Defending Against Sybil Attacks in Anonymizing Networks

Chandhana S D
Prathyusha Institute of Technology and Management
Thiruvallur, Tamilnadu.

Mary Anita E A
Prathyusha Institute of Technology and Management
Thiruvallur, Tamilnadu.

Abstract - An anonymity network enables users to access the Web while blocking any tracking or tracing of their identity on the Internet. Tor is open-source anonymity software free to public use. This type of online anonymity moves Internet traffic through a worldwide network of volunteer servers. Anonymity networks prevent traffic analysis and network surveillance or at least make it more difficult. Website administrators routinely rely on IP-address blocking for disabling access to misbehaving users, but blocking IP addresses is not practical if the abuser routes through an anonymizing network. As a result, administrators block all known exit nodes of anonymizing networks, denying anonymous access to misbehaving and behaving users alike. To address this problem, we present Nymble, a system in which servers can "blacklist" misbehaving users, thereby blocking users without compromising their anonymity. Our system is thus agnostic to different servers definitions of misbehavior.

Keywords - *Anonymous blacklisting, privacy, rate-limited, revocation.*

I. INTRODUCTION

Anonymizing networks allow users to access Internet services privately using a series of routers to hide the client's IP address from the server. Anonymizing networks such as Tor route traffic through independent nodes in separate administrative domains to hide a client's IP address. Tor is open-source anonymity software free to public use. Tor software conceals the user's location and/or usage. In order to use Tor, users must run onion routing. This technology encrypts and then rebounds communications onto a network of relays run by volunteers throughout the world. Users who want their Internet searches to remain private make use of anonymity networking. Within the Tor network, Internet traffic is sent to various routers, one at a time. An anonymous P2P communication system is a peer-to-peer distributed application in which the nodes or participants are anonymous or pseudonymous. Anonymity of participants is usually achieved by special routing overlay networks that hide the physical location of each node from other participants. Unfortunately, some users have misused such networks under the cover of anonymity; users have repeatedly defaced popular Web sites such as Wikipedia. Since Web site administrators cannot blacklist individual malicious users' IP addresses, they blacklist the entire anonymizing network. Such measures eliminate malicious activity through anonymizing networks at the cost of denying anonymous access to behaving users.

In pseudonymous credential systems, users log into Websites using pseudonyms, which can be added to a blacklist if a user misbehaves. Unfortunately, this approach results in pseudonymity for all users, and weakens the anonymity provided by the anonymizing network. Anonymous credential systems employ group signatures.

Basic group signatures allow servers to revoke a misbehaving user’s anonymity by complaining to a group manager. Servers must query the group manager for every authentication, and thus, lacks scalability. Traceable signatures allow the group manager to release a trapdoor that allows all signatures generated by a particular user to be traced; such an approach does not provide the backward unlinkability, where a user’s accesses before the complaint remain anonymous.

Backward unlinkability allows for subjective blacklisting, where servers can blacklist users for whatever reason since the privacy of the blacklisted user is not at risk. In contrast, approaches without backward unlinkability need to pay careful attention to when and why a user must have all their connections linked, and users must worry about whether their behaviors will be judged fairly.

Subjective blacklisting is also better suited to servers such as Wikipedia, where misbehaviors such as questionable edits to a Webpage, are hard to define in mathematical terms. In some systems, misbehavior can indeed be defined precisely. For instance, double spending of an “e-coin” is considered misbehavior in anonymous e-cash systems following which the offending user is deanonymized. Unfortunately, such systems work for only narrow definitions of misbehavior; it is difficult to map more complex notions of misbehavior onto “double spending” or related approaches. With dynamic accumulators a revocation operation results in a new accumulator and public parameters for the group, and all other existing users’ credentials must be updated, making it impractical. Verifier-local revocation (VLR) fixes this shortcoming by requiring the server (“verifier”) to perform only local updates during revocation. Unfortunately, VLR requires heavy computation at the server that is linear in the size of the blacklist.

