Towards Measuring Resilience in Anonymous Communication Networks

Fatemeh Shirazi[1], Claudia Diaz[2], Ciaran Mullan[1], Joss Wright[3], and Johannes Buchmann[1]

[1] Technische Universität Darmstadt, Department of Computer Science
Hochschulstraße 10, 64289 Darmstadt, Germany
{fshirazi,cmullan,buchmann}@cdc.informatik.tu-darmstadt.de

[2] KU Leuven ESAT/COSIC, iMinds
Kasteelpark Arenberg 10 Leuven, Belgium
claudia.diaz@esat.kuleuven.be

[3] Oxford Internet Institute, University of Oxford
1 St. Giles, Oxford. OX1 3JS, UK
joss.wright@oii.ox.ac.uk

Abstract. Most prior work on anonymous communications has focused, to a large extent, on achieving, measuring, and attacking anonymity properties. There are, however, several other properties of importance in anonymous communication networks, such as their performance and their robustness to denial of service attacks, that have received less scrutiny. To our knowledge, no practical measure of the *resilience* of an anonymous communication network against active attackers has yet been proposed.
In this work we propose a metric for quantifying the resilience of anonymous communication networks towards active adversaries with the power to disable selected nodes.

Keywords: Resilience, Anonymous Communication Networks, Tor

1 Introduction

Anonymous communication networks constitute one of the main building blocks for protecting online privacy, as they enable users to anonymously browse the web, share content, or communicate with each other. To ensure adoption by users, anonymous communication networks must provide robust anonymity properties, strong resilience against network failures and attacks, and acceptable performance levels. In order to evaluate these characteristics we need to formalize the required properties and to have a means to measure to which extent the properties are satisfied by a given network design. Whilst performance can be measured using standard techniques, it is very challenging to provide practically useful metrics for resilience and anonymity.

Notions of resilience that exist in the literature are typically applied to networks such as the Internet [1–6]. However, anonymous communication networks

differ from standard communication networks in several respects, notably a different network topology, randomized routing policies, and, typically, higher latency and lower overall bandwidth. Thus a new metric is needed. Existing anonymity metrics [7, 8] are of limited practical use for low-latency anonymous communication networks, as their security is typically evaluated towards non-global adversaries [9]. Anonymity is however not the focus of this work; see [10] for an overview.

This paper introduces a new, practical metric for measuring the resilience of anonymous communication networks to active attacks. Our metric takes into account the different features and constraints of anonymous routing policies. Because of the randomization involved in routing policies, it makes sense to measure resilience in terms of the expected quality of communications. We are specifically concerned with measuring resilience against active adversaries aiming to degrade as much as possible the expected quality of service. We model this by considering an adversary that has the capability to take down a set of nodes of his choosing.

We conceive three attack strategies. The first strategy (*naive approach*) is simply to target nodes with the highest bandwidth. Because the criticality of a node depends not only on its bandwidth but also on routing policy constraints, this strategy is not necessarily optimal. We consider two further strategies (*greedy approach* and *optimal approach*), which require more computational effort but optimize the attack taking into account the routing policy. Although we ignore the effects of natural network failures, they could easily be encompassed in our metric.

The paper is organized as follows. In Section 2 we describe the workings of anonymous communication networks and discuss existing metrics for measuring resilience. In Section 3 we introduce our notion of resilience for anonymous communication networks. Section 4 details the methodology followed to experimentally measure the resilience of a sample node set from the Tor network.

2 Background

We give a brief introduction to low-latency anonymous communication networks, and in particular the Tor network [11], which is the most popular ACN. We also discuss existing metrics for resilience in both standard and anonymous communication networks and discuss their limitations.

2.1 Anonymous Communication Networks

The purpose of an anonymous communication network (ACN) is to enable users to establish anonymous communication channels over an open network. ACN users conceal the destination of their communications towards local observers (e.g., their ISP), as well as their identity towards the destination (e.g., a website). ACNs also allow two users to hide the fact that they are communicating with each other. Typically, an ACN is composed of a set of routers (also called *nodes*)

that constitute an overlay network. In most cases communications are relayed over multiple routers to achieve anonymity from the node operators.

