Local Anonymity: A Metric for Improving User Privacy in Tor

Simina Brânzei
Cheriton School of Computer Science
University of Waterloo
sbranzei@cs.uwaterloo.ca

Tariq Elahi
Cheriton School of Computer Science
University of Waterloo
mtelahi@cs.uwaterloo.ca

Ian Goldberg
Cheriton School of Computer Science
University of Waterloo
iang@cs.uwaterloo.ca

ABSTRACT

In anonymous communication networking, entropy is a popular metric for measuring the average-case difficulty of linking two communicating parties. This paper proposes an alternate view of anonymity, rooted in the observation that global measures of anonymity do not necessarily provide accurate information about the anonymity of an individual user. Such differences can arise due to the characteristics of user needs and local network conditions. We introduce the notion of local anonymity to account for the anonymity of individual users and different sub-regions of the network. Based on this idea, we propose a relay selection algorithm in Tor, introduce a metric for computing the anonymity of individual relays, and discuss the choice of parameters, possible attacks, and defence strategies.

Categories and Subject Descriptors

K.6.5 [**Security and Protection**]: Anonymous communication; C.2.2 [**Computer Communication Networks**]: Network protocols

General Terms

Algorithms, Security

Keywords

Local anonymity, Privacy preserving, User-defined privacy preference, Relay selection algorithms

1. INTRODUCTION

Anonymity-preserving communication networks, such as the popular Tor network [3], provide privacy-preserving capabilities for users who wish to remain anonymous on the Internet. Given that the amount of traffic on such networks is increasing over time [11], anonymity networks are becoming more attractive for both adversaries [10] and security researchers alike, albeit for different reasons.

The main metric used to measure anonymity is entropy [2]. Entropy is an increasing function of the number of users on the network, and is generally believed to capture the difficulty that an adversary would have when trying to identify a particular user. High entropy values can only be attained when the number of users is large and sustainable, and so anonymity networks constantly try to increase their user base.

Tor is composed of relays, also referred to as routers or nodes, which are used to forward traffic through the network. Relays can be flagged as "guard" relays, "exit" relays, both, or neither, by flags set in their descriptors. Prior to using the network, each user is given a set of three guard relays to use over a fixed period of time (currently a month). The user's Tor client constructs an encrypted circuit over an entry, a middle and an exit relay. The client uses one of the three relays in its guard set for the entry position, chooses a suitable exit node from the set of exit relays, and finally chooses a middle node from the remaining relays. Clients discover relays through the directory service. The network provides some security assurances that the data is accurate using a consensus-based approach over trusted directory authorities. Relays can be operated by any individual or organization around the world. The risk, however, is that they can be owned or monitored by adversarial entities which may leak private information, and so relay selection algorithms must be carefully designed to avoid user exposure.

Most relay selection algorithms [3, 6, 7] use global measures to quantify the anonymity of all users on the network. We observe some fundamental problems with this approach. To reduce the complexity of deanonymization, an adversary can look only at subsets of the network; however, global measures do not provide any information about the anonymity level of the subsets, which we refer to as local anonymity. In order for deanonymizing to still be nontrivial, a relay selection algorithm should ensure high local anonymity. Furthermore, algorithms based on global measures also do not provide guarantees with respect to the anonymity level of a particular user. Indeed, it is conceivable that many users receive little protection even though the measure value may be high. This problem becomes much more severe when users with high privacy requirements are among the affected ones.

Our contributions in this paper are as follows. We observe a fundamental problem with entropy based schemes, namely that they fail to provide a measure of anonymity at the local level, and do not measure an individual user's anonymity on the network. We then observe the existence of a tension between anonymity and quality of service. Based on this observation, we formulate a notion of local anonymity and design a novel relay selection algorithm that attempts to provide local anonymity, while respecting the user's anonymity and performance requirements at all times. We provide discussions of parameter choices that would suit deployment in the current version of Tor, as well as the robustness to attack of our proposed scheme.

