International Journal of Future Generation Communication and Networking
Vol. 8, No. 4 (2015), pp. 1-16
http://dx.doi.org/10.14257/ijfgcn.2015.8.4.01

A Framework for Analyzing Anonymous Network Topology

Tianbo Lu[1,2], Shixian Du[1], Yang Li[1], Peiyuan Dong[1] and Xiaoyan Zhang[1]

[1]*School of Software Engineering, Beijing University of Posts and Telecommunications, 100876, Beijing, China*
[2]*Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada*
lutb@bupt.edu.cn, shixian2011@126.com

Abstract

Nowadays, Internet privacy becomes more and more important and sensitive, and has been one of the latest buzz words to hit the Internet world. In response to protect the privacy of Internet users, a variety of privacy enhancing technologies (PETs) have emerged. As one of privacy enhancing technologies, anonymous communication has been extensively studied from various aspects by researchers. Node churn is frequent in anonymous network, which make it difficult to maintain a stable network topology, therefore it is necessary to get more insight into the overall anonymous network topology. In this work, in order to better grasp the present situation of research, we investigate existing organizations and universities which study anonymous communication from the perspective of network topology, and related projects. Then we survey related papers to anonymous communication published in recent years, which focus on the analysis of node selection (especially in the Tor). In addition, we present a framework of anonymous communication network topology, and a visual illustration of this analysis that shows the progression of the research of network topology and node discovery in anonymous communication. Finally, some related problems and follow-up study are presented to be studied deeply in future.

Keywords: *Anonymous communication, node selection, DHT, network topology*

1. Introduction

Nowadays, Internet privacy becomes more and more important and sensitive, and has been one of the latest buzz words to hit the Internet world. In response to protect the privacy of Internet users, a variety of privacy enhancing technologies (PETs) has emerged, such as anonymous communication. Anonymous communication can hide the identity and relationship of both sides of communications, which is primarily concerned with user identity confidentiality issues in network communication. User identity refers to host IP address, email address, and other related information. Anonymity is the indistinguishability of communication entities which cannot be identified from an entity collection (also called anonymity set). Anonymity set includes senders' anonymity set and receivers' anonymity set. The former set can disjoint, overlap, and intersect the latter set. If the size of anonymity set is bigger, and the probability of sending is more similar to the probability of receiving, then the anonymity is stronger.

At present, anonymous communication systems which have been developed varies widely from node discovery, node selection to transporting messages. The typical systems include Tor, Crowds [1], Tarzan, I2P, and so on. In particular, Tor is the most popular and widespread system, which is the object of numerous academic research and testing platform in anonymous communication. Existing work focuses on the research of these systems from various aspects and challenges, especially

ISSN: 2233-7857 IJFGCN

addressing the key issues of anonymity and performance in communication. For example, it is wise to provide anonymous routes with low latency and jitter for anonymous web browsers, while providing anonymous file transfer applications with high bandwidth (but not necessarily low latency or jitter) paths. In short, it's necessary to research how to tailor anonymity properties and performance characteristics according to users' specific communication requirements.

Moreover, many anonymous systems (such as Onion Routing, Crowds, Freedom, MorphMix and Tarzan) have achieved based on P2P, while these systems are more dynamic in nature for many nodes (called relays) are in the network for a short time. When a new node enters the network, it needs other nodes in the network to build path. And when the node exits network, the user in the path has to build a new path. One problem is the node joining and exiting also need the other nodes in the network, while the changing anonymous set make the problem more difficult. Meanwhile, node performance is different from each other. Based on barrel principle, it may cause the problem that poor performance of node can degrade the efficiency of anonymous path, even if the performance of other nodes on the path is very well. To summarize, how to select node is crucial to build anonymous path.

Based on the above analysis, this paper firstly introduces and analyzes major organizations and universities which study anonymous communication from topology and node discovery. We present a framework of anonymous network topology, and a visual illustration of this analysis that shows the progression of the researches of network topology and node discovery in anonymous communication.

The rest of the paper is organized as follows. In Section III, we provide relevant publications related to topology and node discovery, and analyze current trends. In the following, we discuss the methods of node selection, locate random relays, and so on. At last, we briefly describe the research challenges specific to the domain of network topology and node discovery in anonymous communication.

2. Related Organizations Introduction

Anonymous network topology has attracted the world's research organizations' interest. It is found that a large fraction of Tor's volunteer-operated relays are located in the Germany, United States, France, and Netherlands [2]. In order to grasp the present situation of research of anonymous communication, major organizations and universities which put more emphasis on network topology and node discovery in anonymous communication are described in this section, such as Georgetown University, NRI, University of Waterloo, Technische Universit¨at Dresden, SRI International. For the subsequent in-depth research, we give a summary of our investigations, and present a structured framework of anonymous network topology related organization (as the Figure 1 shows). Through the framework, we can easily find what major research agencies conduct this research, which related projects they have, and which related papers they have published recently. In the following, major organizations are introduced.