II. EXISTING METHODOLOGY

A secure system called Nymble was designed to provide anonymous authentication, backward unlinkability, subjective blacklisting, fast authentication speeds, rate-limited anonymous connections, revocation auditability and also addresses sybil attack. Servers can blacklist anonymous users without the knowledge of their IP addresses while allowing behaving users to connect anonymously. This system ensures that users are aware of their blacklist status and they disconnect immediately if they are blacklisted. Any number of anonymizing networks relies on the same system, blacklisting anonymous users regardless of their anonymizing network(s) of choice.

Website administrators rely on IP-address blocking for disabling access to misbehaving users, but this is not practical if the abuser routes through Tor. As a result, administrators block all Tor exit nodes, denying anonymous access to honest and dishonest users alike. To address this problem, we present a system in which (1) honest users remain anonymous and their requests unlinkable; (2) a server can complain about a particular anonymous user and gain the ability to blacklist the user for future connections; (3) this blacklisted user's accesses before the complaint remain anonymous; and (4) users are aware of their blacklist status before accessing a service. The algorithms used are RSA and MAC. The advantages of this system are privacy of the blacklisted users is maintained; cryptographic functions are used for security and also prevent malicious attack.

2.1 RSA algorithm

The RSA algorithm is used to provide confidentiality. The RSA algorithm is named after Ron Rivest, Adi Shamir and Len Adleman, who invented it in 1977 [RIVE78]. The basic technique was first discovered in 1973 by Clifford Cocks [COCK73] of CESG (part of the British GCHQ) but this was a secret until 1997. The patent taken out by RSA Labs has expired. The RSA cryptosystem is the most widely-used public key cryptography algorithm in the world. It can be used to encrypt a message without the need to exchange a secret key separately. The RSA algorithm can be used for both public key encryption and digital signatures. Its security is based on the difficulty of factoring large integers.

2.2 MAC algorithm

Message Authentication Code algorithm is to provide authentication to message. A MAC algorithm, sometimes called a keyed (cryptographic) hash function, accepts as input a secret key and an arbitrary-length message to be authenticated, and outputs a MAC. The MAC value protects both a message's data integrity as well as its authenticity, by allowing verifiers (who also possess the secret key) to detect any changes to the message content. MACs differ from digital signatures as MAC values are both generated and verified using the same secret key. This implies that the sender and receiver of a message must agree on the same key before initiating communications, as is the case with symmetric encryption. For the same reason, MACs do not provide the property of non-repudiation offered by signatures specifically in the case of a network-wide shared secret key: any user who can verify a MAC is also capable of generating MACs for other messages. In contrast, a digital signature is generated using the private key of a key pair, which is asymmetric encryption. Since this private key is only accessible to its holder, a digital signature proves that a document was signed by none other than that holder. Thus, digital signatures do offer non-repudiation.

III. PROPOSED METHODOLOGY

The existing methodologies have some drawbacks. If the message frequency is very high at a node, the NM forwards only the simple messages. Latency is more and hence the speed gets reduced. In proposed methodology, the credential system is secured by the hashing algorithm. The algorithms used are RSA for confidentiality and MD5 for authentication. The advantages are that it is more secure, increased speed so that latency will be less and the NM (Nymble Manager) can forward all the messages even if the frequency of messages is very high at a node.

3.1 MD5 algorithm

MD5 algorithm was developed by Professor Ronald L. Rivest in 1991. According to RFC 1321, “MD5 message-digest algorithm takes as input a message of arbitrary length and produces as output a 128-bit "fingerprint" or "message digest" of the input …The MD5 algorithm is intended for digital signature applications, where a large file must be "compressed" in a secure manner before being encrypted with a private (secret) key under a public-key cryptosystem such as RSA.” Figure 1 shows the MD5 algorithm structure. MD5 is simple to implement, and provides a "fingerprint" or message digest of a message of arbitrary length. It performs very fast on 32-bit machine.

