Análise de Segurança para a Descoberta, Bloqueio e Rastreamento de Tráfego Malicioso sobre a Rede Tor

Marcelo I. P Salas, Paulo Lício de Geus

[1]Laboratory of Security and Cryptography (LASCA), Institute of Computing
University of Campinas (UNICAMP)
Av. Albert Einstein, 1251, Room 84, Campinas, Brazil - CEP 13083-852

{mpalma, paulo}@lasca.ic.unicamp.br

Abstract. *Tor is an overlay network that provides anonymous communication on the Internet for TCP applications. This network serves hundreds of thousands of users, allowing them to decide when they wish to be identified or not. Although this network is substantially used to circumvent Internet censorship in countries under dictatorial regimes, the anonymity offered by the Tor network also supports, in a way, access to hidden services (e.g. Silk Road 2.0) for selling drugs, pedophilia, trafficking people, among others. This service that ensures privacy also hides a whole new side of violence, allowing botnets to go undercover, send SPAM, perform distributed denial of service attacks (DDoS), among other cybercrimes. This project proposes the research of methods and techniques to detect and classify malicious traffic in order to block potential threats and the development of techniques to trace the origin of malicious code over the Tor network. The goal is to design and implement a solution to the growing problem of malicious traffic on this network, researching forensic techniques and methods to try and protect the Tor network from malicious traffic, whilst also trying to preserve the privacy and anonymity of non-malicious traffic.*

Resumo. *Tor é uma rede de sobreposição que fornece comunicação anónima na Internet para aplicações TCP. Esta rede atende centenas de milhares de usuários, permitindo-lhes decidir quando desejam identificar-se ou não. Apesar de ser substancialmente utilizada para contornar censura na Internet em países sob regimes ditatoriais, esta rede de anonimato dá suporte, de certo modo, ao acesso a serviços ocultos (por exemplo Silk Road 2.0) fornecendo vendas de drogas, pedofilia, tráfico de pessoas, entre outros. O serviço que garante a privacidade também esconde, por trás todo um lado oculto de violência, possibilitando a proteção de botnets, envio de SPAM, ataques distribuídos de negação de serviço (DDoS), entre outros cibercrimes. Neste contexto, o presente projeto de pesquisa propõe o estudo de métodos e técnicas para detecção e classificação de tráfego malicioso, bloqueio de possíveis ameaças e desenvolvimento de técnicas de rastreamento da origem do código malicioso sobre a rede Tor. O objetivo é projetar e implementar uma solução ao crescente tráfego de código malicioso sobre esta rede, pesquisando técnicas forenses e métodos para proteger a rede Tor do tráfego malicioso, preservando a privacidade e anonimato do tráfego não malicioso.*

1. Introdução

Tor (anteriormente um acrônimo para *The Onion Router*) é uma rede de sobreposição que fornece comunicação anônima na Internet para aplicações TCP [Ling et al. 2015a]. De código aberto, esta rede atende centenas de milhares de usuários, transportando terabytes de informação cifrada, permitindo-lhes decidir quando desejam identificar-se ou não, evitando rastreamento dos seus dados online e protegendo a privacidade das suas atividades online contra tentativas de adversários de encontrá-los e destruí-los [Simons 2014].

Com mais de 6.700 servidores [Metrics 2015], a rede Tor é propensa a transportar mais de 30 vezes o tráfego malicioso em comparação com servidores que não são parte desta rede [Akamai]. Assim, o dinamismo de Tor torna a tarefa de bloquear o tráfego malicioso em um trabalho complexo para os voluntários [Ling et al. 2015a]. Este problema abre a possibilidade que os voluntários sejam legalmente processados pelo tráfego que circula por seus roteadores.

Infelizmente, os atacantes estão utilizando Tor por causa da sua proteção da privacidade nas comunicações, obtido como descrito a seguir. Através de uma aplicação, Tor seleciona, geralmente, 3 roteadores[1] da sua rede e constrói uma rota anônima utilizando um subconjunto desses roteadores. O tráfego entre o atacante e o destino é retransmitido ao longo desse percurso de forma cifrada e impossibilita o rastreamento, já que cada roteador (*onion Tor* ou *relay*) apenas conhece seu emissor anterior e receptor posterior das mensagens. Por último, o roteador de saída, atua como um *proxy*, comunicando-se diretamente com o destino de forma clara e anônima, sendo uma das poucas opções para os pesquisadores de analisar o tráfego de saída Tor.