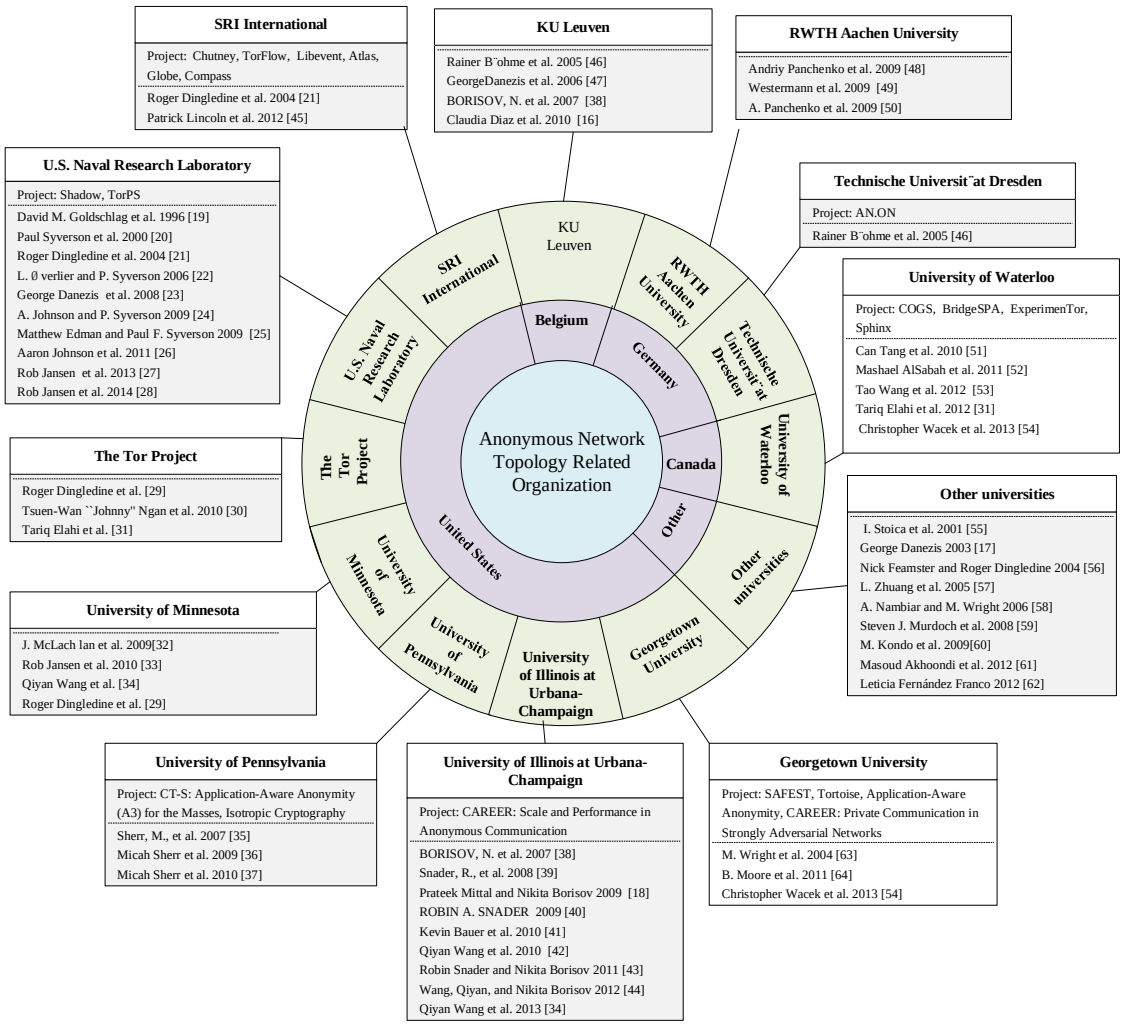


Figure 1. Anonymous Network Topology Related Organization Framework

2.1 U.S. Naval Research Laboratory

Naval Research Laboratory (NRL), especially its Center for High Assurance Computer Systems (CHACS) [3], is interested in anonymous and route-trusted communications. Emphasis will be placed on network topology and structure, routing protocols, secure distribution of network information, *etc*. The details of its theory and design include mathematical modeling and analysis of anonymity and routing security in communication systems; building blocks, cryptographic primitives, and designs for anonymous and route-secure communication; and incentive analysis of route-secure anonymous communication systems.

Major researchers in CHACS include Paul Syverson, Aaron Johnson, and so on. CHACS have conducted lots of research on these aspects and even developed some related technologies, methods and tools, such as Shadow [4] and the Tor Path Simulator (TorPS) [5].

2.2 SRI International

SRI International has done many researches in terms of network topology, especially displaying relay information. For example, various related components and softwares [6, 7] have been developed.

2.3 Georgetown University

Georgetown's Department of Computer Science includes Information Retrieval Lab, InfoSense, and SecurityLab. The SecurityLab [8] which is managed by Profs. Micah Sherr and Clay Shields, carries out research in various topics, such as privacy enhancing technologies, secure distributed systems, and network security. The SecurityLab's current projects about anonymous network topology mainly include SAFEST, Tortoise, Application-Aware Anonymity, and Private Communication in Strongly Adversarial Networks.

2.4 University of Waterloo

In Canada, University of Waterloo's Cryptography, Security, and Privacy Research Group, that is CrySP [10], focus on the research of cryptography, privacy-preserving communications networks, and so on. Especially for privacy-preserving networks, CrySP has developed an experimentation platform for research into large-scale privacy enhancing technologies, namely CrySP RIPPLE Facility. It consists of PIR system, ACN system, and CRS system. They respectively make research into private information retrieval, anonymous communications networks, and censorship resistance.

CrySP have developed various software and frameworks of anonymous network topology, such as COGS, Bridge SPA, ExperimentTor, Sphinx.

2.5 Technische Universit¨at Dresden

As one group of Technische Universit¨at Dresden in Germany, Privacy and Data Security's research topics [11] include multimedia security, channel coding theory, and privacy-enhancing technologies. Especially, the topic of privacy-enhancing technologies mainly study identity management and anonymous communication. TU Dresden's researchers mainly have studied anonymous technologies in circuit and packet switched networks of various topologies.

Technische Universit at Dresden current project associated with anonymous network topology is "Anonymität.Online", that is AN.ON. AN.ON can provide an anonymity service for the Internet which is the only one in the world that implements the cascade architecture, and even be considered as particularly secure.

2.6 Other

There are many other universities which actively conduct the topology research in anonymous communication.

University of Illinois at Urbana-Champaign's researchers in the area of anonymous communication is supported in part by an NSF grant CNS 0953655, namely "CAREER: Scale and Performance in Anonymous Communication"[12].

University of Pennsylvania is supported by an NSF grant CNS 0831376, that is "CT-S: Application-Aware Anonymity (A3) for the Masses" [13].

Lehigh University is supported by an NSF grant CNS 1117701, namely "Nets: Small: Optimal Design of Anonymous Network Systems under Resource Constraints" [14].

In addition, Beijing University of Posts and Telecommunications in China, and KU Leuven Computer Security and Industrial Cryptography (COSIC) group [15] in Belgium, actively conduct the topology research in anonymous communication,

3. Anonymous Network Topology Framework

It is well known that a network's virtual structure is referred as topology. Network topologies have strong effect on both anonymity and overhead in anonymous communication systems. There are two classification methods. From the perspective of the complexity of node connection, it can largely be divided into cascade topology, fully

connected topology (free topology), stratified topology, and stratified restricted topology [16, 17]. In cascade topology network, sender can select fixed path to transmit messages. In fully connected topology network, sender can select arbitrary length path to send messages. In general, fully connected topology network has stronger anonymity than cascade topology network, but it is often too complex and costly. On the other hand, it can be partitioned into structured Peer-to-Peer (P2P) topology and unstructured topology in terms of directory server. Examples of structured P2P topology include Chord and Pastry [18]. The typical unstructured topology is Tor, which is also a fully connected topology.