The remainder of this paper is structured as follows: We consider measures of general anonymity in Section 2 and our proposed local anonymity metric in Section 3. We characterize our adversary in Section 4 and discuss the role of user choice in Section 5. We introduce our algorithm in Section 6 and discuss its parameters in Section 7. Deployment of the proposed scheme is discussed in Section 8 and attacks are discussed in Section 9. Finally we conclude with a discussion of future directions in Section 10.

2. MEASURING ANONYMITY

One of the simplest definitions of anonymity is the following.

Anonymity is the state of being not identifiable within a set of objects, the anonymity set (Pfitzmann and Köhntopp, [4]).

For example, a sender can be anonymous within a set of potential senders, which form his sender anonymity set. Similarly, a recipient would be anonymous within a set of potential recipients. The anonymity set is a compelling notion, however it suffers from several shortcomings as a measure of anonymity. For example, it does not account for probability fluctuations in the anonymity set, making certain senders and receivers more likely than others. Thus, in reality two anonymity sets of equal size can provide different guarantees of anonymity [12].

The standard information-theoretic measure of anonymity is entropy, which was proposed by Serjantov and Danezis [5] and Díaz *et al.* [2].

Formally, let $N = \{1, \ldots, n\}$ be the set of users of a system. For each message m, let $P_m(j)$ denote the probability that the sender (receiver) of the message is user j. Then the entropy of the distribution of users in N with respect to m is:

$$H_m = -\sum_{j=1}^{n} P_m(j) \lg \left(P_m(j)\right).$$

Entropy is a global measure of uncertainty, and captures the idea that the distribution of users is an important consideration when computing anonymity in a system.

3. LOCAL ANONYMITY IN Tor

We now introduce local anonymity as an important measure when quantifying the anonymity of a system. We argue that any global measure of anonymity suffers from some fundamental shortcomings. First, global measures fail to give a particular user any guarantees. Such a user may be completely exposed to the adversary, even though the global measure indicates that the system offers very good anonymity.

Second, an adversary can easily focus only on a sub-region of the network, either because of being bounded, or because that sub-region appears to be most promising. In systems that rely only on entropy to measure anonymity, it is easily conceivable that both these shortcomings hold. Intuitively, the entropy of the system should remain high even when restricting to sub-regions of the network, or even to a single relay. This way an adversary would find it hard to identify a user no matter at what level they may analyze the network.

The observation of global measures failing to provide guarantees to individual users has also been made by other researchers, who, like us, argued for local anonymity [12]. We note, however, that their focus is different, as they introduce a theoretical model of source and destination hiding as a proxy for local anonymity. On the other hand, we are interested in a broad definition of local anonymity which encompasses the requirement that a particular user is unidentifiable, and we tailor our analysis to achieving local anonymity in the Tor network. Syverson [9] also describes problems with the entropic view of anonymity. He argues that, for a measure to be useful, it should reflect the effort that an adversary would have to exert before being able to deanonymize the users, and entropy does not measure that.

Towards this end, we equate local anonymity in Tor with anonymity at a given relay.

Definition 1 (Local Anonymity). *The level of anonymity at a relay equals the number of unique indistinguishable users accessing that relay at the same time; i.e., the size of the relay's anonymity set.*

We note that existing relay selection schemes do not attempt to ensure that individual users are not alone at a given relay. Distinguishing a unique user by an adversary who decides to monitor such a relay becomes relatively easier, compared to the case where cover traffic is provided by having the relay being assigned to several users at the same time. Based on this observation, we propose a relay selection algorithm in Tor that incorporates this notion of local anonymity, while spreading users over the entire set of relays.

4. ADVERSARIAL MODEL

The adversary we defend against is modelled as external, static, passive, and partial. Being external implies that he does not control relays in the network nor can he influence their operation to add, drop, delay, or otherwise alter the flow of traffic. Also, the adversary can not see the internal state or data of any relay, and may only monitor the connection between nodes and the edges of the network.

The adversary has partial visibility of the network at any given time. The size of the visible network may depend on the position of the adversary in a portion of the underlying physical network that he controls or can gain access to. Being static, however, the particular relays that are visible at any given time are not in the control of the adversary.

