. The centralized setting of G is required to initialize the public/private keys. In many scenarios, K is a constant, e.g., $K = 3$ in Tor [1], [2]. Even in wired communications, the use of onion routers significantly reduces the throughput, and thus, the value of K should not be greater than three for DTNs. The value of L is tunable by users for each message transmission request, and $L \leq G$ holds because letting $L > G$ has no effect. As a rule of thumb, the larger the value of L, the better the delivery rate. Note that our analysis in Section 4 implies that the traceable rate and source/destination anonymity are independent from L, although they slightly affect these metrics in simulations. Thus, in the following, we discuss how users can select a proper value of L based on their performance requirements subject to given acceptable message overhead.

Recall that n is the number of nodes in a network and c is the number of compromised nodes. For given parameters K and G specified by the administrator, the function of the message overhead, denoted by $C(L, K, G, n)$, is defined by

$$C(L,K,G,n) \leq \begin{cases} LG(K+1) & \text{for OGR} \\ n & \text{for AE} \\ 2nLG(K-1) & \text{for ZBAR} \\ nLG(K+1) & \text{for FAR.} \end{cases} \quad (24)$$

Let M be the acceptable number of forwarded messages. The desirable value of $L \leq G$ to maximize the delivery rate can be obtained by introducing the number of copies L, subject to $C(L, K, G, n) \leq M$.

7 Related Work

7.1 DTN Routing

Epidemic routing [17] is a flooding-like message forwarding scheme that allows nodes to copy a message at every contact. While this approach maximizes the delivery rate and minimizes the delay when buffer constraint is not considered, it incurs a large amount of overhead. A ticket-based protocol, e.g., spray-and-wait [18], balances the trade-off between the performance and control overhead by limiting the number of copies of a message. Based on how the tickets are controlled, there are two types of spray-and-wait protocols: source and binary spray-and-wait. In the source spray-and-wait protocol, the source node has L tickets and consumes one ticket by forwarding a message at every contact. Thus, the source can duplicate up to L copies of a message.

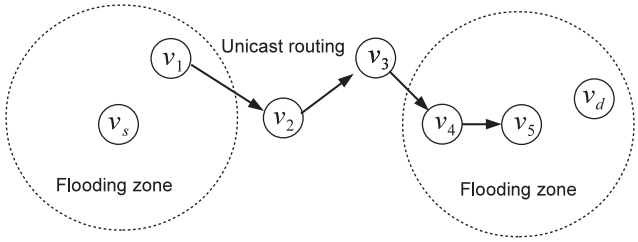


Fig. 21. An example of zone-based routing.

In the binary spray-and-wait, the source node with L tickets gives $L/2$ tickets at the first node it has a contact with. That is, every node with a message consumes half a ticket at every contact. To improve the message delivery with limited tickets, probabilistic analysis based on knowledge oracles [19], e.g., past contact history, queueing, and traffic demand, is incorporated to improve the delivery rate [20] and/or reduce the redundant message forwarding [21]. Depending on what metric a system administrator likes to emphasize the most, such as the average delay and worst-case delay, a suite of utility functions are proposed in [22].

7.2 Anonymous Routing for Ad Hoc Networks

Anonymous routing protocols in ad hoc networks are divided into either onion-based [3], [4], [5], [6] or location-based protocols [6]. In onion-based routing, the layered encryption, with different sets of secret keys, is applied to sensitive data and/or routing information. This data structure forces traffic to travel through a set of *onion routers* so that each layer of the onion can be peeled off, one by one, for the destination node to obtain the message. Onion routers neither store a network log, nor know who is communicating with whom. For the protocols in this category [3], [4], [5], [6], an onion is generated by adding encrypted layers during the route discovery phase.