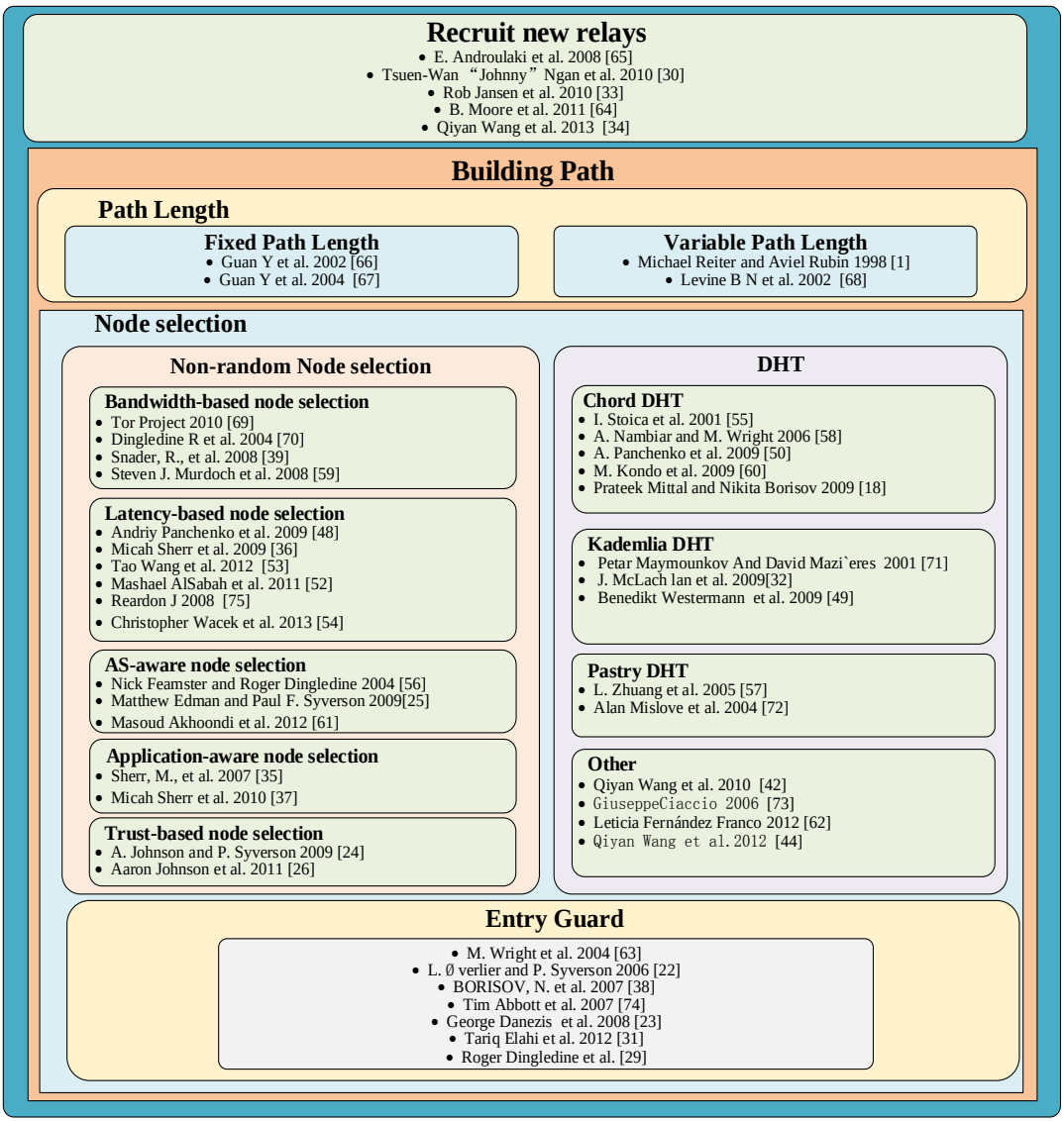


Figure 2. A Classification Framework of Anonymous Network Topology

This section mainly describes the classification framework of anonymous network topology our proposed, is shown in Figure 2. This framework is relatively new: the papers covered in this survey have been generally published since the year of 2001. Besides, all the related papers and references between them are put together to a thorough literature map (as Figure 3 shows). The prevailing fraction of these papers focuses on two sections: joining nodes and building anonymous path. Based on this, the framework mainly solves the two issues: how to recruit new relays and how to build anonymous path. In detail, in front of the communication of two sides, building a path of anonymous communication

should be gotten through first. Building path always be complicated, and in terms of network topology, its path size and node selection always be focused on. In another word, there is how many nodes are used to build anonymous path, and how to select node from the set of anonymous systems node, especially the selection of entry node. After building the path, both sides of communication can communicate with this path.

