Using Traffic Analysis to Identify Tor Usage – A Proposed Study

Authors Name/s per 1st Affiliation (Author)
line 1 (of Affiliation): dept. name of organization
line 2: name of organization, acronyms acceptable
line 3: City, Country
line 4: e-mail address if desired

Authors Name/s per 2nd Affiliation (Author)
line 1 (of Affiliation): dept. name of organization
line 2: name of organization, acronyms acceptable
line 3: City, Country
line 4: e-mail address if desired

Traditional attacks against anonymous routing systems aim to uncover the identities of those involved, however, a more likely goal of attackers is to block or degrade the network itself, discouraging participation and forcing vulnerable users to communicate using less secure means. Since these anonymous networks operate on known protocols and employ strong encryption it is difficult to distinguish them from regular traffic. This paper proposes a method for identifying traffic belonging to anonymous networks by examining their communication patterns. If successful the method would enable the identification of Tor usage and thus allow for more directed attacks and possible user identification.

Keywords: traffic analysis, low-latency anonymity, forensics, machine learning, heuristics, traffic classification, censorship, interception

I. INTRODUCTION

The Internet, in providing an inexpensive broadcast medium for the publication of political information, criticism or challenging ideas, has revolutionised public debate [1]. Despite the potential of the platform, the publication of less desirable political speech, criticism or challenging ideas still carries with it great risk, with many incidents leading to well publicised arrest [2][3][4]. Beyond the level of prosecution, comes the threat of censorship by government and malicious attackers [5][6][7][8].

In response to the rising level of threat, a number of systems have been proposed which use cryptography to provide censorship resistance and anonymity [8][9]. One of the major players in this arena is the second generation Onion Router commonly referred to as 'Tor'. Tor utilises onion routing to provide low latency anonymity to network participants, whilst also providing a level of resistance to censorship based blocking or filtering [10].

Despite such measures, it is postulated that through traffic analysis techniques an attacker may be able to identify Tor connections and thus target them for further attack. Such an attack may discourage usage of the Tor network as it is reliant in part on large numbers of users for the protections it provides [10]. This paper proposes a statistical methodology for identifying Tor traffic which may be utilised as a basis for such attacks.

II. ANONYMOUS ROUTING

The first anonymous network system was proposed by Chaum who included the use of public key cryptography and a centrally located server known as a 'mix' [11]. The theory behind the Chaum approach was that by wrapping the information and address of a destination in a message and encrypting it with the mix's public key, that only the mix system is able to read the message [11]. Eventually, the base of these ideas lead to the process of Onion routing which was first described in a patent in 2001.

As with the earlier 'mix' approach, onion routing utilises public key cryptography to provide anonymity, however its major advancement was in the area of packet routing. The onion method requires that messages are repeatedly encrypted then passed through several nodes in a network known as onion routers. Each of these routers is capable of removing a single layer of encryption to reveal the next set of delivery instructions. The layering based approach acts to prevent individual nodes obtaining the not only the message itself but also the identity of the sender and intended recipient. In this model each internal node is not aware if it is connected to the original requester or is merely another routing node, nor is any internal node able to determine the content of the message.

