Analyzing the Gold Star Scheme
in a Split Tor Network

Benedikt Westermann[1], Pern Hui Chia[1], and Dogan Kesdogan[1,2]

[1] Q2S⋆, NTNU, 7491 Trondheim, Norway
[2] Chair for IT Security, FB5, University of Siegen, 57068 Siegen, Germany
{westermann,chia,kesdogan}@q2s.ntnu.no

Abstract. Tor is an anonymity network and two challenges in Tor are (i) to overcome the scalability problems of Tor's current network information distribution scheme, and (ii) to motivate users to become operators of nodes. Several solutions have been proposed to address these challenges. We investigate the ramifications of combining two seemingly promising proposals, i.e., splitting the Tor network into several sub-networks (for better scalability), while using the Gold Star scheme (for motivating users to become node operators). Through simulation, we show that the sub-networks are likely to end up in a state of highly imbalanced division of size and bandwidth. This threatens the security and worsens the scalability problem of Tor. We identify the ratio of nodes given a gold star and the fact that a gold star is solely awarded based on a node's bandwidth, being highly skewed in practice, as two factors that contribute to an imbalanced split. We explore several potential mitigating strategies and discuss their strengths and shortcomings.

Key words: Tor, Incentive Schemes, Gold Star, Split Network

1 Introduction

Anonymous communication deals with concealing who is communicating with whom and is an important building block for privacy enhancing technologies. One of the most popular anonymity networks is Tor [6]. Here, two actively discussed problems are the issue of *scalability* and the challenge *to motivate more users to become operators of Tor relays*. The scalability problem of Tor stems from its current information distribution scheme, which provides every user the full view of the entire network. In [12], the authors predicted that more bandwidth will be used to distribute the network information than for the actual anonymization process in the near future. Various approaches have been proposed to improve the scalability of Tor, most often by limiting the number of relays a user needs to know (i.e., a partial view of the network). Danezis and Syverson investigated the impact of providing users only a partial view of the

---

anonymity network in [5] and highlighted the problems in doing so. They stated that *"[..] while scaling such systems [as Tor assuming a partial view] can maintain adequate anonymity in the face of route fingerprinting, splitting the network outright may be more desirable"*[5, p.156]

Interestingly, even though splitting the network seems to be a clean and straight-forward solution, there can be ramifications, such as an unintended competition among the sub-networks, e.g., for high bandwidth nodes, or, as pointed out in [5], a malicious entity trying to influence the split to take advantage of it.

In this paper, we analyze the consequences of splitting the Tor network in the presence of the gold star (GS) incentive scheme, proposed in [17] to motivate more users to operate a relay. We show that splitting the Tor network while using the GS scheme in individual sub-networks is likely to give a highly imbalanced division of relays and total available bandwidth. This threatens not only the users' anonymity, but also worsens the scalability of Tor. We identify the ratio of relays given a GS and the bandwidth based GS policy as two factors contributing to an imbalanced split. We explore several potential mitigation strategies and put forward some recommendations.

Our paper is structured as follows. In Section 2, we first detail on prior works related to the scalability issue and incentives schemes in Tor. Next, we describe our simulation design, scenarios, assumptions and dataset used in Section 3, and present the simulation results in Section 4 together with a series of robustness checks to our simulation model. We explore and discuss about several potential mitigation strategies in Section 5 before concluding.

2 Background and Related Work

The general idea of Tor is to hide a user within a large set of users, the so-called anonymity set [19], which is the set of all participants. Tor routes the data of an *Tor client* through several *Tor relays* using layered encryption, and thereby hides the relation between the Tor client and the data receiver from other parties, such as the ISPs. In Tor, a path through the network is known as a *circuit*. The relays used in a circuit are selected automatically by an Tor client. Each circuit is capable of anonymizing multiple TCP connections simultaneously.

2.1 Distributing the Network Information

Being an overlay network, Tor needs to inform the Tor client about different relays in the network. Tor does this currently by providing a global view of the network to all Tor clients. The global view, stored in a data structure called *consensus document*, is generated and distributed as follows. Each relay is required to upload a detailed description of itself (referred to as *descriptor*) to all known *directory authorities*. Every hour, the directory authorities agree on the state of the network based on the descriptors they received and publish an hourly consensus document. The document (together with new or updated descriptors

of different relays) is downloaded by the *directory mirrors*. Each Tor client in turn downloads the consensus document and new (or updated) descriptors of individual relays from a directory mirror, or directly from a directory authority if no known directory mirror is available. This causes a quadratic distribution cost which does not scale [26, 12].

To overcome the problem, several proposals have been made, mainly with the use of a *distributed hash table* (DHT). An early proposal, which use a DHT to distribute the network information, was presented in the first version of Tarzan [9]. In [8], however, the authors substituted the DHT with a gossip algorithm, which provides a full view of the network to the client applications. The importance of having a full view in Tarzan was later shown by Danezis and Clayton [4]. Nevertheless, having a full view, Tarzan faces a similar scalability issue as Tor. Another approach to distribute information using a DHT was proposed with Salsa [16] by Nambiar and Wright. But it was shown by Mittal and Borisov in [13] that Salsa is not as secure as claimed due to the information leakage introduced by redundant lookups in the DHT. Westermann et al. used a DHT based on Kademlia [11] to lookup nodes in an anonymity network [26]. Contrary to Salsa, only the servers are present in the DHT. The users use a small set of servers which they trust to perform node lookups. Additionally, the results are not immediately used to build a connection. This prevents timing correlations. In [18], Panchenko et al. showed that this approach does not provide enough security in big networks as an attacker can significantly bias the node selection.

A more recent approach, named Torsk, was proposed by McLachlan et al. [12]. Torsk is also based on a Kademlia DHT, but it uses additional certificates to verify that a node is the legitimate holder of a key based on the technique proposed in [22]. Furthermore, each node maintains a signed list of lookup buddies. A circuit is built by iteratively asking the buddies of the last node in a circuit to lookup a random ID. The lookup returns a list of verifiable nodes within the close neighborhood of the queried ID. The client then randomly selects a node from the list to extend the circuit. Another DHT based approach called Nisan was proposed by Panchenko et al. [18]. Nisan relies on a chord ring [21] for node lookups. Contrary to a classic chord lookup, nodes are asked for their finger table instead of the closest nodes to a given value. By doing so, a single node cannot learn about the ID a client is looking for. To provide more protection against active attacks by colluding nodes (e.g., by only returning a finger table with colluding malicious nodes), the authors proposed a bound check of the finger table. Yet, Wang et al. [23] showed that both Torsk and Nisan leak information, allowing the attacker to reduce the users' anonymity.

