TOKYO METROPOLITAN UNIVERSITY

A Framework for Anonymous Routing in Delay Tolerant Networks

Kazuya Sakai, Tokyo Metropolitan University
Min-Te Sun, National Central University
Wei-Shinn Ku, Auburn University
Jie Wu, Temple University

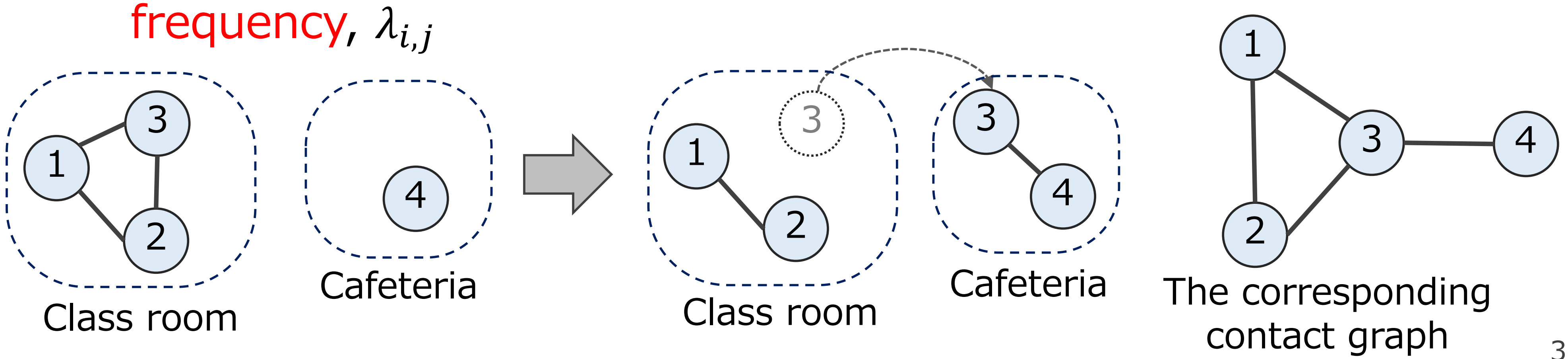
In ICNP 2017
Oct. 10-13, 2017

Outline

1. Introduction
2. Related Works
3. A Framework for Anonymous Routing (FAR)
4. Analyses
5. Simulations
6. Conclusions

1. Introduction

- Delay tolerant networks (DTNs)
  - Intermittently disconnected, opportunistic transmission, store-and-carry forwarding
  - The delay is not concern as long as a message is delivered within time constraint
- The network model of a DTN
  - A graph representation is contact-based
  - The link weight between two nodes is defined by contact frequency, $\lambda_{i,j}$



Introduction (Cont.)

- Anonymous communications
  - Protect the privacy of end hosts
  - Prevent from traffic analyses
  - Intermediate nodes never know where a packet comes from and goes to
- Applications
  - Critical communications, e.g., battlefields

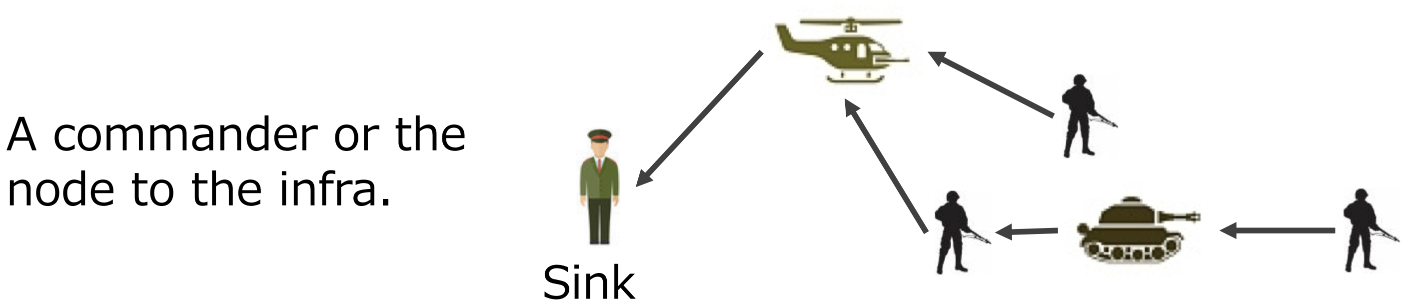
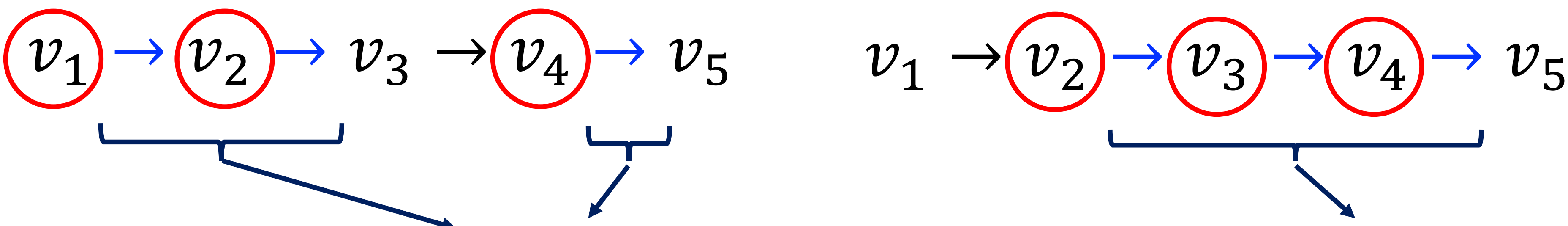