The Second Generation Onion Router (Tor) improves on the initial design by utilising a form of Onion routing that employs telescoping circuits to provide low latency anonymity [9]. A connection through the Tor network constitutes a circuit, which hops over several Tor nodes known as relays. Relays that allow connections directly from Tor clients are known as bridges and relays that deliver messages from the frontier of the Tor network are known as exit nodes. When a relay joins the Tor network, it trades symmetric cipher keys with its neighbours using public key cryptography. These keys are then used to encrypt communications and are destroyed when the session is closed to prevent replay attacks. These processes are illustrated below in which a new client is shown connecting to the Tor network.
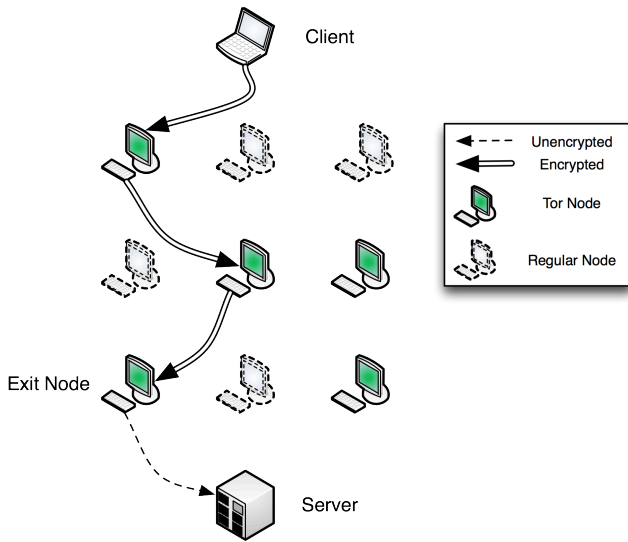
Figure 1.   Connecting through the Tor network

III.   RELATED WORK

The theory behind using traffic analysis to identify an individual type of network traffic has been used previously within a range of contexts. Zhang categorises these efforts into three primary techniques namely exact matching, heuristic matching and machine learning [12]. In the case of Tor, the exact matching techniques are unworkable as the network can operate on any available port and utilises encryption to obfuscate communications. This leaves heuristic and machine learning matching techniques for consideration as methods for Tor traffic identification.

A. Heuristic Techniques

Heuristic matching techniques look for patterns in conversations between nodes to make inference about their relationships or particular roles. The increasing burden of peer to peer (P2P) applications and their shift towards encryption motivated the development of heuristics based techniques for identifying P2P traffic. Early techniques included identifying known properties of P2P networks such as the simultaneous use of both UDP and TCP protocols and the utilisation of a solitary connection to transfer high volumes [13].

This approach has been refined in the works of Perenyi [14] and John & Tafvelin [15], both of which attempt to improve matching accuracy by expanding the scope of matching parameters and eliminating false positives using exact matching techniques. Such methods have also been successfully applied to identify traffic belonging to network worms [16] and the Skype protocol [17].

B. Machine Learning Techniques

Machine learning techniques encompass algorithms that evolve behaviours by training against sets of empirical behaviour, this includes the classification of network communications by observing variables such as packet size and inter packet arrival time. There is a significant volume of work related to traffic analysis utilising machine learning techniques for classification, demonstrating a wide variety of algorithms for this purpose. The first use of machine learning to categorise traffic flows appears in McGregor et al. [18]. A detailed analysis of the attributes that can be used for machine learning and an attempt at coarse grained classification using an Expectation-Maximisation (EM) algorithm are demonstrated. The same technique is also employed in Soule et al. using histograms for finer grained classification [19]. The EM algorithm is again used in Zander et al. [20] and Erman et al. [21], with the latter also comparing this algorithm favourably against a Naive Bayes classifier.

Moore and Zuev's work demonstrates the usage of a supervised Naive Bayes algorithm to classify traffic flows [22]. This paper focuses on many of the most commonly used Internet protocols while Bonfiglio et al. uses the technique for identifying traffic belonging to the commercial Skype application [17].

Herrmann and Wendolsky utilise Bayesian networks to fingerprint visited websites accessed through Privacy Enhancing Technologies (PET), including Tor [23]. This technique performed poorly when applied to the Tor network, but it suggests that Tor traffic has particular characteristics that distinguish it from many existing PETs. It makes a particularly useful observation: "The most frequent packet sizes in the Tor traffic dumps are, in descending order, 1500, 52, 638, 52, 638 and 1150 bytes, accounting for 87.6% of all Tor packets."