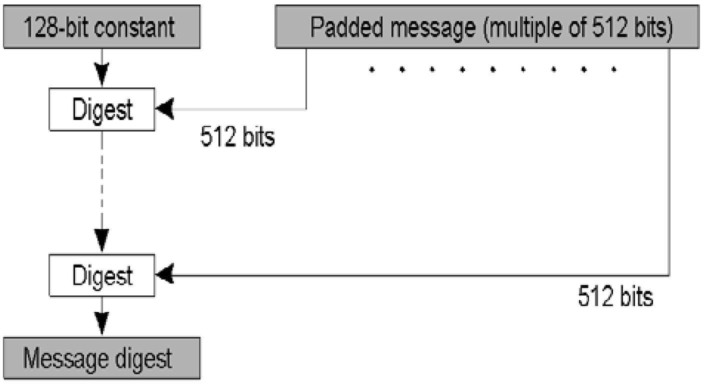


Figure 1: MD5 Algorithm Structure Implementation Steps in MD5

Step 1: Append padding bits.
Step 2: Append length.
Step 3: Initialize MD buffer.
Step 4: Process the message in 16 word blocks.
Step 5: Output (message digest).

The advantages of MD5 algorithm are the generation of a digest is very fast and the digest itself is very small and can easily be encrypted and transmitted over the internet. It is very easy and fast (and therefore cheap) to check some data for validity. The algorithms are well known and implemented in most major programming languages, so they can be used in almost all environments.

IV. PERFORMANCE EVALUATION

Figure 2 shows the size of the various data structures. The X-axis represents the number of entries in each data structure—complaints in the blacklist update request, tickets n the credential (equal to L, the number of time periods in a linkability window), nymbles in the blacklist, tokens and seeds in the blacklist update response, and nymbles in the blacklist (For example, a linkability window of one day with five minute time periods equates to L = 288). Credential in this case is about 59 KB. The size of a blacklist update request with 50 complaints is roughly 11 KB, whereas the size of a blacklist update response for 50 complaints is only about 4 KB. The size of a blacklist (downloaded by users before each connection) with 500 nymbles is 17 KB.

In general, each structure grows linearly as the number of entries increases. Credentials and blacklist update requests grow at the same rate because a credential is a collection of tickets which is more or less what is sent as a complaint list when the server wishes to update its blacklist. In our implementation, we use Google’s Protocol Buffers to (un)marshal these structures because it is cross-platform friendly and language-agnostic.

Figure 3a shows the amount of time it takes the NM to perform various protocols. It takes about 9 ms to create a credential when L = 288. Note that this protocol occurs only once in every linkability window for each user wanting to connect to a particular server. For blacklist updates, the initial jump in the graph corresponds to the fixed overhead associated with signing a blacklist. To execute the update blacklist protocol with 500 complaints, it takes the NM about 54 ms. However, when there are no complaints, it takes the NM on average less than a millisecond to update the daisy.

Figure 3b shows the amount of time it takes the server and user to perform various protocols. These protocols are relatively inexpensive by design, i.e., the amount of computation performed by the users and servers should be minimal. For example, it takes less than 3 ms for a user to execute a security check on a blacklist with 500 nymbles. Note that this figure includes signature verification as well, and hence, the fixed-cost overhead exhibited in the graph. It takes less than a millisecond for a server to perform authentication of a ticket against a blacklist with 500 nymbles.

Every time period (e.g., every five minutes), a server must update its state and blacklist. Given a linking list with 500 entries, the server will spend less than 2ms updating the linking list. If the server were to issue a blacklist update request with 500 complaints, it would take less than 3 ms for the server to update its blacklist.