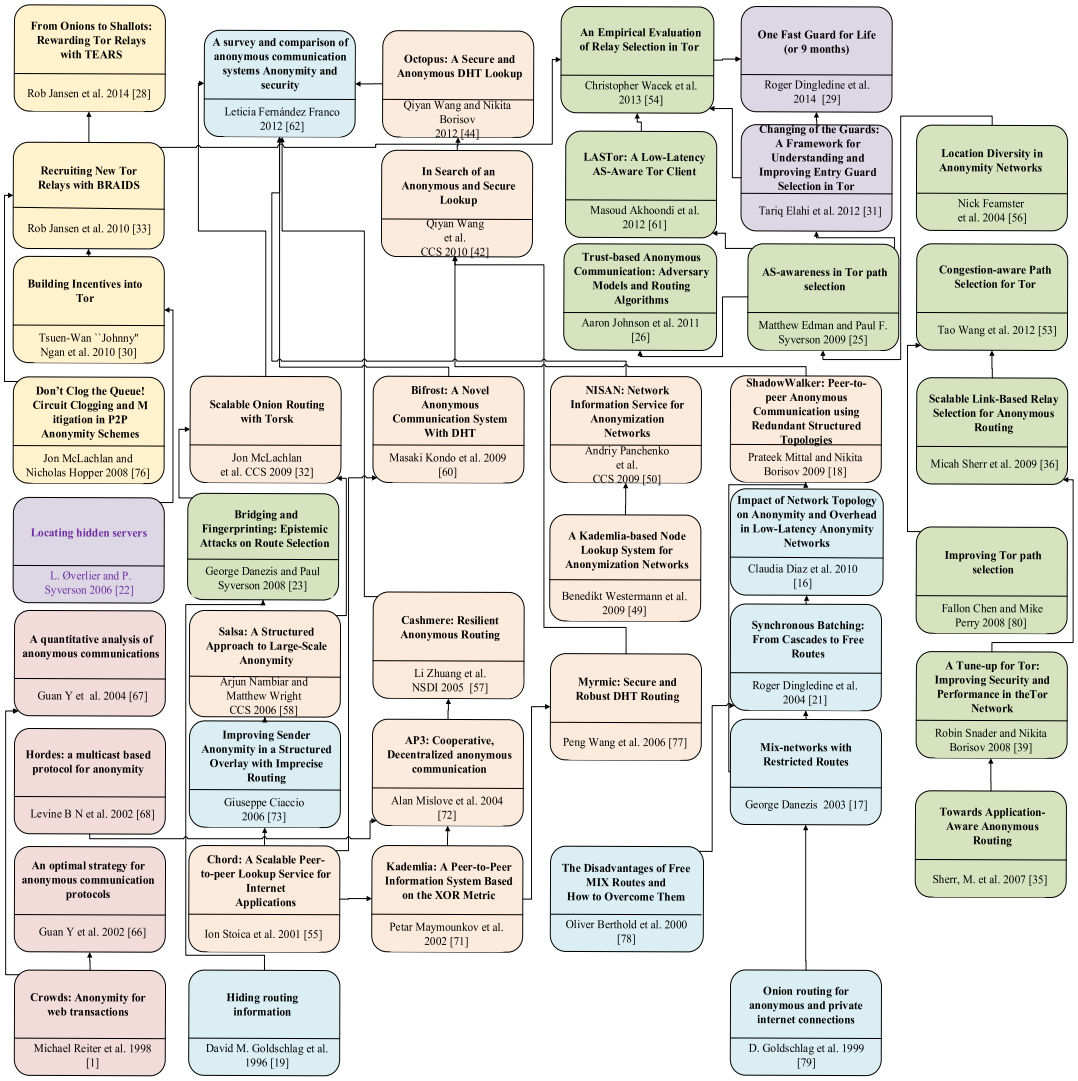


Figure 3. Anonymous Network Topology Framework's Literature Map

4. Recruit New Relays

During the process of anonymous communication, there are many problems, such as dynamic joining and leaving of nodes. Frequent node churn make it difficult to maintain a stable network topology. Moreover, anonymity set consists of numerous nodes, and the size of anonymity set has a direct effect on the anonymity and performance of system. For example, Tor is made of voluntary node, its performance depend on its volunteers to donate their resources. However, the resource can't meet the ever-increasing demand on Tor. Joining new node can solve the problem, because new node expands the diversity and size of the anonymity set to improve system performance and anonymity. But the sufficient incentives for volunteers to run relays are currently lacking.

In recent years, various incentives to join anonymous network is proposed. Back in 2008, Androulaki *et al.*, [65] introduced a hybrid payment system that combined the

Peppercoin Micropayment system with "one use" electronic cash, proposed one theoretical design mechanism, Payment for Anonymous Routing (PAR), which can provide economic incentives for network participants. During the mechanism, a centralized bank released coins to clients, clients should pay coins when using every circuit.

In 2010, Ngan et al. [30] proposed "gold star" mechanism, that the Tor directory servers measure the performance of relays and grant the satisfactory relays in the directory with "gold star", that the better performance the relays have, the higher chance the relays are selected to build circuit.

Later, Jansen et al. [33] proposed BRAIDS, which encourage users to run Tor relays by introducing relay-specific tickets for service accounting. The tickets are embedded into Tor cells to request some sort of service，such as low-latency service, high-throughput service, and normal service.

Further, unlike above incentive and e-cash approaches that need centralized banks, Moore et al. [64] proposed decentralized incentive scheme, which Tortoise enforced low universal rate limit on individual client connections at the network's ingress points to achieve speedup. Recently, unlike recruiting relays, Wang et al. [34] propose user reputation system, rBridge, which uses an introduction-based mechanism to recruit new relays.

5. Build Anonymous Path

Although one node join the anonymous system，it does not mean the node must be used, but open the possibility of building anonymous path. The most section in building the path is anonymous routing algorithm. Anonymous routing algorithm is the core of anonymous communication systems, because it determines the whole system performance and security. Furthermore, the number of users decides anonymity degree that anonymous system can offer, while whether users use the anonymous system depend on system performance, so it's necessary to propose better routing algorithms to design better anonymous systems.

Before building the path, it's necessary to determine the anonymous path length, which is subject to network topology, and can be divided into variable length path and fixed length path.

A. Fixed length path

In terms of fixed length path, sender makes the decision to select all nodes to build path. Onion-Routing I and Freedom use fixed-length path [66,67]. In Mix cascaded topology, anonymous path is fixed. The initiator only decides which anonymous path is selected in anonymous communication. Although its length is fixed, its latency is still incalculable because intermediate nodes can't be sure its delay is short.

B. Variable length path

For variable length path (such as Crowds [1], Hordes [68], and AP3), sender only selects the first node, and can't determine the path length. For example, Crowds use the forwarding probability way that each node select next node to forward message or directly send the message to receiver depending on the forwarding probability. Thus the path length varies from each node forwarding probability. In extreme circumstances, the length may be infinitely long, then the latency will be incalculable.

6. Node Selection

Routing algorithm can divide into two sections: how many nodes and which nodes can be used to build anonymous path. In other words, they can be called node selection and path construction. Node selection can affect the bandwidth utilization ratio and throughput, and mainly solves the problem that how safely, rationally and efficiently select relay node. Path construction mainly solves the problem that how to use the selected nodes quickly

and effectively construct path. During node selection algorithms, different proposals have been presented from varying elements. Broadly speaking, node selection can divide into two classes: based on node characteristics (such as bandwidth) and based on link characteristics (such as latency). But the classification can't summarize all existing selection algorithms. Considering many network metrics can describe the performance of an anonymous path, such as bandwidth, latency, network jitter, loss, etc. existing algorithms can be subdivided into six classes, as follows.

6.1 Bandwidth-based Node Selection

Bandwidth is one of critical node properties. Especially in Tor, it select nodes based on bandwidth [69, 70]. In other words, the higher bandwidth the node have, the higher the probability of selecting the node is. In general, Tor relays report their bandwidth capabilities by themselves, and sender select relays based on bandwidth. There are some malicious relays that deliberately exaggerate their bandwidth to do some sort of attack, such as capturing a large fraction of paths. In other words, malicious node can report so higher bandwidth than actual bandwidth that many paths use the node, then the path is compromised. In turn, it cause insecurity and unreliable performance delivered to Tor users.

Many researchers have attempted to improve the performance problem. For example, Snader and Borisov [39] propose an opportunistic bandwidth measurement algorithm to improve Tor selection algorithm. Moreover, they also propose a tunable path selection that different users vary different sets of node from their preferences for anonymity versus performance. Then, Murdoch et al. [59] explore Tor current path selection algorithms with one Tor path simulator, such as bandwidth-weighted algorithm and uniform selection path algorithm.