Hidden Markov Models (HMMs) are first used as a traffic analysis technique in HMM profiles for network traffic classification [24]. The primary identification characteristics for use with this algorithm are packet size and inter-packet arrival times. With refinement this algorithm is used with increasing accuracy in Wright et al. [24] and Dainotti et al. [25]. HMMs are also used in Bernaille et al. to discover distinguishing characteristics of traffic flows, rather than specifically as a classifier [26].

Clustering algorithms group observations into subsets based on similar characteristics. They have been used in a number of traffic classification techniques with the K-Means clustering technique being the most prominent. K-Means cluster analysis appears in Bernaille et al. [26] and Erman et al. [27][28]. The use of the Density Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm appears alongside K-Means and Autoscan algorithms [29].

Other algorithms used for traffic classification include Nearest Neighbour and Linear Discriminant Analysis [30], Normalized Threshold [31]. The use of the Gaussian Mixture Model to identify applications and identities inside SSH tunnels was demonstrated later in Dusi et al.[32].

C. Comparing Techniques

It is difficult to say which technique would be most suitable for identifying Tor traffic. In the literature, no particular consensus has been reached and no two papers use the same baseline data set for their technique and so no comparison can be made [33]. Although there has been some comparison of machine learning algorithms [34], no paper has published their data set due to privacy concerns. It is clear that choosing the right technique will require some experimentation and comparison of the effectiveness of each algorithm.

IV.   PROPOSED APPROACH

The proposed experiment will employ a quasi-experimental research method with empirical learning elements, employing techniques common to traffic analysis papers. This includes the capturing of packet traces, development of a matching technique/algorithm and documentation of efficacy and results. The first phase of the experiment covers the data capturing phase, while the second phase concludes the traffic analysis technique.

A.   Assumptions

It is assumed that the usage patterns exhibited by individual users will be smaller than the communications characteristics that will lead to the identification of anonymous and censorship resistant networks. Thus there is no need to obtain a large sample of regular network traffic from varying user profiles.

As of the current implementation, Tor network traffic is readily distinguishable by looking at the handshake packets. It is likely that this weakness will be addressed in a future version of the Tor protocol as it is recognised as a design goal in Dingledine and Mathewson [10]. For this reason, this proposal focuses on traffic analysis techniques that are content agnostic.

B.   Variables

The data capturing stage will be affected by a number of variables that will influence the accuracy of the chosen matching algorithms. These include:

- System Performance
- Network Performance
- Application Protocol
- Quality of the Anonymous Network
- Caching

An attempt will be made to reduce the impact of these variables by capturing packets generated by an isolated simulation of the Tor network, ensuring only vital network applications are running and executing software inside a virtual machine which can be rolled back to a pre-established control state. In this manner the signal to noise ratio and performance of the test network will be controlled in order to ensure minimum influence from identified impacting variables.

C.   Materials and Method

The experiment will consist of two phases, first a simulation phase which involves the capture of packets belonging to a Tor network, the second will be a comparison and documentation of the results of using several techniques to identify Tor traffic. Fig. 2 shows the placement of applications and relevant data flows for the simulation phase.
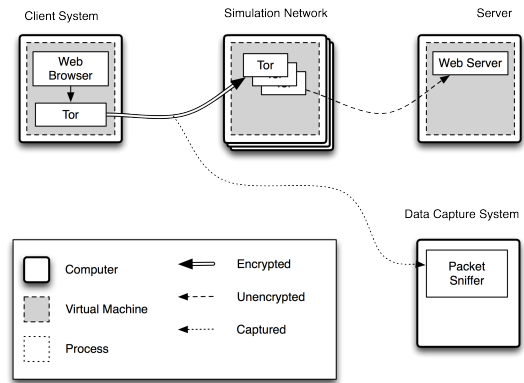

Figure 2.   Proposed expiremental setup

D.   Data Collection & Algorithm Selection

The data collection phase will be conducted by running a series of simulations designed to mimic real world traffic. These simulations will be controlled so that the data captured will only contain Tor traffic in the training and testing data sets, and non Tor encrypted traffic in the data set for determining the false positive rate.