An example of this type of adversary is an Internet Service Provider and the operator that controls it. This type of entity is also referred to as an Autonomous System (AS). He can view all traffic that flows on his own network but can not view traffic on links that fall outside his infrastructure. He does not know the internal state of the computers connecting to the network. Similarly, he does not control the transmission of data on the network originating and destined for these same computers. At most he can shape the traffic to maintain service levels but does so at a coarse level. He is, however, able to monitor and record data all transmissions of computers attached to his network. He may request transmission logs from other ISPs, where such an arrangement exists, but does so under constraints that require positive identification of the records being requested and disallows blanket requests. Given the legal constraints on sharing customer records, this is a reasonable assumption.

We further refine our adversary as a corrupt employee, or a planted spy, at a Network Operations Center (NOC) where several AS networks join and are managed. This individual

seeks to compile information about communication patterns of individuals of interest to sell or to use as leverage. He does not have a free hand in monitoring everything at the NOC, because of rules concerning separation of duty and confidentiality, but can serendipitously monitor the traffic on a few AS networks at once.

The motivation for protecting against this adversary is that a potent threat to the privacy of Tor users is from their ISP / AS operator who may, through customer records and pervasive traffic-logging, link traffic and identity trivially. Real world examples include the "Great Firewall" of China, or AT&T in the U.S. [13].

5. USER CHOICE

Algorithms sometimes compromise between performance and anonymity. They try to achieve a common default, without differentiating between users who require high anonymity and users who require less protection but may have higher standards for the quality of service.

A prime example of this is the relay selection algorithm found in the Tor network. By fixing the anonymity vs. performance trade-off, the Tor designers aim to (*i*) provide reasonable performance to encourage clients to use Tor, since large numbers contribute to increasing everyone's anonymity, and (*ii*) maintain the health of the network, as measured by parameters such as performance, number of relays, and number of exit nodes.

However, these compromises can result in suboptimal experiences for many users. For example, users with high bandwidth needs do not always get high-performance relays. One way to address this problem is to get users' input as to their privacy and performance preferences. Algorithms such as Tunable Tor [7] and WEIGHTED [6] give users the ability to choose their own anonymity vs. performance trade-off. The advantage is that users can fine tune their experience according to their needs. They attempt to ensure that a user's experience matches the user's choice, and also to maintain the anonymity level provided by the network as a whole. However, these algorithms do not provide assurances in the sense of local anonymity as defined in Section 3 and so a user can often be assigned to relays with poor anonymity scores (see below).

In creating a local anonymity aware scheme, we rely on the following observation: in general, anonymity and quality of service (QoS) are at odds with one another. High anonymity is positively influenced by the existence of a large number of users accessing a resource, while high quality of service, as measured by bandwidth, latency, and jitter, for example, is negatively influenced by many users accessing the same resources.

Given this relationship, we model each user by her preference between the two. Let $\theta \in [0, 1]$ be a parameter illustrating the user's requirement. A value of $\theta = 0$ expresses maximum privacy considerations, while $\theta = 1$ reflects maximal QoS requirements. In reality, a good algorithm should ensure that all users obtain a minimal acceptable level for both anonymity and quality of service.

6. OUR ALGORITHM: LocaTor

In this section we introduce LocaTor, a relay selection algorithm that is local anonymity aware and which attempts to also provide individual guarantees with respect to anonymity and performance.

Let $\mathcal{L}$ and $\mathcal{Q}$ be two functions defined on the set of relays and taking real values, such that $\mathcal{L}(R)$ represents the anonymity score of relay R, while $\mathcal{Q}(R)$ represents the QoS score. In Section 7.3 we discuss ideas for instantiating the $\mathcal{L}$ and $\mathcal{Q}$ functions. At all times, LocaTor maintains two copies of the relay list, $\mathcal{A}$ and $\mathcal{B}$, sorted by anonymity scores and quality of service, respectively.