Fig. A battlefield communication

Attack Model

- The tracing attacks and node deanonymization
- Tracing attacks
  - Adversaries try to identify a path (or a set of links) along which packets traveled
  - $P_{trace} = \frac{1}{\eta^2} \sum_{i=1}^{C_{seg}} \left(c_{seg,i}\right)^2$, where η is the path length



Case 1: $P_{trace} = \frac{2^2+1^2}{4^2} = \frac{5}{16}$

Case 2: $P_{trace} = \frac{3^2}{4^2} = \frac{9}{16}$

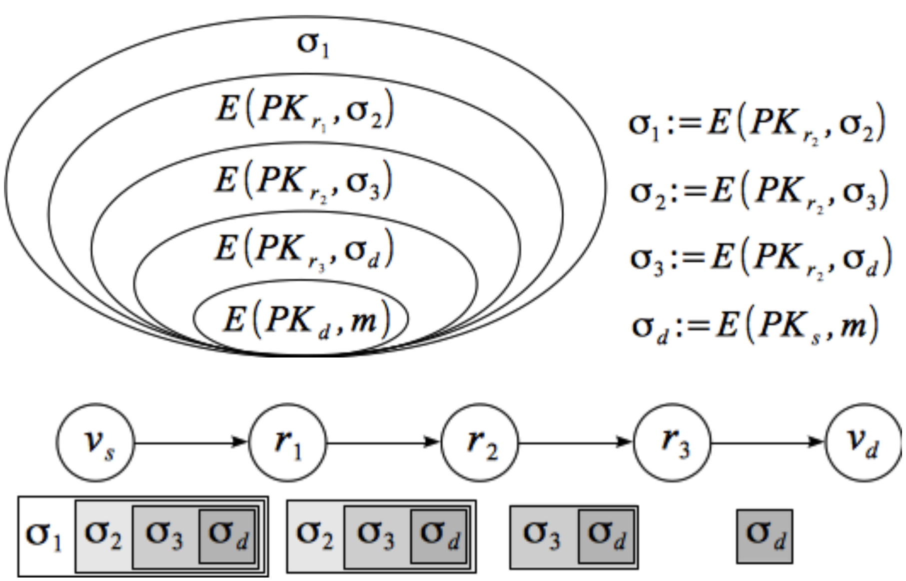
The Attack Model (Cont.)

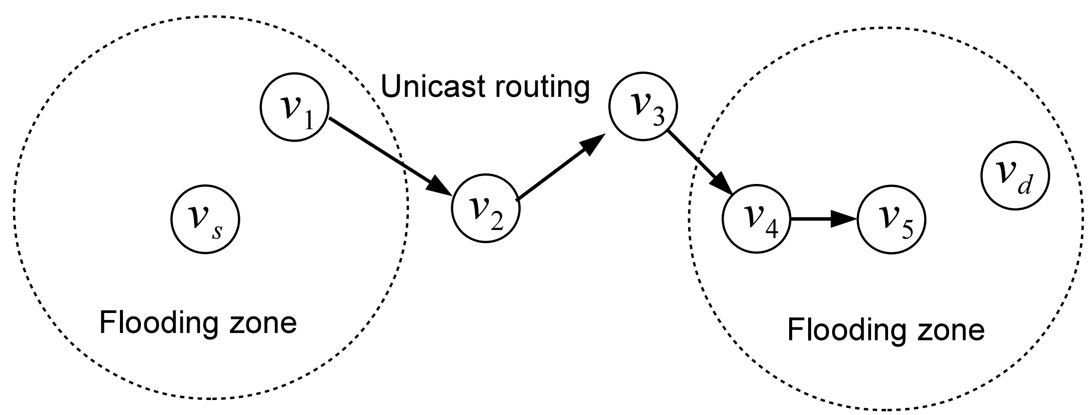
- The node deanonymization
  - Adversaries try to identify the source and destination nodes
- Anonymity
  - The state not being identifiable among an anonymous set
  - An entropy-based metric, $-\sum_{\forall i \in \phi} p_i \log(p_i)$
  - Application-dependent
  - Example: a bit string 01XX

