22. EXIT NODE REPUDIATION FOR ANONYMITY NETWORKS

JEREMY CLARK, PHILIPPE GAUVIN,
AND CARLISLE ADAMS

I. INTRODUCTION

Recruiting volunteers to act as node operators in anonymity networks can be a daunting task. Although setting up a network node is becoming a simpler task, there remains the serious question of liability for the forwarding of unlawful communications such as terrorist threats, child pornography, or hate speech. In cryptography, repudiation means disclaiming responsibility for an action.[1] Cryptographers have proposed anonymity network protocols that would allow network node operators to avoid undue liability for illegal communications that have been anonymized by the network.[2] However, current research only allows the owner of a node to prove that he or she is not the originator of the message *if asked*. Given the ease with which digital evidence can be destroyed, it is unlikely that investigators would ask a suspect node operator for his or her cooperation.

1. Alfred J. Menezes, Paul C. van Oorschot and Scott A. Vanstone, *Handbook of Applied Cryptography* (Boca Raton: CRC Press, 1997), 4.

2. P. Golle, "Reputable mix networks," *Fourth Workshop on Privacy Enchancing Technologies: Proceedings of PET 2004*, in *Lecture Notes in Computer Science*, volume 3424 (Berlin: Springer-Verlag, 2005): 51–62.

It is much more probable that a warrant for the search, and possibly the seizure, of the node will be acquired without the knowledge of the node operator. It is therefore imperative that network designers understand the circumstances under which a warrant will be issued and how networks could be designed to avoid the disincentive of seized servers.

By their nature, data sent through an anonymity network will appear to have originated from the last server in the chain—the exit node. This situation illustrates two salient problems with online anonymity: it can deter or prevent law enforcement from identifying users who behave unlawfully online, and it creates liability issues for the innocent operator of the exit node who could be erroneously accused of being the perpetrator.

These problems have a seemingly easy solution: the anonymity network could simply reveal the identity of the sender. Before considering the plausibility of this, one must understand how anonymity networks of this type work. We present an overview of online anonymity and demonstrate that the decentralized structure of anonymity networks complicate this simple solution. Furthermore, we consider whether anonymity networks in theory should be able to reveal the sender's identity and what unintended consequences may result from this. We focus our attention on a solution to the second of the two problems—that the threat of a long legal process and equipment seizure is a deterrent to voluntarily operating a server in an anonymity network. We examine the legal protections against unreasonable search and seizure afforded by the Canadian Charter and laws in other jurisdictions, and propose a legally informed cryptographic protocol to allow exit nodes to repudiate any data originating from a different Internet protocol address than its own. In essence, we propose to allow exit nodes to prove, in addition to not knowing the identity of the sender and without the need for the node owner to intervene, that they are not themselves the originators of a communication.

II. PRELIMINARIES

A. A Technical Definition of Anonymity

Anonymity can mean different things in different contexts. In the field of security and privacy, anonymity requires two necessary conditions that together are sufficient for anonymity:

> P1: an anonymous action is not linkable to the identity of the actor, and P2: two anonymous actions performed by the same actor are not linkable to each other.

If the proposition P1 is false, actions are associated with the actor's identity, and the identifier is considered veronymous (a Latin portmanteau for "true name"[3]).

3. Carlisle Adams, "A classification for privacy techniques," *University of Ottawa Law & Technology Journal: special issue on anonymity, privacy, and identity* 3, no. 1 (2006): 35–52.

In this case, two disparate actions performed by the same actor would be linkable to the actor's identity and are thereby linkable to each other. This implies that proposition P2 is false whenever P1 is. If P1 is true and P2 is false, then actions can be linked to a common identifier that is not the actor's true identity. This is referred to as pseudonymity ("alternate name"). P1 is necessary for pseudonymity.

We now consider what "the identity of the actor" is in an online world. Pseudonymous identifiers are pervasive online. A self-assigned identifier is a digital pseudonym used to access features on a Web service (i.e., a screen name, user name, or e-mail address). A server-assigned identifier is a unique identifier used to monitor users (i.e., a cookie or spyware). The anonymity afforded by anonymity networks does not extend to either of these categories of identifiers. Rather, it deals with transport-layer identifiers—specifically Internet protocol (IP) addresses. When a device is online, it is reachable through its unique IP address. An IP address does not necessarily correspond to a single computer; it could, for example, identify the gateway to a network of computers. At best, IP addresses tie actions from this device together, and, therefore, could be considered pseudonymous. However, if the holder of an IP address is revealed (e.g., through self-disclosure in the holder's traffic or by the holder's Internet service provider), then the IP address could become a veronymous identifier. Anonymity networks unlink a user's actions from her IP address.

The anonymity afforded by an anonymity network is important even if the user does not reveal her full identity during communications. An IP address can be augmented with other personally identifiable information (PII), such as a search query for a relative or the revelation of a postal code, and aggregating enough information can be used to reduce the user's privacy and possibly uncover her true identity. Datamining and geo-location[4] are examples of this privacy threat.

