The Anonymizing Networks of Blocking Misbehaving Users

[1]G.Naga Mallika, [2]K.Rajini Kumari, [3]T.Rajesh
[1,2,3]Dept. of CSE, A.S.R College of Engineering & Tech., Tetali, Tanuku, AP, India

Abstract
Anonymizing networks such as Tor allow users to access internet services privately by using a series of routers to hide the client's IP address from the server. The success of such networks However has been limited by users employing this anonymity for abusive purposes such as defacing popular web sites. Web site administrators routinely rely on IP-address for blocking or 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 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 server definitions of misbehavior servers can blacklist users for whatever reason, and the privacy of blacklisted users is maintained.

Keywords
Anonymzing Networks, IP, Blocking

I. Introduction

Anonymizing networks such as Tor route traffic through independent nodes in separate administrative domains to hide a client's IP address. Unfortunately, some users have misused such networks-under the cover of anonymity, users have repeatedly defaced popular Web sites. 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 other words, a few"bad apples" can spoil the fun for all. Subjective blacklisting is also better suited to servers such as Wikipedia, where misbehaviours 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.

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.

A. Anonymity and Pseudonymity

Some of the networks commonly referred to as "anonymous P2P" are truly anonymous, in the sense that network nodes carry no identifiers. Others are actually pseudonymous: instead of being identified by their IP address, nodes are identified by pseudonyms such as cryptographic keys. For example, each node in the MUTE network has an overlay address that is derived from its public key. This overlay address functions as a pseudonym for the node, allowing messages to be addressed to it. In Frenet, on the other hand, messages are routed using keys that identify specific pieces of data rather than specific nodes; the nodes themselves are anonymous.

The term anonymous is used to describe both kinds of network because it is difficult—if not impossible—to determine whether a node that sends a message originated the message or is simply forwarding it on behalf of another node. Every node in an anonymous P2P network acts as a universal sender and universal receiver to maintain anonymity. If a node was only a receiver and did not send, then neighboring nodes would know that the information it was requesting was for itself only, removing any reasonable deniability that it was the recipient (and consumer) of the information. Thus, in order to remain anonymous, nodes must ferry information for others on the network.

B. Anonymizing Networks — Presentation Transcript

Different types of anonymous networks, how they work, the advantages and weaknesses of each anonymous communication network. Everyone's daily life, People access to Internet to do business, to find job, to contact friends, to pay bills. Internet has become another utility like water and electricity, which plays more and more significant role in every day life. With the impact of Internet on society, people became more sensitive regarding privacy issues in the Internet. They realized that they leave all kinds of traces and personal information while surfing websites and exchanging emails. In some cases, people do not want others know what they are talking. So, encryption like Pretty Good Privacy (PGP) was introduced, eavesdropping on content becoming very difficult. However, preserving privacy means not only the content of messages, but also hiding routing information which means who is talking to whom. Unfortunately, the Internet was not designed with anonymity in mind; in fact, one of the original design goals was accountability. IP packet which is one of the most important infrastructure protocols in network contains lots of fingerprint. Demand and interest in anonymous network has increased recently for many reasons. The material or its distribution is illegal or incriminating. Music and movie files sharing in peer-to-peer network applications, e. g. Kazaa, Bit Torrent. Material is legal but socially deplored, embarrassing or problematic in the individual's social world. For example, people may openly discuss personal stuff which would be embarrassing to tell many people about, such as sexual problems. Fear of retribution. (Whistleblowers, unofficial leaks, and activists who do not believe in restrictions on information or knowledge), Censorship at the local, organizational, or national level. Cisco designed and deployed packet content filtering equipments in every ISP access points mainland China. The TCP connection will be reset if it contains susceptive domain name, IP address or even key words. Personal privacy preferences such as preventing tracking or data mining activities MAC and IP address can be used to identify one device. Furthermore these persistent addresses can be linked to physical persons, seriously compromising their privacy. People can use anonymity in different purpose, good or bad.

C. Back Ground Survey

Achieving anonymity in a network is very difficult. Encryptions are used to protect data's confidentiality, while anonymity means protect both data and participants in this communication.

