Article

Tor Hidden Services: a systematic literature review

Diana L. Huete Trujillo [1,‡], and Antonio Ruiz-Martínez [2,*]

[1] University of Murcia, Department of Information and Communications Engineering; dianalissette.huete@um.es
[2] University of Murcia, Department of Information and Communications Engineering; arm@um.es
* Correspondence: arm@um.es
‡ These authors contributed equally to this work.

Abstract: Anonymous communications networks were born to protect the privacy of our communications, preventing censorship and traffic analysis. The most famous anonymous communication network is Tor. This anonymous communication network provides some interesting features, among them, we can mention user's IP location or Tor Hidden Services (THS) as a mechanism to conceal the location of servers, mainly, web servers. THS is an important research field in Tor. However, there is a lack of reviews that sump up main findings and research challenges. In this article we present a systematic literature review that aims to offer a comprehensive view on the research made on Tor Hidden services presenting the state of the art and the different research challenges to be addressed. This review has been developed from a selection of 57 articles and present main findings and advances regarding Tor Hidden Services, limitations found, and future issues to be investigated.

Keywords: Tor; hidden services; onion services; systematic literature review; survey

1. Introduction

Tor [1] is currently the most broad low-latency anonymous communication network and is used everyday by many people to protect their privacy and overcome censorship [2, 3].

The main function that provides Tor to any citizen is, through a set nodes in the Tor network, protect his/her anonymity by hiding his/her network IP address [1]. The security and anonymity of Tor has been deeply studied in many works and, although, different types of attacks are possible [3,4], the research community is working in providing solutions to the different attacks [5–7].

But Tor goes beyond protecting citizen's network identity. Since 2004 it offers the possibility of protecting the location of services through a mechanism that so far was named Tor Hidden Services (THS) and that, currently, is named as Onion Services. Thus, anyone can provide anonymous services.

These services can be clearly identified because their addresses are finished in .onion. According Tor metrics, there are more than 150,000 of these services [8]. Any of them requires that any citizen executes the rendezvous protocol to contact to the service in an anonymous way. Since its introduction the number of THS offered has exponentially increased [8]. Users use them to access services of different categories such as Adult, Drugs, Politics, security, anonymity, etc. Cybercriminals make use of THS to perform illegal activities [9]. These kind of activities are related to counterfeit credit cards, drugs, pornography and weapons. But, although it is used for these activities, which are not desired, as Snowden said "Tor is a critical technology, not just in terms of privacy protection, but in defense of our publication right" [10].

The whole set of services that are accessed in this way with other services from other anonymous networks that follow the same approach form what it is named as Dark Web [11]. This dark web has become famous because many illicit services has been offered

 doi:10.20944/preprints202107.0600.v1

through darknet markets [12]. However, other services that can be used by journalists and to provide freedom of speech are being offered by human rights and whistle-blowing organizations such as Wikileaks and Goldballeaks [13]. Furthermore, there is an important amount of THS providing links to services of the surface Web. Namely, to social networks, web content management, news, and adult content [11].

As mentioned previously, due to its popularity and the services it offers, Tor has been broadly studied and many issues regarding it has been analysed. It has been analysed both from user's point of view [14] and a technical point of view [5,11,13]. Being the technical studies the most common. In a taxonomy presented by Saleh et al. [5], they classify Tor research topics in deanonymization, performance analysis and architectural improvements, and path selection. Within these topics, we can found THS. But, in this analysis, they only cover 4 papers. These services are key within the Tor network and we consider they deserve to be studied into detail. However, we have not found a survey that tries to sum up what we have learnt about these hidden services and to identify what research issues still need to be addressed or improved. Therefore, we decided to analyse the state of the art of THS. For this purpose, we decided to address this issue through a systematic literature review (SLR) where we have analysed THS from different point of views. Our methodology is based on SLR because it follows a well-defined methodology which makes that the research made in a specific issue can be exhaustive, fair and repeatable [15]. Thus, the paper presents what we have discovered after analysing the research on THS during the period of 2006 to 2019. Namely, we have analysed 57 papers and this paper explains how we develop the survey, the questions we asked, and the answers we have obtained regarding the state of the art of THS.