{ 0100, 0101, 0110, 0111 }

2. Related Works

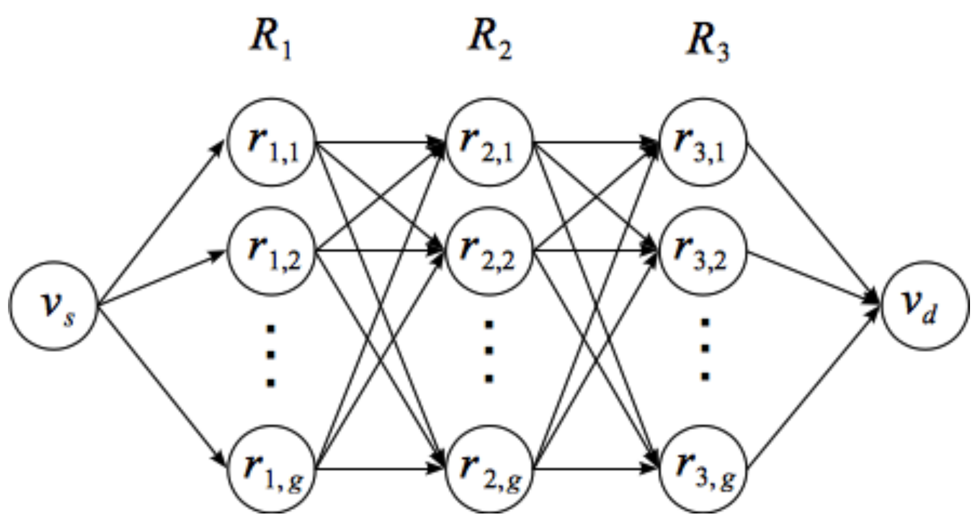
- Onion-based routing, e.g., Tor
  - Layered encryptions are applied to a message
- Flooding-based [Mobihoc’05], zone-based [TDSC’07]
  - The node’ privacy is protected by flooding (or partial flooding)
  - They are designed for ad hoc networks





Anonymous Routing in DTNs

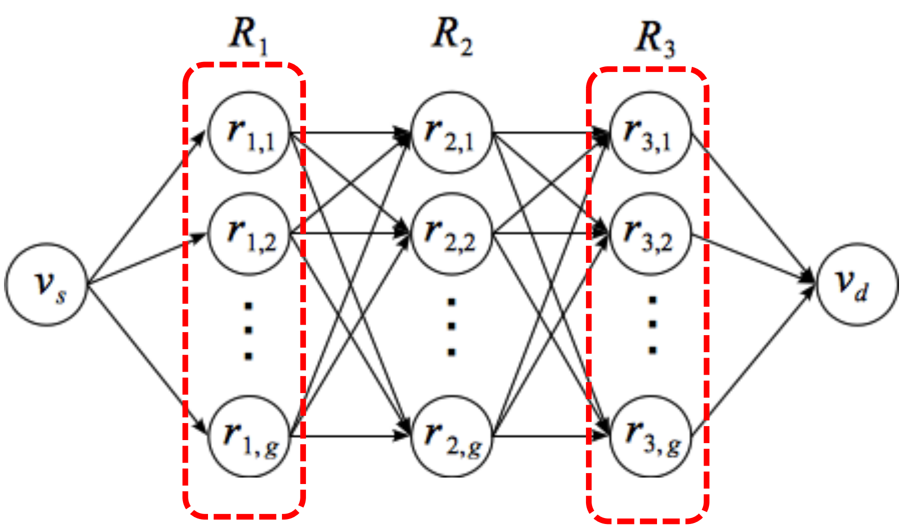
- Only a few protocols have been designed for DTNs
  - e.g., onion-based and threshold-based
- Onion-group routing (OGR) [ICDCS'16]
  - A set of nodes forms a group for faster delivery
  - Any of the nodes in a group can peels of the encrypted layer

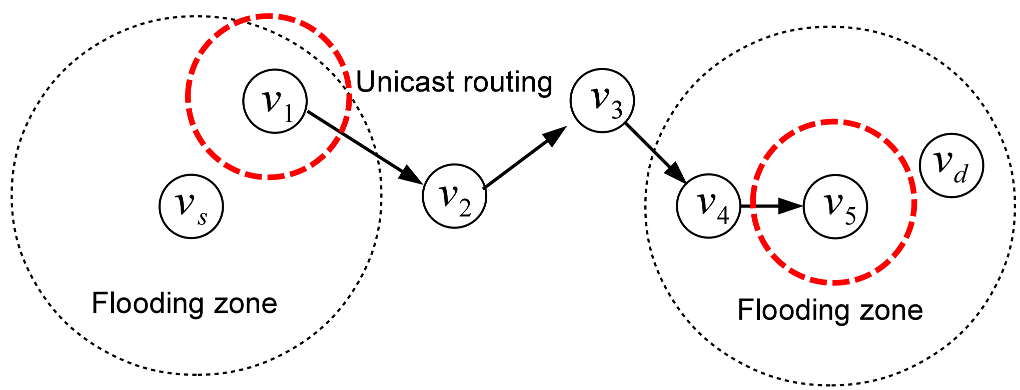


[ICDCS'16] K. Sakai, M.-T. Sun, W.-S. Ku, J. Wu, and F. Alanazi, "An Analysis of Onion-Based Anonymous Routing for Delay Tolerant Networks," In ICDCS, pp. 609-618, 2016.