B. An Analogy of an Anonymity Network

In order to illustrate how an anonymity network works, consider Bob who is very flattered when he realizes that someone has left him an anonymous valentine. His secret admirer, Alice, knew that leaving the note on Bob's desk was too risky—she might be seen—so she decided to ask her trustworthy friend, Charlie, to assist her. The initial idea was that Alice would give Charlie the valentine, and Charlie would leave it on Bob's desk. In this case, Charlie is acting as a proxy for Alice, and Charlie decides to announce publicly that he will be acting as a

4. Venkata N. Padmanabhan and Lakshminarayanan Subramanian, "An investigation of geographic mapping techniques for internet hosts," *Proceedings of SIGCOMM 2001*, in *ACM SIGCOMM Computer Communication Review* 31, no. 4 (2001): 173–185. See also James A. Muir and P. C. van Oorschot, "Internet geolocation and evasion," *TR-06-05* (Carleton University: Technical Report, 2006): 1–22, http://cs.smu.ca/~jamuir/papers/TR-06-05.pdf.

go-between for anyone wanting to send anonymous valentines. To Charlie's surprise, a large number of co-workers emerge to take him up on the deal. Alice is also happy with this news. She knows that if she is seen giving Charlie a valentine, she will be just one in a large group of potential senders.

There are a few complications, though. If Alice is seen giving the valentine to Charlie, someone could later recognize it when it is in Bob's possession. To prevent this, Alice hides the valentine in an envelope, and Charlie, in the privacy of his office, opens the envelope and accordingly forwards the valentine found inside it. However, Charlie must also be careful in the order in which he distributes the valentines he has received. For example, if, immediately upon receiving a sealed envelope from a sender, Charlie ducks into his office and then promptly places a valentine on someone's desk, it is easy to deduce who is sending a valentine to whom. Instead, Charlie spends the day collecting a batch of envelopes, and then at the end of the day, he takes them all out of their envelopes, shuffles them, and distributes them in a different order than he received them.

This process works if Charlie is trustworthy. However, trust can also be distributed to more than one person. For example, Alice could put her valentine to Bob in an envelope and write another trusted co-worker's name on it. She could then put this envelope in a second envelope with Charlie's name on it. She gives the package to Charlie, who opens the first envelope and learns that the envelope should be given to Deborah. Deborah opens the second envelope and finds the valentine for Bob. In this case, neither Charlie nor Deborah know that Alice is sending a valentine to Bob. Charlie knows that Alice sent a valentine to someone care of Deborah, and Deborah knows that Bob received a valentine from someone care of Charlie. As long as they do not collude with each other, no one can link Alice and Bob together.[5] This model can be expanded with an arbitrary number of proxies, and the only requirement for anonymity is that at least one is trustworthy.

C. Anonymity Networks

Online, the role of Charlie and Deborah are played by nodes, which forward Internet data between a user and a recipient. Many Web sites log the IP addresses of users who visit their site, and the use of a proxy server hides the user's IP address from the Web site. However, an IP address is not hidden if someone sees the traffic before it reaches the proxy server. In this case, the eavesdropper knows the source of the packets (the user), the destination (the proxy server), and if they open the packets, they can learn the ultimate destination (the recipient). An example of an entity who could easily log this information is an Internet service provider (ISP). This privacy threat is plausible: in June 2006 Canadian

5. It is possible for Charlie to open both envelopes, but this is a shortcoming of the analogy, not the technology: opening digital envelopes requires a secret key that only the intended recipient possesses.

ISP Bell Sympatico announced to its customers, in response to expectations that the federal government would revive an Internet surveillance bill, that it may "monitor or investigate content or your use of your service provider's networks and to disclose any information necessary to satisfy any laws, regulations or other governmental request."[6]

To protect the final destination of data from an eavesdropper, the destination can be placed into a digital "envelope" by using cryptography. Some proxy servers offer an encrypted channel to their users using the transport layer security (TLS) protocol.[7] This prevents an eavesdropper, like an ISP or someone with access to a user's network, from discovering the final destination based on the content of the message. However, if the eavesdropper could see both the traffic entering and leaving the proxy server (an entity with access to both ISP and Web site logs), they could link messages using simple timing analysis. For example, if every time an unreadable encrypted packet comes in from a given user and immediately a packet is sent from the proxy to a certain recipient, it can be reasonably deduced what recipient a user is communicating with. To prevent this, the proxy can take a batch of data from multiple users and reorder it before forwarding it. By sending traffic through a chain of such mix proxies, no one proxy will know both the original source and the final destination of the data. Every proxy server in the chain would have to collude to break the sender's anonymity, and as the sender herself could operate one of these servers, she can guarantee her own anonymity without trusting anyone else.

Mix proxies that use a random permutation to remove order-based correspondence between an input and output message set and cryptography to remove content-based correspondence were first proposed in 1981 by David Chaum.[8] A network of mix nodes is shown in Figure 1. Many modern anonymity networks are based on the idea of sending traffic through several of these specialized servers, although variations on how the servers work exist. Anonymity networks have been proposed to anonymize email[9] and Web traffic.[10]

---

6. M. Hammond, "Big brother watching you surf?" *The Globe and Mail*, June 27, 2006.

7. Or its predecessor, secure sockets layer (SSL).

8. David Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," *Communications of the ACM* 24, no. 2 (1981): 84–88.

9. George Danezis, Roger Dingledine, and Nick Mathewson, "Mixminion: design of a type III anonymous remailer protocol," *Proceedings of the 2003 IEEE Symposium on Security and Privacy* (2003): 2–15.