6.2 Latency-based Node Selection

Considering that node's actual bandwidth is unknown, in other words, it's difficult to prove the node's self-reported bandwidth is true, bandwidth–based selection can cause some problems. Moreover, some nodes in the anonymous system are in a constant state of congested or unused, so that it not only leaves loopholes for attackers, but also degrades network performance. To solve these problems, latency-based path selection has been presented. Latency in anonymous path includes transmission delay, queuing delay and propagation delay. Transmission delay refers to the amount of time required to push all of the packet's bits into the wire. Queuing delay is the delay caused by a packet that it may wait in a queue until handled. Propagation delay, that is link latency, is the amount of time it takes for the packet to travel from the sender to the receiver along the anonymous path. Moreover, latency is the selection criteria of linked-based path selection. Put another way, the lower the delay is, the higher priority the link that is selected have.

In 2009, Panchenko and Renner [48] propose novel path selection based on actively measured the latencies of anonymous paths in terms of round-trip times (RTTs), and using passive observations of throughput to estimate available link-wise capacities.

Sherr et al. [36] present link-based relay selection which can provide more flexible routing and anonymous path with low latency and network jitter, and also introduce virtual coordinate system. The Euclidean distance between nodes in the virtual coordinate system is regarded as metric of latency. To protect the anonymity, both parties of communication don't take part in virtual coordinate system, only the set of nodes do. Before selecting path, sender firstly calculates the nodes' distance in the coordinate system, then estimate the latency of potential path.

In 2012, Wang et al. [53] points out the neglected problem easily that Tor path selection algorithm don't consider the current load of nodes. Unlike redesigning some congestion control methods [52,75], they lay emphasis upon identifying and avoiding congested

nodes. Congestion-aware path selection algorithm has come up that avoid selecting very congested circuits to decrease node latency. In detail, to decrease latency as an indicator of congestion, sender firstly use a combination of lightweight active and opportunistic measurements that means opportunistically sample RTTs across potential path and subtract the lowest recorded RTT to obtain the overall latency of the path, infer node latency to judge the state of nodes whether appear congested, and select nodes which don't appear congested.

Until recently, Wacek et al. [54] improve the Sherr's algorithm that combine Tor algorithm with Sherr et al. algorithm. Clients still use the Tor's relay selection to select k candidate paths, and then estimate the latency of each k candidate paths, finally select the path with lowest latency as the ultimate path. In addition, they indicate that if the value of k is 3, it can provide the best trade-off of performance and the time spent identifying the best path.

6.3 AS-aware Node Selection

Low-latency anonymous system is more vulnerable to timing attacks and statistical correlation attack. To mitigate the attack and improve anonymity, one method is location diversity by making each path more complex, such as spreading over multiple jurisdictions, not in the same autonomous system (AS). Therefore, the aware of autonomous system should be taken into account during node selection.

As is a network or a collection of networks under mutual administration that shares the same routing methodology, and can independently operate network. In 2004, Feamster and Dingledine [56] consider location independence metric, and argue that node selection algorithms that look only at IP prefixes likely inefficiently implement location independence, in which selected paths are subject to compromise anonymity by a single AS. They have found that different node selection algorithm help minimize the chance that entry path and exit path traverse the same AS, and the longer mix path has, the smaller parts a single AS can observe.

In 2009, Edman and Syverson [25] further study path inference algorithms, suggest that although Tor's node is increasing continuously, the probability of AS observing the ends of path is higher than previously thought. In turn, AS-aware node selection algorithm is proposed to resist AS level attackers.

In 2012, Akhoondi et al. [61] present tunable node selection algorithm that user can set a value between 0 to 1 to balance the anonymity and latency. Moreover, they estimate the geographic locations of clients and nodes using an IP geolocation database to decrease the latency. Besides, the algorithm can predict Internet routing between nodes and clients to resist correlated attack.

6.4 Application-aware Node Selection

To meet some specific applications' requirements, clients should have application-aware to select nodes. In 2007, Sherr et al. [35] firstly propose A3 to meet application-specified criteria. Later, A3 have been implemented by them [37], and they design a declarative language (A3LOG) that make applications to compactly specify path instantiation and node selection.

6.5 Trust-based Node Selection

In anonymous network, node frequently enter and exit the network, so that the trust between nodes become more and more important because it is often associated with the user status and information privacy protection. Trust-based path selection is taking trust into account selecting nodes.

In 2009, Johnson and Syverson [25] regard the probability that the adversary fails to compromise nodes as trust. They propose node trust model, and first explicitly use trust to

redesign node selection strategies, but only consider correlation attacks then. Later, they further identify the importance of trust, consider different users have different distributions on trust, then propose a novel trust-based node selection algorithm which can protect anonymous system from attacking a significant fraction of the network [26].

Besides these above node algorithms, the simplest algorithm is uniform random selection that the node is selected from a set of candidate nodes with uniform probability. In short, the probability that each node is selected to build anonymous path is same.

7. Locate Random Relays

Locate random relay also refers to secure lookup node. Existing locating relay methods are largely based on various distributed hash tables (DHT). DHT provides a lookup service similar to a hash table. In other words, DHT can extract some information in the file (usually a filename) as key to be unique values by hash function, and stores the (key, value) pairs, then any participating node can efficiently retrieve the value associated with a given key. Typical DHT algorithms include Chord, CAN, Pastry, Kademlia [71], and Tapestry. It's important to locate random nodes in anonymous communication systems. Unlike Tor, there are some anonymous communication systems which are based on Distributed Hash Table (DHT), such as Salsa [58], AP3 [72], NISAN [50], Torsk [32], Bifrost, Cashmere ShadowWalker, and Octopus [44].

Based on Chord[55], there are NISAN, Salsa, Bifrost which have implemented. Salsa uses a specifically designed secure lookup over a custom DHT to select nodes, which based on a Chord-like DHT that maps nodes to a point in an ID-space corresponding to the hash of their IP address. The identities are based on hashes of the nodes' IP addresses which are organized in a tree structure. However, Salsa's lookup may suffer selective DoS [38]. In 2009, Kondo et al. [60] design Bifrost system, which separate a node management layer (NML) realized by Chord from anonymous communication layer with multiplex encryptions to avoid anonymous route affecting node. NML can not only use the Finger Table of Chord to search the next-hop node, but also assign backup node once relay node seceded.

Based on Kademlia DHT and Myrmic [77], Torsk is designed by McLachlan et al in 2009. In Torsk, clients don't need to communicate with directory server, and combine its' lookup with root verification, buddies and cover traffic.

Based on Pastry DHT, Zhuang et al. [57] in 2005 design Cashmere, which use mix idea and multi-hop routing, and select a set of nodes in overlay namespace to solve the issues of node churn and improve the stability of anonymous path. The set of nodes is called a relay group, each node in the group can be a mix, and share a public/private key pair. Anonymous path can be built by various relay groups.