The rest of this paper is structured as follows. In Section 2, we briefly present how THS works. Next, we present related work on Tor and THS 3. In Section 4, we explain the methodology we have followed to develop the SLR. Then, in Section 5 we present the results obtained and a discussion about them. Finally, we conclude the paper in Section 6.

2. Tor Hidden Services in a nutshell

Tor provides anonymity for a receiver or a service provider through what is known as Onion Services or Hidden Services (HS) [1].

For a user to communicate with a Hidden Service, twice as many nodes are required as those required to make an anonymous communication to a site outside the Tor network (Surface web), three chosen by the HS and three chosen by the user, where the third of these it becomes a rendezvous point (RP) that in simple words relays client's encrypted messages to the services and vice versa.

With a little more detail, a user who wants to connect or visit a hidden service must know his onion address, a string of 16 alphanumeric characters. Then, the client searches for information in a directory node (HSDir) that is basically a distributed hash table (DHT) that contains the descriptors of each HS, and finds an introduction point (IP) that the hidden service has previously defined. Simultaneously, client also chooses a node that will function as RP and builds a circuit to it. Once that client connects to one of these nodes, the IP transmits a notification to the hidden service, that contains the address of the rendezvous point. When the Hidden Service recognizes the RP, it creates a new circuit to it, and sends a new message to the client through the IP, in order to this closes the initial connection with the IP, and both of them keep the circuit built towards the RP, the communication is finally established. Both the IPs and RP have three intermediate jumps between them and user, and between them and the server. Therefore, none knows the identity of each part. Finally, the communication of the HS and user travels through 6 hops and use the RP as an intermediary, so they do not identify each other.

More details on Tor protocol, directory server, rendezvous protocol can be found in Tor's specifications [16] and in its design document [1].

 doi:10.20944/preprints202107.0600.v1

3. Related Work

In this section we present some related work on Tor. Namely, we will focus on Tor Hidden Services and previous surveys.

In the literature, to the best of our knowledge, there is no a comprehensive survey focused directly on THS. In spite of this fact, we can mention that there are several reviews of anonymous communication systems (ACS), two of them are focused on Tor and, finally, there is a survey focused on the deanonymization of hidden services.

We can start the review of the different surveys on ACS with the work of Ren and Wu [17], where they review the main techniques in the field of ACS, including Tor and also include a quick review of Hidden Services. In their work, the main topics covered address security issues and they conclude by exposing that the main vulnerabilities for a HS in Tor is the selection of the first and last node in the communication path.

Next, we can consider AlSabah and Golberg survey research on Tor [18], where they examine its design, identify weaknesses and deficiencies, present a classification of the directions that the research is taking in these areas. As for HSs, their greatest findings are again in security matters. In this work, they conclude by stating that improvements are needed in the design of THS. They mainly expose the problems with malicious services directory and problems with services that accept "anonymous" payments with Bitcoins. They also make a tour of documents that mention the ease of that a Command and Control (C&C) server may be protected behind THS.

Alidoost and Ruiz-Martínez [2] also published in 2017 a systematic review of the literature on anonymous communication systems. They conduct a review based on the collection of information in 7 academic databases and selected 203 papers for its analysis. Of these papers, 9 covered issues related to Hidden Services, where aspects related to security, main attacks, proposals to strengthen security, analysis and measurement of anonymity were addressed. Being the main topic the deanonymization of hidden services. Finally, from this SLR, to point out that, from all the papers analysed, the paper with more references was a paper related to THS [19], which is focused on trawling for HS. This show the research interest in this kind of services.

More recently, we can find the work done by Saleh et al. [5], where a review focused on Tor is presented. This review classifies the collected articles into three main groups: de-anonymization, path selection, and improvement and performance analysis. Of all the articles that analyze Tor, only 9% are related to Hidden Services and most of these are focused on de-anonymization, indicating that relays and traffic are the most susceptible factors.

As last general review on Tor we can mention Basyoni et al.'s [20] review where they analyse traffic analysis attacks on Tor and make an evaluation on how practical these attacks can be made on real-time in the Tor network. They point out that many of the de-anonymization attacks on THS, based on Tor's threat model, assume that there are one or multiple malicious Tor relays. This model is also considered in flow watermarking attacks. Its analysis provides a classification of 9 attacks analysed but they not cover into detail THS.