Once the data collection phase is complete, a number of different algorithms will be chosen to classify Tor traffic. These will be compared using a number of criteria to rate effectiveness. These will include the accuracy of the matching technique and the number of packets needed to make a confident identification.

V.   CONCLUSION & ONGOING RESEARCH

The effectiveness of anonymous routers relies on their ability to provide covering traffic through a network of participants. Tor achieves this effectively through a number of design goals intended to encourage participation, as well as paying attention to non technical issues which affect its adoption.

The proposed research seeks to identify potential flaws in the nature of anonymous routing systems, by demonstrating that the characteristics that provide effective anonymity also make these anonymous protocols easily distinguished from regular protocols. An anonymous network that is easily identifiable is more easily attacked by those who wish to discourage participation, or eliminate anonymity.

A number of techniques both statistical and heuristic will be evaluated for the purpose of classification of Tor routing flows. The results from this analysis could serve as a basis for enhancing the security of Tor and other anonymous networks.

REFERENCES

[1]   M. Bonchek, "From Broadcast to Netcast: The Internet and the Flow of Political Information," *en.scientificcommons.org,* Jan 1 1997.

[2]   C. Coonan, "Chinese couple sue Yahoo! in US over torture case, The Independent," The Independent, 2007.

[3]   A. Howard, "Blogger Arrests," World Information Access, 2008.

[4]   J. Rutwitch, "Vietnam bloggers arrested over China shirt protest," Reuters, 2009..

[5]   C. Bolan and P. Hannay, "Freegate: A Defence against the Pending Censorship of Australia?," in *Third International Symposium on Human Aspects of Information Security & Assurance*, Piraeus, Greece, 2009.

[6] J. Crandall, D. Zinn, M. Byrd, E. Barr, and R. East, "Conceptdoppler: A weather tracker for internet censorship," *14th ACM Conference on Computer and Communication Security (CCS),* 2007.

[7] J. Karlin, S. Forrest, and J. Rexford, "Nation-State Routing: Censorship, Wiretapping, and BGP," *Arxiv preprint arXiv:0903.3218,* 2009.

[8] J. Smart, K. Tedeschi, D. Meakins, P. Hannay, and C. Bolan, "Subverting National Internet Censorship - An Investigation into existing Tools and Techniques," in *The 6th Australian Digital Forensics Conference*, Perth, Western Australia, 2008.

[9] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," *Proceedings of the 13th conference on USENIX Security Symposium-Volume 13,* p. 21, 2004.

[10] R. Dingledine and N. Mathewson, "Design of a blocking-resistant anonymity system," *Citeseer,* Jan 1 2008.

[11] D. Chaum, "Untraceable electronic mail, return addresses, and digital pseudonyms," 1981.

[12] M. Zhang, "State if the art in traffic classification: A research review," in *PAM Student Workshop*, 2009.

[13] T. Karagiannis, A. Broido, and M. Faloutsos, "Transport layer identification of P2P traffic," *Proceedings of the 4th ACM SIGCOMM conference on Internet measurement,* pp. 121-134, 2004.

[14] M. Perenyi, "Identification and Analysis of peer-to-peer traffic," *Journal of Communications,* vol. 1, 2006.

[15] W. John and S. Tafvelin, "Heuristics to classify internet backbone traffic based on connection patterns," *Information Networking,* 2008.

[16] A. Lazarevic, L. Ertoz, V. Kumar, A. Ozgur, and J. Srivastava, "A comparative study of anomaly detection schemes in network intrusion detection," *Proceedings of the Third SIAM International Conference on Data Mining,* pp. 25–36, 2003.

[17] D. Bonfiglio, M. Mellia, M. Meo, D. Rossi, and P. Tofanelli, "Revealing skype traffic: when randomness plays with you," *ACM SIGCOMM Computer Communication Review,* vol. 37, p. 48, 2007.

