Analyzing the Effectiveness of DoS Attacks on Tor

[Fast Abstract]

Bernd Conrad and Fatemeh Shirazi
TU Darmstadt, Germany
{bconrad,fshirazi}@cdc.informatik.tu-darmstadt.de

ABSTRACT

Anonymous communication networks (ACNs) are one of the main building blocks of protecting online privacy. However, such systems have often to deal with performance problems and are also not resilient against attacks. We survey the effect of denial-of-service (DoS) attacks on the performance and anonymity of the well known ACN Tor. Two types of DoS attacks are simulated to investigate the effectiveness of the attacks and the differences between various strategies on applying these attacks are analyzed.

Categories and Subject Descriptors

C.2.4 [**Computer-Communication Networks**]: Computer-Communication Networks; C.2.0 [**Computer- Communication Networks**]: General—*Security and protection*

General Terms

Experimentation, Performance, Security

Keywords

Anonymous Communication, Denial-of-Service Attack

1. INTRODUCTION

When users communicate over the Internet, an adversary eavesdropping the packets sent via the Internet is able to identify communication partners regardless of encryption. Conventional end-to-end encryption is only applied to the payload of a packet and is therefore not able to hide the source and destination IP addresses contained in the header of each IP packet. In Tor [8], currently the most widely used ACN, messages are forwarded using a network of relays, which seeks to conceal the correlation of communication partners. DoS attacks are a common network attack that aim at flooding a target with messages, e.g., connection requests, at a rate that the target is not able to process in time. This will lead to the service being unusable for all users as long as the DoS attack persists. We investigate the potency of different DoS attacks against the Tor network in terms of performance and anonymity. Furthermore, we evaluate to what extent different strategies can be utilized to optimize certain aspects of DoS attacks, in particular the selection of targets. We survey two types of DoS attacks by simulating them with a Tor simulator and by evaluating and analyzing the results of the simulations.

SIN '14 Sep 09-11 2014, Glasgow, Scotland Uk
ACM 978-1-4503-3033-6/14/09 ...$15.00
http://dx.doi.org/10.1145/2659651.2659707.

2. BACKGROUND

The purpose of an ACN is to enable users to establish anonymous communication channels over an open network. Typically, an ACN is composed of a set of routers (*nodes*) constituting an overlay network and messages are relayed over multiple routers to achieve anonymity. Tor is a low latency ACN with currently over 200,000 daily users[3]. Rendering the service useless for all users is critical when talking about ACNs like Tor, since users that are presented with a high latency and low throughput may be driven away from the network and therefore no longer anonymously communicating, which could just be what an adversary intends. DoS attacks may use well known protocol or application flaws and exploits them to crash the target or exhaust all processing power without investing unnecessary amount of resources. Significantly reducing the performance of Tor will lead to retransmission of messages, due to unreliable message transport and congestion, and also drive users away from the network and therefore no longer anonymously communicating. Messages that need to be retransmitted may give an adversary additional information that is needed to identify communication partners or potentially opening new attack windows for the adversary, while users leaving the network will reduce the anonymity set; both lead to a weaker degree of anonymity. To compromise the anonymity of a Tor user an adversary has to at least control both the entry guard and the exit relay of the user's circuit; which enables her to perform a traffic analysis attack matching traffic patterns and linking the user with his destination[4]. To increase the probability of compromised circuits, the so called selective DoS[7, 5, 6] can be used. In contrast to traditional DoS attacks, which most of the time aim at completely disabling a service, this attack does not intend to drive users away from the service. The adversary launches a selective DoS attack to force the system to enter a less secure state, often without the user even noticing it[5]. Controlled relays will selectively deny service if the circuit is not compromised. This will force users to retransmit messages as well as Tor to rebuild circuits to give the adversary more chances to deanonymize them. The aim of an selective DoS attack is to maximize the number of compromised circuits by killing of uncompromised circuits and forcing new circuits to be build.

3. SIMULATION OF DOS ATTACKS

In this section the results of simulations of two different DoS attacks on Tor are presented: a naive DoS attack that targets the performance of the Tor network and a slight variation of the selective DoS attack that targets the anonymity of the Tor network. The Shadow simulator[9, 1], a discrete-event simulator, was used to simulate both attacks.