Finally, Nepal's et al. [21]' review, it is the only paper that makes a review a particular issue on THS: attacks schemes for revealing hidden server's identities. This paper is from 2015 and they analyse three kind of methods of attacks: manipulating Tor cells method, cell counting-based method, and, padding cell method. However, the paper is not focused on reviewing the literature.

As we have seen throughout this section, Tor has devoted an important attention in the research community. Within Tor Hidden Services is an important issue and generated attention. However, so far there is no survey that analyse this issue. Therefore, a survey on THS would be useful to compile the lessons learned from its appearance.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 July 2021 doi:10.20944/preprints202107.0600.v1

Table 1: Results provided by each database

Database	Number of documents
Google Scholar	1770
Web of Science	105
Scopus	339

4. Systematic Literature Review Methodology

Our review of the literature of Tor Hidden Services is based on a Systematic Literature Review (SLR). According to [22], SLRs start by defining a review protocol that specifies the research questions that will be addressed and the methods to develop the review. A SLR is based on a defined search strategy, which will be documented, and that aims to gather as much relevant literature as possible to apply it explicit inclusion and exclusion criteria to evaluate each primary research work.

In our SLR we have collected literature regarding Tor Hidden Services (THS) with the aim of answering the following research questions:

4.1. Research questions

Our SLR on THS aims to know the state of the art regarding THS, the research areas that are being studied, the reasons for its study and open issues. In order to satisfy this goal we have defined the following research questions:

1. What are the main research areas and the main findings regarding THS?
2. What are the limitations that research work present and in which research lines there is not enough research regarding THS?
3. Are there significant advances in THS during the latest years?
4. What are the main cited articles in the area of THS?
5. Is there any relationship between launching OR or updates in Tor project with the research made in THS?
6. What are the main future subjects or problems to be investigated regarding THS?

4.2. Search process

The processed followed is depicted in Figure 1. Next, we explain it into detail.

To obtain the different research works that have been used in this research work, we have retrieved articles from three databases: Google Scholar, Web of Science, and Scopus. We performed the query during the months of June/July 2019 using the following search terms:

- +Tor "Onion | Hidden Services | Server"
- Tor "Hidden Services" OR TOR "onion services"
- ("Tor" AND "Hidden services") OR ("Tor" AND "onion services")

These search terms have been adapted to the different databases.

Taking into account that THS were born in 2004 and that we are focused is quite specific, we decided that is not needed to apply a filter regarding the years to be covered. The results we have obtained after each search in the databases is shown in Table 1.

4.3. Inclusion and exclusion criteria

To the results we have obtained with the query in the different databases, we have applied a set of inclusion and exclusion criteria to decide which research works will be finally analysed.

As inclusion criteria we have considered:

- The document has to be published in a journal or a conference.
- The research presented in the paper has to be focused on THS.
- The article must be written in English.

As exclusion criteria we have considered:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 July 2021 doi:10.20944/preprints202107.0600.v1