10. Roger Dingledine, Nick Mathewson, and Paul Syverson, "Tor: The second-generation onion router," in *Proceedings of the 13th USENIX Security Symposium* (2004): 303–320. *Also* see O. Berthold, H. Federrath, and S. Köpsell, "Web MIXes: a system for anonymous and unobservable internet access," *Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability*, in *Lecture Notes in Computer Science* 2009 (2001): 115–129.

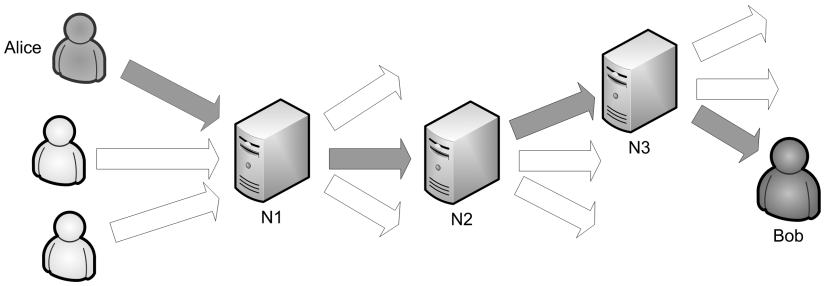

FIGURE 1. ALICE SENDS AN ANONYMOUS MESSAGE TO BOB THROUGH A NETWORK OF THREE MIX NODES. TO BOB, IT APPEARS THE MESSAGE ORIGINATED FROM THE EXIT NODE AND NOT FROM ALICE.

D. Revisiting the Motivating Problem

A seemingly simple solution to the problem of determining the originator of unlawful anonymous messages would be for the final node to reveal from whom it received the data, and law enforcement could iteratively trace the data back to the original sender. However, this would require the servers to store server logs, and server logs have no inherent integrity—they can be easily modified or forged. A further complication is that anonymity networks deliberately stretch across multiple countries. Even if server logs were reliable, an international effort would be required to subpoena the required data.[11] Alternatively, anonymity networks could be legally compelled to encrypt the identity of all participants and leave the decryption key to this information in escrow with law enforcement. However, the political viability of this situation seems dismal, as it closely parallels the proposed Clipper chip in the United States during the 1990s, which was met with a political backlash that ensured it was never adopted. Concern has also been raised that provisions created to facilitate the prosecution of heinous online crimes, like the distribution of child pornography, could also be used for

11. International law enforcement agreements already exist. As stated in "Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations" (July 2002) Computer Crime and Intellectual Property Section, Criminal Division, United States Department of Justice, http://www.justice.gov/criminal/cybercrime/searching.html#searchmanual:

> "To secure preservation, or in emergencies when immediate international assistance is required, the international Network of 24-hour Points of Contact established by the High-tech Crime Subgroup of the G8 countries can provide assistance. This network, created in 1997, is comprised of approximately twenty-eight member countries, and continues to grow every year. Participating countries have a dedicated computer crime expert and a means to contact that office or person twenty-four hours a day. See *generally* Michael A. Sussmann, *The Critical Challenges from International High-Tech and Computer-Related Crime at the Millennium*, 9 Duke J. Comp. & Int'l L. 451, 484 (1999)."

anti-democratic purposes in prohibitive nation-states or in less clear-cut situations, such as civil disputes over copyright infringement.

The situation forces us to choose between the right to online anonymity and the efficacy of criminal prosecution. It is not an easy decision. Although we have focused thus far on the costs to society, online anonymity has benefits as well. It offers privacy protection to whistle-blowers, victims of abuse, political advocates in oppressive nation states, military and intelligence agencies, individuals seeking information in confidence, or simply citizens concerned with how easily personal data can be aggregated in an online world. It is our expectation that public opinion on online anonymity will converge to a position similar to that of cryptography—that the benefits outweigh the danger.

In the meantime, we turn our focus to the perhaps secondary but more imminent legal concern for anonymity networks: server operators in anonymity networks could face anything from the seizure of equipment to the threat of criminal prosecution as a result of unlawful data they did not originate. German police have recently seized Tor servers that (presumably unwittingly) served to anonymize a child porn ring's communications.[12] It is important to understand the circumstances under which a lawful seizure can be instigated before constructing technical measures to prevent such seizures.

III. COMPUTER-RELATED SEARCH AND SEIZURE

Even if server operators are not the originators of "bad" communications, this does not resolve the basic fact that once a server is seized in an investigation, it may take well over a year before the judicial system processes the evidence and exonerates the server operator. To avoid the stress, hassles, and expense of a seized computer, the designers of anonymity networks should be concerned with search and seizure procedures along with exonerating node operators from guilt.