Mittal et al. [14] proposed ShadowWalker – another scheme to address the scalability problem in anonymity networks. The nodes for a circuit are selected by performing a random walk through the network. Only the last nodes in the random walk are used for circuit building so to improve performance and anonymity. Route capture attacks and manipulation to select malicious nodes are countered with the so-called *shadows*, which maintain and attest the finger table of shadowed nodes independently. But Schuchard et al. showed in [20] that

ShadowWalker is not as secure as claimed. Yet, they noted that the impact of their attacks can be mitigated by modifying the parameters and the consensus requirements slightly.

In [15] Mittal et al. move towards a new direction to overcome the scalability problem with the use of *private information retrieval* (PIR) techniques. The authors suggested two different solutions based on PIR. The first solution utilizes the current directory servers for the distribution of the network information. Here, Tor clients download a small block of descriptors of (untrusted non-authoritative) directory server. Due to PIR, the directory server does not learn which block has been downloaded by the Tor client. The second solution relies on the client's *guards*, being the trusted entry points to the Tor network, to fetch the descriptors for a circuit. Both solutions have in common that a Tor client only downloads a small set of descriptors. Thereby, PIR ensures that only the Tor client knows which descriptor has been downloaded. The authors show that both solutions scale sufficiently to overcome the Tor's scalability problem.

An analytical study was performed by Danezis and Syverson in [5] extending from the work in [4]. The authors analyzed the impact of providing only a partial view to individual users in anonymity networks. Their analysis shows that with a partial view scheme, the anonymity set can be drastically reduced by just knowing two nodes in a path. They also discussed the potentials of splitting the Tor network (favoring it over a partial view scheme) but noted the importance of a secure split to avoid exploitation by malicious parties. Related to the problem of a secure split is a work of Dingledine and Syverson [7] where the authors worked on the problem of building reliable mix cascade networks. They suggested assigning mixes to cascades in an unpredictable but verifiable fashion based on random inputs from all mixes. This helps among others to prevent malicious mixes from targeting a specific cascade.

Considering the various approaches proposed, their complexity and the potential attacks, it seems that splitting the Tor network into several sub-networks is an interesting option to investigate more throughly.

2.2 Motivating the Users to become relay Operators

Another challenge in anonymity networks is to motivate enough users to become a node operator. In Tor, relays are mostly operated by volunteers who hardly get any benefits for doing so. Despite a growing number of users, it is difficult to find enough independent relays and operators in Tor [17]. Several incentive schemes have been proposed to address this challenge and can be generally categorized into two classes: *incentive-by-money* and *incentive-by-performance*.

Incentive-by-money schemes include JonDonym's payment system [24, 25], PAR [2] and XPay [3]. We omitted the details of these schemes here, instead, we focus on incentive-by-performance schemes.

Two incentive-by-performance schemes are the *gold star scheme* [17] and Braids [10]. The basic idea of the *gold star scheme* is to prioritize the traffic of useful relays by assigning a gold star (GS) to a fraction of the most useful relays. In [17] the authors assigned a GS to the 87.5% best performing relays. A relay

having a GS (abbreviated as GS-relay) is entitled to extend a prioritized circuit to another GS-relay. If the whole circuit consists of only GS-relays, the traffic is prioritized resulting in an improved performance, e.g., lower response times or a higher bandwidth. One disadvantage of the GS scheme is that a prioritized circuit can only be initiated by a GS-relay. Therefore, the anonymity set for such a circuit is limited to the GS-relays only.

On the other hand, Braids proposes to distinguish the traffic in three different classes: *low latency*, *high throughput*, and *normal* [10]. In order to route the traffic in the low latency or the high throughput class, a client has to provide a relay with *tickets*. Tickets are distributed freely to all users, but the number is limited and bound to an expiry date. By enabling the relays to convert the collected tickets into new tickets, operators can route more of their own traffic in the low-latency or high-throughput class than the non-contributing Tor clients. Contrary to the GS scheme, prioritized traffic does not necessarily stem from a relay.

3 Simulation Design

In this paper, we investigate by simulations the outcomes of combining the two viable strategies: *splitting the Tor network* and *the GS scheme*. We detail on the simulation design, assumptions and dataset used in this section.

3.1 Basic Assumptions and Simulation Scenarios

We focus on the scenario where the Tor network is split into two sub-networks[2] and the relays are incentivized to contribute through separate GS schemes in individual sub-networks. Our simulation builds on two basic assumptions:

A1: GS policy is publicly known. Ngan et al. [17] suggested that the GSs are assigned (by directory authorities) in the consensus document. As the consensus document is publicly available, we assume that operators can learn about the details of the GS policy and can determine if he can obtain a relay in a particular sub-network reliably.

A2: GS performance is estimable from average bandwidth of GS-relays. We estimated the GS performance within a sub-network using the *average observed bandwidth* of GS-relays. We checked the suitability of the *average observed bandwidth*, obtainable from the descriptors of the relays, as a performance estimator by measuring its correlation with the *average download time*[3]. We obtained a Pearson correlation factor of -0.71 (with 280 data points corresponding to the daily average values from 03/27 to 12/31/2010), indicating a strong (negative) linear correlation.

We model the operators to decide if they should switch to another sub-network with the following decision rules:

---

[2] We present the results of splitting into more than 2 sub-networks in the appendix.

[3] Average download time measures the time Tor clients need to download files of different sizes and is available on the *Tor Metrics Portal* [1].

– If a relay has no GS and can become a GS-relay in a particular sub-network, it switches to this sub-network. We basically assume that the GS is an incentive for relays and they are eager to get one.
– If a relay can get a GS in multiple sub-networks, it chooses the sub-network that provides the best service according to its objective.