The Issues of The Existing Solutions

- Onion-Based
  - Slow, less message overhead
  - All the members of the first/last onion groups knows the source/destination node
- Zone-based (has not been tailored to DTNs)
  - Relatively fast, more message overhead
  - The first/last proxy knows the source/destination node





The Contributions of This Paper

- Our goals
  - Improving the untraceability and anonymity
  - For given security requirements and cost constraint, we try to maximize the performance of anonymous routing
- The contributions of this paper
  - Designing Anonymous Epidemic (AE), Restricted Epidemic Routing (RER), and Zone-Based Anonymous Routing (ZBAR)
  - Designing a Framework of Anonymous Routing (FAR), which subsumes all the AE, RER, ZBAR, and OGR
  - Modeling the closed-form solutions for the traceable rate and node anonymity

3. Framework for Anonymous Routing

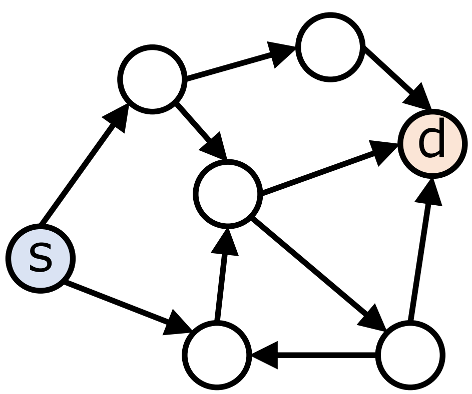
Ad hoc networks	DTNs
Flooding-based [Mobihoc'03]	AE and RER
Zone-based [TDSC'07]	ZBAR
Onion-based [Tor]	OGR [ICDCS'16]

} FAR

[Mobihoc'03] J. Kong and X. Hong, "ANODR: Anonymous on Demand Routing with Untraceable Routes for Mobile Ad-hoc Networks," In Mobihoc, 2013.
[TDSC'07] X. Wu and E. Bertino, "An Analysis Study on Zone-Based Anonymous Communication in Mobile Ad Hoc Networks," IEEE TDSC, 2007.
[ICDCS'16] K. Sakai, M.-T. Sun, W.-S. Ku, J. Wu, and F. Alanazi, "An Analysis of Onion-Based Anonymous Routing for Delay Tolerant Networks," In ICDCS, 2016.

Anonymous Epidemic (AE)

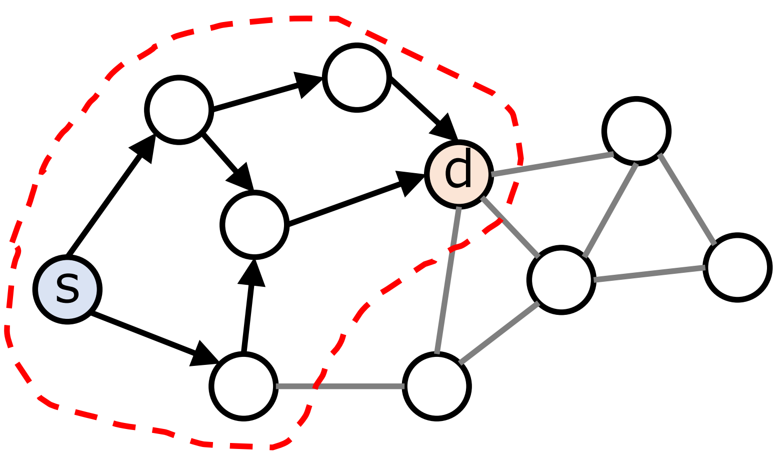
- AE (v_s, v_d, m, T)
  - Message m is encrypted using public key PK_d
  - The source node v_s sends $\sigma \leftarrow E(PK_d, m)$ to the destination v_d
  - Only v_d can decrypt ciphertext σ
  - Routing terminates when the end-to-end deadline T expires



- No onion relays between v_s and v_d
- All the n nodes forwards σ

Restricted Epidemic Routing (RER)