Unfortunately, the Internet was not designed with anonymity in mind; in fact, one of the original design goals was accountability. In packet switching network, every IP packet contains a header to describe the packet itself: The header contains Identification—contains an integer that identifies the current datagram. This field is used to help piece together datagram fragments; Time-to-Live—maintains a counter that gradually decrements down to zero, at which point the datagram is discarded. This keeps packets from looping endlessly; Source Address—specifies the sending node; Destination Address—specifies the receiving node. What more is, there is plenty of useful information within packet for network analysers to identify communication between two parties. This information includes source port, destination port, sequence number, window size. So the first step to anonymize communication is to encrypt the data in the packet, change source IP address, modify port number and Time-to-Live value to hide the fingerprint of initiator. However, these methods are not enough to counter network traffic analysers. More sophisticated anonymous methods are desired. Terminology Based on previous papers in this field, researchers proposed a set of precise terminologies. These definitions might help researchers invents new word with same meaning. I am going to use these terminologies in later sections. Anonymity: Anonymity of a subject from an attacker's perspective means that the attacker cannot sufficiently identify the subject within a set of subjects, the anonymity set. Unlinkablity: Unlinkablity of two or more items of interest (IOIs, e.g., subjects, messages, actions, ...) from an attacker's perspective means that within the system (comprising these and possibly other items), the attacker cannot sufficiently distinguish whether these IOIs are related or not. Unobservability: Unobservability of an Item Of Interest (IOI) means undetectability of the IOI against all subjects uninvolved in it and anonymity of the subject (s) involved in the IOI even against the other subject (s) involved in that IOI.

D. Pseudonymity

Pseudonymity is the use of pseudonyms as identifiers. Taxonomy According to the architecture and usability, anonymity communication can be classified into four categories: High latency, Low latency, Central Email relay, Web proxy Distributed and N/A Tarzan/Tor Pseudo-distributed. Central/High latency: There is a central server that provides anonymity service to clients, for example email relay service like anon.penet.fi and MixMaster. Central/Low latency: Clients can send requests to the central server, the server modify the packet and resend these requests to destinations. For instance, Anonymizer and Safe Web are such type of service. Pseudo-Distributed/High Latency and Distributed/ High Latency: Due to the volatile of distributed networks and interaction like AJAX and Flash between user and server is desired nowadays, we are not interested in these categories.

D. Pseudo-Distributed/Low Latency

Clients have to download network structure information from well known servers to start anonymity communication. One famous example is The Onion Router. Distributed/High latency: There is no central server to store information of anonymous network. Every node within network is equal to others. Tarzan and Morph Mix are such type of implementation. Basically, high latency systems provide stronger anonymity, but are unsuitably slow for Internet browsing and most types of Internet interactive applications. Centralized architecture is more vulnerable than distributed system; because of it is single point of failure. In this paper, I am going to introduce all of these categories, but focus on the analysis and comparison of implementations of low latency with pseudo-distributed and distributed architecture. Threat model and limitation Adversary may launch different types of attack towards anonymous network. They are Message coding attack: In this attack, messages that do not change their coding can by traced through the network by pattern matching. . , Message length attack: This attack examines the length of a message as it travels through the network. Replay attack: An attacker replays data packets and waits that target node processes the same packet repeatedly, therefore enabling the attacker to correlate incoming and outgoing packets.

E. Collusion Attack

This happens if a certain number of involved parties collide to break the anonymity of connections. Flooding attack: Anonymity is usually achieved with respect to a certain group. In this attack, an adversary floods the system to separate certain messages from the group. Message volume attack: In this attack, it is tried to detect an end-to-end connection by observing the message volume at the endpoints. Timing attack: A timing attack tries to observe the duration of a connection by correlating its establishment or release at the possible endpoints.

F. Profiling Attack

1. A profiling attack tracks users over long-term periods. It is basically a combination of the timing and message volume attack over a long time. First two types of attack can be prevented from re-encrypting message when transmitting between nodes. Padding and chunk methods can be applied to make all messages flatten as well. Maintain a temporary data base to manage processed message can be used to prevent replay attack. Decentralized network architecture can effectively prevent Flooding attack. Message volume, timing and profiling attack are traffic analysis attack from network wide scale. Researchers has proved that no matter how good the anonymity performed by the network, persistent communication between two users will eventually be detected just by observing the edges of the network and correlating the activity at the two ends of the communication. As long as attackers are selected uniformly at random to be a part of active set and sessions can be identified across path reformations, the degree of anonymity of any sender will degrade under attack. However, to achieve network wide scale analysis attack or black box attack is costly.

2. It is almost impractical to spend a large amount of time to spot the relation between sender and receiver in Internet scale which contains millions of hosts and interconnected globally. Based on this assumption, anonymize network is possible. A lot of implementation has been developed recently. In next section, I am going to introduce and analysis these different implementations.