The location-based protocol [6] preserves the anonymity of end hosts by making their locations ambiguous. For instance, ZAP [6] selects two proxies for delegate, source and destination nodes as shown in Fig. 21. While unicast routing is used in the communications between two proxies, anonymous flooding is applied to the communications within anonymous zones where a proxy and source node or destination node are located. By doing this, the source and destination nodes are not identifiable within the zone. The definition of a zone can be defined by a topology-based zone, such as the number of hops from a node, or by a geographical area including one of the end points.

7.3 Anonymous Routing Protocols in DTNs

The most relevant research is the anonymous routing protocol design in DTNs. ALAR [8] preserves the location privacy of a source node by dividing a message into several segments, and then forwarding them via different neighbors. However, this approach hides the location but not the identity of the source node. A natural approach to preserving node anonymity involves the use of proxies, such as onion routers or pivot. Based on the threshold secret sharing [23], TPS [11] routes a message through at least τ groups out of s groups, and the last intermediate node serves as a pivot. The difference between TPS and onion-based routing is that the layered encryption is not performed, and thus, the pivot knows the identity of the destination. To the best of our knowledge, the most viable protocols at this moment are group onion-based protocols, such as ARDEN [9] and OGR [10] in which a set of nodes share a secret key to form an onion group, and any node in the same group can encrypt/decrypt the corresponding layer of an onion.

7.4 Anonymity Models

While we apply the quantitative metrics, including the traceable rate, the node anonymity, and the path anonymity, to the anonymous routing in DTNs, some works evaluate anonymous communications in a different way. In [24], the prior to posterior probability-based analysis is proposed, where the difference between the probability of a subject being identified before and after an adversary observes protocol output is discussed. In [25], the distinguishable-based privacy model is proposed, where an anonymity experiment is defined. For given protocol output, the probability that an adversary distinguishes two subjects is defined as the adversary's advantage. An anonymous protocol is said to preserve the privacy when the advantage is bounded by a certain factor.

In our scenario, we consider the compromise attack, where some nodes in a network are compromised and an adversary can monitor message transmissions from a compromised node to another node. Unfortunately, nodes (or paths) are distinguishable with a non-negligible probability under such an attack. For example, when a randomly selected node is compromised, the source (or destination) node is identified with probability $1/n$, where n is the number of nodes in the network, which is not negligible. Thus, we believe that neither the prior to posterior probability-based nor indistinguishability-based analyses are appropriate for our scenario.

8 Conclusion

In this paper, we first construct anonymous Epidemic and zone-based routing protocols for DTNs by porting the existing solutions designed for ad hoc networks. Then, we design a framework for anonymous routing that subsumes all the Epidemic, zone-based, and onion-based routing. By tuning parameters, the proposed FAR enjoys the advantages of these protocols, but at the same time offsets disadvantages. With this design, FAR accommodates compatibility problems among DTNs with different routing policies, and thus, it can be deployed to DTNs with different security and anonymous requirements with ease. In addition, quantitative analyses are studied in terms of node and path anonymity as well as traceable rate. Furthermore, the extensive simulations resulting from randomly generated graphs as well as one of the well-known real traces called CRAWDAD dataset Cambridge/haggle demonstrate that the proposed scheme outperforms the existing solutions. Moreover, simulations and numerical results are compared and validated by each other. We believe our framework serves the foundation of anonymous routing for many types of contact-based networks.

Acknowledgments

This research has been funded in part by the JSPS KAKENHI Grant Number JP17K12675, by the Taiwan MOST grants 107-PFA2-550-448, 107-2218-E-001-006, and 107-2221-E-008-082-MY2, and by the U.S. National Science Foundation grants IIS-1618669 (III) and ACI-1642133 (CICI).

REFERENCES

[1] D. Goldschlag, M. Reed, and P. Syverson, "Hiding routing information," in *Proc. 1st Int. Workshop Inf. Hiding*, 1996, pp. 137–150.

[2] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," in *Proc. 13th Conf. USENIX Secur. Symp.*, 2004, pp. 21.