AP3 is similar to Crowds, and perform a stochastic expected-length random walk. However, the lookup mechanisms in Salsa and AP3 lack anonymity, so that adversary can infer the path structure and compromise user anonymity [42].

In 2009, Mittal and Borisov propose ShadowWalker [18], which uses special secure lookup protocol for redundant structured topology. That is, one node want to find an identifier ID, it will query the closest node to the ID in the finger table for its finger, which is the closest preceding node for ID.

8. Entry Guard

In general, the first relay in anonymous path is often called entry node. In all relay nodes, only entry node knows the communication initiator, and malicious entry nodes can make selective DoS attack more powerful so that many anonymous paths will be compromised [38, 74], therefore how to select entry node is of great significance to protect sender. In Tor, entry node is generally selected from node set with high stability (such as long time online) and bandwidth to cope with predecessor attack [63], locating hidden

service, and statistical profiling. In general, clients select three entry guards to apply to all circuits, and reselect new entry guard at intervals of 30 days to 60 days.

In 2006, Øverlier and Syverson [22] has researched entry nodes to protect hidden servers. The parameters in the nodes include the size of entry node set, performance, trust, and so on. They not only propose backup a longer list of entry nodes than normal entry node set to cope with the normal set unavailable, but also put forward layering entry nodes and select a small fixed number of nodes to always regard as entry nodes.

In 2012, Elahi et al. [31] design a simulation framework called Changing of the Guards (COGS) to provide quantitative data about Tor's entry guard selection algorithms with an empirical analysis. They suggest that natural churn and guard rotation are main factors affecting guard selection. The best balance between making guard nodes diverse and avoiding selecting malicious nodes as guard nodes should be further study.

Recently, Dingledine et al. [29] propose a single fast entry node rather than selecting three entry guards. While select one guard node can make clients poor performance, raising the guard node bandwidth threshold they proposed solve the problem. In general, the lifespan of one guard node is from 30 days to 60 days, but they set the guard rotation period as 9 months to increase the time to first compromise.

9. Conclusion

Frequent node churn make it difficult to maintain a stable network topology, therefore it is necessary to get more insight into the overall anonymous network topology and its' nodes, because systematic anonymous network topology knowledge can help better cope with various attackers. This paper present a structured framework of anonymous network topology related organization, then present the classification knowledge system of anonymous network topology, including recruiting new node, building path, and so on. These frameworks provide a basic research overview to research network topology and node selection. Based on this, concrete inspection solutions to the issues of anonymity and performance in anonymous communication should be in-depth studied.

In detail, there are still many problems to be solved in the topology research. Although there are many improved algorithms for node selection and entry guard proposed under the frame of theory, the practicality of these algorithms to be applied in live anonymous system is unknown, then it's urgent that develop new method to model Tor more realistically. In addition, it's worth exploring new node selection algorithm to defend themselves against various attacks, such as predecessor attack and routing capture attack. At last, some measures of improving anonymity tend to increase system loads, add long latency to the processing of service, and reduce efficiency. It's certainly worth considering that how to balance anonymous systems' anonymity and performance.

Acknowledgment

This work is supported by the following programs: the National Natural Science Foundation of China under Grant No.61170273; the China Scholarship Council under Grant No.[2013]3050; Open Project Foundation of Information Technology Research Base of Civil Aviation Administration of China (NO. CAAC-ITRB-201201);2010 Information Security Program of China National Development and Reform Commission with the title "Testing Usability and Security of Network Service Software".

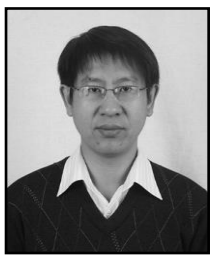
References

[1] M. Reiter and A. Rubin, "Crowds: Anonymity for web transactions", ACM Transactions on Information and System Security, vol.1, no.1, **(1998)**, pp.66-92.
[2] Tor Metrics, https://metrics.torproject.org/bubbles.html#country.
[3] U.S. Naval Research Laboratory, Center for High Assurance Computer Systems, http://www.nrl.navy.mil/itd/chacs/.
[4] The Shadow Simulator, http://shadow.github.io/.
[5] The Tor Path Simulator (TorPS), http://torps.github.io/.
[6] Safer Warfighter Communications (SAFER), http://www.darpa.mil/OpenCatalog/SAFER.html.
[7] https://www.torproject.org/getinvolved/volunteer.html.en#project-atlas.
[8] SecurityLab at Georgetown University, https://security.cs.georgetown.edu/.
[9] CAREER: Private Communication in Strongly Adversarial Networks, http://www.nsf.gov/awardsearch/showAward?AWD_ID=1149832&HistoricalAwards=false.
[10] Cryptography, Security, and Privacy Research Group, https://crysp.uwaterloo.ca/research/.
[11] http://www.inf.tudresden.de/index.php?node_id=447&ln=en.
[12] CAREER: Scale and Performance in Anonymous Communication, http://nsf.gov/awardsearch/showAward?AWD_ID=0953655.
[13] CT-S: Application-Aware Anonymity (A3) for the Masses, http://nsf.gov/awardsearch/showAward?AWD_ID=0831376.
[14] Nets: Small: Optimal Design of Anonymous Network Systems under Resource Constraints, http://www.nsf.gov/awardsearch/showAward?AWD_ID=1117701&HistoricalAwards=false.
[15] Computer Security and Industrial Cryptography, http://www.esat.kuleuven.be/cosic/?page_id=13.
[16] C. Diaz, S. J. Murdoch and C. Troncoso, "Impact of Network Topology on Anonymity and Overhead in Low-Latency Anonymity Networks", In the Proceedings of the 10th Privacy Enhancing Technologies Symposium (PETS 2010), **(2010)**; Berlin, Germany.
[17] G. Danezis, "Mix-networks with Restricted Routes", In the Proceedings of Privacy Enhancing Technologies workshop (PET 2003), **(2003)**.
[18] P. Mittal and N. Borisov, "ShadowWalker: Peer-to-peer Anonymous Communication using Redundant Structured Topologies", In the Proceedings of the 2009 ACM Conference on Computer and Communications Security, **(2009)**; Chicago, Illinois, USA.
[19] D. M. Goldschlag, M. G. Reed and P. F. Syverson, "Hiding routing information", In Ross Anderson, editor, Information Hiding: First International Workshop, **(1996)**; Springer-Verlag.
[20] P. Syverson, G. Tsudik, M. Reed, and C. Landwehr, "Towards an Analysis of Onion Routing Security", In Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability, **(2000)**.
[21] R. Dingledine, V. Shmatikov and P. Syverson, "Synchronous batching: From cascades to free routes", In Proceedings of the Privacy Enhancing Technologies Workshop, **(2004)**; Springer LNCS 3424.
[22] L. Øverlier and P. Syverson, "Locating hidden servers", In Proceedings of the 2006 IEEE Symposium on Security and Privacy, **(2006)**.
[23] G. Danezis and P. Syverson, "Bridging and Fingerprinting: Epistemic Attacks on Route Selection", In the Proceedings of the Eighth International Symposium on Privacy Enhancing Technologies, **(2008)**; Leuven, Belgium.
[24] A. Johnson and P. Syverson, "More anonymous onion routing through trust. In 22nd IEEE Computer Security Foundations Symposium, **(2009)**; New York.
[25] M. Edman and P. F. Syverson, "AS-awareness in Tor path selection", In the Proceedings of the 2009 ACM Conference on Computer and Communications Security, **(2009)**; Chicago, Illinois, USA.
[26] A. Johnson, P. Syverson, R. Dingledine and N. Mathewson, "Trust-based Anonymous Communication: Adversary Models and Routing Algorithms", In the Proceedings of the 18th ACM conference on Computer and Communications Security, **(2011)**.
[27] R. Jansen, A. Johnson and P. F. Syverson, "LIRA: Lightweight Incentivized Routing for Anonymity//NDSS, **(2013)**.
[28] R. Jansen, A. Miller, P. Syverson and B. Ford, "From onions to shallots: Rewarding Tor relays with TEARS[J]", HotPETS, **(2014)**.
[29] R. Dingledine, N. Hopper, G. Kadianakis and N. Mathewson, "One Fast Guard for Life (or 9 months)", **(2014)**.
[30] T.-W. Ngan, R. Dingledine and D. S. Wallach, "Building Incentives into Tor", In the Proceedings of Financial Cryptography (FC '10), **(2010)**.
[31] T. Elahi, K. Bauer, M. AlSabah, R. Dingledine and I. Goldberg, "Changing of the Guards: A Framework for Understanding and Improving Entry Guard Selection in Tor", In the Proceedings of the Workshop on Privacy in the Electronic Society (WPES 2012), **(2012)**; Raleigh, NC, USA.
[32] J. McL. lan, A. Tran, N. Hopper and Y. Kim, "Scalable onion routing with torsk, ACM CCS, **(2009)**.
[33] R. Jansen, N. Hopper and Y. Kim, "Recruiting New Tor Relays with BRAIDS", In the Proceedings of the 2010 ACM Conference on Computer and Communications Security, **(2010)**; Chicago, Illinois, USA.