3. Centralized approaches Anonymizer and Safe Web Anonymizer is an Internet privacy company, founded in 1995. Anonymizer offers variety of services include proxy server, encrypted email and so on. Anonymizer access Internet on behalf of real user. Some dynamic content like JavaScript, Java applets and Flash are filtered out, since data exchange which may lead to information leakage are needed for these dynamic applications. The Safe Web is another web proxy service similar to Anonymizer. The main difference is Safe Web provides SSL encryption between proxy and clients. Obviously, both two services are out date in the modern web environment which is full of interactions like AJAX and Flash. 4

5. From security point of view, centralized service is weak as well. First of all, we should have a doubt how can we trust proxy service? Does your service neutral? How can you prove your service is not compromised when clients are accessing? Second weakness is adversary can identify requestor using the message volume attack. For example, adversary can analysis proxy's in and out traffic to match messages which client has the same packet size. Third, client mostly browses web pages through links within page. Adversary may extracts this information and trace users' pattern using machine learning technology. Fourth, centralized server is a single point of failure. Adversary may launch a denial of service attack to take down proxy. Or receiver can simply block all packets transmitted from proxy. Crowds Crowd is a proposed anonymity network that gives probable innocence in the face of a large number of attackers. Crowd is important as it introduced the concept of users blending into a crowd of computers, and many of the concepts used in later systems. The main idea behind Crowds anonymity protocol is to hide each user's communications by re-routing packets randomly within a group of users. User can register himself to a Crowd. All the other users are notified then. The request message from initiator can be sent out from randomly selected node within Crowds. From outside world, it is difficult to tell who the real initiator is. For example, node 5 registers itself and retrieves information from server to join a group. Node 5 routes request to randomly selected node 3. Node 3 randomly decides relay this message or send out. In this case, node 3 decides to relay to node 1. Node 1 decided to route message to node 2. Finally, node 2 sends out the request on behalf of node 5. The advantage of Crowd over Anonymizer is even if one or a few number of nodes are subverted, the system can still provide some degree of anonymity. The assumption of Crowd is adversary can not eavesdrop on the traffic within group. Since, attacker can count on the time of a particular message go through each node. The node which has highest rank could be initiator. The registration server is a single point of failure as well. Distributed approaches Chaum's MIX David Chaum is a famous cryptographer. He not only invented Blind digital signature protocols, but also introduced a type of anonymous network—MIX in 1981.A lot of modern implementations of anonymous network are based on MIX model.

6. The main idea is to use a network of proxy servers. First, initiator randomly chooses a path within MIX to route message. Based on the selected proxy nodes on path, initiator encrypts the message using corresponding nodes' RSA public key. When routing, each proxy can decrypt the most out layer of the message. Finally, the message is sent out to destination server until all layers are striped out by corresponding nodes. The return message can route back to initiator in reverse way. In MIX, each node only has idea about its predecessor and successor. Actually, not only sender anonymity but also receiver anonymity is proposed by Chaum. The hidden receiver can wrap his address in RSA public keys as the same way we did. The wrapped packet can be distributed to users who wan to request this hidden service. User can use this wrapped packet as header of his own message to send. I will discuss this in later section. To improve anonymity, junk can be added to message while routing to hide its fingerprint and make all messages in same size. To make adversary more difficult to trace message, each node may delay message, and send messages in batch. Node can send dummy messages randomly to make eavesdropper confuse. However, there are still several weaknesses have been identified in the classic MIX network. MIX provides low level of protection on unlink ability. The nature property of RSA brings such weakness. Active attacker can inject a duplicated message into MIX node to find out the route which contains two same messages. Since MIX nodes have no idea about the payload in each message, they can not detect active attack. A lot of modern anonymity network systems are based on the idea of MIX, such as onion routing, Mixminion, Tarzan… In following sections, I am going to analysis and compare these different implementations. The Onion Router (Tor) Tor is the implementation of second generation onion routing based on classic MIX. It was originally sponsored by US Naval Research Laboratory, and then became a project of Electronic Frontier Foundation. Theoretically, Tor is not a fully distributed system. Since there are some central servers take the responsibility to maintain the topology of the anonymous network. Tor works like Napster. Architecture and Design Tor is a circuit-based low-latency anonymous communication service. Tor software acts as an agent on user's computer to relay http requests and responses. In this case, Tor can be used to go through firewall. For instance. First, the request will trigger Tor's local proxy software to fetch a table of current state and addresses of active onion routers from directory service.