Neste contexto, o presente projeto de pesquisa propõe uma análise de segurança para o estudo e implementação de métodos e técnicas para detecção e classificação de tráfego malicioso, bloqueio de possíveis ameaças e desenvolvimento de técnicas de rastreamento da origem do código malicioso sobre a rede Tor. O objetivo é projetar e implementar soluções ao crescente tráfego de código malicioso sobre esta rede, pesquisando técnicas forenses e ferramentas para protegê-la do tráfego malicioso, preservando a privacidade e anonimato dos usuários. A combinação de técnicas (p.ex. análise de padrão do tráfego malicioso e análise forense de ataques) e ferramentas (p.ex. IDSs[2] e analisadores de tráfego) coadjuvarão no desenvolvimento de uma plataforma para avaliar em tempo real possíveis ameaças circulando pela rede Tor, ajudarão no desenvolvimento de técnicas para a análise forense dos ataques e colaborarão na prevenção de futuras ameaças de segurança contra a Internet no Brasil e no mundo [Martins et al.].

Dadas as políticas da *Digital Millennium Copyright Act* (DMCA), que monitora constantemente os roteadores de saída Tor e envia notificações contra o compartilhamento de materiais com direitos autorias. Nossa arquitetura encaminhará o tráfego do roteador de saída através de um firewall para outro roteador de entrada (*guard onion*). Desta forma, haverá condições de bloquear tráfego capaz de possível imputação de responsabilidade legal sobre a universidade. Lamentavelmente, o reencaminhamento do tráfego gerará atraso na comunicação entre os usuários e os servidores, no entanto este é um custo quase

[1] A configuração padrão da rede Tor é composta por 3 roteadores: um de entrada (*entry guard*), outro de saída (*exit router*) e um roteador intermediário. Existem outras configurações que serão descritas nos capítulos seguintes.

[2] Intrusion Detection Systems, i.e. sistemas de detecção de intrusão

sempre presente ao se pesquisar tráfego malicioso em ambientes sensíveis.

Pretende-se enfatizar três frentes de pesquisa, envolvendo: **i) análise de tráfego malicioso em Tor**, o que inclui identificação, classificação e bloqueio deste tráfego utilizando técnicas e ferramentas; **ii) botnets**, com o objetivo de detectar, analisar e bloquear suas atividades maliciosas sobre a rede Tor, tais como DDoS, SPAM, roubo de informação e outros [Wang et al. 2015]; e **iii) Rastreamento de tráfego malicioso sobre Tor**, com o objetivo de tentar localizar os centros de C&C das botnets e outros servidores na *Deep Web*. O desenvolvimento da plataforma permitirá fazer análise estatística de tráfego malicioso sobre a rede Tor. Isto será possível pela utilização de ferramentas de IDS e análise de reconhecimento de padrões de tráfego para botnet, SPAM e outros. Além de mitigar tráfego malicioso, projeta-se implementar mecanismos para identificar novas ameaças, bloqueá-las e tentar estimar futuros cenários. De certo modo, isto permitirá categorizar o tráfego malicioso sobre a rede Tor.

2. Proposta

O propósito do projeto de pesquisa é melhorar a segurança de Tor e ajudar a reduzir a ciberdelinquência desta rede de anonimato e privacidade através de técnicas de análise de tráfego de um roteador de saída Tor.

Isto inclui a pesquisa para configurar uma rede de captura e reencaminhamento do tráfego, além dos sistemas de análise, monitoração, localização e proteção contra o tráfego malicioso. Para alcançar esse objetivo, propõe-se utilizar ferramentas como analisadores de tráfego, IDS e técnicas de aprendizado de máquina para capturar comportamento malicioso. Além disso, é necessário desenvolver uma ferramenta de identificação e bloqueio de tráfego malicioso [Cavalcante et al. 2014], que possa ser utilizada por voluntários da rede Tor. Enquanto esses recursos são compatíveis para os objetivos propostos, podem também ser de interesse para outras redes de privacidade e anonimato.

Esta seção apresenta as atividades já realizadas, referentes a cada uma das etapas citadas no capítulo anterior, ou seja, coleta, análise e classificação, e bloqueio e rastreamento de tráfego malicioso em Tor, seguindo nossa propostas descrita na Figura 1.

2.1. Coleta

A seguir, serão apresentados os métodos de coleta de tráfego da rede Tor utilizados. Estas estão classificadas por amostras de fontes públicas, fontes privadas e amostras coletadas por analisador de tráfego de rede.

Fontes Públicas. Foram obtidas amostras de tráfego malicioso e benigno de páginas Web, tais como "Contagiodump"[3], "Codeplex"[4], "Netresec"[5], "Pcapr"[6], "Packetlife"[7] e "wireshark"[8]. Estas páginas publicam novos exemplares continuamente. Neste caso, faremos o processo de obtenção deles contínuo durante o período do doutorado.