We distinguish between two different objectives: *performance-maximizing* and *anonymity-maximizing*. If a GS-relay is performance-maximizing, it chooses to join the sub-network that provides the best performance, i.e., one with the highest *average observed bandwidth*. If a GS-relay is anonymity-maximizing, it chooses to join the sub-network having the most GS-relays. It is important to note that a prioritized circuit can only be initiated by a GS-relay and therefore the sender anonymity set is limited only to the set of GS-relays.

We augment the simulation dynamics with two additional decision factors, which serve to test the robustness of our basic model. First, we consider the case where a relay switches to another sub-network only if the expected improvement, according to his objective is higher than a threshold γ. If the improvement is lower than γ, the relay will switch only with a probability equals to the ratio of improvement over γ. This models a cost for switching (e.g., extra configuration effort) and the reluctance to switch if the expected improvement in performance (or anonymity set) is small. We consider also a random decision factor, where an relay will with probability θ ignore the usual decision rules and switch to a random network (even when it cannot retrieve a GS there). This models the bounded rationality of the operators, especially in the event that the GS scheme is hard to predict in practice.

3.2 Simulation Details

Network State. To mimic the real life scenario, we ran our simulation based on Tor's actual network state as given in the consensus document published on 09/30/2010 at 6pm (GMT). The document describes 2136 active relays, each of which comes with a different *observed bandwidth*.

Figure 1(a) depicts the CDF of the observed bandwidth, that appears in the descriptors of the individual relays. The distribution is skewed: the top 17% of the relays actually provide 83% of the *total available bandwidth*.

To evaluate the effect of the skewed bandwidth distribution, we repeated most of our simulation scenarios assuming a uniform bandwidth distribution (see Figure 1). The uniform distribution was constructed such that the mean equals the *average observed bandwidth*. The number of relays equals 2136. We investigated also the effect of network size by scaling up the original network to have 10 times as many relays.

Simulation Flow. In the initialization phase of each simulation run, the relays are first ranked according to their *observed bandwidth* and assigned into different sub-networks in turns. Let $\{r_1, \ldots, r_N\}$ be the set of relays ranked in decreasing *observed bandwidth*, and let $\{S_1, \ldots, S_M\}$ be the set of sub-networks available, then the set of relays assigned to sub-network S_j is given by $\{r_i \mid i$
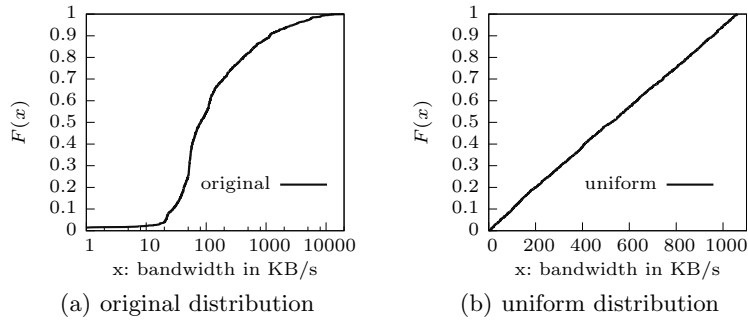
(a) original distribution   (b) uniform distribution

Fig. 1. Bandwidth distributions among the relays

repeat
   pick a random relay
   determine the sub-network relay is currently in, say S_k
   determine if relay will switch to another sub-network
   if relay switches **then**
      $stable_k \leftarrow 0$
   else
      $stable_k \leftarrow stable_k + 1$
   end if
   $round \leftarrow round + 1$
until $\forall j \in [1, M]$ $stable_j > STABILITY$, or $round > ROUND_{max}$

Fig. 2. The Simulation Flow.

mod $M = j$}. This gives us a near-balanced division of relays and bandwidth in the initial state. The simulation then proceeds as shown in Figure 2. Note that a relay is randomly selected per round (independent of the sub-networks) to check if it will switch to another sub-network given different objectives, θ and γ values. We repeated each simulation run 30 times to obtain the average outcomes.

4 Results

We present the findings from our simulation in the following.

4.1 Basic Simulation Model

We started our simulation by assuming that the operators are performance-maximizing, i.e., they will switch to another sub-network given a better performance, no matter the extent of improvement of the new sub-network. This simplistic scenario enables us to gauge the basic consequences of splitting the Tor network while having the GS scheme in place. We investigate the robustness of our findings with several extended models in Section 4.3.

We first simulated the case where GS ratio (GSR) equals 87.5% as proposed in [17]. We assumed here that the Tor network is split into two sub-networks (A and B). Figure 3(a) depicts the size and total available bandwidth of sub-network A during the course of a simulation run. The simulation began with a near-balanced state in terms of size and total available bandwidth. In the first few thousand rounds, however, sub-network A started to attract more and more relays from sub-network B. This initial rush caused sub-network A to be highly dominating both in terms of size and bandwidth, and could lead to a collapse of sub-network B. At one point, sub-network B consisted of 875 relays (40%) and had mere 3% share of the total available bandwidth (35.45 MB/s). There is a considerable risk that sub-network B will stop to be functional. With a low share of bandwidth, it can only support few Tor clients. This in turn causes a small anonymity set and may drive away the Tor users.

The initial rush was followed by a reversal in switching direction. By inspecting the simulation log, we found that this was led by the low-to-medium bandwidth relays which began to realize that they could obtain a GS in the sub-network B. The migration of the low-to-medium bandwidth relays caused the medium bandwidth relays to gradually lose their GS (due to a fixed GSR), and thus followed suit. The simulation ended with the stability count being reached, i.e., when no relay switched for consecutive 2400 rounds. In the final state, sub-network A consisted of only 558 relays (26.12%) but had a large share of the total available bandwidth, 972.1 MB/s (87.9%). Meanwhile, sub-network B consisted of 1578 relays (73.88%) but was providing only 133.81 MB/s (12.1%). The distribution of relays and total available bandwidth was highly uneven. The distribution of exit bandwidth, considering the flags of the relays in the consensus document, corresponds roughly to the case of total available bandwidth. Sub-network A and B have 87.3% and 12.7% of the total exit bandwidth respectively. The situation is slightly different with respect to guard relays. Both sub-networks have a similar share of the total number of guard relays (sub-network A has a 51% share), but the division of guard bandwidth is again highly uneven (the guards in sub-network A actually provide 93.0% of the total guard bandwidth).