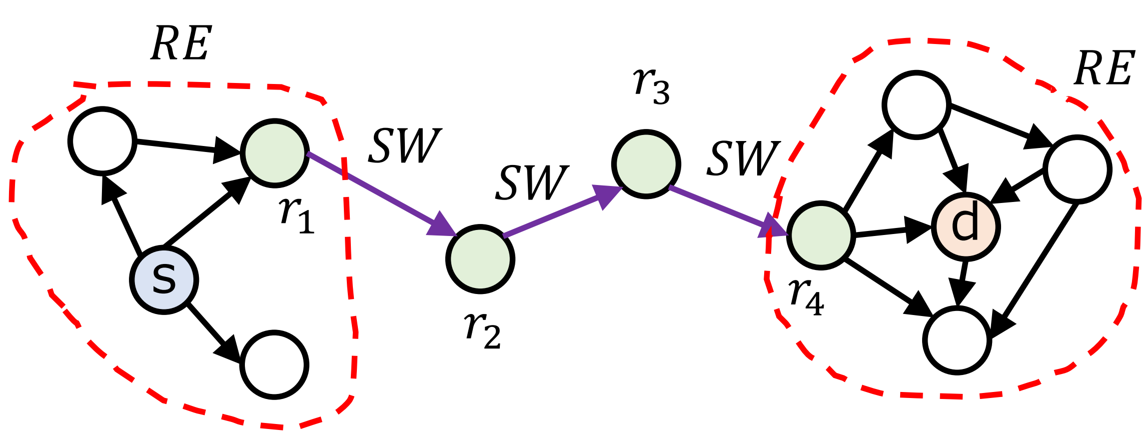
- RER (v_s, v_d, σ, T)
  - A epidemic zone is controlled by the deadline, $t \leftarrow -\frac{\ln(1-\tau)}{\lambda}$,
    - τ is the probability of the next node/destination receiving σ
    - λ is the average contact frequency
    - The idea: v_d will receive σ within t with probability τ
  - Routing terminates when either T or t expires
  - The destination can be a group of nodes



- Neither Euclidian distance nor TTL can be used for DTNs
- τ $(0.95 \leq \tau \leq 0.99)$ is set by the simulation

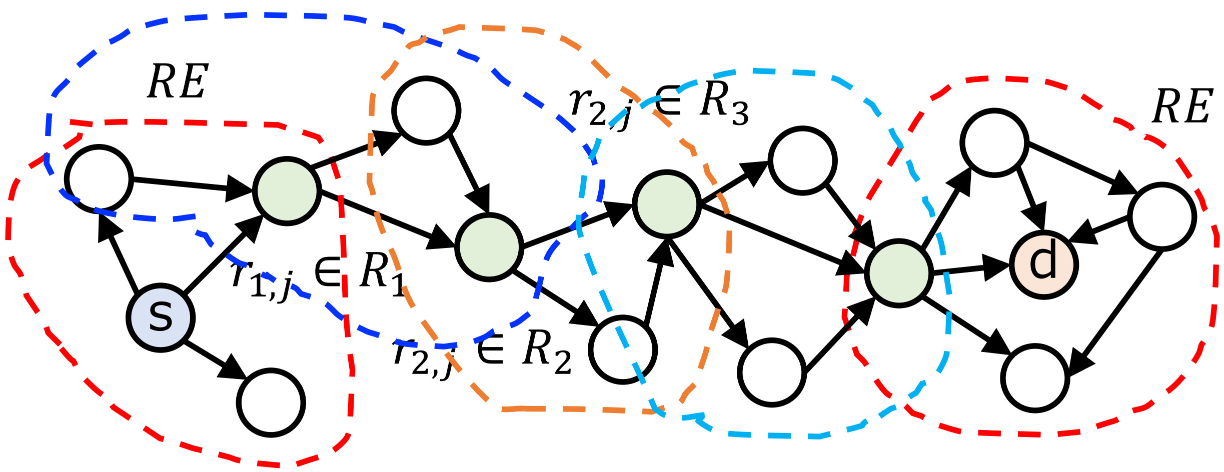
Zone-Based Anonymous Routing (ZBAR)

- ZBAR (v_s, v_d, m, T)
  - The source and destination proxies are randomly selected within their proximity
  - A set of K onion relays are randomly selected
  - An onion is created by the layered encryption
  - Forwarding modes
    - Restricted epidemic (RE) between v_s (or v_d) and its proxy
    - Spray-and-wait (SW) between onion relays



A Framework for Anonymous Routing (FAR)

- The idea
  - The use of a set of onion group relays
  - The use of epidemic-like forwarding
- FAR
  - A set of K onion groups are selected and an onion is created
  - An encrypted message is forwarded by RER



Security and Performance of FAR

- Security
  - Each previous/next is anonymous among the nodes in the zone
  - The first and last onion relays are indistinguishable from the intermediate onion relays
- Performance
  - More message overhead due to RER
  - Slow delivery due to a set of onion relays
- => For given security requirements and cost constraint, FAR tries to maximize the performance by tuning its parameters

Making A Framework

- FAR
  - An extreme case behaves as either AE, RER, ZBAR, and OGR
- Parameterizing
  - v_s, v_d : the source and destination nodes
  - m : the message
  - K : the number of onion relays
  - L : the number of message copies
  - G : the size of an onion group
  - F : a set of forwarding modes
    - A forwarding mode can be either RE or SW
  - T, τ : the end-to-end deadline and the zone deadline

FAR Subsumes AE, ZBAR, and OGR

- FAR with a particular configuration serves on either AR, ZBAR, or OGR
- AE
  - $K = 0$, $L = null$, $G = null$, $F = \{RE\}$
- ZBAR
  - K, G can be any integer, $L \leq G$, $F = \{RE, SW, SW, \ldots, SW, RE\}$
- OGR
  - K and G can be any integer, $L \leq G$, $F = \{SW, SW, \ldots, SW\}$