Some ACNs such as Crowds [12] are *random routed*, meaning that subsequent hops in the communication route are decided independently by the last node in the current path. In this paper however, we focus on the (more common) *source-routed* ACNs. In these networks, the initiator of the communication selects, according to the ACN *routing policy*, the set of relays that form the anonymous channel as well as their position in the route. Examples of such deployed ACNs are Tor [11] and I2P [13].We will focus on Tor as an example to illustrate our resilience metric.

To communicate anonymously through Tor, users first obtain a list of available Tor nodes from the *directory server*, together with their contact details such as IP address and port, public keys, and other relevant information such as flags that denote the capabilities of the routers. Users then randomly select three nodes according to Tor's routing policy. The first node is called the *entry* or *guard node*, the second is the *middle node*[4], and the last the *exit node*.

While all routers can potentially act as middle nodes, only some nodes are flagged as guard and/or exit nodes. Each node decides for itself whether it wants to be an exit node. Guard nodes were introduced to limit the chances of an adversarial router being selected as entry node [14], considering that the adversary wins if he can compromise a single communication. In the current Tor design each user is assigned a small number of guard nodes. A node is flagged to be a guard node based on superior bandwidth and mean uptime. By default, a users' guard node set is periodically changed every 30-60 days. Nodes that are both guard and exit nodes are called *guard-exit nodes*. In a recent sample consensus, 2012-10-19 21:00:00, around 20% of the Tor routers were flagged as exit nodes, and 20% as guard nodes. Guard-exit nodes constituted slightly less than 10% of the total.

A further development in the routing policy is to disallow a communication to pass through two nodes with the same /16 subnet IP address. Hence the underlying topology is not a complete graph. Finally, due to performance considerations, Tor's routing policy does not simply select each possible route with the same probability: preference is given to high-bandwidth nodes, and the likelihood that nodes are chosen for certain positions depends on the ratios of overall guard and exit node bandwidths. A recent study [15] has shown that as much as 20% of all Tor traffic flows over as few as seven high bandwidth nodes, resulting in potential critical points of attack.

2.2 Resilience Metrics

Resilience of a communication network can be thought of as 'the ability to provide and maintain an acceptable level of service in the face of various faults and challenges to normal operation' [16]. Resilience is often measured in terms of network connectivity and failure rates [17, 6, 4]. For example, Dolev et al. [18] define

[4] There is no special name for nodes in the middle of the communication path.

the resilience of a network as the relation between the number of connected node pairs and the total number of nodes.

Existing resilience metrics are however not appropriate for ACNs, as these differ from standard communication networks in various important ways. The underlying topology of ACNs often forms a complete graph, hence connectivity is not such a useful metric. Modeling failure rates is of interest, but for ACNs we are also concerned with resilience towards attacks. Active adversaries may compromise nodes or perform denial of service attacks to degrade the ACN service. Routing policies in standard networks are usually deterministic, aimed at minimizing cost factors and finding shortest available routes. ACN routing policies on the other hand are randomized to provide anonymity properties. Hence, a resilience metric for ACNs should take into account randomness in the routing policy.

In recent years, resilience has been studied in several ACNs such as Cashmere [19], Salsa [20], and Hydra Onion [21]. Resilience in these works is however measured with respect to random failures, while active adversaries are not considered. Borisov et al. [22] investigated reliability in ACNs in the face of *selective DoS attacks* in which the adversary aims at reducing the anonymity of the network. Therefore, the adversary only performs a DoS attack when he cannot compromise the communication in order to force the user to reconnect through a new circuit and thereby improving his chances for compromising the communication. Their proposed metric is aimed at evaluating the impact of these attacks on anonymity rather than on resilience. Therefore, for our metric, investigating resilience in face of attacks such as DoS attacks, the adversary's goal is complementary to their adversary model. However, when using the metrics for optimizing the routing policy of an ACN, both metrics address different aspects. Borisov et al. [23] investigated optimizing the routing policy in the face of selective DoS attacks in order to maintain anonymity.

3 Resilience in Anonymous Communication Networks

We now present our metric for measuring resilience of anonymous communication networks (ACNs). The metric does not consider connectivity, but rather measures the expected loss in the quality of communications (modeled as a random variable) when the adversary takes out a number of nodes.