Given the preference θ of a particular user, LocaTor maps θ to a relay R_θ in the anonymity list. We discuss some ideas for such mappings in Section 7.1. LocaTor then selects a window of relays, denoted W_θ, surrounding R_θ in the anonymity-sorted list $\mathcal{A}$. The relays in W_θ are sorted by QoS, after which those deemed to have unacceptable QoS are marked as unavailable; this approach is already implemented by today's Tor, which never uses the slowest 1/8 of the nodes. Finally, an exit and middle relay are chosen at random from the remaining relays in the window W_θ. If either an exit or middle with acceptable QoS does not exist in W_θ, the window size is increased and the search is started over for the missing relay, until both a middle and exit have been selected. The pseudocode is given in Algorithm 1.

Algorithm 1 Relay Selection with LocaTor

```
exit ← NULL; // Exit relay
middle ← NULL; // Middle relay
// A is the anonymity-sorted relay list (using L)
Map θ to relay R_θ in A;
// Sample mapping: Position(R_θ) = θ · (Len(A) − 1)
WS ← Default-Window-Size;
while (middle = NULL) or (exit = NULL) do
   W_θ ← Window of size WS in A of relays around R_θ;
   // B is the QoS-sorted relay list (using Q)
   Locate W_θ in B;
   Remove relays with low QoS from W_θ;
   if (W_θ has exit relays) then
      exit ← Random-Exit-Relay (W_θ);
   end if
   if (W_θ has middle relays) then
      middle ← Random-Middle-Relay (W_θ);
   end if
   if (middle = NULL) or (exit = NULL) then
      Increase(WS);
   end if
end while
return (exit, middle);
```

A feature of LocaTor is that the sorted relay list is highly dynamic, since the rank of a relay can vary significantly over time depending on how many users are using it and what their bandwidth requirements are. In particular, high-bandwidth relays can support much more variability in the number of users, and so can be situated anywhere in the anonymity- and QoS-sorted lists. Alternatives to our selection mechanism, such as using a Gaussian distribution around R_θ, instead of the window W_θ, are discussed in Section 7.4.

7. PARAMETERS

Having described the algorithm of LocaTor, we now discuss the choice of parameters suitable for implementing the algorithm within Tor.

7.1 Mapping User Requirements to Relays

The user specifies her preference on the anonymity versus quality of service scale using parameter $\theta \in [0, 1]$. When selecting a relay for such a user, we need a mapping function to determine where in the list of relays one should make the selection from. A simple linear mapping can be computed as follows:

$$Position(R_\theta) = \theta \cdot (Len(\mathcal{A}) - 1),$$

where R_θ is the relay around which the algorithm searches and $Position(R_\theta)$ is the relay's position in the anonymity-sorted list $\mathcal{A}$. Under this mapping, $\theta = 0$ corresponds to a relay currently situated at the top of $\mathcal{A}$, that is $Position(R_0) = 0$. On the other hand, $\theta = 1$ is mapped to the bottom relay in $\mathcal{A}$, which is situated at index $Len(\mathcal{A}) - 1$. We note that such a linear mapping is equivalent to the user directly choosing the rank of the relay they want to be situated on. However, since users should not be aware of low-level details of the implementation, a preference is a much more user-friendly option to specify.

Local anonymity guarantees cannot be absolute, since they depend on the number of users on the network and their preferences. In our context, the preference θ is equivalent to a ranking, which specifies how the anonymity requirement of a user is related to that of others. Clearly such preference specifications are most useful when there exist users with different values of θ. Also, in reality, the value of θ should be discretized, such that users would only choose among several options. Some possibilities for implementing user selection include showing several different activities, such as email, banking, internet surfing, and asking the user to select the ones they intend to perform. Alternatively, users could be asked to identify themselves with a persona, such as being a whistleblower versus someone who wants to help the Tor project by providing resources to the network. Finally, they could directly specify the tradeoff of anonymity versus QoS, by selecting one of the following options:

- *Performance* $\gg$ *Anonymity*
- *Performance* $>$ *Anonymity*
- *Performance* $\approx$ *Anonymity*
- *Anonymity* $>$ *Performance*
- *Anonymity* $\gg$ *Performance*.