A. Constitutional or Supra-Statutory Protections

The Supreme Court of Canada in *R. v. Genest* discussed the balancing of interests involved in the state's right of search and seizure versus an individual's right to privacy:

> The privacy of a man's home and the security and integrity of his person and property have long been recognised as basic human rights, enjoying both an impressive history and a firm footing in most constitutional documents and international instruments. But much as these rights are valued they

12. John Oates, "German police seize Tor servers," *The Register*, September 11, 2006, http://www.theregister.co.uk/2006/09/11/anon_servers_seized/.

> cannot be absolute. All legal systems must and do allow official power in various circumstances and on satisfaction of certain conditions to encroach upon rights of privacy and security in the interests of law enforcement, either to investigate an alleged offence or to apprehend a lawbreaker or to search for and seize evidence of crime. The interests at stake are compelling. On the one hand the security and privacy of a person's home and possessions should not be invaded except for compelling reasons. On the other hand society, represented by its organised institutions, also has an undeniable and equally powerful interest in effectively investigating crime and punishing wrongdoers. The task of balancing these conflicting interests is a matter of great importance and of considerable difficulty; but it must be attempted, and so far as possible, for the health of civil liberty and law enforcement alike, satisfactorily performed.[13]

Such a balance is struck, at least nominally, in most western nations. Section 8 of the Canadian Charter of Rights and Freedoms stipulates, "Everyone has the right to be secure against *unreasonable* search and seizure" (emphasis added).[14] Section 8 not only provides the basic rights of individuals, but also serves as a constraint against unreasonable search and seizure by the state. Not only does it restrain agents of the state, but, due to the constitutional nature of the Charter, it also protects against the erosion of Canadian civil liberties through the enactment of privacy invasive legislation.

The United States also provides constitutional protections against search and seizures by the State through the Fourth Amendment to its Constitution:

> The right of the people to be secure in their persons, houses, papers, and effects, against *unreasonable searches* and seizures, shall not be violated, and *no Warrants shall issue, but upon probable cause*, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized [emphasis added].[15]

In Europe, similar protection is accorded through Article 8 of the European Convention on Human Rights, which states the following:

1. Everyone has the right to respect for his private and family life, his home and his correspondence.
2. There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the

13. *R. v Genest*, (1989), 45 CCC (3d) 385 (SCC) at 388, citing Polyvios G. Polyviou, *Search and Seizure: Constitutional and Common Law* (London: Duckworth, 1982) at vii.

14. Part I of the Constitution Act, 1982, being Schedule B to the Canada Act 1982 (.K), 1982, c. 11. (the "Charter").

15. U.S. Const. am. 4.

> economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.[16]

Doubt as to the meaning of "correspondence" was eliminated in the Charter of Fundamental Rights of the European Union,[17] which provides that everyone has the right "to respect for his or her private and family life, home *and communications* [emphasis added]."[18]

These European Union documents do not, in and of themselves, ensure the protection of Member States' citizens. They have, however, been implemented by most Member States with the European Commission taking enforcement action against those that have lagged behind.[19] Although not all European nations have a written constitution in which to implement these rights, some, such as the United Kingdom, have nevertheless provided "enhanced protection" to privacy with reference to European Convention on Human Rights.[20]

B. Search and Seizure

The definition of a "search" is simple enough. The Supreme Court of Canada,[21] for example, noted in *R. v. Wise*, "If the police activity invades a reasonable expectation of privacy, then the activity is a search."[22] One has to keep in mind, however, that the Charter only applies to governmental entities, as espoused in Section 32 of the Charter. Private individuals may be found to be acting as state agents in certain situations. In *R. v. M. (M.R.)*, the Supreme Court considered

16. Convention for the Protection of Human Rights and Fundamental Freedoms, Rome 4 November 1950 as amended by Protocol 11.

17. [2000] OJ C 364/8, 18 December 2000.

18. *Ibid.* art. 7.

19. For more on this topic, see Douwe Korff, *EC Study on Implementation of Data Protection Directive: comparative summary of national laws* (Cambridge: Human Rights Centre, September 2002).

20. *Ibid.* 8–9.

21. Due to space constraints, we, being Canadian, have focused our analysis on the Canadian situation. We note, however, that search and seizure legal schemes are substantially similar in other free and democratic jurisdictions. In the United States, for example, if government conduct does not violate the "reasonable expectation of privacy," it does not constitute a "search," and warrants are issued upon the establishment of "probable cause." (*Illinois v Andreas*, 463 U.S. 765, 771 (1983).) Unfortunately, in some jurisdictions it may not be possible to avoid the risk of search and seizure or even imprisonment for the hosting or use of anonymizing networks. Some would argue that the United States is such a jurisdiction since the passing into law of the Protect America Act of 2007, which allows for the warrantless wiretapping of international communications. We offer no opinion other than the fact that our proposed solution would render such wiretapping unreasonable,as the information sought by the State would be unavailable in properly configured nodes.