[3] S. Basagni, K. Herrin, D. Bruschi, and E. Rosti, "Secure pebblenets," in *Proc. ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2001, pp. 156–163.

[4] Y. Zhang, W. Liu, and W. Lou, "Anonymous communications in mobile ad hoc networks," in *Proc. IEEE INFOCOM*, 2005, pp. 1940–1951.

[5] J. Kong and X. Hong, "ANODR: Anonymous on demand routing with untraceable routes for mobile ad-hoc networks," in *Proc. ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2003, pp. 291–302.

[6] X. Wu and E. Bertino, "An analysis study on zone-based anonymous communication in mobile ad hoc networks," *IEEE Trans. Depend. Secur. Comput.*, vol. 4, no. 4, pp. 252–265, Oct.–Dec. 2007.

[7] D. Goldschlag, M. Reed, and P. Syverson, "Onion routing," *Commun. ACM*, vol. 42, no. 2, pp. 39–41, 1999.

[8] X. Lu, P. Hui, D. Towsley, J. Pu, and Z. Xiong, "Anti-localization anonymous routing for delay tolerant network," *Comput. Netw.*, vol. 54, no. 11, pp. 1899–1910, 2010.

[9] C. Shi, X. Luo, P. Traynor, M. H. Ammar, and E. W. Zegura, "ARDEN: Anonymous networking in delay tolerant networks," *Ad Hoc Netw.*, vol. 10, no. 6, pp. 918–930, 2012.

[10] K. Sakai, M.-T. Sun, W.-S. Ku, J. Wu, and F. S. Alanazi, "An analysis of onion-based anonymous routing for delay tolerant networks," in *Proc. IEEE Int. Conf. Distrib. Comput. Syst.*, 2016, pp. 609–618.

[11] R. Jansen and R. Beverly, "Toward anonymity in delay tolerant networks: Threshold pivot scheme," in *Proc. IEEE Mil. Commun. Conf.*, 2010, pp. 587–592.

[12] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, "CRAWDAD dataset cambridge/haggle (v. 2009–05-29)," May 2009. [Online]. Available: http://crawdad.org/cambridge/haggle/20090529

[13] W. Gao, G. Cao, A. Iyengar, and M. Srivatsa, "Supporting cooperative caching in disruption tolerant networks," in *Proc. IEEE Int. Conf. Distrib. Comput. Syst.*, 2011, pp. 151–161.

[14] K. Sakai, M.-T. Sun, W.-S. Ku, and J. Wu, "A framework for anonymous routing in delay tolerant networks," in *Proc. IEEE Int. Conf. Netw. Protocols*, 2017, pp. 1–10.

[15] C. Díaz, S. Seys, J. Claessens, and B. Preneel, "Towards measuring anonymity," in *Proc. 2nd Int. Conf. Privacy Enhancing Technol.*, 2002, pp. 54–68.

[16] G. Vakde, R. Bibikar, Z. Le, and M. Wright, "EnPassant: Anonymous routing for disruption-tolerant networks with applications in assistive environments," *Secur. Commun. Netw.*, vol. 4, no. 11, pp. 1243–1256, 2011.

[17] A. Vahdat and D. Becker, "Epidemic routing for partially-connected ad hoc networks," Duke Univ., Duhram, NC, USA, Tech. Rep. CS-200006, 2000.

[18] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: An efficient routing scheme for intermittently connected mobile networks," in *Proc. SIGCOMM Workshop Delay-Tolerant Netw.*, 2005, pp. 252–259.

[19] S. Jain, K. Fall, and R. Patra, "Routing in a delay tolerant network," in *Proc. Conf. Appl. Technol. Archit. Protocols Comput. Commun.*, 2004, pp. 145–158.

[20] C. Liu and J. Wu, "An optimal probabilistic forwarding protocol in delay tolerant networks," in *Proc. ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2009, pp. 105–114.

