Anonymity Analysis of Onion Routing in the Universally Composable Framework

Joan Feigenbaum
Yale University
Joan.Feigenbaum@yale.edu

Aaron Johnson
Naval Research Laboratory
aaron.m.johnson@nrl.navy.mil

Paul Syverson
Naval Research Laboratory
syverson@itd.navy.mil

ABSTRACT

We present the formalization and analysis of a practical paradigm for general anonymous communication using standard cryptographic primitives. Specifically we present a probabilistic analysis of onion routing in a black-box model of anonymous communication in the Universally Composable framework. Full statements of results and proofs can be found in the full paper [7].

1. INTRODUCTION

Onion routing is the paradigm for the most widely used and widely deployed anonymous communications systems. The Tor onion routing network currently comprises some three thousand nodes worldwide and has about a half million daily users [10]. The original NRL onion routing networks processed tens of thousands of circuits per day even in the 1990s [12], and the Freedom Network [8] had hundreds of nodes and tens of thousands of users in the early 2000s a few years before Tor was first deployed.

Despite this success, formal anonymity analyses of onion routing have been slow to develop. A primary reason for this was that onion routing was designed to be practical. Until recently, we have thus been forced to choose between theoretically well-grounded paradigms, such as mixing or DC-nets, and practical paradigms without solid theoretical basis. Note that although mix networks and DC-nets have been implemented and used, there are stricter inherent limits to their usability than in onion routing, and thus their implementations have never had more than a few hundred concurrent users [4, 15]. Though onion routing was certainly developed with security in mind, the research community has struggled to create definitions and models that simultaneously had provable results and actually captured anything like onion routing.

Feigenbaum et al. provided an early formalization of onion routing using I/O automata in 2007 [5] and analyzed its anonymity[1]. However, their formalization did not use standard cryptographic tools. In addition, their results were limited to a possibilistic characterization of security. So, e.g., they could not reflect the difference between when a hundred possible senders of a message are equally likely and when one in particular is the sender with probability .99. They later provide a probabilistic analysis of a similar system [6], but this analysis is not tied to their earlier formalization.

In this work we present a solution to these problems. First, we present a *black-box model* of anonymous communication using an ideal functionality in the Universally Composable (UC) framework[3]. This model gives the basis for the kind of probabilistic anonymity analysis performed in [6] in a standard cryptographic framework.

Second, we prove a relationship between this model and the earlier I/O-automata formalization [5]. This allows results in the old formalization to be applied to the new model and vice versa. In addition, it gives an example of how the results of our anonymity analysis can be applied to many protocols (and in particular those that *UC-emulate* our functionality) using the black-box model.

Third, we update the set of anonymity results from [6] to apply to the new model. Those results quantify how much the adversary can gain in identifying users by exploiting knowledge of their behavior, and, among other things, show that the worst-case anonymity with an adversary that controls a fraction b of the routers is comparable to the best-case anonymity against an adversary that controls a fraction $\sqrt{b}$.

In the interest of space, we refer the reader to [7] for related prior work. However, we highlight here the work of Backes et al. [1] that builds on our black-box model. In that work, the authors formally describe an onion-routing protocol within the UC-framework. They present a different ideal functionality and prove that it is UC-emulated by their protocol. This ideal functionality provides desirable privacy properties such as forward secrecy. However, it does not automatically imply anonymity, and for these properties, they show UC-emulation by their protocol of the ideal functionality we give herein. Thus the results of the probabilistic anonymity analysis that we obtain in our model apply to their work.

2. MODEL

We describe our analysis of onion routing in terms of an ideal functionality in the UC framework [3]. We use such a functionality for three reasons: First, it abstracts away the details that aren't relevant to anonymity. Second, the UC framework provides the notion of UC emulation, which captures exactly when our analysis applies to a cryptographic protocol. Third, it immediately suggests ways to perform similar analyses of other anonymous-communication protocols that may not strictly provide this functionality.