[34] Q. Wang, Z. Lin, N. Borisov and N. J. Hopper, "rBridge: User Reputation based Tor Bridge Distribution with Privacy Preservation", In the Proceedings of the Network and Distributed System Security Symposium - NDSS'13, **(2013)**.
[35] M. Sherr, B. T. Loo and M. Blaze, "Towards Application-Aware Anonymous Routing", In: USENIX Workshop on Hot Topics in Security, **(2007)**.
[36] M. Sherr, M. Blaze and B. T. Loo, "Scalable Link-Based Relay Selection for Anonymous Routing", In the Proceedings of Privacy Enhancing Technologies, 9th International Symposium, **(2009)**.
[37] M. Sherr, A. Mao, W. R. Marczak, W. Zhou, B. T. Loo and M. Blaze, "A3: An Extensible Platform for Application-Aware Anonymity", InNetwork and Distributed System Se-curity Symposium, **(2010)**.
[38] N. Borisov, G. Danezis, P. Mital and P. Tabriz, "Denial of service or denial of security?", In Proceedings of the 14th ACM conference on Computer and communications security, **(2007)**; New York, USA.
[39] R. Snader and N. Borisov, "A Tune-up for Tor: Improving Security and Performance in the Tor Network", In: 15th Annual Network and Distributed System Security Symposium, **(2008)**.
[40] R. A. Snader, "Path Selection for Performance and Security-improved Onion Routing", Doctor thesis, University of Illinois at Urbana-Champaign,**(2009)**.
[41] K. Bauer, J. Juen, N. Borisov, D. Grunwald, D. Sicker and D. McCoy, "On the optimal path length for Tor", HotPets in conjunction with Tenth International Symposium on Privacy Enhancing Technologies, **(2010)**; Berlin, Germany.
[42] Q. Wang, P. Mittal and N. Borisov, "In search of an anonymous and secure lookup: attacks on structured peer-to-peer anonymous communication systems",Proceedings of the 17th ACM conference on Computer and communications security, **(2010)**.
[43] R. Snader and N. Borisov, "Improving Security and Performance in the Tor Network through Tunable Path Selection", IEEE Transactions on Dependable and Secure Computing, vol.8, no.5, **(2011)**.
[44] Q. Wang, and N. Borisov, "Octopus: A secure and anonymous dht lookup", Distributed Computing Systems (ICDCS), IEEE 32nd International Conference on, **(2012)**.
[45] P. Lincoln, I. Mason, P. Porras, V. Yegneswaran, Z. Weinberg, J. Massar and D. Boneh, "Bootstrapping Communications into an Anti-Censorship System", 2nd USENIX Workshop on Free and Open Communications on the Internet, **(2012)**.
[46] R. Böhme, G. Danezis, C. Diaz, S. Köpsell and A. Pfitzmann, "On the PET workshop panel mix cascades versus peer-to-peer: is one concept superior?", Privacy Enhancing Technologies, **(2005)**; Springer Berlin Heidelberg.
[47] G. Danezis and R. Clayton, "Route fingerprinting in anonymous communications[C]//Peer-to-Peer Computing", Sixth IEEE International Conference, **(2006)**.
[48] A. Panchenko and J. Renner, "Path selection metrics for performance-improved onion routing", In Proceedings of the 2009 Ninth Annual International Symposium on Applications and the Internet (Washington, DC, USA, 2009), IEEE Computer Society, **(2009)**.
[49] B. Westermann, A. Panchenko and L. Pimenidis, "A kademlia-based node lookup system for anonymization networks", Advances in Information Security and Assurance, Springer Berlin Heidelberg, **(2009)**, pp.179-189.
[50] A. Panchenko, S. Richter and A. Rache, "Nisan: Network information service for anonymization networks", ACM CCS, **(2009)**.
[51] C. Tang and I. Goldberg, "An improved algorithm for Tor circuit scheduling", Proceedings of the 17th ACM conference on Computer and communications security. ACM, **(2010)**.
[52] M. AlSabah, K. Bauer, I. Goldberg, D. Grunwald, D. McCoy, S. Savage and G. M. Voelker, "DefenestraTor: Throwing out windows in Tor", In PETS **(2011)**.
[53] T. Wang, K. Bauer, C. Forero and I. Goldberg, "Congestion-aware Path Selection for Tor", In the Proceedings of Financial Cryptography and Data Security, **(2012)**.
[54] C. Wacek, H. Tan, K. Bauer and M. Sherr, "An Empirical Evaluation of Relay Selection in Tor", In the Proceedings of the Network and Distributed System Security Symposium, **(2013)**.
[55] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek and H. Balakrishnan, "Chord: a Scalable Peer-to-peer Lookup Service for Internet Applications", In Proc. of ACM SIGCOMM, **(2001)**.
[56] N. Feamster and R. Dingledine, "Location Diversity in Anonymity Networks", In the Proceedings of the Workshop on Privacy in the Electronic Society, **(2004)**; Washington, DC, USA.
[57] L. Zhuang, F. Zhou, B. Y. Zhao and A. Rowstron, "Cashmere: Resilient Anonymous Routing", In Proceedings of the 2nd Conference on Symposium on Networked Systems Design and Implementation, **(2005)**.
[58] A. Nambiar and M. Wright, "Salsa: A structured approach to large-scale anonymity", ACM CCS, **(2006)**.
[59] S. J. Murdoch and R. N. M. Watson, "Metrics for Security and Performance in Low-Latency Anonymity Systems", In Privacy Enhancing Technologies Symposium, **(2008)**.
[60] M. Kondo, S. Saito, K. Ishiguro, H. Tanaka and H. Matsuo, "Bifrost: A Novel Anonymous Communication System with DHT", In Second International Workshop on Reliability, Availability, and Security, **(2009)**.
[61] M. Akhoondi, C. Yu and H. V. Madhyastha, "LASTor: A Low-Latency AS-Aware Tor Client", In the Proceedings of the 2012 IEEE Symposium on Security and Privacy, **(2012)**.