3.1 Setting

All simulations were run with Shadow version v1.9.0-dev on Ubuntu 12.10. All required information are extracted from a randomly chosen consensus from the archived consensus files of February 2013. As base for the hosts file, the consensus file *2013-02-05-20-00-00-consensus* is chosen. All files were retrieved from the Tor Metrics Portal web page[2].

3.2 Simulating a Naive DoS Attack Strategy

The adversary aims at reducing the overall performance of the Tor network employing a so-called naive strategy. The attack is simulated by removing the targeted nodes from the simulated Tor network. In order to perform the DoS attack, the adversary has to first choose a set of Tor relays for the attack. The adversary adds the relay with the highest advertised bandwidth that he is capable of attacking, in terms of bandwidth, to the list of targets. The remaining bandwidth of the adversary is then used to repeat this process, until no suitable relays are left or all of the adversary's bandwidth is used up. Simultaneously DoS attacks against each of the relays in the list is applied, effectively removing them from the Tor network. To evaluate the influence of the adversary's available bandwidth on the effect of the DoS attack on Tor's performance, various strong adversaries are assumed; capable of taking out 5%, 10%, 15%, 20%, and 25% of the total bandwidth of the downscaled Tor network.

3.2.1 Evaluation

We run for each of the 5 different types of adversaries, a set of 6 repeated simulations. A set of 6 simulations was run to measure the performance of the downscaled Tor network in absence of an attack. The mean simulation time was around 9 hours. It is important to mention that each single simulation of a set of simulations produced slightly different results, caused by the way that Tor's node selection algorithm chooses circuit participants. Clients joining the virtual network choose a potentially different set of three entry guards each time the simulation is repeated. Therefore it is likely that a client chooses another entry guard for all its circuits compared to other simulations of the same hosts file. This leads to a greater diversity of performance values due to the different path that are chosen through the network, which actually somewhat models the real Tor network. Hence, all sets of results of a set of simulations are aggregated and then evaluated as one. This means for example, that the sets of time to last byte values for a web client coming from 6 different simulations are aggregated into one set. We measured two types of performance measurements: time to last byte and time to first byte. Both show that all adversaries are able to impact the performance of the downscaled Tor network, however with sometimes surprising results.

Time to Last Byte.

The 25% adversary is able to score good results in terms of reduced performance as expected. While the network that is not under attack is able to provide a mean download time of 71.96 seconds for the 5 MiB file, the network that is under the attack of a 25% adversary performing a DoS attack, is only able to provide a mean download time of 90.18 seconds. This is an increase of 25.13% in the mean download time.The 5%, 15% and 20% adversaries show somewhat close results in terms of increased mean download time ranging from 18.0% to 15.96%. Surprisingly, the 10% adversary was able to score the best results in terms of reduced performance, slightly more than the 25% adversary, with an increase of 27.81% of the mean download time compared to the network that is not under attack. The download times for the 5 MiB file by bulk clients show similar results. The above mentioned behavior of the network can also be observed in all other results of the different client's mean download time, with the exception of the download of the 50KiB file by the performance clients. Somehow, the network that is under the attack of the 25% adversary is able to score the best results in terms of mean download time, an increased performance of 2.55% compared to the network that is not under attack. Furthermore, all other adversaries are also not able to score significantly increases in the download time, which most likely is caused by the small size of the file. Once more, the 10% adversary is able to score the best results in terms of reduced performance with a 7.89% increase in the mean download time.The results for the 320KiB web request shows, that the mean download time of relatively small files is only slightly increased. The 10% adversary is able to score a 13.89% increased mean download time, closely followed by the 25%, 20% and 15% adversary with a 8.19% to 10.68% increase. Moreover, the 25% adversary is able to score by far the highest increase in the mean download time with 35.15%, this is also the maximum increase of all set of results in terms of mean download time. The other adversaries are able to score an increase in the mean download time between 2.25% for the 5% adversary up to a 8.86% increase for the 20% adversary, however closely followed by the 10% and 15% adversary. Those results can potentially be explained with Tor's node selection algorithm. If the adversary luckily chose relays, which represented valuable relays of one of the three classes of entry guard, middle relay or exit relay, the performance impact on the network was higher than in cases where less important relays were chosen. This may explain why the 10% adversary was able to score overall good results, closely followed by the 25% attacker.