Let U be the set of users with $|U| = n$. Let Δ be the set

[1]Note that while Mauw et al. [11] and Camenisch and Lysyanskaya[2] perform formal analyses of anonymity protocols, those protocols do not use persistent circuits. This particular feature of onion routing affects its anonymity analysis. Camenisch and Lysyanskaya, moreover, do not analyze anonymity.

upon receiving $d \in \Delta$ from $u \in U$

$$x \leftarrow \begin{cases} u & \text{with probability } b \\ \perp & \text{with probability } 1-b \end{cases}$$

$$y \leftarrow \begin{cases} d & \text{with probability } b \\ \perp & \text{with probability } 1-b \end{cases}$$

send (x, y) to the adversary

Figure 1: Black-box ideal functionality $\mathcal{F}_{OR}$

of destinations. Let R be the set of onion routers. Let $\mathcal{F}_{OR}$ be the ideal functionality. $\mathcal{F}_{OR}$ takes the set $A \subseteq R$ of compromised routers from the adversary at the beginning of the execution.[2] Let $b = |A|/|R|$. The black-box functionality is given in Figure 1.

When user u forwards his input from the environment to $\mathcal{F}_{OR}$, the functionality checks to see if it is some $d \in \Delta$. If so, $\mathcal{F}_{OR}$ notifies the adversary of the connection and includes the source with probability b and the destination with probability b.

To analyze the anonymity provided by the ideal functionality, we make two assumptions about the inputs from the environment. First, we assume that the environment selects the destination of user u from a distribution p^u over Δ, where we denote the probability that u chooses d as p^u_d. Second, we assume that the environment sends a destination to each user. Note that these assumptions need not be made when showing that a protocol UC-emulates $\mathcal{F}_{OR}$.

We refer to the combination of the adversary model, the assumptions about the environment, and the ideal functionality as the *black-box model*.[3]

Our ideal functionality models anonymous communication over some period of time. It takes as input from each user the identity of a destination. For every such connection between a user and destination, the functionality may reveal to the adversary the identity of the user, the identity of the destination, or both. The adversary captured in our model is computationally bounded, controls a fixed set of routers, and can actively attack the protocol. We note that we include only information flow to the adversary in this functionality rather than try to capture the type of communication primitive offered by onion routing because our focus is analyzing anonymity rather than defining a useful anonymous-communication functionality. This model is reminiscent of the general model of anonymous communication used by Kesdogan et al. [9] in their analysis of an intersection attack.

3. RESULTS

The black-box model just given captures the information relevant to anonymity that the adversary can infer from his observations of onion routing—namely, the observed users, the observed destinations, and the possible connections between the two. Users in onion routing choose a sequence of routers, or a *circuit*, to route their traffic. Revealing the user corresponds in onion routing to the first router in the circuit being compromised, revealing the destination corresponds to the last router being compromised, and revealing when the two are paired reflects the presence of timing attacks [16]. In this way, we abstract away from much of the design specific to onion routing so that our results apply both to onion routing and to other low-latency anonymous-communication designs.

We can in fact tie our model to the guarantees of onion routing by showing it reveals as much information about users' communication as the earlier I/O-automata onion routing protocol [5]. The analysis of that model identifies the user states that are information-theoretically indistinguishable to the adversary. The black-box model we provide here is a valid abstraction of that formalization because, under the following probability measures on executions, it preserves the anonymity properties.

Let users in the I/O-automata model choose the routers in their circuits uniformly at random from R and choose the destination according to user-specific distributions p^u. Given these circuits and a set of adversary automata A, [5] identifies an equivalence class of circuit and destination choices with respect to which, for every pair of members in the class, a bijection exists between their executions such that paired executions are indistinguishable. Let the indistinguishable executions thus paired have the same probability, conditional on the circuit and destination choices.

Given this measure, the black-box model that abstracts the I/O-automata model has the same user set U, the same destination set Δ, an adversary parameter of $b = |A|/|R|$, and the same destination distributions p^u. We can then prove a theorem showing that each posterior distribution of the adversary on the destinations of users has the same probability under both the I/O-automata model and its black-box model. See [7] for details.