Figure 3(b) presents the CDFs of relays' bandwidth in separate sub-networks in the final state. Most notable is the absence of the medium bandwidth relays represented by the horizontal line in the mid area of the CDF of sub-network A.

Such an imbalanced split increases the risk of a user being deanonymized: the higher the fraction of bandwidth some relays provide, the higher is their probability of being selected as end-points of a circuit. Tor cannot provide any anonymity against an operator who holds the first and the last position in a circuit. An attacker with some high-bandwidth relays could enter sub-network B, where they can provide a higher fraction of bandwidth more easily, to attempt deanonymizing the users.

Independent of the risk to anonymity, the scalability problem is also worsened in sub-network B which has many relays, but on average provide only 87 KB/s (with a maximum of 330 KB/s).
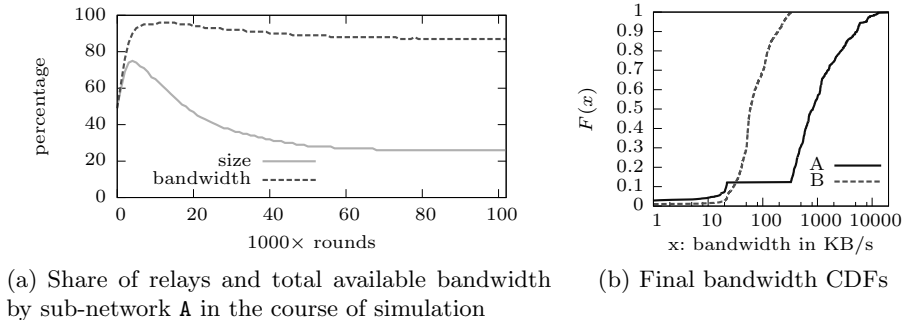
(a) Share of relays and total available bandwidth by sub-network A in the course of simulation

(b) Final bandwidth CDFs

Fig. 3. Simulating the basic model with GSR=87.5% as originally proposed in [17]

(a) Impact on size
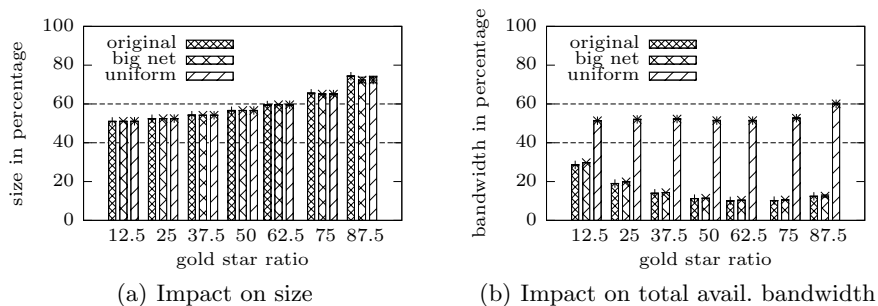
(b) Impact on total avail. bandwidth

Fig. 4. The effect of GSR, number of relays and bandwidth distribution on the share of size and total available bandwidth. Both figures plot the state of the larger sub-network

4.2 Probing the Reasons of an Uneven Split

Different GSRs. We then simulated the basic model with other GSRs, ranging from 12.5% to 87.5%, to investigate the effect of GSR on the distribution of relays and total available bandwidth among the sub-networks. We found that the GSR has a significant effect on network size, as shown in Figure 4(a). The higher the GSR, the more uneven is the size of the sub-networks. This is mainly by the fact that the number of relays, which have a potential incentive to switch to a better sub-network, increases with a higher GSR.

Increased number of relays. We re-ran the simulation with a scaled-up Tor network with 10 times as many relays (following the same bandwidth distribution) as in the consensus document. As shown in the Figure 4, an increased number of relays has little or no effect to the simulation outcomes.

Bandwidth distribution. Next, we considered a hypothetical set of relays where the relays' bandwidth follows a uniform distribution. The benefits of this can be seen in Figure 4. Although the sub-networks become more uneven in terms of size as the GSR increases, we see that with a uniform bandwidth distribution, the share of the total available bandwidth between the 2 sub-networks maintains at an even ratio (around 50%-60%).

Summary. The results show that the GSR has a significant impact on the distribution of relays among the sub-networks. The bandwidth distribution of relays, on the other hand, affects the share of the total available bandwidth.

As relays are rewarded solely based on the bandwidth they and their peers provide, we observed that the high-bandwidth relays prefer to gather in the same sub-network. Given the bandwidth distribution is skewed, the self-sorting of the high-bandwidth relays causes a highly uneven division of total available bandwidth. Interestingly also, an increased number of relays has only minimal effect on the distribution of relays and bandwidth.

4.3 Extended Models for Robustness Check

Switching Costs. Previously, we have assumed that the operators would prefer to switch to another sub-network whenever there is a slight improvement in performance. In practice, however, switching incurs a cost. Besides the configuration effort, other reasons, such as a good prior experience, perceived anonymity or trust for some specific set of relays, could render an operator to prefer remaining in his sub-network. To model such inertia, we simulated a threshold based switching strategy. Consider a threshold of γ, an operator switches to another sub-network (with certainty) if the improvement in performance δ is greater than or equal to γ. Meanwhile, if the improvement δ is smaller γ, we model that the operator to switch with a probability of δ/γ.

We plot the simulation outcomes in Figure 5. Comparing to the outcomes when GSR = 50% and 87.5% in Figure 4, we observe that the effect of having a switching threshold is minimal. A switching cost delays but does not deter the operators from switching for self-interests. Our previous findings on the effect of different GSR and bandwidth distribution also remain applicable. The counterintuitive 'delays-but-does-not-deter' result could be partly due to the fact that we have used a probabilistic decision rule when the expected improvement in performance is below a specific threshold instead of a clear-cut switch or not decision. Yet, we note that the eventual switching of relays is also attributed to the dynamics of performance improvement. Specifically, the switching of a GS-relay (with either a deterministic or probabilistic decision) increases the bandwidth of the destined sub-network, while decreasing the performance of the origin sub-network. This increases the expected improvement in performance for relays remaining in the origin sub-network, inducing them to switch in turn.