4. Analyses

- Analyses
  - The traceable rate
  - The node anonymity
- The attack scenario
  - The compromise attack: a node is physically compromised, and the transmission of a message is monitored
  - Compromised nodes are randomly chosen by the uniform distribution

The Traceable Rate

- The traceable rate, $P_{trace} = \frac{1}{\eta^2}\sum_{i=1}^{C_{seg}}\left(c_{seg,i}\right)^2$, where η is the path len.
- The closed form solution
  - The traceable rate is computed for individual message, and thus G and L do not affect
  - A path can be traced by the reserve order from the destination

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

  - where n is the number of nodes and c is the number of compromised nodes
  - $\epsilon_1 = \sum_{i=1}^{\eta}\left(\frac{c}{n}\right)^{i-1} \cdot \left(1-\frac{c}{n}\right)$ and $\epsilon_2 = \eta\left(\frac{c}{n}\right)^{\eta}$

The Node Anonymity

- The entropy of a system: $-\sum_{\forall i \in \phi} p_i \log(p_i)$
  - Where ϕ is a suspicious set of nodes
- The maximal entropy
  - If a node is not compromised, it is anonymous within $n - c$ nodes

$$H_{max} = \sum_{\forall \text{nodes} \in \phi} \frac{1}{n-c} \log\left(\frac{1}{n-c}\right)$$

- The entropy of a node
  - If a node is compromised, its entropy equals to 0
  - Otherwise, it is still anonymous within $n - c$ nodes

$$H_{\phi'} = \sum_{\forall \text{nodes} \in \phi'} \left(1 - \frac{c}{n}\right) \frac{1}{n-c} \log\left(\frac{1}{n-c}\right)$$

- The anonymity:

$$D(\phi') = \frac{H_{\phi'}}{H_{max}} = 1 - \frac{c}{n}$$

The Message Cost

- Parameters
  - K : the number of onion relays
  - L : the number of message copies
  - G : the size of an onion group
  - n : the number of nodes
- The cost function

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

- For given acceptable cost M, the delivery rate can be maximized by increasing G and L $(L \leq G)$ with subject to $C(L,K,G,n) \leq M$

5. Simulations

- Protocols
  - FAR, AE, ZBAR, and OGR [ICDCS'16]
- Parameters
  - Group size G, Num of onion routers K, Num of copies L
- Metrics
  - The delivery rate, traceable rate, and node anonymity
- Two scenarios
  - Randomly generated graphs
  - Real traces with the CRAWDAD dataset

The Delivery Rate

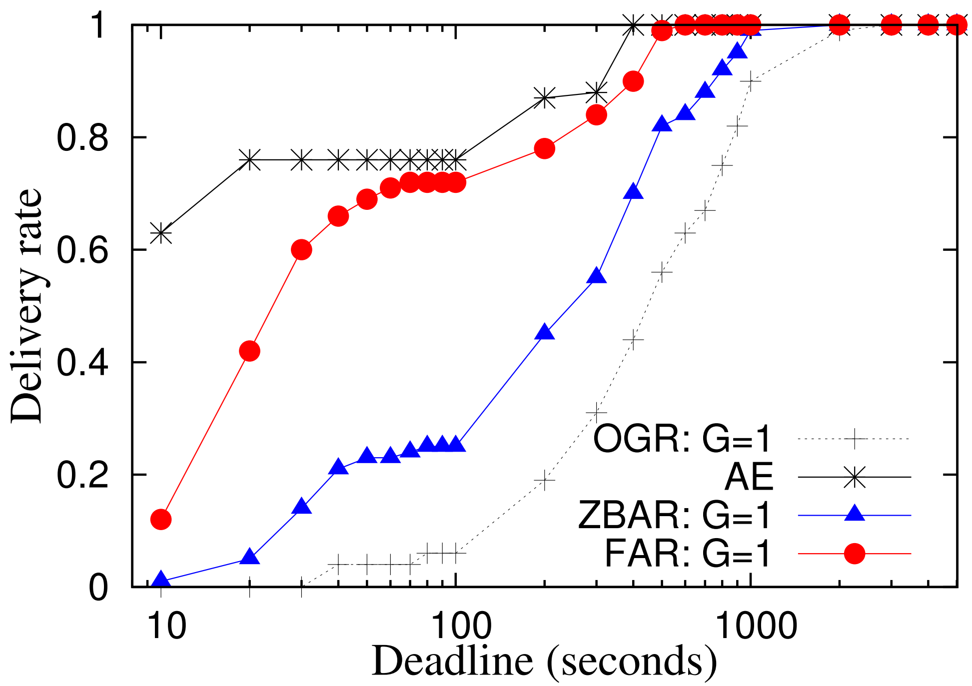


Fig. the delivery rate
Cambridge traces (a small and dense network with 12 iMotes)