7. In the second step, local proxy randomly selects three onion routers as a path to route the request, like classic MIX did. In this case, the route is Alice OR1 OR2 OR6 which is fully encrypted. OR6 sends the plain text request to WS1. From WS1's point of view, the message is sent from OR6, not Alice. To send back respond packet. OR6 encrypt the packet using its private key and send the encrypted packet to its predecessor OR2. Again, OR2 encrypts this encrypted message using its private key and send it to its predecessor OR1. Alice can decrypt this packet using corresponding public keys. Tor also provides "hidden server". Hidden service could allow Tor users to set up a website where people publish material without worrying about censorship. 1. Server picks some introduction points and builds circuits to them 2. Server advertises his hidden service "XYZ.onion" at the DB 3. Alice hears "XYZ.onion" exists, and she requests information from DB 4. Alice writes a message with rendezvous point to hidden server through introduction point 5. Alice and hidden server validate one-time secret. Tor circuits established between Alice and hidden server nobody would be able to determine who was offering the site, and nobody who offered the site would know who is visiting. To provide hidden service, an external server is desired to store some information of the service. Clients can retrieve this information from server to send request. Security Analysis The directory server acts as an http server, clients can fetch current network state and router list. Currently there are 3 directory servers globally and maintain a uniform directory list. The node who wants to be an onion router and be added to OR list in directory server has to be approved by directory server administrator. The ORs should periodically update state information and validate themselves using keys with directory servers. Directory is cached similar to Tapestry mechanism in each Onion Router to avoid the performance bottle neck. The network topology is relatively static. Adversary can take down key onion routers and directory servers, or eavesdrop on them. Dummy messages are used.

8. Cover network traffic, but dummy messages bring too much over head. Since every node participated node has to send out dummy messages. Many Internet applications directly send out DNS requests without invoke Tor's local proxy. This may leak some clues to eavesdropper, so usually Tor is companied with Privacy. After 0.2.0.1-alpha, Tor has implemented its own DNS service. Peer to peer anonymous network are proposed to solve weaknesses in centralized anonymity network. Tarzan and Morph

Mix Tarzan is a peer-to-peer MIX network. It was introduced by researchers from MIT and New York University in 2002. Tarzan provides anonymous service at transport level, so applications build on the top can transparently use anonymous communication. Morph Mix is an implementation similar to Tarzan Architecture and Design .The main difference between Tarzan and Tor is no directory server and core routers at all in Tarzan. All participants within the network are equal. To join the system, new node must contact at lest one already known node within the system to retrieve its neighbour nodes list. Then, directly ping these neighbours to validate and retrieve more neighbours until a threshold is met. To find new neighbours, gossip algorithm is used. When transmitting, the node needs to select some nodes. Since malicious nodes are often in same IP space. So, instead of randomly select node, Tarzan selects nodes based on the assumption that malicious nodes are in one IP subnet mainly rather than distributed world wide. Tarzan uses a three-level hierarchy: first among all known /16 subnets, then among /24 subnets belonging to this 16-bit address space, then among the relevant IP addresses. The traffic mimic is introduced in Tarzan network to protect against information leakage. Every node asks some nodes within network to exchange traffic mimics. Meaningful messages can be inserted into these cover traffics.

II. Existing System

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.

III. Proposed System

We present a secure system called Nymble, which provides all the following properties: anonymous authentication, backward unlinkability, subjective blacklisting, fast authentication speeds, rate-limited anonymous connections, revocation auditability (where users can verify whether they have been blacklisted), and also addresses the Sybil attack to make its deployment practical. In Nymble, users acquire an ordered collection of nymbles, a special type of pseudonym, to connect to websites. Without additional information, these nymbles are computationally hard to link, and hence using the stream of nymbles simulates anonymous access to services.

Websites, however, can blacklist users by obtaining a seed for a particular Nymble, allowing them to link future nymbles from the same user — those used before the complaints remain unlikable. Servers can therefore blacklist anonymous users without knowledge of their IP addresses while allowing behaving users to connect anonymously. Our system ensures that users are aware of their blacklist status before they present a Nymble, and disconnect immediately if they are blacklisted. Although our work applies to anonymizing networks in general, we consider Tor for purposes of exposition. In fact, any number of anonymizing networks can rely on the same Nymble system, blacklisting anonymous users regardless of their anonymizing network(s) of choice. 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. Interest in anonymous P2P systems has increased in recent years for many reasons, ranging from the desire to share files without revealing one's network identity and risking litigation to distrust in governments, concerns over mass surveillance and data retention, and lawsuits against bloggers.