Bounded Rationality / Complex GS Policy. Our basic model assumes that all operators could estimate the policy for GS-relays and determine if they can retrieve a GS in individual sub-networks. In practice, it may not be trivial for all operators to do so (bounded rationality). A complex (or hidden) GS policy could also cause many operators to be uninformed (or indifferent) of the 'better' sub-network. We investigated whether the problem of uneven split remains considering a random decision factor where an operator ignores the usual decision rules and joins a sub-network randomly with probability θ.

Figure 6 shows that when a majority of relays decide randomly, the division becomes more even. A more even split is expected when θ is large, since the

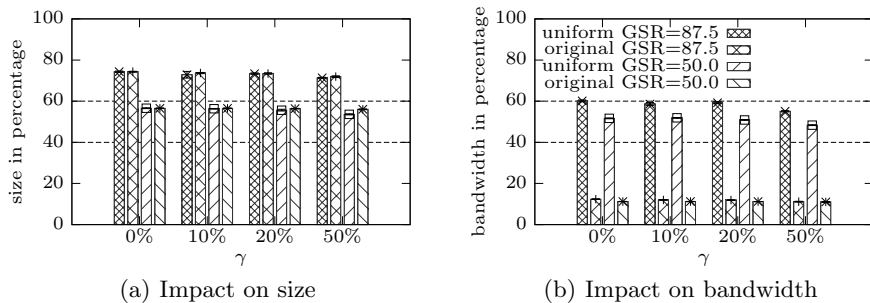(a) Impact on size   (b) Impact on bandwidth

Fig. 5. The effect of threshold based switching on the distribution of relays and total available bandwidth. Both figures plot the state of the larger sub-network.

bigger sub-network has a higher chance to lose a relay due to the uniform selection of a relay from all relays. Additionally, if many relays decide randomly, the effect of self-sorting among the relays (high-bandwidth relays prefer to be with high-bandwidth peers) reduces. The relays in a sub-network become more heterogeneous in terms of bandwidth and this contributes to an even division of total available bandwidth.

While the problem of uneven split seems to disappear as θ increases, we note that this may not be desirable. A large θ in Tor translates into a GS policy that is hidden or too complex to be predictable by the relay operators. This in turn takes away the incentives for becoming a relay operator, defeating the purpose of implementing the GS scheme in the first place.
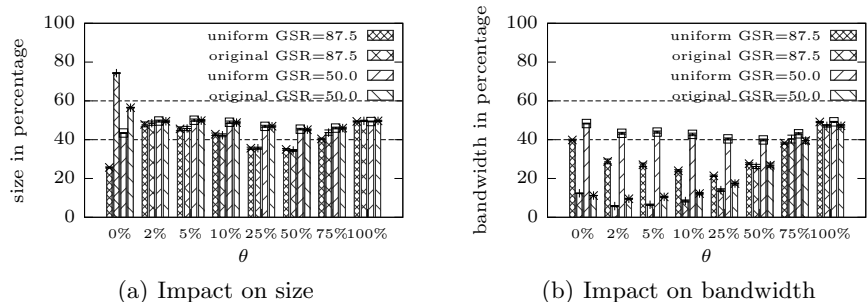

(a) Impact on size   (b) Impact on bandwidth

Fig. 6. The effect of random switching decision on the distribution of relays and total available bandwidth. The figures show the sub-network with the lower share of total available bandwidth.

Anonymity-maximizing Operators. So far we have assumed that Tor's users are performance-maximizing, not considering the anonymity that a sub-network provides. We regard this as a reasonable assumption for relays not having a GS (denoted as NGS-relays) and the Tor clients as long as the network

split is not too extreme[4]. With several hundreds of thousands of users in today's Tor network, the anonymity set can be seen as sufficiently large. The situation is, however, different for GS-relays as their anonymity set is limited to only the GS-relays. When an attacker observes a prioritized circuit, he knows that one of the GS-relays in a particular sub-network has initiated it, thus it can be more critical for the GS-relays to consider the anonymity factor. We study the effect of the anonymity-maximizing objective of GS-relays here. For NGS-relays, the objective remains to switch to a sub-network where it can retrieve a GS. Meanwhile, we model the GS-relays to prefer switching to a sub-network with the largest anonymity set (i.e., one with the highest number of GS-relays). If two sub-networks have the same number of GS-relays, a GS-relay chooses the sub-network with the best performance.

Figure 7 depicts that the course of a simulation run where relays are all anonymity-maximizing. The simulation started with an initial rush to one sub-network (hereafter, sub-network A). The over-crowding in sub-network A caused the low-to-medium bandwidth GS-relays to lose their GS and to switch in the reverse direction in order to (re)gain a GS. The migration of low-to-medium bandwidth relays in turn caused the medium bandwidth relays to also gradually lose their GS (due to a fixed GSR) and thus followed suit. Thus far, this has been similar to the case as shown in Figure 3(a). The situation, however, started to differ when the migration of medium bandwidth relays caused sub-network B to have a higher number of GS-relays (i.e., a larger anonymity set). This made even the high bandwidth GS-relays to prefer joining sub-network B as it would provide better anonymity. However, the arrival of high bandwidth GS-relays in sub-network B caused the low-to-medium bandwidth GS-relays to start losing their GS again, and decided to return to sub-network A. The same process then repeated itself, which explains the oscillating nature of the share of total relays and total available bandwidth.

Notice that at the extreme cases during the course of simulation, the smaller sub-network has only a $<5\%$ share of total available bandwidth and can thus be expected to support only few Tor clients. This hints on a small anonymity set and may drive away the remaining Tor relays and Tor clients. Thus, the risk of a failed network split remains even with anonymity-maximizing relays.