22. *R v Wise* [2002] 70 CCC (3d) 193 (SCC).

whether a school vice-principal conducting a search of a school student in the presence of a police officer was in fact a "state agent" bound by the Charter. The court formulated a test to be followed in determining whether someone is a state agent under section 8 of the Charter: "Applying the test to this case, it must be determined whether the search of the appellant would have taken place, in the form and in the manner in which it did, but for the involvement of the police."[23] As the primary purpose of the search was enforcing school discipline, the majority found that the vice-president was not a state agent in this case. There was no violation of Section 8 of the Charter in this case. This case is important as it means that an employer seizing an employee's computer, or an ISP conducting its own investigation of suspicious communications as per their user agreement may not be bound by the Charter (or similar legislation in other countries).

In seeking to obtain a warrant, enforcement agencies in Canada must, in addition to clearly defining what is sought by the warrant, establish that "there are *reasonable grounds* to believe [that what is sought by the warrant] *will afford evidence* with respect to the commission of an offence, or will reveal the whereabouts of a person who is believed to have committed an offence"[24] (emphasis added). A prudent anonymity network designer will therefore wish to ensure that the server will not produce any information that would have probative value with respect to the commission of an offence. The easiest way would be to give the police the ability to confirm for themselves whether or not the node is the originator of a communication. If the node is not the originator and the network does not allow the collection of evidence with respect to the whereabouts of the suspect, it may be unreasonable for the police to seize the server. This is precisely the design approach we will take in the next section.

IV. EXIT NODE REPUDIATION (ENR)

In this section, we propose a protocol that allows exit nodes in an anonymity network to prove that the traffic that they forward on behalf of other users does not originate from their IP address. We term our solution *exit node repudiation* (ENR).

23. *R. v M. (M.R.)* [1998] 129 CCC (3d) 361 (SCC).

24. An Act respecting the criminal law, R.S.C. 1985, c. C-46, as amended, art. 487 (1) (b). Similarly, law enforcement officers must, in the United States, draft a sworn statement that explains the basis for their belief that the search is justified by probable cause that that contraband, evidence, fruits, or instrumentalities of crime exist in the location to be searched. See "Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations" (July 2002), Computer Crime and Intellectual Property Section, Criminal Division, United States Department of Justice, http://www.justice.gov/criminal/cybercrime/searching.html#searchmanual.

Previous technical research on the issue of dealing with unlawful messages has predictably forked between providing traceability for messages and providing measures that allow the anonymity network to prove it did not originate the messages without revoking anonymity. A representative work on the traceability side[25] presents a scheme that allows for the selective tracing of a single message in an anonymity network without revealing the origin of other messages. Alternatively, mathematical proofs can be constructed to prove that an output set of an anonymity network is a perfect random bijection of the input set without revealing the permutation,[26] a property termed *robustness*, which indirectly proves that the servers are not responsible for any data in the output set unless if they contributed a message to the input set. However, robustness proofs are burdensome and not very practical.

A. Defining Exit Node Repudiation

Phillipe Golle suggests a weaker but computationally feasible form of robustness termed *near-reputability*:

> An anonymity network is near-reputable for demarcation function f, batch output B, and set of players P_B if there exists a subset of the batch output $f(B) \subseteq B$ such that each message in $f(B)$ can be proven to have originated from some player $p \in P_B$ without revealing which one.[27]

We expect legal enforcement action to be levied against the exit nodes of an anonymity network and not the anonymity network as a whole. As a result, we prefer a definition of a near-reputable exit *node*. However, it is not sufficient for a node to have near-reputability by extension of operating in a near-reputable network. If all the nodes behave correctly and the exit node is in P_B, then this definition will suffice; but these assumptions are too strict. First, a major incentive to operating a node is the ability to mix in your own traffic (this way, you can ensure that one node in the network operates correctly), and so requiring the set of exit nodes to be disjoint from P_B is not ideal. Second, we expect some nodes will not behave correctly, whether maliciously or as a result of unintentional data corruption. Thus, we must consider the case that a message is not in $f(B)$ (i.e., is in $B - f(B)$). Such a message may have originated from the exit node in question, or it may have originated from any other node in the network.

25. Luis von Ahn and others, "Selectively traceable anonymity," *Sixth Workshop on Privacy Enchancing Technologies: Proceedings of PET 2006*, in *Lecture Notes in Computer Science* 4258 (Berlin: Springer-Verlag, 2006): 208–222.

26. Markus Jakobsson, Ari Juels, and Ronald L. Rivest, "Making mix nets robust for electronic voting by randomized partial checking," *Proceedings of the 11th USENIX Security Symposium* (2002): 339–353.

27. Golle, "Reputable mix networks," 55 (n. 2).

The situation is ambiguous and offers plausible deniability to all nodes. We have chosen to tighten definition 4.1.2 so that the consequent can be affirmed:

> An exit node is g-reputable for batch output $\boldsymbol{B}$, demarcation function $\boldsymbol{g}$, and subset of all players $\boldsymbol{g(P) \subset P}$ if every message can be proven to have originated from a player in $\boldsymbol{g(P)}$ without revealing which one. Exit node repudiation (ENR) is the further condition that the only player in $\boldsymbol{P - g(P)}$ is the exit node itself.