Time to First Byte.

The results for the download time of the first byte show unexpected network behaviors. In all cases, the network under the effect of the 25% adversary is always able to provide a faster mean download time of the first byte compared to the network that is not under attack. In case of the both 5MiB file downloads, the network with a 20% and 15% adversary also shows a faster mean values of time to first byte. For the 320KiB, 50KiB and 1MiB file download, the adversaries, except the 25% adversary, show a slight increase in the mean download time, with no real correlation to the adversaries invested bandwidth. There are multiple possible explanations for the described network behavior. Removing certain high bandwidth relays from the network will force clients to choose different entry guard lists as well as middle and exit relays. Circuits consisting of the new relays, which potentially were not used so frequently prior to the attack, may be able to provide a better performance due to a better

geographical distribution of the participating relays, reducing the overall latency of a circuit. This will reduce the time that is required to construct a circuit, which would explain, why only the time to first byte values are affected and not the total download time. The time to first byte values are only lower because the circuit construction time is included, however, subsequent data is susceptible to the same bandwidth limitations and congestion caused by the attack as for the other adversaries, and therefore not able to further benefit from the slightly lower latency. If we look at the relays that are removed by the 25% adversary, this argumentation makes sense. Two high bandwidth exit guards are removed, which are frequently used as exit relays.

Discussion.

At first glance, it may make sense to choose a naive strategy to take out the most valuable relays in regard to their bandwidth. However, as the results have shown, it is even more important to find the relays that are valuable to the network in regard to circuit building. Taking down a high bandwidth exit relay that is only exiting to some less used services, may not impact the performance as much as a high bandwidth entry guard that is taken down. Furthermore, taking down certain relays may also boost the performance for some clients, at least for small sized files. However, it is unclear if this effect is only present due to the relatively small Tor network in this set of simulations and a greater geographical diversity of relays may void this effect. It is also important to identify other aspects that may influence the performance of the Tor network, since the chosen relays, geographical location and the bandwidth alone are not able to explain all network behaviors. Therefore, it is unclear if the increase of the available bandwidth of an adversary, which is able to find a near optimally strategy to select valuable target, also results in a linearly reduced performance of the Tor network. In case of the naive strategy, this is not the case. The 10% adversary is in most cases able to score better or at least equal results in terms of increased download times compared to the 25% adversary, with the exception of the 1MiB file download.

3.3 Simulating a Selective DoS Attack

In order to compare the effectiveness of the different strategies in regard to their impact on the degree of anonymity, we need to define a way to measure, or at least estimate, changes of the degree of anonymity that is provided by Tor. The state of all created circuits can be utilized to efficiently evaluate the effect of the selective DoS attack on the degree of anonymity. Therefore, this way of measuring anonymity also seems to be the appropriate approach to evaluate the results of this set of experiments, in particular since the goal of the adversary is to maximize the number of compromised circuits. There are three different types of circuits that have to be considered; *honest circuits*: all relays are honest; *compromised or fully compromised circuits*: at least a controlled entry guard and exit relay; and *killed circuits*: at least one controlled relay, but no controlled entry guard and exit relay at the same time. Using this classification, the effectiveness of the different strategies can be compared by evaluating the simulation results in regard to the fraction of honest, compromised and killed circuits.

3.3.1 Simulation of the Attack

Simulating the selective DoS attack is performed in two phases: generating a list of circuits and then using this list to passively simulate the attack. To generate a list of "realistic" circuits, the Shadow simulator is used to run repeated simulations of Tor. Before the attack can be simulated, a set of relays that are controlled by the adversary has to be chosen. Note, that the adversary always has to control at least two relays in order to be able to deanonymize a circuit, whereas at least one needs to be able to serve as entry guard and one as exit relay, otherwise the attack will not produce any compromised circuits. The simulation features four main classes of router distribution strategies, whereas each class may further introduce different distribution:Guard and Exit; Guard, Exit and Middle; Guard, Exit and Exit Guard; Guard, Exit, Middle and Exit Guard. Utilizing the given strategy, which specifies the distribution of the adversarial bandwidth to the different types of relays, the relays of the respective class are chosen somewhat similar to the naive approach.There is only one exception, in contrast to a network level based DoS attack, the attacker of a selective DoS attack has to remain undetected or his relays will be removed from the Tor network. In addition, it makes sense to assume, that the relays with the highest available bandwidth of each class are either operated by the Tor developers or another well known honest provider. Therefore, when choosing the appropriate relays for this attack, following the naive approach, the topmost relays of each class will be ignored. After choosing a set of controlled relays, the attack is passively simulated.