Moreover, we derive earlier anonymity results [6] within our the new black-box model. These results describe *relationship anonymity* [13, 14], which is obtained when the adversary cannot identify the destination of a user. The adversary can infer a probability distribution for a user's destination given the adversary's observations. We use the probability assigned to the correct destination as our anonymity metric. We summarize those results:

- We show that a standard approximation to our metric provides a lower bound on it.
- We show that the worst case for anonymity over other users' behavior is when every other user either always visits the destinations the user is otherwise least likely to visit or always visits his actual destination. The former will be the worst case in most situations.
- We give an asymptotic expression for our metric in the worst cases. The limit of this expression in the most common worst case with an adversary controlling a fraction b of the network is equal to the lower bound on the metric when the adversary controls a larger fraction $\sqrt{b}$ of the network. This is significantly worse than the standard analysis suggested, and shows the importance of carefully considering the adversary's knowledge of the system.
- We consider anonymity in a more typical set of user distributions in which each user selects a destination from a common Zipfian distribution. We show that, as the user population grows, the anonymity approaches the lower bound.

[2]The adversary compromises routers only because a compromised user has no anonymity and is effectively removed from the set of users U for purposes of deanonymization.

[3]Some readers may only be familiar with "black box" as indicating black-box access to some cryptographic primitives used as a starting point to achieve some other desired functionality. Here we show how, for purposes of anonymity analysis, we need only consider a black-box abstraction.

4. REFERENCES

[1] Michael Backes, Ian Goldberg, Aniket Kate, and Esfandiar Mohammadi. Provably secure and practical onion routing. In *Proceedings of the 25th IEEE Computer Security Foundations Symposium (CSF 2012)*, 2012. Forthcoming.

[2] Jan Camenisch and Anna Lysyanskaya. A formal treatment of onion routing. In *Proceedings of CRYPTO 2005*, pages 169–187, 2005.

[3] Ran Canetti. Universally composable security: A new paradigm for cryptographic protocols. Cryptology ePrint Archive, Report 2000/067, 2000. http://eprint.iacr.org/2000/067.

[4] Roger Dingledine and Nick Mathewson. Anonymity loves company: Usability and the network effect. In Ross Anderson, editor, *Fifth Workshop on the Economics of Information Security (WEIS 2006)*, June 2006.

[5] Joan Feigenbaum, Aaron Johnson, and Paul Syverson. A model of onion routing with provable anonymity. In *Financial Cryptography and Data Security: 11th International Conference, FC 2007*, pages 57–71, 2007.

[6] Joan Feigenbaum, Aaron Johnson, and Paul Syverson. Probabilistic analysis of onion routing in a black-box model (extended abstract). In *Proceedings of the 2007 ACM Workshop on Privacy in Electronic Society (WPES 2007)*, pages 1–10, 2007.

[7] Joan Feigenbaum, Aaron Johnson, and Paul Syverson. Probabilistic analysis of onion routing in a black-box model. *ACM Transactions on Information and System Security (TISSEC)*, 2012. Forthcoming, draft available at http://arxiv.org/abs/1111.2520.

[8] Ian Goldberg and Adam Shostack. Freedom network 1.0 architecture and protocols. White paper, Zero Knowledge Systems, Inc., October 2001.

[9] Dogan Kesdogan, Dakshi Agrawal, and Stefan Penz. Limits of anonymity in open environments. In *Proceedings of the 5th Information Hiding Workshop (IH 2002)*, pages 53–69, 2002.

[10] Loesing et al. Tor metrics portal. https://metrics.torproject.org/.

[11] Sjouke Mauw, Jan Verschuren, and Erik de Vink. A formalization of anonymity and onion routing. In *Proceedings of the 9th European Symposium on Research in Computer Security (ESORICS 2004)*, pages 109–124, 2004.

[12] Onion routing archives. http://www.onion-router.net/Archives.html.

[13] Andreas Pfitzmann and Marit Hansen. Anonymity, unobservability, and pseudonymity: A consolidated proposal for terminology. Draft, July 2000.

[14] Vitaly Shmatikov and Ming-Hsui Wang. Measuring relationship anonymity in mix networks. In *Proceedings of the 5th ACM Workshop on Privacy in the Electronic Society (WPES 2006)*, pages 59–62, 2006.

[15] Paul Syverson. Why I'm not an entropist. In *Seventeenth International Workshop on Security Protocols*. Springer-Verlag, LNCS, 2009. Forthcoming.

[16] Paul Syverson, Gene Tsudik, Michael Reed, and Carl Landwehr. Towards an Analysis of Onion Routing Security. In *Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability*, pages 96–114, 2000.