ENR divides the set of all players into two subsets: the exit node in question, and everyone else. Our proposed solution will query an algorithm to determine if a message originated from the set of "everyone else." If the algorithm returns true, the message is proven to not have originated from the exit node. If the algorithm returns false, the message is proven not to have originated from the exit node. This definition is perfectly precise and resolves any ambiguity over the exit node's actions. If accused of originating a message, the message is either repudiable or nonrepudiable. This definition presumes that the anonymity network will only output messages if they properly conform to a protocol and drop everything else. It also excludes the exit node from serving as an exit node for its own anonymous messages; however, it can still originate anonymous messages and serve as an entrance or intermediary node to ensure the integrity of the chain.

B. A Nonmathematical Overview of ENR

We will consider the following participants in our solution: Alice who wishes to send an anonymous message to Bob through three nodes in an anonymity network (the exit node we refer to as $\mathbf{N_3}$), and an issuing authority who we will call Justine. The ultimate goal of this protocol is to provide $\mathbf{N_3}$ with the ability to prove that Alice's message did not originate from its own IP address. To accomplish this, we will employ digital credentials that were proposed by Stefan Brands for identity management.[28] Digital credentials are similar to a digital certificate in that they enclose attributes in a signed document. However, these attributes can be selectively hidden or disclosed in a fine-grained manner. Moreover, the presentation of a digital credential cannot be linked to its issuance on the basis of the issuer's signature or other cryptographic materials contained in the credential.

The protocol begins with Alice contacting Justine for a digital credential that encloses her IP address. We allow law enforcement, for whom the proofs are ultimately intended, to assume the role of Justine or delegate it to an entity it trusts. Justine is free to choose the most trustworthy method she is aware of for

28. Stefan A. Brands, *Rethinking public key infrastructures and digital certificates: building in privacy* (Cambridge, MA: MIT Press, 2000), http://www.credentica.com/the_mit_pressbook.html. See also Stefan Brands, "A technical overview of digital credentials," (February 20, 2002), http://www.cypherspace.org/credlib/brands-technical.pdf.

verifying Alice's IP address. Verifying the integrity of an IP address is a nontrivial problem; however, it persists even if law enforcement is given the ability to trace a message as that message will be traced to an IP address that will need to be resolved to an identity.

Justine creates the credential in cooperation with Alice. Both Alice and Justine use private keys in this protocol: Justine to sign the credential and Alice to ensure that she will be the only person able to use the credential. During the interactive creation of the credential, Alice can blind the credential—a process that makes it unrecognizable to Justine without destroying the integrity of the IP address in the credential or Justine's signature on the credential. Later, Alice will show her credential to $\mathbf{N_3}$ and Bob without fully revealing the attribute inside it. Either can use Justine's public key to verify that the credential was issued by her and is intact. However, because of the blinding process, $\mathbf{N_3}$ or Bob can show Alice's credential to Justine, and Justine will not be able to determine that it is the same credential she issued to Alice. Therefore Alice is anonymous to $\mathbf{N_3}$ and Bob due to the properties of the anonymity network, and she is anonymous to Justine due to the properties of the digital credential.

With Alice's digital credential alone, any attribute in it cannot be determined by anyone unless Alice actively participates in a showing protocol. To reveal an attribute, Alice claims that the credential contains a certain value, and then proves it does by showing a mathematical relationship that depends on her private key and on a random challenge (chosen by a publicly verifiable method). This proof is unforgeable by anyone without Alice's secret key, and because it is in response to a random challenge, the credential and proof cannot reused together.

To complete the protocol, Alice appends her credential and a proof to her messages. The proof does not reveal the attribute in the credential, Alice's IP address, as this would break her anonymity. Instead it proves a property of her credential: that it is not equal to the exit node's IP address. The scheme is shown in Figure 2.

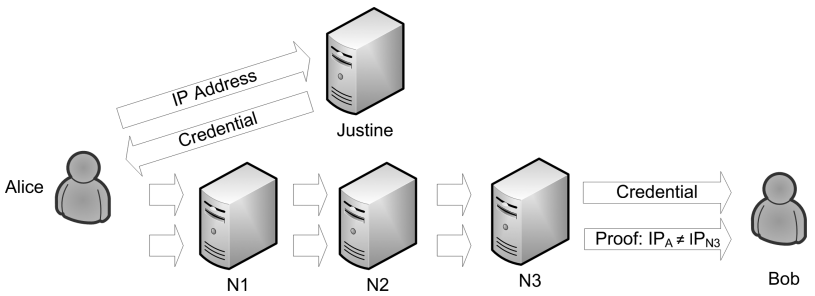

FIGURE 2. EXIT NODE REPUDIATION USING DIGITAL CREDENTIALS: ALICE IS ISSUED AN ANONYMOUS CREDENTIAL AND OFFERS A SIGNED PROOF THAT HER IP ADDRESS IS NOT EQUAL TO THE EXIT NODE'S IP ADDRESS. THIS CAN BE VERIFIED BY THE EXIT NODE AND BY THE RECIPIENT.

The remainder of this section will detail the cryptographic nature of the solution. It is included for completeness and is intended for computer scientists. It may be skipped over by those without knowledge of cryptographic primitives, as the high-level description above should suffice for understanding the design of the system.