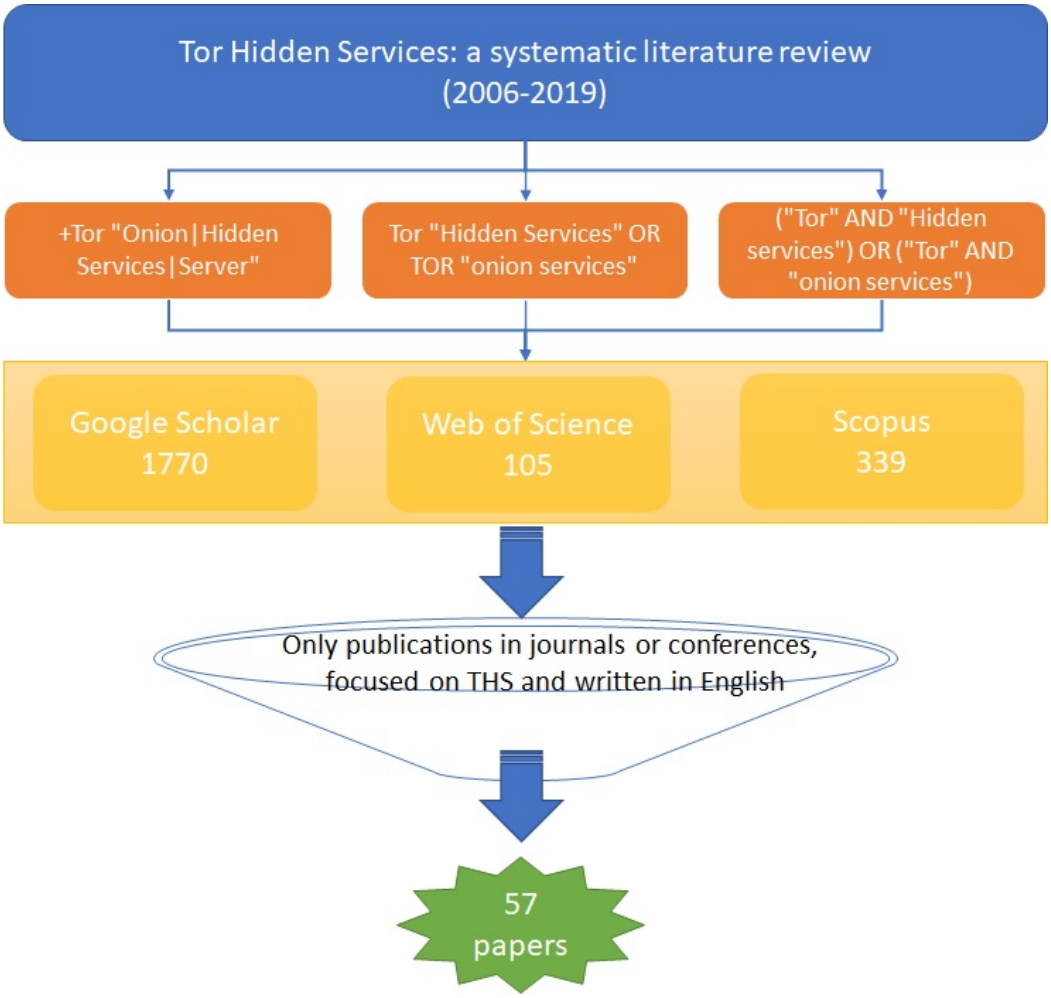

Figure 1. SLR process

- Book chapter.
- Patents.
- Citations.
- Research works that cover privacy and anonymity issues in anonymous communications systems that only cover THS in a general way.
- Technical reports.

After applying these criteria, the number of papers was reduced to 57. The list of articles is shown in Table A1 in Appendix A.

4.4. Research works and data analysis

Once we have applied all inclusion and exclusion criteria to the papers, with the resulting collection of research works, that is, 57 papers, we have gathered from each research work the following information: bibliographic information, number of references, number of cites, and type of research work (review or new research work). To each paper, we have also assigned a set of relevant keywords that allow us to identify main research topics.

When we collected all this data, we analysed the papers and answered the different research questions.

5. Results and Discussion

Based on the analysis made of the different papers, the answers for the different research questions are provided next.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 July 2021 doi:10.20944/preprints202107.0600.v1

Table 2: Classification of the papers in research topics

Research field	Research papers	N. of papers
Content classification	[23], [24], [25], [26], [9], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38]	17
Security	[39],[40], [41], [42], [43], [44], [45], [46], [47], [48],[49], [50], [51], [52], [53], [54], [55], [56], [57], [21], [58], [59], [60], [61], [62], [63], [64], [19]	28
Performance and deficiencies	[65], [66], [67], [68], [69]	5
Changes in the design	[69], [70], [71], [61], [72], [73]	6
Discovery and measurement	[74], [75], [32], [33], [34], [37], [38], [68], [23], [30], [31], [19]	12
Others	[76], [9]	2

Table 3: Classification of the papers in the security area

Focus of the research	Research papers	N. of papers
Attacks to THS	[40], [41], [42], [43], [44], [45], [47], [48], [50], [51], [52], [53], [54], [56], [57], [21], [58], [63], [49], [19]	20
Prevention against attacks	[39], [45], [55], [56], [59], [60], [61], [62], [64]	9

5.1. What are the main research areas and the main findings regarding THS?

In the analysis of the selected papers on THS, we have observed that they study them from different point of views and approaches. Namely, we have identified six main research areas (see Table 2): content classification, security, performance and deficiencies, changes in their design, discovery and measurement, and, finally, others that are not covered by previous ones. In Figure 2, we can observe the classification and the papers that contribute to more than area. Next, in the following sections, we analyse into more detail each area.