[3] http://contagiodump.blogspot.com.br

[4] http://sysdoccap.codeplex.com/wikipage?title=System\%20Overview\%20Document\%20Scenario\%20Captures

[5] http://www.netresec.com/?page=PcapFiles

[6] http://www.pcapr.net/home

[7] http://packetlife.net/captures/

[8] https://wiki.wireshark.org/SampleCaptures\#Sample_Captures

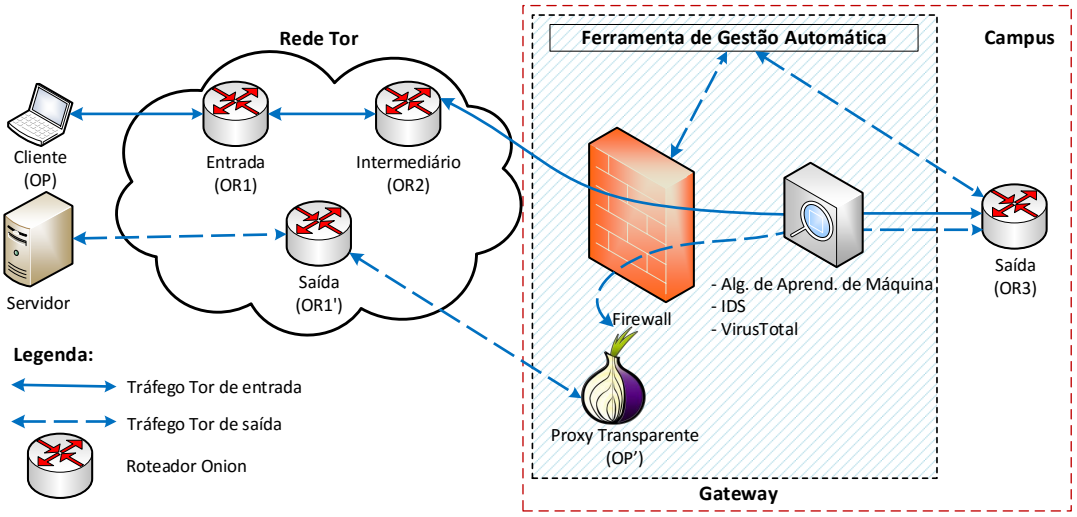

Figure 1. Arquitetura da plataforma de Coleta de Tráfego Malicioso para a rede Tor [Ling et al. 2015b].

Fontes Privadas. Compostas por uma combinação de amostras de tráfegos maliciosos e benignos, pertencentes ao banco de dados ISOT [Net 2011]. O tráfego malicioso inclui amostras de botnets Zeus, Storm e Waledac. Outras fontes de tráfego benigno foram obtidas de dois diferentes bancos de dados, um a partir do projeto *"Traffic Lab"* do centro de pesquisa Ericsson da Hungria [ICIR] e outro do *"Lawrence Berkeley National Lab (LBNL)"* [Orebaugh et al. 2006].

Coleta por Analisador de Tráfego de Rede. Será implementada uma interface de coleta de amostras de tráfego malicioso e benigno fazendo uso de um roteador conectado à rede Tor, de preferência para um roteador de saída (*exit router*).

2.2. Análise e Classificação

Foi desenvolvido um sistema de análise dinâmica, em ambiente emulado (honeypot), para analisar o padrão de comportamento dos *malware* nas plataformas Windows XP e 7. Para isso, foi selecionado um conjunto de 125 amostras de malware pertencentes a 21 famílias dentre vírus, cavalos de Troia e botnets, e as 125 aplicações mais baixadas do repositório CNET. Foram aplicados dois algoritmos de aprendizado de máquina, *Support Vector Machine* (SVM) e *Boosting* com Árvores de Decisão. Utilizou-se a técnica *Principal Component Analysis* (PCA) para evitar a Maldição da Dimensionalidade e eliminar as características que não contribuíssem para a separação das classes.

O software malicioso é caracterizado por um comportamento complexo e diversificado, que vai desde simples modificações de recursos do sistema até atividades vinculadas a redes externas. Variantes de malware de uma mesma família compartilham padrões de comportamento comuns, tais como o uso de mutações de sua carga útil, modificações de arquivos de sistema específicos, uso não típico de bibliotecas do sistema e acesso a redes externas, caracterizadas pelos seus endereços IP.

Assim, nosso objetivo foi explorar esses padrões comuns para análise automática com vistas à detecção de malware baseada no seu comportamento perante o sistema op-

eracional Windows, versões XP e 7.

A técnica de Boosting mostrou ser superior a SVM para este caso, alcançando melhores resultados tanto para o Windows 7 quanto para o Windows XP, com acurácias de até 92% e 94%, respectivamente.

Esta pesquisa demostrou que a utilização de poucas amostras combinadas com PCA para redução de dimensionalidade permite obter resultados satisfatórios e aumenta a credibilidade dos resultados da pesquisa. Além disso, é possível aplicar esta metodologia para análise de *malware* recuperado do tráfego malicioso da rede Tor. Também esperamos obter *malware* para outros plataformas, tais como Android.