Additionally, we simulated the case where there is a mix of performance- and anonymity-maximizing relays. Let the fraction of performance-maximizing relays be ϕ. With GSR $= 87.5\%$ and $\phi = 40\%$, we observed that the oscillating nature of the share of size and total available bandwidth disappeared. Meanwhile, when GSR $= 50\%$ and $\phi = 40\%$, we observed the oscillating outcomes occasionally but not when $\phi \geq 60\%$. This holds for both simulation cases using the original bandwidth distribution and a uniform distribution. Most interestingly, when $40 \leq \phi \leq 60\%$, high bandwidth relays no longer gather in a single sub-network

---

[4] When the distribution of relays is extremely uneven, there is a good chance that the smaller sub-network collapses as users see their anonymity threatened.
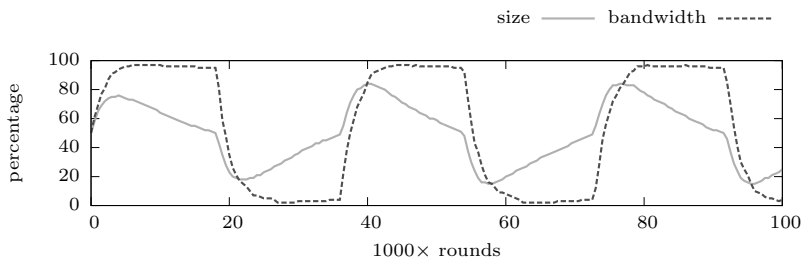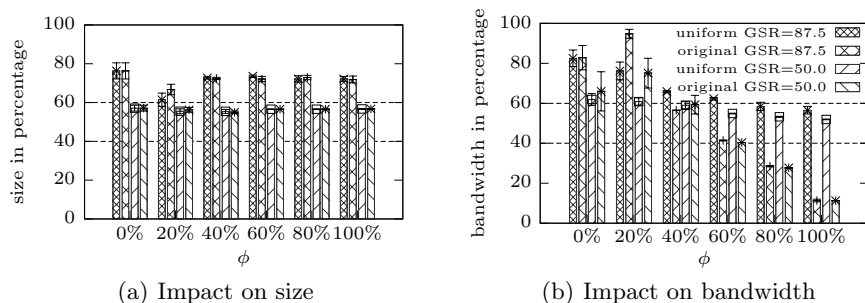
Fig. 7. The course of simulation where all relays are anonymity-maximizing.

(a) Impact on size

(b) Impact on bandwidth

Fig. 8. The effect of having a mixture of performance- and anonymity-maximizing relays. ϕ denotes the fraction of performance-maximizing relays. With a low ϕ, the confidence intervals are big, indicating an oscillating nature of the simulation outcomes.

5 Discussion

A highly imbalanced split of the Tor network has multiple serious implications. First, the scalability problem, being the original motivation for a split of the Tor network, can become worse in a sub-network that is large in size but has a low share of total available bandwidth. Secondly, the higher the fraction of bandwidth some relays provide, the higher is their chance of being the endpoints of a Tor circuit. In the event of an uneven split, the risk of a user being deanonymized is higher within the low-bandwidth sub-network as malicious relays can enter the sub-network, where they can provide a higher fraction of bandwidth more easily.

Global GS Scheme. An interesting question is whether an uneven split will still occur considering a global GS scheme, rather than separate GS schemes in individual sub-networks. In the worst case of an uneven split, all GS-relays will gather in one sub-network with the number of relays equals:

$$\lceil N \cdot GSR \rceil + \left\lfloor \frac{(1 - GSR) \cdot N}{M} \right\rfloor,$$

where N is the number of all relays and M is the number of sub-networks.

We observed the worst case outcome in our simulation as soon as one sub-network provides better performance and anonymity than the others, indepen-

dent of the GSR, the relays' objective, and the underlying bandwidth distribution. A global GS scheme therefore does not help the situation.

Fixed sub-network. We investigated the possibility of having a fraction of relays that do not switch from their assigned sub-networks (either by encouraging them to be cooperative or prohibiting them to switch at all). Figure 9 shows the outcomes where a fraction σ of Tor relays, selected in descending order of bandwidth or randomly from all Tor relays, do not switch from their assigned sub-networks. An even (40-60%) share of relays and total available bandwidth is only possible by fixing the sub-networks for the top 10% high-bandwidth relays. However, an even split is not achievable even with $\sigma = 50\%$ if relays are selected randomly to have fixed sub-networks. An alternative is to assign all Tor operators into sub-networks (randomly) while disallowing self-switching completely (see the case when $\sigma = 100$r).

Fixing the sub-networks for some percentage of the top high-bandwidth Tor relays, or (randomly) assigning all relays to sub-networks, are hence two possible solutions. These require an effective way to force a relay to stay in one sub-network. Additionally, assigning the ORs to one sub-network can raise multiple concerns. For instance, whether the Tor operators would be discouraged if their volunteering effort is 'punished' by not being able to choose their preferred sub-network freely. There may also be questions, e.g., on fairness, transparency, and security, if the assignment of sub-networks by a centralized authority is not completely random. A way to address such questions can be found in [7], where the authors proposed assigning mixes to cascades in a unpredictable but verifiable fashion. However, their approach deals mainly with mixes and cascades. Porting it to the problem of a fair and secure assignment of relays to sub-networks may warrant further investigation.
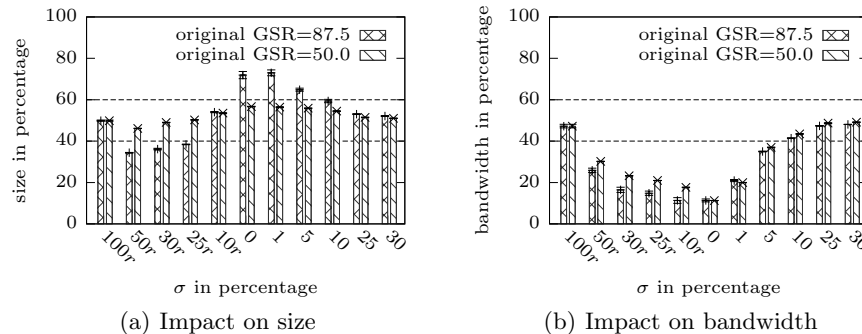

(a) Impact on size   (b) Impact on bandwidth

Fig. 9. The effect of having a fraction σ of relays, selected randomly (denoted with a suffix 'r') or in descending order of bandwidth, that do not switch from their assigned sub-network.

An appropriate GSR. We note that a GSR of 87.5% is not optimal in a split network setup as it leads to an uneven share of relays. Our simulations

indicate that a GSR of roughly $\frac{1}{M}$, where M equals the number of sub-networks, seems to be a good choice to avoid a uneven split.[5] A lower GSR trades off the risk of an uneven split with a reduced anonymity set for the GS-relays.