5.1.1. Security

As shown in Table 2, security is the main area of research regarding THS. In this area, on the one hand, there are articles that cover specific attacks for Denial-of-Service (DoS) or deanonymize THS, and on the other hand, others are focused on preventing any of these attacks. Table 3 show articles according to the focus of the paper.

In Table 2 we classify the papers covering security into these categories. As we can see most of them are focused on attacks. We can also point out that there are also papers that cover both possible attacks and mechanisms to prevent them as in[45,55,56].

Attacks to THS

Overlier and Syverson [45] are the first to document an attack on THS. They study the first version of the Tor hidden service protocol and claim that one of the main vulnerabilities in Tor is the selection of the first and last nodes in the communication path. Generally speaking, if an opponent can see the edges of a Tor circuit, then they can confirm who is communicating.

For this attack to be successful, the attacker has to control at least one node within the network and after multiple connection attempts, he or one of these malicious Onion Routers will be chosen as the first node in the rendezvous circuit established by the hidden service. From there, the attacker sends specific traffic patterns from the client to determine if one of its nodes was finally chosen as part of the circuit of interest. Once the attacker has

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 July 2021 doi:10.20944/preprints202107.0600.v1

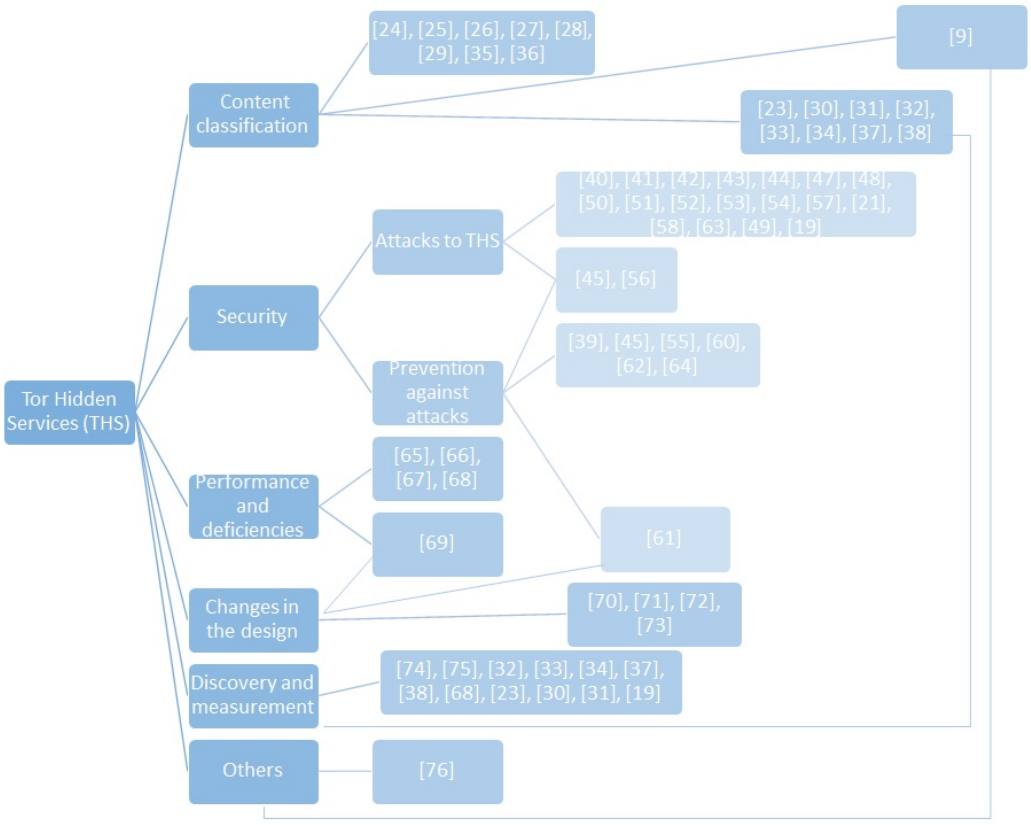

Figure 2. Relationship between papers and areas

identified that his node is the first incoming OR of the server, the identity of the server is revealed.