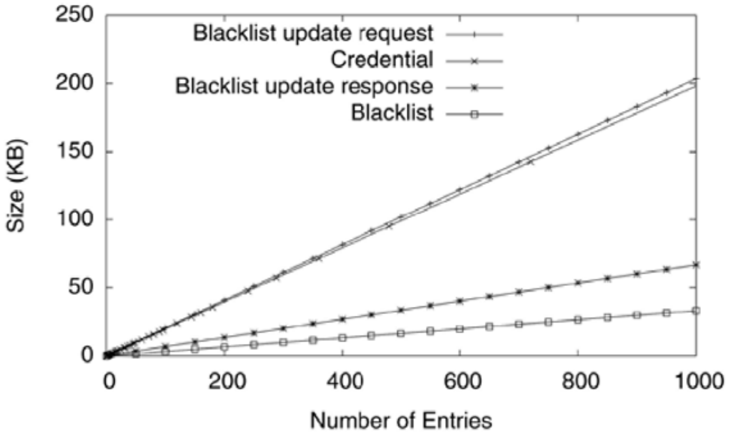
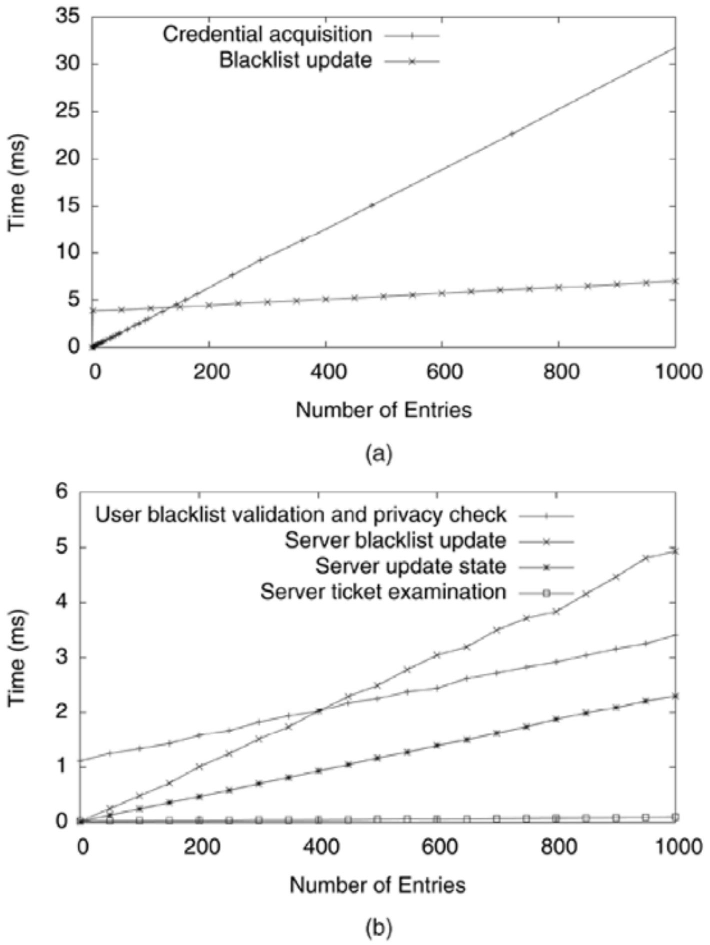


Figure 2: The Marshaled Size of Various Nymble Data Structures

The X-axis refers to the number of entries complaints in the blacklist update request, tickets in the credential, tokens and seeds in the blacklist update response, and nymbles in the blacklist.



(a)

(b)

Figure 3: Nymble's performance at (a) the NM and (b) the user and the server when performing various protocols. (a) Blacklist updates take several milliseconds and credentials can be generated in 9 ms for the suggested parameter of L = 288. (b) The bottleneck operation of server ticket examination is less than 1 ms and validating the blacklist takes the user only a few ms.

V. CONCLUSION

The major issues in anonymizing networks are misbehaving user access and blacklisting the misbehaving users without knowing their IP addresses. The paper proposes a system which can be used to add a layer of accountability to any publicly known anonymizing network. This system allows websites to selectively block users of anonymizing networks such as Tor. Servers can blacklist misbehaving users while maintaining their privacy and these properties can be attained in a way that is practical, efficient, and sensitive to the needs of both users and services. This work will increase the mainstream acceptance of anonymizing networks such as Tor, which has, thus far, been completely blocked by several services because of users who abuse their anonymity.