IV. Motivation for Anonymity

There are many reasons to use anonymous P2P technology; most of them are generic to all forms of online anonymity. P2P users who desire anonymity usually do so as they do not wish to be identified as a publisher (sender), or reader (receiver), of information. Common reasons include:

1. The material or its distribution is illegal or incriminating.
2. Material is legal but socially deplored, embarrassing or problematic in the individual's social world (for example, anonymity is seen as a key requirement for organizations like Alcoholics Anonymous).
3. Fear of retribution (against whistleblowers, unofficial leaks, and activists who do not believe in restrictions on information or knowledge).
4. Censorship at the local, organizational, or national level.
5. Personal privacy preferences such as preventing tracking or data mining activities.

A particularly open view on legal and illegal content is given in the philosophy Behind Freenet. Governments are also interested in anonymous P2P technology. The United States Navy funded the original onion routing research that led to the development of the Tor network, which was later funded by the Electronic Frontier Foundation and is now developed by the non-profit organization The Tor Project, Inc.

V. Anonymous Blogging

Anonymous blogging is one widespread use of anonymous networks. While anonymous blogging is possible on the non-anonymous internet to some degree too, a provider hosting the blog in question might be forced to disclose the blogger's IP address (as when Google revealed an anonymous blogger's identity). Anonymous networks provide a better degree of anonymity. Flogs in Freenet, Syndie and other blogging tools in 12 Pand Osiris sps are some examples of anonymous blogging technologies. One argument for anonymous blogging is a delicate nature of work situation. Sometimes a blogger writing under his/her real name faces a choice between either staying silent or causing harm to himself, his colleagues or the company he works for another reason is risk of lawsuits. Some bloggers have faced multi-million dollar lawsuits that were later dropped completely; anonymous blogging provides protection against such risks.

VI. Conclusions

A comprehensive credential system called Nymble, which can be used to add a layer of accountability to any publicly known anonymizing network, Serves can blacklist misbehaving users while maintaining their privacy, and we show how these properties can be attained in a way that is practical, efficient and sensitive to the needs of both users and services.

References

[1] Electronic Frontier Foundation (2005), Retrieved March 5, 2008.
[2] Julien Pain, editor (2005), Retrieved January 23, 2008.
[3] Russell D. Hoffmann (1996),"Transcript of a radio interview".
[4] John Pilger (2002). ZNet article, retrieved 2008.

[5] A. Baggio,"Wireless sensor networks in precision agriculture", in ACM Workshop on Real-World Wireless Sensor Networks (REALWSN 2005), Stockholm, Sweden, June 2005.
[6] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, J. Anderson,"Wireless Sensor Networks for Habitat Monitoring", in First ACM Workshop on Wireless Sensor Networks and Applications, Atlanta, GA, USA, September 2002.
[7] G. Fuchs, S. Truchat, F. Dressler,"Distributed Software Management in Sensor Networks using Profiling Techniques", in 1st IEEE/ACM International Conference on Communication System Software and Middleware (IEEE COMSWARE 2006): 1st International Workshop on Software for Sensor Networks (SensorWare 2006), New Dehli, India, January 2006.

Mrs.K.RAJANI KUMARI, well known Author and excellent teacher Received M.Tech (CSE) from Andhra University, Visakhapatnam. She is working as a H.O.D for the Department of M.Tech., Computer Science & Engineering, Akula Sree Ramulu College of Engineering & Technology, Tanuku. She has vast teaching experience in various engineering colleges. To her credit couple of publications both National&International conferences /journals. Her area of Interest includes Data Warehouse and Data Mining, information security, Data Communications & Networks and other advances in Computer Applications. She has guided many projects for Engineering Students.

Mrs.G.Naga Mallika is a student of A.S.R College of Engineering & Technology, Tanuku. Presently she is pursuing her M.Tech (C.S.E) from this college and she received her Master of Computer Applications (M.C.A) from IGNOU University, New Delhi in the year 2009. Her area of interest includes Computer Networks and Object oriented Programming languages in Computer Applications.

Mr.T.Rajesh, an efficient teacher, received M.Tech (CSE) from JNTU Kakinada is working as an Assistant Professor in Department of C.S.E, A.S.R College of Engineering and Technology, Tanuku. He has 5 years of teaching experience. He has published many papers in both National & International Journals. His area of Interest includes Data Communications & Networks, Database Management Systems, C Programming and other advances in Computer Applications.