2.3. Outras atividades

Os Testes de Segurança permitem avaliar as vulnerabilidades em aplicações e serviços frente a diversos tipos de ataques e descobrir novas vulnerabilidades antes que sejam exploradas por atacantes [Salas and Martins 2012]. Nesta pesquisa, estas técnicas permitiriam injetar diversos tipos de ataques contra os roteadores Tor e seus serviços ocultos, permitindo encontrar vulnerabilidades para poder rastrear serviços ilícitos da *Deep Web*, bloquear o tráfego malicioso das botnets ou injetar ataques nos roteadores de saída.

Entre as técnicas de Testes de Segurança que podem ser utilizadas estão: Testes de Penetração e Injeção de Falhas. Os Testes de Penetração (TP) emulam ataques, com o objetivo de revelar vulnerabilidades. Estes testes são automatizados pelo uso de ferramentas denominadas *vulnerability scanners* (VS). Já a Injeção de Falhas é uma técnica que pode ser utilizada para avaliar aspectos de dependabilidade dos sistemas de computação, podendo ser implementada em hardware ou software, para emular anomalias, defeitos ou erros no sistema alvo e observar seu comportamento sob um ambiente estressante.

A vantagem de usar Injeção de Falhas com Testes de Penetração é que a primeira permite maior cobertura de ataques. Neste contexto, as falhas são introduzidas por um injetor—software responsável por injetar falhas no sistema—antes ou durante a execução. Assim, é possível modificar os ataques dinamicamente para encontrar diversas vulnerabilidades utilizando diferentes ambientes de testes, i.e. analisar as vulnerabilidades em função do tipo de serviço que presta o roteador Tor (entrada, intermediário, saída, diretório, Ponto de Encontro e Ponto de Introdução). Por enquanto, o principal objetivo de atacar roteadores de saída é bloquear o tráfego malicioso, ao passo que o de atacar serviços ocultos é revelar os endereços IP.

Para utilizar estes ataques por Injeção de Falhas, utilizaremos testes constituídos por dois conjuntos de entrada: a carga de trabalho (workload) e a carga de falhas (faultload) [Hsueh et al. 1997]. A primeira representa as entradas usuais do sistema, que servem para ativar suas funcionalidades, enquanto a segunda representa as falhas a serem introduzidas no sistema. Para isso, contamos com um conjunto de 36 tipos de ataques que serão testados contra um dos ambientes Tor.

References

Akamai. The q2 state of the internet security report. `www.stateoftheinternet.com/security-report`.

Cavalcante, T., Rocha, A., and Carvalho, A. (2014). Large-scale micro-blog authorship attribution: Beyond simple feature engineering. In *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, pages 399–407. Springer.

Hsueh, M.-C., Tsai, T. K., and Iyer, R. K. (1997). Fault injection techniques and tools. *Computer*, 30(4):75–82.

ICIR. Lbnl/icsi enterprise tracing project. `http://www.icir.org/enterprise-tracing/`.

Ling, Z., Luo, J., Wu, K., Yu, W., and Fu, X. (2015a). Torward: Discovery, blocking, and traceback of malicious traffic over tor. *Information Forensics and Security, IEEE Transactions on*, 10(12):2515–2530.

Ling, Z., Luo, J., Wu, K., Yu, W., and Fu, X. (2015b). Torward: Discovery, blocking, and traceback of malicious traffic over tor. *IEEE Transactions on Information Forensics and Security*, 10(12):2515–2530.

Martins, Gilbert B, S. E., de Freitas, R., and Feitosa, E. Estruturas virtuais e diferenciação de vértices em grafos de dependência para detecção de malware metamórfico.

Metrics, T. (2015). Tor metrics. `https://metrics.torproject.org/`.

Net, H. (2011). The honey project france chapter. `http://www.honeynet.org/chapters/france`.

Orebaugh, A., Ramirez, G., and Beale, J. (2006). *Wireshark & Ethereal network protocol analyzer toolkit*. Syngress.

Salas, M. I. P. and Martins, E. (2012). *Metodologia de testes de seguranÃ§a para anÃ¡lise de robustez de Web services por injeÃ§Ã£o de falhas (Security testing methodology for robustness analysis of Web services by fault injection)*. PhD thesis.

Simons, J. W. (2014). A rede secreta. `http://www.publico.pt/tecnologia/noticia/a-rede-secreta-1673221`.

Wang, A., Mohaisen, A., Chang, W., and Chen, S. (2015). Delving into internet ddos attacks by botnets: Characterization and analysis. In *IEEE International Conference on Dependable Systems and Networks (DSN)*.