3.3.2 Evaluation

The simulation of the selective DoS attack was conducted in two phases. For phase one, a set of 5 simulations were run. Results of this set of simulations were parsed.The parsed circuits are either one-hop or three-hop circuits, whereas only the latter are important for the simulation of the attack and serve as a baseline for phase two. The minimum and maximum values of the relevant three-hop circuit only differ by a value of 1.01%. All simulations of phase one were run using the original unmodified hosts file as input. The difference in the simulation time is caused by external factors, e.g., other simulations running on the same machine at that time. The mean simulation time was between 9 and 10 hours. For phase two, we assumed an adversary that is able to invest a bandwidth of 10% (12810 KiB/s) of the total bandwidth provided by the downscaled Tor network. The respective distributions, which will be listed in the following, were utilized to select a set of controlled relays. This set was then used to simulation the selective DoS attack and determine the fraction of compromised, killed and honest circuits for each of the 5 different lists of circuits from phase one.

Entry Guard, Exit Relay.

This class features three different strategies to distribute the bandwidth: 1:1, 1:2 and 2:1 (entry:exit). This class is the potentially most rewarding class of relays in terms of compromised circuits, since a malicious entry guard that is chosen by a client for all of his circuits is able to contribute to a high amount of compromised circuits. An equal distribution (1:1) of the bandwidth to both the set of entry guards and exit relays resulted in the best value in terms of compromised circuits in this distribution class with a fraction of 1.176% compromised circuits, 27.11% killed circuits and 71.71% honest circuits. Both other strategies, which distribute the bandwidth unevenly either favoring the entry guard (2:1) or the exit relay (1:2), are not capable of scoring

as good results as the 1:1 distribution. The 1:2 distribution shows a decrease of 14.7% (- 0.1731) in terms of the fraction of compromised circuits, a decrease of 2.32% (- 0.63) in terms of killed circuits and an increase of 1.1% (+ 0.81) in terms of honest circuits, compared to the 1:1 distribution. The 2:1 distribution is also only able to compromise a slightly higher fraction of circuits than the 1:2 distribution, however, the fraction of killed circuits is significantly higher with a value of 28.75%, furthermore the fraction of honest circuits is only 70.18%. Those two values represent the best results in terms of fraction of killed circuits and fraction of honest circuits for all of the selective DoS simulations that were performed. This shows that it is preferable for the adversary to choose a single high bandwidth relay instead of choosing multiple relays with equally low bandwidths.

Entry Guard, Exit Relay, Middle Relay.

This class features two different distributions: 1:1:1 and 1:1:2 (entry:exit:middle). The latter distribution was chosen due to the fact that the set of middle relays is quite large with its 51 relays. Both distributions only show relatively low results in terms of compromised circuits with 0.735% and 0.689%, respectively. However, both are able to score acceptable values in terms of killed circuits, whereas the 1:1:2 distribution was able to score a slightly better value with 27.79%. This makes sense since a higher fraction of the adversary's bandwidth was directed to the middle relays.

Entry Guard, Exit Relay, Exit Guard.

This class features one distribution: 2:2:1 (entry:exit:exit guard). This distribution was chosen due to the fact that the set of exit guards is relatively small (only 7 relays) compared to all other relay classes. Controlling a great fraction of bandwidth of a small set of relays may lead to the detection of the adversary. This distribution was able to score the best results of all simulated distributions in terms of compromised circuits. With a fraction of 1.246% of compromised circuits, the 2:2:1 distribution is able to score an increase of 6% (+ 0.07) of the fraction of compromised circuits compared to the 1:1 distribution. However, the fraction of killed circuits is decreased by 0.44% (- 0.12), leading to an increase of 0.1% (+ 0.05) of the fraction of honest circuits. Evaluating the circuits from the first phase of the simulation reveal why this distribution is able to score the best results in terms of compromised circuits: exit guards are the relays which are most frequently used for exit relays, but due to their high bandwidth are also used for entry guard, leading to an overall good distribution of bandwidth.