Alternative GS criteria. On top of an even distribution of relays, it is also necessary to have the relays distributed across the sub-networks independently from the bandwidth they provide, to ensure an even distribution of total available bandwidth. This is, however, not possible if a GS is granted based on the relay's bandwidth only, as high-bandwidth relays will gather in the same sub-network. Given that the GS scheme is introduced to motivate more users to become relay operators, using other requirements for awarding the GS will work. By having several independent requirements, relays in a sub-network can be more heterogeneous with respect to bandwidth.

To test our intuition, we re-ran the simulations using the basic model and assigned a value x_i to every relay r_i, which was sampled from a random variable X. In practice, x_i could be computed from any suitable properties, including the uptime, location or reputation of a relay. We then measured the *usefulness* of a relay by combining the relay's x_i value and its bandwidth, as shown in Equation 1, to decide if a relay r_i is eligible for a GS. In Equation 1, p_i denotes the fraction of relays providing less bandwidth as the relay r_i, and f_i is the fraction of relays that have a lower x value than r_i.

$$u_i = \omega \cdot p_i + (1 - \omega) \cdot f_i \tag{1}$$

The weights for p_i and f_i was controlled using the variable ω. We used two different distributions of X: (i) a uniform distribution, $X \sim \mathcal{U}(0, N)$, and (ii) a heavy-tailed distribution constructed based on the skewed bandwidth distribution in Figure 1(a). We note that the x_i value of each relay is drawn independently of its bandwidth.

Figure 10 shows that when ω is low (i.e., when the usefulness of a relay depends largely on the random value x_i), the division of total available bandwidth is even. The sub-networks also have a similar share of guards and exit relays, both in terms of number and bandwidth (not depicted). Meanwhile, as ω increases (i.e., as the usefulness depends more on the bandwidth of a Tor relay), the problem of an uneven distribution of bandwidth arises. This highlights that the GS criteria should not be solely dependent on the bandwidth, which is highly skewed in practice. We suggest to assign a GS based on the usefulness of a relay which can be a combination of multiple bandwidth-independent properties to ensure a good mix of relays with heterogeneous bandwidth in each sub-network.

It is important to note that the relays with the highest u-values will again gather in one particular sub-network. However, the impact can be minimized with a careful selection of factors contributing to the usefulness, u measure. For example, by having a usefulness measure that is distributed uniformly among the relays, we can expect the effect of the self-sorting to be less prominent.

---

[5] Simulation outcomes for $M = 3$ and 4 sub-networks are included in the appendix. For example, in Figure 11, one can see that for $M = 4$, the largest sub-network gets about 30% close to the $\frac{1}{4}$ share of relays.

This has been exemplified by the hypothetical scenario where there is a uniform distribution of bandwidth among the relays, as shown in Figure 4.
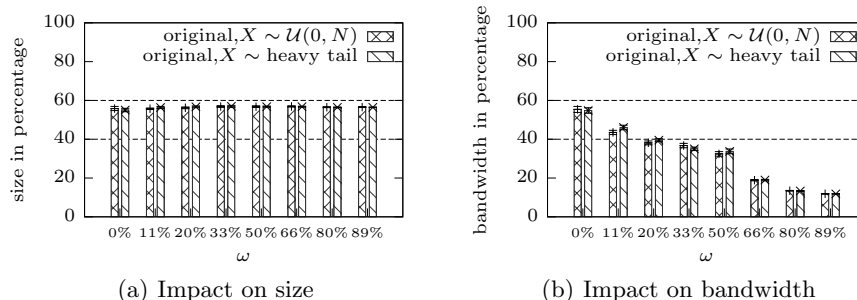

(a) Impact on size   (b) Impact on bandwidth

Fig. 10. Varying the dependence of the GS criteria (i.e., the usefulness measure) on the relay's bandwidth. With $\omega = 0\%$, the GS criteria is independent of the relay's bandwidth, while with $\omega = 100\%$ it depends on the relay's bandwidth only.

6 Conclusions

In this paper, we have analyzed the consequences of applying the Gold Star (GS) scheme in a split Tor network. While our simulation model has abstracted away the Tor clients, guards and exit relays for simplicity purposes, we have refrained ourselves from unrealistic assumptions besides taking into consideration a large number of different simulation scenarios.

We showed that applying the GS scheme directly in the setting of a split Tor network can lead to extremely imbalanced sub-networks both in terms of the share of relays and total available bandwidth. This threatens the users' anonymity and worsens the scalability problem of Tor.

In search of mitigation measures, we identified the ratio of relays given a GS (GSR) to be the main factor of an uneven distribution of relays across the sub-networks. By decreasing the GSR to $\frac{1}{M}$, where M is the number of sub-networks, we observed a near-balanced division of relays into the sub-networks.

Meanwhile, fixing the sub-network of some percentage of high-bandwidth relays or assigning all relays randomly, may represent two solutions for an even distribution of bandwidth across the sub-networks. Yet, while technically viable, fixing the sub-network of some or all relays can raise multiple concerns, including on respecting the contributors' choice and fairness.

A self-regulating solution can be achieved by changing how Tor would assign a GS to a relay. We showed that the imbalanced division of total available bandwidth can be addressed by designing a different set of GS criteria, for example by measuring the *usefulness* of a relay based on multiple bandwidth-independent properties, to improve the heterogeneity of relays in individual sub-networks.