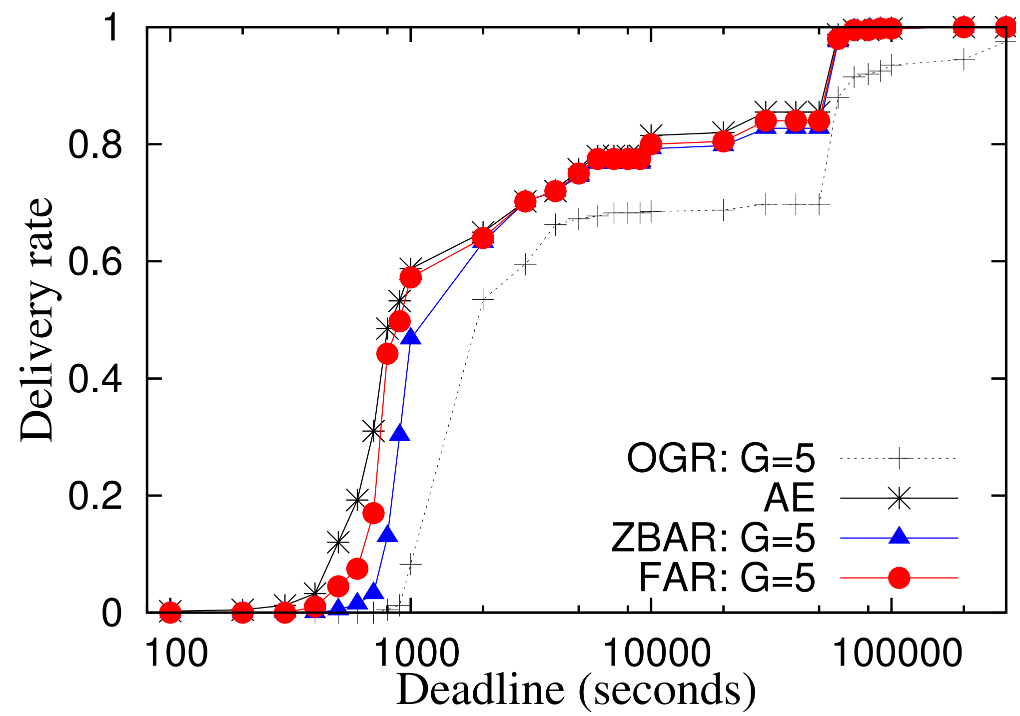


Fig. the delivery rate
Infocom'05 traces (a medium size network with 41 iMotes)

Traceable Rate

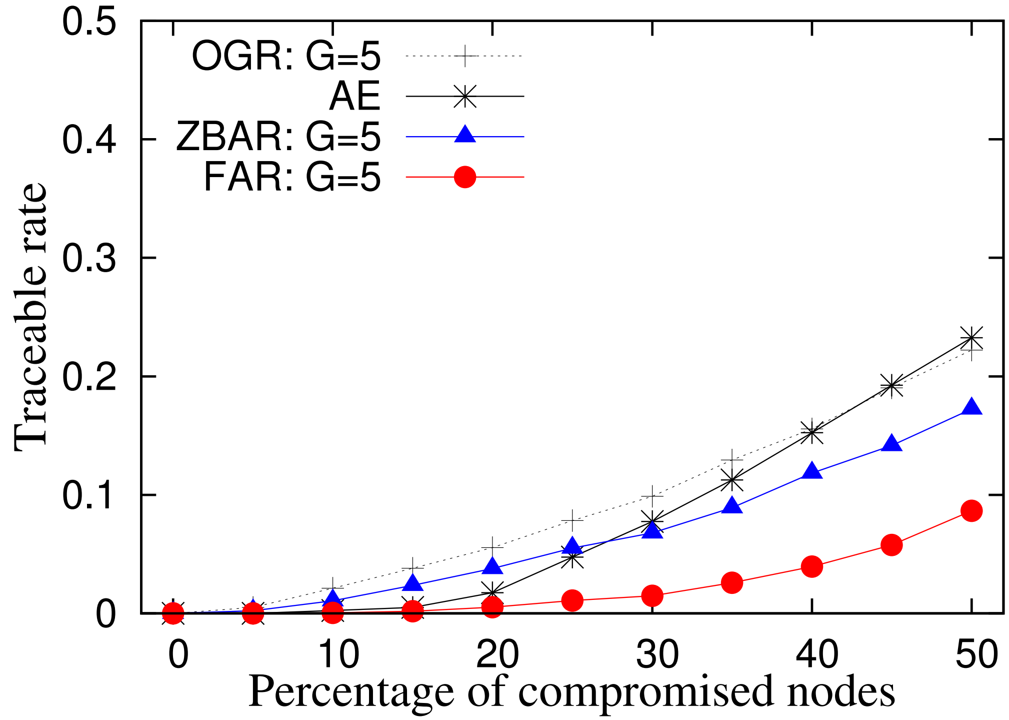


Fig. Traceable rate.