REFERENCES

1. Camenisch.J and Lysyanskaya.A, "An Efficient System for Non-Transferable Anonymous Credentials with Optional Anonymity Revocation," Proc. Int'l Conf. Theory and Application of Cryptographic Techniques (EUROCRYPT), Springer, pp. 93-118, 2001.
2. Camenisch.J and Lysyanskaya.A, "Dynamic Accumulators and Application to Efficient Revocation of Anonymous Credentials," Proc. Ann. Int'l Cryptology Conf. (CRYPTO), Springer, pp. 61-76, 2002..
3. Camenisch.J and Lysyanskaya.A, "Signature Schemes and Anonymous Credentials from Bilinear Maps," Proc. Ann. Int'l Cryptology Conf. (CRYPTO), Springer, pp. 56-72, 2004.
4. Dingledine.R, Mathewson.N, and Syverson.P, "Tor: The Second- Generation Onion Router," Proc. Usenix Security Symp, pp. 303- 320, Aug. 2004.
5. Holt.J.E and Seamons.K.E, "Nym: Practical Pseudonymity for Anonymous Networks," Internet Security Research Lab Technical Report 2006-4, Brigham Young Univ., June 2006.
6. Johnson.P.C, Kapadia.A, Tsang.P.P, and Smith.S.W, "Nymble: Anonymous IP-Address Blocking," Proc. Conf. Privacy Enhancing Technologies, Springer, pp. 113-133, 2007.
7. Lysyanskaya.A, Rivest.R.L, Sahai.A, and Wolf.S, "Pseudonym Systems," Proc. Conf. Selected Areas in Cryptography, Springer, pp. 184-199, 1999.
8. C. Cornelius, A. Kapadia, P.P. Tsang, and S.W. Smith, "Nymble: Blocking Misbehaving Users in Anonymizing Networks," Technical Report TR2008-637, Dartmouth College, Computer Science,Dec. 2008.
9. G. Ateniese, D.X. Song, and G. Tsudik, "Quasi-Efficient Revocation in Group Signatures," Proc. Conf. Financial Cryptography,Springer, pp. 183-197, 2002.
10. D. Boneh and H. Shacham, "Group Signatures with Verifier-Local Revocation," Proc. ACM Conf. Computer and Comm. Security,pp. 168-177, 2004.
11. E. Bresson and J. Stern, "Efficient Revocation in Group Signatures,"Proc. Conf. Public Key Cryptography, Springer, pp. 190-206,2001.
12. J. Camenisch and A. Lysyanskaya, "An Efficient System for Non-Transferable Anonymous Credentials with Optional Anonymity Revocation," Proc. Int'l Conf. Theory and Application of Cryptographic Techniques (EUROCRYPT), Springer, pp. 93-118, 2001.
13. D. Chaum, "Showing Credentials without Identification Transfeering Signatures between Unconditionally Unlinkable Pseudonyms," Proc. Int'l Conf. Cryptology (AUSCRYPT), Springer, pp. 246-264, 1990.
14. D. Chaum and E. van Heyst, "Group Signatures," Proc. Int'l Conf. Theory and Application of Cryptographic Techniques (EUROCRYPT), pp. 257-265, 1991.
15. R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The Second- Generation Onion Router," Proc. Usenix Security Symp., pp. 303- 320, Aug. 2004.
16. J.R. Douceur, "The Sybil Attack," Proc. Int'l Workshop on Peer-to-Peer Systems (IPTPS), Springer, pp. 251-260, 2002.
17. B.N. Levine, C. Shields, and N.B. Margolin, "A Survey of Solutions to the Sybil Attack," Technical Report 2006-052, Univ. of Massachusetts, Oct. 2006.
18. P.P. Tsang, M.H. Au, A. Kapadia, and S.W. Smith, "Blacklistable Anonymous Credentials: Blocking Misbehaving Users without TTPs," Proc. 14th ACM Conf. Computer and Comm. Security (CCS '07), pp. 72-81, 2007.