[62] L. F. Franco, "A survey and comparison of anonymous communication systems: Anonymity and security", [J], **(2012)**.
[63] M. Wright, M. Adler, B. N. Levine and C. Shields, "The Predecessor Attack: An Analysis of a Threat to Anonymous Communications Systems", ACM Transactions on Information and System Security (TISSEC), vol.4, no.7, **(2004)**, pp.489–522.
[64] B. Moore, C. Wacek and M. Sherr, "Exploring the Potential Benefits of Expanded Rate Limiting in Tor: Slow and Steady Wins the Race With Tortoise", In Annual Computer Security Applications Conference, **(2011)**.
[65] E. Androulaki, M. Raykova, S. Srivatsan, A. Stavrou and S. M. Bellovin, "PAR: Payment for anonymous routing", InPETS '08: Proceedings of the 8th International Symposium on Privacy Enhancing Technologies, **(2008)**.
[66] Y. Guan, X. Fu, R. Bettati and W. Zhao, "An optimal strategy for anonymous communication protocols//Distributed Computing Systems", Proceedings. 22nd International Conference on. IEEE, **(2002)**.
[67] Y. Guan, X. Fu, R. Bettati and W. Zhao, "A quantitative analysis of anonymous communications", Reliability, IEEE Transactions on, vol.53, no.1, **(2004)**, pp.103-115.
[68] B. N. Levine, C. Shields, "Hordes: a multicast based protocol for anonymity", Journal of Computer Security, vol.10, no.3, **(2002)**, pp.213-240.
[69] Tor Path Specification. Tor Project.http://gitweb.torproject.org/tor.git?a=blob_plain;hb=HEAD;f=doc/spec/path-spec.txt. **(2010)**.
[70]R. Dingledine, N. Mathewson and P. Syverson, "Tor: The second-generation onion router", Naval Research Lab Washington DC, **(2004)**.
[71] P. Maymounkov and D. Mazieres, "Kademlia: A peer-to-peer information system based on the xor metric", IPTPS, **(2001)**.
[72] A. Mislove, G. Oberoi, A. Post, C. Reis, P. Druschel and D. S. Wallach, "Ap3: Cooperative, decentrialized anonymous communication", ACM SIGOPS European Workshop, **(2004)**.
[73] G. Ciaccio, "Improving Sender Anonymity in a Structured Overlay with Imprecise Routing", In the Proceedings of the Sixth Workshop on Privacy Enhancing Technologies, **(2006)**; Cambridge, UK.
[74] T. G. Abbott, K. J. Lai, M. R. Lieberman and E. C. Price, "Browser-based Attacks on Tor", In Proceedings of the 7th International Conference on Privacy Enhancing Technologies, Berlin, Heidelberg, **(2007)**; Springer-Verlag.
[75] J. Reardon, "Improving Tor using a TCP-over-DTLS tunnel", [J], **(2008)**.
[76] J. McLachlan and N. Hopper, "Don't clog the queue: Circuit clogging and mitigation in P2P anonymity schemes", In Proceedings of FC, **(2008)**.
[77] P. Wang, I. Osipkov, N. Hopper and Y. Kim, "Myrmic: Secure and robust dht routing", Tech. Rep. University of Minnesota DTC Research Report, **(2006)**.
[78] O. Berthold, A. Pfitzmann and R. Standtke, "The disadvantages of free MIX routes and how to overcome them", In H. Federrath, editor, Designing Privacy Enhancing Technologies: Workshop on Design Issue in Anonymity and Unobservability,**(2000)**; Springer-Verlag.
[79] D. Goldschlag, M. Reed and P. Syverson, "Onion routing for anonymous and private internet connections", Communications of the ACM(USA), vol.42, no.2, **(1999)**, pp.39-41.
[80] F. Chen and M. Perry, "Improving Tor path selection", https://gitweb.torproject.org/torspec.git/blob_plain/HEAD:/proposals/151-path-selection-improvements.txt, **(2008)**.

Authors



Tian-Bo Lu, he was born in Guizhou Province, China, 1977. He is an Associate professor in School of Software Engineering, Beijing University of Posts and Telecommunications, China. His technical interests include information security and computer network.

Shi-Xian Du, she is a graduate in School of Software Engineering, Beijing University of Posts and Telecommunications, China. Her research interests include information and network security, anonymous communication, and P2P computing.

Yang Li, he was born in Hunan Province, China, 1978. He is a PhD and his technical interests include information security, distributed computing and P2P network.

Pei-Yuan Dong, she is a lecturer in School of Software Engineering, Beijing University of Posts and Telecommunications, China. Her research interests include network security and Cyber Physical System.

Xiao-Yan Zhang, she was born in Shandong Province, China, 1973. She is an Associate professor in School of Software Engineering, Beijing University of Posts and Telecommunications, Beijing, China. Her technical interests include software cost estimation and software process improvement.