[21] W. Gao, Q. Li, and G. Cao, "Forwarding redundancy in opportunistic mobile networks: Investigation and elimination," in *Proc. IEEE INFOCOM*, 2014, pp. 2301–2309.

[22] A. Balasubramanian, B. Levine, and A. Venkataramani, "DTN routing as a resource allocation problem," in *Proc. Conf. Appl. Technol. Archit. Protocols Comput. Commun.*, 2007, pp. 373–384.

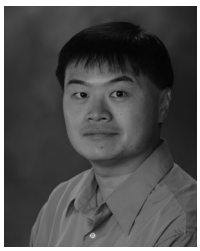
[23] A. Shamir, "How to share a secret," *Commun. ACM*, vol. 22, no. 11, pp. 612–613, 1979.

[24] S. Oya, C. Troncoso, and F. Pérez-González, "Back to the drawing board: Revisiting the design of optimal location privacy-preserving mechanisms," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2017, pp. 1959–1972.

[25] M. Backes, A. Kate, P. Manoharan, S. Meiser, and E. Mohammadi, "AnoA: A framework for analyzing anonymous communication protocols," *J. Privacy Confidentiality*, vol. 7, no. 2, 2017, Art. no. 5.

Kazuya Sakai (S'09-M'14) received the PhD degree in computer science and engineering from the Ohio State University, in 2013. He is currently an associate professor with the Department of Electrical Engineering and Computer Science, Tokyo Metropolitan University. His research interests include the area of information and network security, wireless and mobile computing, and distributed algorithms. He received the IEEE Computer Society Japan Chapter Young Author Award 2016. He is a member of the IEEE and ACM.

Min-Te Sun (S'99-M'02) received the BSc degree from National Taiwan University, the MSc degree from the Indiana University, Bloomington, and the PhD degree in computer and information science from the Ohio State University. He is a professor with the Department of Computer Science and Information Engineering, National Central University, Taiwan. His research interests include distributed computing and IoT. He is a member of the IEEE and ACM.



Wei-Shinn Ku (S'02-M'07-SM'12) received the MS degrees in computer science and in electrical engineering from the University of Southern California (USC), in 2003 and 2006, respectively, and the PhD degree in computer science from USC, in 2007. He is a professor with the Department of Computer Science and Software Engineering, Auburn University, Auburn, Alabama. His current research interests include data management systems, data science, cybersecurity, and mobile computing. He has published more than 100 research papers in refereed international journals and conference proceedings. He is a senior member of the IEEE and a member of the ACM SIGSPATIAL.

Jie Wu is the director of the Center for Networked Computing and Laura H. Carnell professor with Temple University. He also serves as the director of international affairs at the College of Science and Technology. He served as chair of the Department of Computer and Information Sciences from the summer of 2009 to the summer of 2016 and as associate vice provost for international affairs from the fall of 2015 to the summer of 2017. Prior to joining Temple University, he was a program director with the National Science Foundation and was a distinguished professor with Florida Atlantic University. His current research interests include mobile computing and wireless networks, routing protocols, cloud and green computing, network trust and security, and social network applications. He regularly publishes in scholarly journals, conference proceedings, and books. He serves on several editorial boards, including the *IEEE Transactions on Mobile Computing*, the *IEEE Transactions on Service Computing*, the *Journal of Parallel and Distributed Computing*, and the *Journal of Computer Science and Technology*. He was general co-chair for IEEE MASS 2006, IEEE IPDPS 2008, IEEE ICDCS 2013, ACM MobiHoc 2014, ICPP 2016, and IEEE CNS 2016, as well as program co-chair for IEEE INFOCOM 2011 and CCF CNCC 2013. He was an IEEE Computer Society distinguished visitor, ACM distinguished speaker, and chair for the IEEE Technical Committee on Distributed Processing (TCDP). He is a CCF distinguished speaker and a fellow of the IEEE. He is the recipient of the 2011 China Computer Federation (CCF) Overseas Outstanding Achievement Award.