C. Key Generation

The key generation protocol run by Justine to establish her public and private key is shown in Algorithm 1. Note that all algorithms are derived from the work of Stefan Brands.[29]

Algorithm 1: Key Generation for J

Input: Public parameter p.
Output: Private key $\langle s_1, s_2 \rangle$ and public key $\langle g_0, g, h \rangle$.

J should:
  Choose random secrets $s_1, s_2 \leftarrow_r \mathbb{Z}_p$.
  Choose random generator $g_0 \leftarrow_r \mathbb{G}_p$.
  Compute $g = g_0^{s_1}$ and $h = g_0^{s_2}$.
end

Public parameter p is a suitably large prime number (e.g., 1024 bits), and $\mathbb{G}_p$ is the set of primitive roots in $\mathbb{Z}_p^*$. Justine's public key is $\langle g_0, g, h, p \rangle$, and s_1 and s_2 are retained as her private key. Note that s_1 and s_2 cannot be recovered from the knowledge of the parameters in the public key without computing a discrete logarithm, a problem assumed to be computationally infeasible for large p.

D. The Issuing Protocol

The issuing protocol is shown in Algorithm 2. The key generation algorithm produces public parameters g and h, which are arranged by Alice into a credential of form $g^x h^\alpha$ where x is Alice's IP address. This credential can be thought of as having secret key α applied to an encrypted attribute x. For every value of x, there is a unique value of α that will produce the same value for the credential. Thus if α is unknown, the value of x is perfectly hidden.

In Algorithm 2, Alice creates the credential, I, and Justine provides a signature certifying that is correct. Note that Justine never sees the value of I itself, so she cannot recognize I when Alice uses it. The signature on the credential, $\langle c, r \rangle$, is more properly a private key certificate;[30] however, we refer to it as a signature for convenience. Once again, Justine does not see , only a blinded version of the values: $\langle \tilde{c}, \tilde{r} \rangle$. The protocol employs a hash function $\mathcal{H}$ that is assumed to be publicly known and cryptographically secure with an output space less than p.

29. Brands, *Rethinking*, 91 (n. 28).
30. Brands, *Technical Overview*, footnote 3 at 17 (n. 28).

Alice employs the hash to send a function of her credential, ***c***, to Justine who calculates a suitable response using the value of ***x***. Should Alice's credential not contain the same value of ***x*** that Justine uses in her response, the signature will not hold.

Algorithm 2: Issuing Protocol

Input: Public Key $\langle g_0, g, h, p\rangle$, A's IP address x, and (known only to J) Private Key $\langle s_1, s_2\rangle$.
Output: Credential I and signature sig(I)= $\langle c, r\rangle$.

A should:
  Choose random secret $\alpha \leftarrow_r \mathbb{Z}_p^*$.
  Compute $I = g^x h^\alpha$.
end

J should:
  Choose random secret $w \leftarrow_r \mathbb{Z}_p$.
  Compute $z = g_0^w$ and send to A.
end

A should:
  Choose random secrets $\beta_1, \beta_2 \leftarrow_r \mathbb{Z}_p$.
  Compute $c = \mathcal{H}(I, (g^x h)^{\beta 1} \cdot g_0^{\beta_2} \cdot z)$.
  Blind c by computing $\tilde{c} = (c + \beta_1) \bmod p$ and send to J
end

J should:
  Compute $\tilde{r} = \widetilde{c}(s_2 + xs_1) + w) \bmod p$ and send to A.
end

A should:
  Verify $z = g_0^{\tilde{r}}(g^x h)^{-\tilde{c}}$.
  Unblind r by computing $r = (\tilde{r} + \beta_2 + c\alpha) \bmod p$.
end

E. The Showing Protocol

Algorithm 3 demonstrates how Alice can generate a signed proof that the attribute in her credential ***x*** is not the same as another attribute ***y***. In this case, ***x*** is her IP address and ***y*** is the IP address of the exit node. The IP address of the exit node must be known by Alice. Although it is more efficient if she knows it a priori, it is possible for $\mathbf{N_3}$ to send its IP address back through the anonymity network to Alice. In anonymity networks like Tor, Alice can choose her own exit node and thus know its IP address.

The showing protocol is based on a challenge-response, where the challenge requires nonce ***n***. The nonce is used to ensure that the credential is not used by anyone other than Alice (i.e., only by those who know the secret key α). If the protocol were not challenge-response, the credential and proof could be replayed together by someone who observed Alice using a credential. We suggest that the nonce be a hash of the message, Bob's IP address, which both $\mathbf{N_3}$ and Bob know, and a large random number collaboratively generated by the nodes in the anonymity network—the latter being published with a timestamp and periodically updated. This does not completely prevent replay attacks, but it severely limits

Algorithm 3: Signed Proof ($x \neq y$)

Input: $\langle g_0, g, h, p\rangle$, I, x, N_3's IP address y, and nonce n.
Output: $\langle a, r_2, r_3\rangle$.