It would be very useful for the implementation of LocaTor, as well as of others that also consider user preference, such as Tunable Tor [7], to conduct a user study that would show what realistic distributions of θ are.

7.2 Window Size

Ideally, the window size should be chosen depending on the distribution of θ. Since the user preference distribution is not currently known, we will assume the uniform distribution. We propose the following window size:

$$WS = \frac{Num\text{-}Relays}{Num\text{-}User\text{-}Choices},$$

where *Num-Relays* represents the number of number of relays in Tor, while *Num-User-Choices* is the number of different values of θ that the user can choose from. For example, given that the number of relays is 2000 and the number of different values of θ is 5, we would set $WS = 400$. This window size is large enough to ensure a good spread when selecting the relay of a given user, yet small enough so that relays within this window size have similar $\mathcal{L}$ and $\mathcal{Q}$ values.

In the rare event that the search inside a given window fails, possibly due to all the relays in that window having unacceptable QoS, the window size is increased by a fixed percentage. A possible update rule is $WS' \leftarrow \frac{3}{2}WS$.

7.3 Scoring Functions

The scoring functions $\mathcal{L}$ and $\mathcal{Q}$ are used to compute the anonymity and quality of service scores for each relay, respectively. The quality of service that a relay offers is often a function of many variables, including bandwidth, queueing latency, and jitter. Here we use a simplified approach and approximate the quality of service $\mathcal{Q}$ using bandwidth, which is already known. In general, it would be desirable to incorporate additional characteristics, such as queueing latency, in the formulas.

Given our definition of local anonymity (Definition 1), we desire to sort the relays by $\mathcal{L}^*(R) = |\text{Anonymity-Set}(R)|$. However, the size of the anonymity set at a relay is not easy to obtain, and so we also approximate the ordering induced by $\mathcal{L}^*$ with one induced by a function $\mathcal{L}$ depending on measures of bandwidth. Thus, given relay R, possible choices for $\mathcal{L}$ and $\mathcal{Q}$ are:

$$\mathcal{L}(R) = Max\text{-}Bandwidth(R) - Current\text{-}Bandwidth(R)$$

and:

$$\mathcal{Q}(R) = Current\text{-}Bandwidth(R)$$

The value $Max\text{-}Bandwidth(R)$ can be obtained from the Tor directory service through the server descriptors which are downloaded periodically by the client.

The $Current\text{-}Bandwidth(R)$ value is not currently available, but can be obtained either (i) by the client, who can opportunistically probe for it, or (ii) from the reports of relay nodes, utilizing an EigenSpeed-like [8] approach to provide a level of assurance of accurate information. The first option has relatively lower communication complexity and does not place more trust in relays than the current Tor network; however, it may provide a limited view of the network, and so the scores computed from collected data might be very different from the actual scores. This can lead to effectively random relay assignments. The second option likely provides a much more accurate view of the network and is more robust to dishonest reports, at the expense of higher communication complexity.

7.4 LocaTor Variants

The version of LocaTor given in Algorithm 1 maps the user's choice, θ, to a relay R_θ, and selects uniformly at random from a window of relays around R_θ. An interesting alternative is to select relays according to a Gaussian distribution over the anonymity-sorted list, $\mathcal{A}$. Under this model, the relay most likely to be selected is R_θ, and so the mean of the distribution would be the index $Position(R_\theta)$. The variance can be defined as $2WS$, which ensures that approximately 68% of the time, a relay is chosen from the window centered at R_θ, and 95% of the time, at a distance of at most $3/2WS$ from R_θ. Whenever the index of the relay selected according to this distribution falls outside the valid range, $[0, \ldots, Len(\mathcal{A}) - 1]$, the selection is repeated.