References

1. Tor metric portal. `http://metrics.torproject.org`. (last visited Feb 2011).
2. E. Androulaki, M. Raykova, S. Srivatsan, A. Stavrou, and S. M. Bellovin. Par: Payment for anonymous routing. In *PETS*, pages 219–236. Springer LNCS, 2008.
3. Y. Chen, R. Sion, and B. Carbunar. XPay: practical anonymous payments for tor routing and other networked services. In *WPES*, pages 41–50. ACM, 2009.
4. G. Danezis and R. Clayton. Route fingerprinting in anonymous communications. In *Peer-to-Peer Computing*, pages 69–72. IEEE Computer Society, 2006.
5. G. Danezis and P. F. Syverson. Bridging and fingerprinting: Epistemic attacks on route selection. In *PETS*, pages 151–166. Springer LNCS, 2008.
6. R. Dingledine, N. Mathewson, and P. F. Syverson. Tor: The second-generation onion router. In *USENIX Security*, pages 303–320. USENIX, 2004.
7. R. Dingledine and P. F. Syverson. Reliable mix cascade networks through reputation. In *FC*, pages 253–268. Springer LNCS, 2002.
8. M. J. Freedman and R. Morris. Tarzan: a peer-to-peer anonymizing network layer. In *CCS*, pages 193–206. ACM, 2002.
9. M. J. Freedman, E. Sit, J. Cates, and R. Morris. Introducing tarzan, a peer-to-peer anonymizing network layer. In *IPTPS*, pages 121–129. Springer LNCS, 2002.
10. R. Jansen, N. Hopper, and Y. Kim. Recruiting new tor relays with braids. In *CCS*, pages 319–328. ACM, 2010.
11. P. Maymounkov and D. Mazières. Kademlia: A peer-to-peer information system based on the xor metric. In *IPTPS*, pages 53–65. Springer LNCS, 2002.
12. J. McLachlan, A. Tran, N. Hopper, and Y. Kim. Scalable onion routing with torsk. In *CCS*, pages 590–599. ACM, 2009.
13. P. Mittal and N. Borisov. Information leaks in structured peer-to-peer anonymous communication systems. In *CCS*, pages 267–278. ACM, 2008.
14. P. Mittal and N. Borisov. Shadowwalker: peer-to-peer anonymous communication using redundant structured topologies. In *CCS*, pages 161–172. ACM, 2009.
15. P. Mittal, F. Olumofin, C. Troncoso, N. Borisov, and I. Goldberg. PIR-Tor: Scalable anonymous communication using private information retrieval. In *USENIX Security*, 2011.
16. A. Nambiar and M. Wright. Salsa: a structured approach to large-scale anonymity. In *CCS*, pages 17–26. ACM, 2006.
17. T.-W. Ngan, R. Dingledine, and D. S. Wallach. Building incentives into tor. In *FC*, pages 238–256. Springer LNCS, 2010.
18. A. Panchenko, S. Richter, and A. Rache. Nisan: network information service for anonymization networks. In *CCS*, pages 141–150. ACM, 2009.
19. A. Pfitzmann and M. Hansen. Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management - a consolidated proposal for terminology, Feb 2008. v0.31.
20. M. Schuchard, A. W. Dean, V. Heorhiadi, N. Hopper, and Y. Kim. Balancing the shadows. In *WPES*, pages 1–10. ACM, 2010.
21. I. Stoica, R. Morris, D. R. Karger, M. F. Kaashoek, and H. Balakrishnan. Chord: A scalable peer-to-peer lookup service for internet applications. In *SIGCOMM*, pages 149–160, 2001.
22. P. Wang, N. Hopper, I. Osipkov, and Y. Kim. Myrmic: Secure and robust DHT Routing. Technical report, Uni. of Minnesota DTC Research, 2006.
23. Q. Wang, P. Mittal, and N. Borisov. In search of an anonymous and secure lookup: attacks on structured peer-to-peer anonymous communication systems. In *CCS*, pages 308–318. ACM, 2010.

24. R. Wendolsky. A volume-based accounting system for fixed-route mix cascade systems. In *Bamberger Beiträge zur Wirtschaftsinformatik und angewandten Informatik*, pages 26–33, Feb 2008.
25. B. Westermann. Security analysis of AN.ON's payment scheme. In *NordSec*, pages 255–270. Springer LNCS, 2009.
26. B. Westermann, A. Panchenko, and L. Pimenidis. A kademlia-based node lookup system for anonymization networks. In *International Conference on Information Security and Assurance*, Jun 2009.

Appendix

We simulated also the scenarios where the network is split into $M = 3$ or 4 sub-networks. As shown in Figure 11, the distribution of relays and total available bandwidth is uneven, same as the case when $M = 2$.
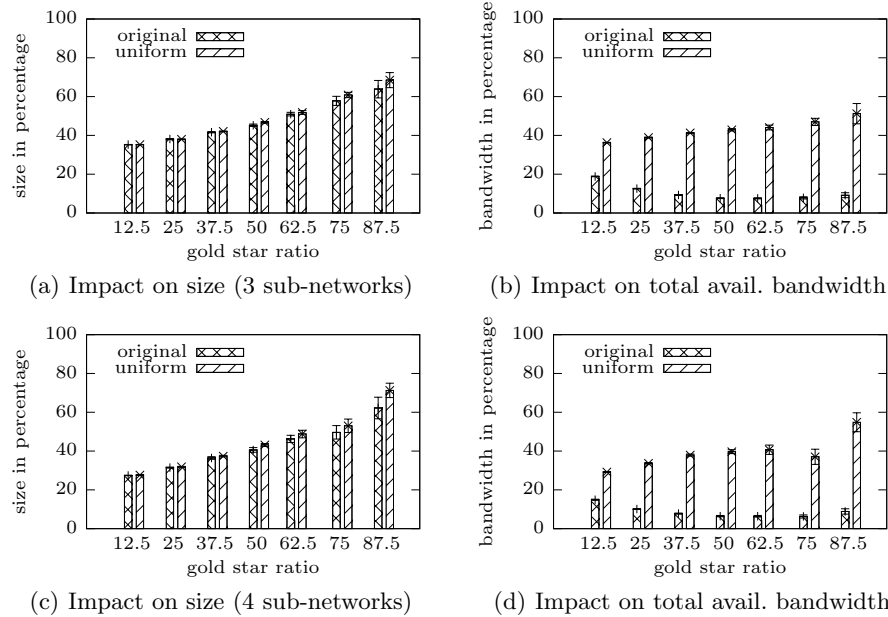

(a) Impact on size (3 sub-networks)

(b) Impact on total avail. bandwidth

(c) Impact on size (4 sub-networks)

(d) Impact on total avail. bandwidth

Fig. 11. The figures show the impact of different GSRs in the basic model when the network is split into $M > 2$ sub-networks. Figures a) and b) show the outcomes of a split into 3 sub-networks, while figures c) and d) show the outcomes of a split into 4 sub-networks. All of them plot the state of the largest sub-network at the end of simulation. The distribution of relays and total available bandwidth is uneven. The largest sub-network attracted more than $\frac{1}{M}$ of the relays but got less than 20% of the total available bandwidth.