3.1 Adversary Model

We consider an adversary whose aim is to decrease overall quality of communication within the ACN. In other words, the adversary is not interested in compromising anonymity, but rather wants to discourage or even prevent use of the ACN. We allow an active adversary to select the nodes to be attacked (in the form of denial of service). We consider up to three node selection strategies that the adversary might use in order to reduce the ACN communication quality. The strategies take into account not only the characteristics of the nodes, but also constraints in the ACN routing policy.

3.2 Resilience Model

Let $G = (V, E)$ be a complete graph. We think of V and E as representing, respectively, the set of ACN routers and the connections between them. Each vertex is labelled with a non-negative real number corresponding to router capacity. We define a *route* in G to be a finite sequence of distinct vertices, and view a communication path from the initiator to the receiver as a route through G.

Definition 1 *Let G be a complete graph. A routing policy on G is a set of rules specifying all admissible routes R through G, together with a probability distribution on R.*

The routing policy thus determines which routes (of all the possible ones) can be used for relaying communications, as well as the probability that a specific route is chosen.

Definition 2 *An anonymous communication network A consists of a complete graph G together with a routing policy on G.*

Note that, for the purposes of measuring resilience, we do not concern ourselves with details of *how* anonymity is provided, but regard this as an implicit part of the routing policy.

Definition 3 *Let A be an anonymous communication network on graph G and let X be a discrete nonnegative random variable measuring some aspect of quality of communication in A. (The quality measure may be in terms of bandwidth, latency, throughput, or some other observable variable of interest.) We define the quality $Q_X(A)$ of A with respect to X as the expected value of X:*

$$Q_X(A) := \sum_x x \cdot Pr[X = x].$$

In our attack model we consider an adversary who wishes to significantly alter (typically, lower) $Q_X(A)$. He does so by taking out a set S of k nodes from the network. We denote such an adversary as $Ad_{X,S}$.

Definition 4 *Let A be an ACN on G and let A_S be the the ACN obtained after removing a set S of k nodes from G. (The routing policy of A_S is induced from A.) We define the resilience of A against $Ad_{X,S}$ to be*

$$\mathcal{R}_{Ad_{X,S}}(A) := \frac{Q_X(A_S)}{Q_X(A)}.$$

Considering an adversary who can take down any k nodes, we define the resilience of A with respect to X to be

$$\mathcal{R}_{X,k}(A) := \min_{S:|S|=k} \frac{Q_X(A_S)}{Q_X(A)}.$$

Thus $\mathcal{R}_{Ad_{X,S}}$ captures resilience against the case when an adversary selects a particular set S of nodes to attack, and $\mathcal{R}_{X,k}$ measures resilience against the optimal attack. Comparing $\mathcal{R}_{Ad_{X,S}}$ and $\mathcal{R}_{X,k}$ gives us insight as to the effectiveness of the adversary's strategy.

We can also use the metric to measure the minimum number of nodes that, if removed, degrade quality by a factor δ. This factor is of interest both to the network designer, to understand the resilience of the network against targeted attacks, and to attackers seeking to conduct efficient attacks against the network.

Definition 5 *The δ-threshold of A with respect to $Ad_{X,S}$ is given by*

$$\mathcal{T}_{\delta,Ad_{X,S}}(A) := \min\{k > 0 : Q_X(A_S) = \delta \cdot Q_X(A)\}.$$

4 Measuring Resilience for a Tor-like Network

We can apply our metric of resilience to a Tor-like anonymous communication network A.

In terms of user experience, bandwidth is a highly critical factor in Tor: any significant decrease in bandwidth will discourage use of the ACN. Moreover, the fact that bandwidth influences even the latency of communication in Tor [24] adds to the significance of bandwidth. In addition, it is worth mentioning that Tor's node selection algorithm is also weighted based on the node's bandwidth. Therefore, we take the *expected* bandwidth of a communication as our quality measure of interest. We define the bandwidth of a route as follows.