Currently, unused relays can be used to construct a circuit. In order to provide a more robust level of local anonymity, constraints must be placed on unused relays—those with zero or near zero anonymity scores. First, unused relays may only be chosen in the middle relay position. This is done so that an adversary viewing this relay's traffic is unable to cross Tor's network boundaries and discover either the client or destination identities. However, an unused guard may service its assigned user in the guard position as usual since doing so does not expose the user more than necessary; the adversary can monitor the client link and over time identify its guard nodes in any event. Second, once a minimum threshold of circuits is active over a relay this restriction is relaxed allowing it to be chosen in the exit role.

8. DEPLOYABILITY

Proposals enhancing Tor must consider ease of deployment if they are to be adopted. The client software would need to be updated to leverage the new algorithm and measurements. This can be deployed to the user base as a new version of the client software as is done currently. Depending on the selected mechanism for collecting current bandwidth from Section 7.3, relays may also need to upgrade; the extent of the changes would depend on the exact method of collecting current bandwidth data from the network. Importantly, the circuit construction and data transfer protocols remain the same as in the current Tor network. While the uptake of LocaTor occurs, and client penetration is not total, the usual problem of user-base partitioning may occur. This is easily addressed using the standard technique of using a global flag in the consensus that indicates when it is safe for clients to start using the new algorithm.

9. ATTACKS

Our adversary, who now works at a Network Operations Centre, attempts to learn his targets' communication partners. This information will be valuable to his paymasters. Although he is able to identify the targets' traffic before it enters the Tor network he is unable to do the same once it has crossed Tor's boundary and must resort to other means of identifying and tracing their traffic. He runs his own Tor client and learns which relays operate within the networks he can access. He now has three avenues of attack.

He can trace the traffic through each hop and link the source and destination together. He may succeed in this if he can successfully link the traffic going into and coming out of the relay at each hop. Unfortunately, he is hindered by the fact that each relay serves more than one circuit at any given time. This is especially true for relays with high anonymity scores where many circuits are being serviced at the same time.

Relatedly, Díaz et al. [1] propose a reduced overhead data link padding (RO-DLP) scheme that seeks to reduce the traffic correlation threat we have described. RO-DLP requires that each outgoing link be active and that data padding be used when there is no data destined for some of the links connected to the relay. In this manner the adversary may not discern the path taken by exiting traffic. We can approximate the behaviour of RO-DLP in LocaTor in situations where the anonymity score of the relays over which a user's traffic flows is very high. Here the sheer amount of traffic over all the relay's links makes them behave in much the same manner that data padding would under RO-DLP. Indeed, RO-DLP can co-exist with LocaTor and together provide higher anonymity assurances across a wider range of scenarios.

The adversary is left to filter the traffic in other ways; for example, by reducing the likely set of preceding and succeeding relays. He can partition the relays into subsets, according to the θ-windows they are found in. He also makes an assumption of the θ value the client may choose, for example that of high anonymity. The adversary would know that if an exit relay is within a particular window, then the preceding relay (the middle node in this case) is most likely to be in that window as well. Unfortunately for the adversary, the window sizes are only an order of magnitude smaller than the size of the entire network. With LocaTor's current parameter choices, a window holds 400 relays. Each relay services a number of circuits and so each is a likely candidate as the preceding node. Coupled with the fact that not all relays within a window may be visible to a particular partial adversary; our attacker does not have an easier time tracing the preceding relay with the knowledge of the exit relay. If he observes the company's exit point communicating with a guard relay the θ windows do not provide him with any information at all since guard relays are independently chosen before any interaction with the relays in windows. Stuck, our attacker may then try to gain some further information by analyzing the sorted relay list itself.

Our attacker looks at the sorted list and tries to reduce the set of relays in a target window by noting their transits into and out of the window. This can be useful if the parties have been communicating over a long duration of time and have been creating circuits frequently with the same θ value. The attacker would then focus on relays that rarely leave the window and focus on tracing the communication over those. However, the likelihood of a relay remaining within a window is uncertain. Even when some relays do stay in the same window over time, it is also possible that those involved in the communication of interest are relays that transit frequently.