A should:
  Choose random secrets $w_1, w_2 \leftarrow_r \mathbb{Z}_p$.
  Compute $a = I^{-w_1} g^{yw_1} h^{w_2}$.
  Compute $c_1 = \mathcal{H}(a, I, y, n)$.
  Compute $\varepsilon = y - x$.
  Compute $\delta = \varepsilon^{-1}$.
  Compute $r_2 = c_1\delta + w_1$.
  Compute $r_3 = c_1\alpha\delta + w_2$.
end

them to the same message and same receiver in the same window of time. This small cost is outweighed by the benefit of a standardized public nonce: Alice can compute the value of the nonce a priori and can create her response without having to exchange any information with $\mathbf{N_3}$.

The proof itself is based on the observation that if **x** (Alice's IP address) and **y** ($\mathbf{N_3}$'s IP address) are different, their difference is nonzero and thus invertible within an appropriate group such that $(x-y)(x-y)^{-1} \equiv 1 \bmod p$. If x and y are the same, the difference is zero, which is noninvertible, leaving δ uncalculated (or zero if the inverse of zero is so defined). However, in the case that $\delta=0$, then r_2 and r_3 would equal w_1 and w_2, respectively, and the verification procedure in Algorithm 4 would fail.

Algorithm 4: Verification Algorithm

Input: $\langle g_0, g, h, p\rangle$, $\langle I, c, r, a, r_2, r_3\rangle$, y, n.
Output: TRUE or FALSE.

N_3 **should:**
  Verify $c = \mathcal{H}(I, g_0^{r_0}(Ih)^{-c})$.
  Compute $c_1 = \mathcal{H}(a, I, y, n)$.
  Verify $I^{r_2} a = g^{r_2 y - c_1} h^{r_3}$.
end

The complete package that Alice delivers to $\mathbf{N_3}$ is $\langle I, c, r, a, r_2, r_3\rangle$. There are three distinct parts to this package: I is the credential; c and r are used to verify Justine's signature on the credential; and a, r_2, and r_3 are Alice's signed proof that x is not equal to y. This verification should be performed by $\mathbf{N_3}$ before releasing the message to Bob. If either verification fails, the circuit should be destroyed. The package can also be forwarded to Bob, who also has all the information needed to verify the correctness of the credential. This is important because it allows law enforcement to satisfy themselves of ENR without requiring any *ex post* interaction with $\mathbf{N_3}$.

The credentials can be independent of what anonymity network Alice wants to use or what message she will send; in fact, they could be used for any online purpose where Alice wants to prove some property about her IP address.

Furthermore, Alice can be issued a large quantity of credentials in bulk, each unique but with the same attribute, at some time prior to using an anonymity service as long as the issuing authority's public parameters are still known when she uses the credential. This changes the efficiency of the issuing protocol from a marginal cost to a fixed overhead cost.

One criticism of our proposed ENR protocols is the validity of **x** in the credential. For example, it is possible for a credential to be issued to a user at one IP address and then used by the same user to send a message from a different IP address. It would also be possible to use a proxy server to interact with credential issuer, so that the proxy server's IP address is encoded into the credential instead of Alice's. In response to this criticism, we note several things. First, Alice has no incentive to try to obscure her IP address from the credential issuer. The only property of her IP address that will be revealed is that it is not equal to $\mathbf{N}_3$'s, and any further proofs about **x** or its properties require Alice's private key. Second, lending and borrowing credentials is the equivalent of using someone else's computer—something that is possible independent of whether an anonymity network is even used. Third, lists of known proxy servers could be compiled and checked against. Finally, as noted previously, the legal alternative to ENR is traceability, and this problem applies equally to it. If a message is traced through an anonymity network to a supposed sender's IP address, there is no guarantee that the IP address is actually the sender's and not that of a proxy server or compromised machine.

V. CONCLUDING REMARKS

Recruiting server operators for anonymity networks is of primary importance to the functionality of the network. Network node servers must therefore not only be easy to set up but there must also be low risks for the node operators themselves for the dissemination of unlawful communications. Evading liability is of little comfort, however, if the node operator's computer is seized by police forces. The anonymous capability of the receiver to verify that the last communicator is not the originator of a message, without revealing the originator's IP address, could actually increase the network's anonymizing capability. Indeed, the threat of breaking the privacy of lawful communications for the purpose of uncovering unlawful ones would decrease.

We note that in Canada, a computer may be seized only if "there are *reasonable* grounds to believe [that the articles] will afford evidence with respect to the commission of an offence, or will reveal the whereabouts of a person who is believed to have committed an offence"[31] (emphasis added). It will be much harder to convince a judge that seizing an exit node will afford evidence of a crime if the

31. Criminal Code of Canada 487 (1) (b).

exit node can prove, without the node owner's knowledge or intervention, that it did not originate the communication and does not harbor information that could be linked to the sender.[32] Exit node repudiation provides a method of retaining the anonymity of the sender while presenting a response to the pertinent question of legal liability for the exit node as well as the practical hassles of equipment seizure. We hope this innovation is helpful in preserving the legality of anonymity networks and decreasing the aversion to volunteer as operators of servers in these networks.

32. This approach is also consistent with the need to establish "probable cause" in preparing a warrant to search and seize a computer under U.S. law. "Probable cause" has been defined by the U.S. Supreme Court as the establishment of "a fair probability that contraband or evidence of a crime will be found in a particular place." (*Illinois v Gates* [1983] 462 U.S. 213 at 238).