Definition 6 *Let A be an ACN on graph G and let $r = (v_1, \ldots, v_m)$ be a route in G. Suppose the vertices $v_1, \ldots, v_m$ have router capacities $c_1, \ldots, c_m$. We define the bandwidth of r as:*

$$BW(r) := \min\{c_1, \ldots, c_m\}.$$

Let R be the set of all admissible routes in G, and let B be the set of bandwidths of all admissible routes. We compute the probability of a communication having a bandwidth b as:

$$Pr[B = b] = \sum_{r \in R : BW(r) = b} Pr[R = r],$$

and define the expected quality of the communication as:

$$Q_B(A) = Exp(B) = \sum_b b \cdot Pr[B = b].$$

Note that we are considering $Q_B(A)$ to represent the expected quality for a single communication. An additional measure of interest is the total bandwidth available to the population of users as a whole.

4.1 Attack Strategies

We conceive three attack strategies that aim to decrease $Q_B(A)$ in a Tor-like network A. Recall our adversary model of Subsection 3.1. Since bandwidth is the critical factor, a first strategy which we call the *naive approach*, is for an adversary to take out the k nodes with the highest capacities. Although intuitively this seems like a good approach due to the constraints in the routing policy, this may not be the best strategy to decrease $Q_B(A)$.

One could consider two further strategies, which we call the *greedy approach* and the *optimal approach*. In the optimal approach, the k nodes to be removed are those that maximize the degradation in quality. To find these k nodes, the attacker needs to test how quality degrades for all the possible combinations in which k nodes are removed. For large networks (such as Tor, having approximately 3000 nodes) this is very costly even for small k. In the more practical *greedy approach*, the attacker finds the best first node to remove, recomputes the routing policy for the remaining network, and repeats the process k times.

4.2 Practical Issues

Due to the intricacies of Tor's routing policy, obtaining the exact bandwidth probability distribution is impractical. However, by sampling the path selection process sufficiently many times, we can obtain a close approximation to this distribution. We can obtain these samples using one of several Tor Simulators such as Shadow [25], Experimentor [26], or COGS [27]. This in turn will enable statistically accurate resilience measurements. Note that to measure resilience we need to remove nodes from the network, which for our purposes can be modelled sufficiently well with a simulator. Moreover, with multiple users communicating through the ACN, some measure for the total available bandwidth needs to be provided to complement our quality measure.

As part of ongoing work, we are in the process of computing resilience measurements using these Tor simulators. Furthermore we plan to evaluate resilience of other existing anonymous communication networks.

References

1. Sterbenz, J.P., Çetinkaya, E.K., Hameed, M.A., Jabbar, A., Shi, Q., Rohrer, J.P.: Evaluation of network resilience, survivability, and disruption tolerance: Analysis, topology generation, simulation, and experimentation (invited paper). Springer Telecommunication Systems (2011)
2. Sterbenz, J., Çetinkaya, E., Hameed, M., Jabbar, A., Rohrer, J.: Modelling and analysis of network resilience. In: Communication Systems and Networks (COMSNETS), 2011 Third International Conference on. (January 2011) 1–10
3. Omer, M., Nilchiani, R., Mostashari, A.: Measuring the resilience of the global internet infrastructure system. In: Systems Conference, 2009 3rd Annual IEEE. (March 2009) 156–162