Entry Guard, Exit Relay, Middle Relay, Exit Guard.

Here we feature two distributions: 1:1:1:1 and 2:2:2:1 (entry:exit:middle:exit guard). The second distribution was chosen with the same argumentation of the small exit guard set as stated before. Both distributions show low fractions of compromised circuits, 0.700% and 0.585%. In addition, the 1:1:1:1 distributions also shows the lowest fractions of killed circuits of all simulated distributions with only 24.93%, which leads to the highest fraction of honest circuits of all distributions with 74.38%. This is due to the fact that the fraction of bandwidth that is available for the respective relays classes is so small, that only very low bandwidth relays could be chosen to be controlled. Hence, those relays are infrequently chosen by Tor's node selection algorithm, leading to the relatively high fraction of honest circuits.

Discussion.

In terms of compromised circuits, the 2:2:1(entry:exit:exit guard) distribution scores the best results, whereas the 2:1(entry:exit) distribution shows the best results in terms of killed circuits while also providing a good fraction of compromised circuits. Including middle relays into the distributions always reduces significantly the fraction of compromised circuits. This makes sense, since middle relays are rarely able to contribute to the fraction of compromised circuits. Middle relays are mostly useful to increase the fraction of killed circuits. However, as seen with the 2:1 (entry:exit) distribution, the adversary is also able to achieve a high fraction of killed circuits without including any middle relays in the distribution, and therefore sacrificing important bandwidth for relays that are potentially not able to either serve as entry nor exit. Distributing the bandwidth to all available relays has also shown to be neither useful for the fraction of compromised circuits nor for the fraction of killed circuits. This may change if a stronger adversary is able to invest more bandwidth into the attack, therefore being able to choose high bandwidth relays of each of those classes. Note, that the adversary, since he is actively participating in the network, has to avoid detection or he will be excluded from the network and no longer able to perform his attack. Therefore, a stronger adversary would not be able to invest all his bandwidth into, for example, only exit and entry relays without risking to be detected. Overall, the simulations showed that an adversary that has the capability to invest 10% of the total bandwidth of the downscaled Tor network is only able to compromise at maximum a fraction of 1.246% of all circuits. However, a fraction of almost 29% killed circuits may lead to an increased fraction of compromised circuits when newly built circuits contain controlled relays, furthermore, an adversary may be able to retrieve better distributions and/or invest more bandwidth potentially leading to a significantly higher fraction of compromised circuits.

4. REFERENCES

[1] The shadow simulator. `http://shadow.github.io/`.

[2] Tor metrics portal: Data. `https://metrics.torproject.org/data.html`.

[3] Tor metrics portal: Users. `https://metrics.torproject.org/users.html`.

[4] K. Bauer, D. McCoy, D. Grunwald, T. Kohno, and D. Sicker. Low-resource routing attacks against tor. In *WPES*, pages 11–20. ACM, 2007.

[5] N. Borisov, G. Danezis, P. Mittal, and P. Tabriz. Denial of service or denial of security? In *CCS*, pages 92–102. ACM, 2007.

[6] N. Danner, S. DeFabbia-Kane, D. Krizanc, and M. Liberatore. Effectiveness and detection of denial-of-service attacks in tor. *TISSEC*, 15(3):11, 2012.

[7] A. Das and N. Borisov. Securing anonymous communication channels under the selective dos attack. In *Financial Cryptography*, pages 362–370. Springer, 2013.

[8] R. Dingledine, N. Mathewson, and P. Syverson. Tor: The second-generation onion router. Technical report, DTIC Document, 2004.

[9] R. Jansen and N. Hooper. Shadow: Running tor in a box for accurate and efficient experimentation. Technical report, DTIC Document, 2011.