[18] A. McGregor, M. Hall, P. Lorier, and J. Brunskill, "Flow clustering using machine learning techniques," *Lecture Notes in Computer Science,* vol. 3015, pp. 205-214, 2004.

[19] A. Soule, K. Salamatia, N. Taft, R. Emilion, and K. Papagiannaki, "Flow classification by histograms: or how to go on safari in the internet," *Proceedings of the joint international conference on Measurement and modeling of computer systems,* pp. 49-60, 2004.

[20] S. Zander, T. Nguyen, and G. Armitage, "Self-learning IP traffic classification based on statistical flow characteristics," *Passive and Active Network Measurement,* pp. 325-328, 2005.

[21] J. Erman, M. Arlitt, and A. Mahanti, "Traffic classification using clustering algorithms," *Proceedings of the 2006 SIGCOMM workshop on Mining network data,* p. 286, 2006.

[22] A. Moore and D. Zuev, "Internet traffic classification using bayesian analysis techniques," *ACM SIGMETRICS Performance Evaluation Review,* vol. 33, pp. 50-60, 2005.

[23] D. Herrmann and R. Wendolsky, "Website Fingerprinting: Attacking Popular Privacy Enhancing Technologies with the Multinomial Naïve-Bayes Classifier," *Proceedings of the 2009 ...,* Jan 1 2009.

[24] C. Wright, F. Monrose, and G. Masson, "HMM Profiles for Network Traffic Classification (Extended Abstract)," VizSEC/DMSEC'04: proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security, Washington, DC, USA, October 29, 2004, co-located with CCS 2004, p. 9, 2004.

[25] A. Dainotti, W. de Donato, A. Pescape, and P. Rossi, "Classification of Network Traffic via Packet-Level Hidden Markov Models," *Proc. of IEEE Global Telecommunications Conference (GLOBECOM 2008),* 2008.

[26] L. Bernaille, R. Teixeira, I. Akodkenou, A. Soule, and K. Salamatian, "Traffic classification on the fly," *ACM SIGCOMM Computer Communication Review,* vol. 36, p. 26, 2006.

[27] J. Erman, A. Mahanti, M. Arlitt, and C. Williamson, "Identifying and discriminating between web and peer-to-peer traffic in the network core," *Proceedings of the 16th international conference on World Wide Web,* p. 892, 2007.

[28] J. Erman, A. Mahanti, M. Arlitt, I. Cohen, and C. Williamson, "Offline/realtime traffic classification using semi-supervised learning," *Performance Evaluation,* vol. 64, pp. 1194-1213, 2007.

[29] J. Erman, A. Mahanti, and M. Arlitt, "Internet traffic identification using machine learning," *GLOBECOM,* 2006.

[30] M. Roughan, S. Sen, O. Spatscheck, and N. Duffield, "Class-of-service mapping for QoS: a statistical signature-based approach to IP traffic classification," *Proceedings of the 4th ACM SIGCOMM conference on Internet measurement,* pp. 135-148, 2004.

[31] M. Crotti, M. Dusi, F. Gringoli, and L. Salgarelli, "Traffic classification through simple statistical fingerprinting," *ACM SIGCOMM Computer Communication Review,* vol. 37, p. 16, 2007.

[32] M. Dusi, F. Gringoli, and L. Salgarelli, "A Preliminary Look at the Privacy of SSH Tunnels," *Computer Communications and Networks,* pp. 1–7, 2008.

[33] H. Kim, U. CAIDA, D. Barman, and M. Faloutsos, "Comparison of Internet Traffic Classification Tools," *ANF Workshop,* vol. 2, 2007.

[34] A. Mohd, "Towards a Flow-based Internet Traffic Classification for Bandwidth Optimization," *International Journal of Computer Science and Security* vol. 3, p. 146, 2009.