4. Cohen, R., Erez, K., ben Avraham, D., Havlin, S.: Resilience of the internet to random breakdowns. Phys. Rev. Lett. **85** (Nov 2000) 4626–4628
5. Molisz, W., Rak, J.: Quality of resilience in ip-based future internet communications. In: Transparent Optical Networks (ICTON), 2011 13th International Conference on. (June 2011) 1–4
6. Wu, J., Zhang, Y., Mao, Z.M., Shin, K.G.: Internet routing resilience to failures: analysis and implications. In: Proceedings of the 2007 ACM CoNEXT conference. CoNEXT '07, New York, NY, USA, ACM (2007) 25:1–25:12
7. Diaz, C., Seys, S., Claessens, J., Preneel, B.: Towards measuring anonymity. In: Proceedings of the 2nd international conference on Privacy enhancing technologies. PET'02, Berlin, Heidelberg, Springer-Verlag (2003) 54–68
8. Serjantov, A., Danezis, G.: Towards an information theoretic metric for anonymity. In: Proceedings of the 2nd international conference on Privacy enhancing technologies. PET'02, Berlin, Heidelberg, Springer-Verlag (2003) 41–53
9. Syverson, P.: Why i'm not an entropist. In: Workshop on Security Protocols. (2009)
10. Andersson, C., Lundin, R.: On the Fundamentals of Anonymity Metrics. Proceedings of the IFIP WG 9.2, 9.6/11.7 Summer School on Risks and Challenges of the Network Society, Karlstad, Sweden, 6-10 Aug 2007 (2007)
11. Dingledine, R., Mathewson, N., Syverson, P.: Tor: the second-generation onion router. In: Proceedings of the 13th conference on USENIX Security Symposium - Volume 13. SSYM'04, Berkeley, CA, USA, USENIX Association (2004) 21–21
12. Reiter, M.K., Rubin, A.D.: Crowds: anonymity for web transactions. ACM Trans. Inf. Syst. Secur. **1**(1) (November 1998) 66–92
13. I2P: Invisible internet project. Available at: http://www.i2p.net/
14. Wright, M.K., Adler, M., Levine, B.N., Shields, C.: The predecessor attack: An analysis of a threat to anonymous communications systems. ACM Trans. Inf. Syst. Secur. **7**(4) (November 2004) 489–522
15. Gilad, Y., Herzberg, A.: Spying in the dark: Tcp and tor traffic analysis. In: Proceedings of the 12th international conference on Privacy Enhancing Technologies. PETS'12, Berlin, Heidelberg, Springer-Verlag (2012) 100–119
16. Xie, L., Smith, P., Banfield, M., Leopold, H., Sterbenz, J.P., Hutchison, D.: Active and programmable networks. (2009) 83–95
17. Albert, R., Jeong, H., Barabasi, A.: Error and attack tolerance of complex networks. Nature **406**(6794) (2000) 378–382
18. Dolev, D., Jamin, S., Mokryn, O., Shavitt, Y.: Internet resiliency to attacks and failures under bgp policy routing. Comput. Netw. **50**(16) (November 2006) 3183–3196
19. Zhuang, L., Zhou, F., Zhao, B.Y., Rowstron, A.: Cashmere: resilient anonymous routing. In: Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation - Volume 2. NSDI'05, Berkeley, CA, USA, USENIX Association (2005) 301–314
20. Nambiar, A., Wright, M.: Salsa: a structured approach to large-scale anonymity. In: Proceedings of the 13th ACM conference on Computer and communications security. CCS '06, New York, NY, USA, ACM (2006) 17–26
21. Iwanik, J., Klonowski, M., Kutylowski, M.: Duo–onions and hydra–onions – failure and adversary resistant onion protocols. In: Proceedings of the IFIP TC-6 TC-11 Conference on Communications and Multimedia Security 2004, Springer Boston, Springer Boston (September 2004)

22. Borisov, N., Danezis, G., Mittal, P., Tabriz, P.: Denial of service or denial of security? In: Proceedings of the 14th ACM conference on Computer and communications security. CCS '07, New York, NY, USA, ACM (2007) 92–102
23. Das, A., Borisov, N.: Securing Anonymous Communication Channels under the Selective DoS Attack. In: Proceedings of Financial Cryptography and Data Security (FC'13). (April 2013)
24. Wang, T., Bauer, K., Forero, C., Goldberg, I.: Congestion-aware Path Selection for Tor. In: Proceedings of Financial Cryptography and Data Security (FC'12). (February 2012)
25. Jansen, R., Hopper, N.: Shadow: Running Tor in a Box for Accurate and Efficient Experimentation. In: Proceedings of the Network and Distributed System Security Symposium - NDSS'12, Internet Society (February 2012)
26. Bauer, K., Sherr, M., McCoy, D., Grunwald, D.: Experimentor: A testbed for safe and realistic tor experimentation. In: Proceedings of the USENIX Workshop on Cyber Security Experimentation and Test (CSET 2011). (August 2011)
27. Elahi, T., Bauer, K., AlSabah, M., Dingledine, R., Goldberg, I.: Changing of the guards: a framework for understanding and improving entry guard selection in tor. In: Proceedings of the 2012 ACM workshop on Privacy in the electronic society. WPES '12, New York, NY, USA, ACM (2012) 43–54