- Infocom’05 traces (a medium size network with 41 iMotes)
- Note: the traceable rate is independent from the value of L

Traceable Rate and Anonymity

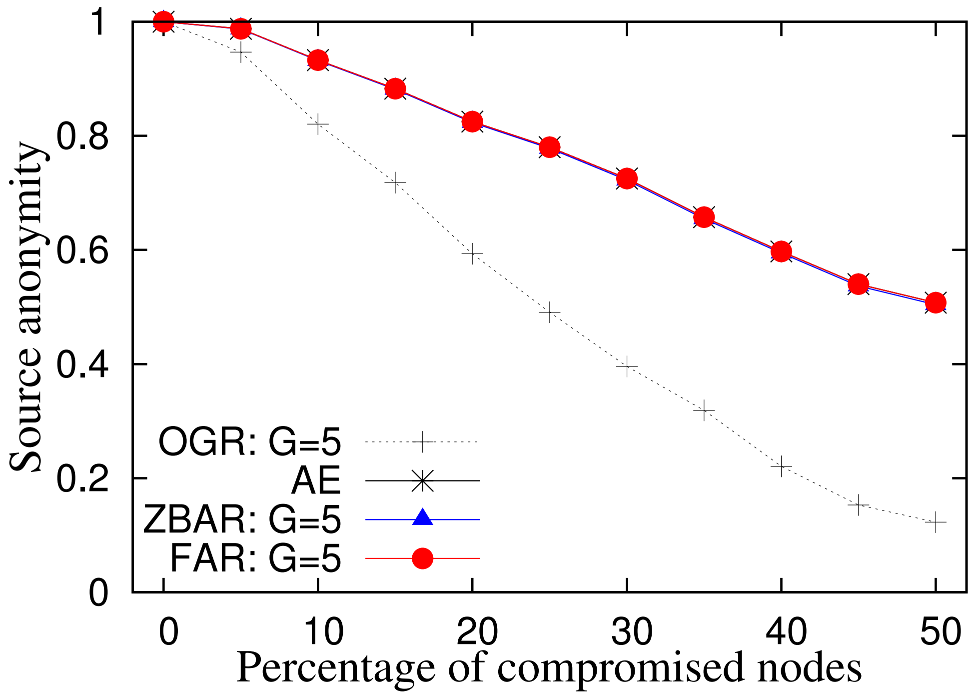


Fig. The source anonymity

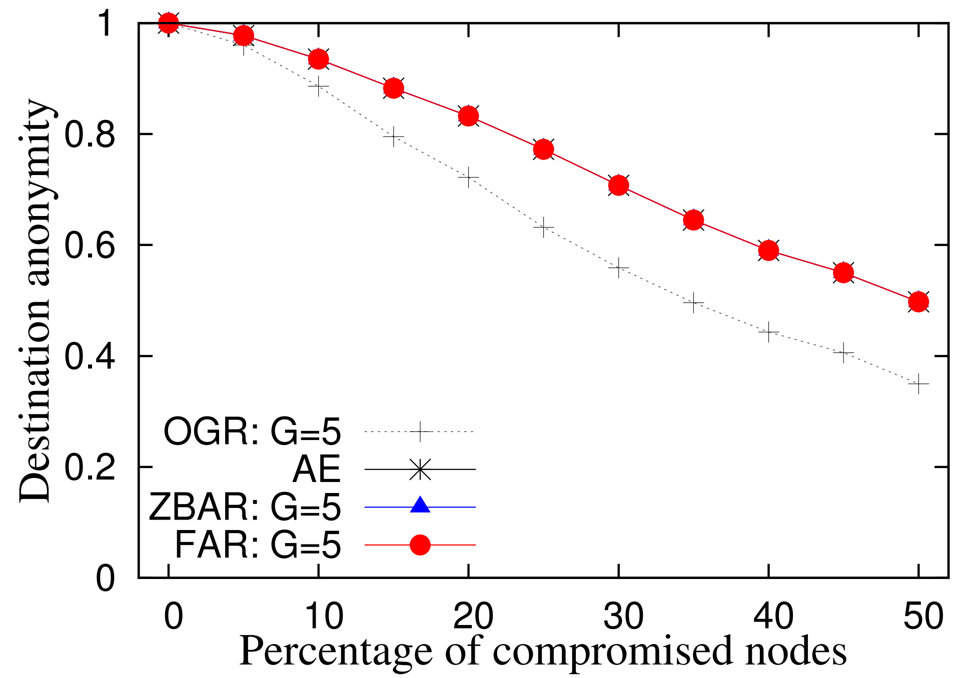


Fig. The destination anonymity

6. Conclusions

- In this paper, we address anonymous routing in DTNs
- Protocols
  - AE, RER, ZBAR, and FAR
- Analyses
  - The traceable rate and node anonymity
  - The message cost
- Simulation
  - Random graphs and the CRAWDAD dataset

Thank you