In summary, the attacker may try to leverage the apparent information leakage from the sorted anonymity list and the θ values and windows. However, we have argued that he is not more likely to succeed in linking communication between his targets and their partners than in the current Tor network.

10. CONCLUSIONS AND FUTURE WORK

In this paper we promoted the benefits of local anonymity awareness, which requires that (i) individual users receive guarantees when using a network, and (ii) all the sub-regions of the network have good anonymity scores (as measured by entropy computed on those sub-regions, for example). Moreover, we designed an algorithm to integrate local anonymity within the Tor network, and discussed some attacks and possible mitigation strategies.

This work can be extended in several ways. We would like to investigate variants of our algorithm and analyze different choices for the anonymity score functions. Ideally, we would like a local anonymity metric that is both accurate and hard to manipulate. In addition, it remains to be determined what are the optimal window sizes and locations, how much randomization is required, and what other attacks can be performed. Finally, we would like to understand at a more fundamental level whether local anonymity

aware algorithms, such as LocaTor, can lead to more accurate measurements of anonymity and how they compare to existing entropy-based approaches.

11. ACKNOWLEDGEMENTS

We are indebted to Kevin Bauer for being an excellent sounding board and Jean-Charles Grégoire, Ryan Henry, Angèle Hamel, Femi Olumofin, Rob Smits, and Tao Wang for their constructive comments of our manuscript. Funding support was provided by NSERC, MITACS, and The Tor Project.

12. REFERENCES

[1] C. Díaz, S. Murdoch, and C. Troncoso. Impact of network topology on anonymity and overhead in low-latency anonymity networks. In *Proceedings of the 10th Privacy Enhancing Technologies Symposium (PETS)*, pages 184–201, 2010.

[2] C. Díaz, S. Seys, J. Claessens, and B. Preneel. Towards Measuring Anonymity. In *Proceedings of the 2nd International Conference on Privacy Enhancing Technologies (PET)*, pages 54–68, 2002.

[3] R. Dingledine, N. Mathewson, and P. Syverson. Tor: The Second-Generation Onion Router. In *Proceedings of the 13th USENIX Security Symposium*, pages 303–320, 2004.

[4] A. Pfitzmann and M. Köhntopp. Anonymity, Unobservability, and Pseudonymity - A Proposal for Terminology. In *Proceedings of the Workshop on Design Issues in Anonymity and Unobservability*, pages 1–9, 2000.

[5] A. Serjantov and G. Danezis. Towards an Information Theoretic Metric for Anonymity. In *Proceedings of the 2nd International Conference on Privacy Enhancing Technologies (PET)*, pages 41–53, 2002.

[6] M. Sherr, M. Blaze, and B. T. Loo. Scalable Link-Based Relay Selection for Anonymous Routing. In *Proceedings of the 9th Privacy Enhancing Technologies Symposium (PETS)*, pages 73–93, 2009.

[7] R. Snader and N. Borisov. A Tune-up for Tor: Improving Security and Performance in the Tor Network. In *Proceedings of the Network and Distributed Security Symposium (NDSS)*, 2008.

[8] R. Snader and N. Borisov. EigenSpeed: Secure Peer-to-Peer Bandwidth Evaluation. In *Proceedings of the 8th International Conference on Peer-to-Peer Systems (IPTPS)*, 2009.

[9] P. Syverson. Why I'm not an Entropist. In *The 17th International Workshop on Security Protocols*, 2009.

[10] The Tor Project. Tor Blog: China Blocking Tor: Round 2, 2010. https://blog.torproject.org/blog/china-blocking-tor-round-two, accessed July 2011.

[11] The Tor Project. Tor Metrics Portal: Users, 2011. http://metrics.torproject.org/users.html, accessed June 2011.

[12] G. Tóth, Z. Hornák, and F. Vajda. Measuring Anonymity Revisited. In *Proceedings of the 9th Nordic Workshop on Secure IT Systems*, pages 85–90, 2004.

[13] Wired. Whistle-Blower Outs NSA Spy Room, 2006. http://www.wired.com/science/discoveries/news/2006-/04/70619, accessed July 2011.