Using the same idea of controlling the server-side input node, Zhang et al. [40] use the web browser to access the target hidden service and generate specific traffic on the circuit. If the controlled OR detects the traffic pattern, the opponent can compare and correlate the time between the web accesses and the generated traffic on the compromised router and detect the actual IP of the server through the inbound node. Other attacks like this one based on the correlation of the traffic increase using the entry node are also described by Elices et al. [43], Ling [44], and Wang et al. [53,54].

Ma and Xu [47] present the same idea of using an entry node and traffic flow, but in this case to discover the identity of the client who wants to communicate with the server, assuming that this THS is malicious.

Another paper showing an attack based on the fact that it is possible to control the entry node is by [46] who mentions that Tor has a TCP/Tor congestion management mechanism. This mechanism causes a data source to stop sending messages, until it receives an acknowledgement of receipt of a previous message. The researchers manage to exploit this to create an attack, by embedding a watermark on the client side and a flow detector module on the inbound node they control, which monitors and analyses all the traffic passing through in order to detect the hidden server's IP.

Elices [42] and Kwon [63] demonstrate that a HS could be identified using fingerprint tracking, created from requests to the application server, disguised as a common user pattern sent by the client. In this attack they assume that the client can listen to the server's communication channel in order to try to detect the inserted fingerprint.

Another paper documenting a fingerprint attack [48] mentions that the larger a .onion site is, i.e. the more content it offers, the more susceptible it is to being tracked; while those that are more difficult to identify and tend to be small and dynamic. Pachenko et al. [50] also analyse fingerprinting techniques and propose their own attack. However, they

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 July 2021 doi:10.20944/preprints202107.0600.v1

conclude that neither their attack nor other existing approaches are adapted when applied in realistic environments.

Other types of attacks are raised [49,51] where researchers exploit the fact that the frequency of clocks in the equipment system is affected by the temperature of the CPU. The attacker via the Tor client periodically sends large amounts of data to the hidden server, which increases its CPU temperature in such a way as to generate a clock tilt pattern, while remotely measuring changes in the clock frequency of a set of candidate systems and trying to detect the matching pattern and thus deanonymize the HS.

The article **Trawling for Tor Services** addresses several points about hidden services, including a study of both DoS and de-anonymisation attacks. They mention that if the attacker can control access to HS descriptors, hidden service activity can be monitored or be completely inaccessible to clients. They also highlight that attacks to discourage large-scale hidden services do not require large amounts of resources. In their research they present the main vulnerabilities found and offer preventive measures, but which they themselves describe as superficial.

Matic et al. [52] discussed how information leaks in the configuration or content hosted on HSs can be used to reveal their location, reviewed a considerable number of .onion sites for URLs or email addresses that point to regular websites to determine if they could be hosted on the same server, investigated HTTP certificates to extract candidate IP addresses, searched for specific HS strings and used search engines to identify candidates hosting similar content.

Some authors [41,58] point out that if an attacker manages to position itself as a false HSDir, it can cause an eclipse attack for denial of service. In general terms, an eclipse attack is a mean of attacking decentralised networks, isolating and hiding a specific target, rather than attacking the entire network [58]. In the case of Tor, if an opponent can control a routing table, he/she can monopolise all the incoming and outgoing extensions of the victims, so the victim will be completely hidden. More specifically, if the first responsible directory is controlled, the attacker can monopolise or block all incoming HS nodes before customers can contact them.

Finally, Nepal et al. [21] review the deanonymisation attacks on hidden services raised in previous articles. These articles have also been addressed in this article concluding that all attack schemes need, at least, one client and one malicious guard node to deanomymize the hidden service. Thus, in all attacks, HS are forced to choose the compromised guard nodes as their entry nodes, if this does not happen none of the attempts to discover the identity of the server will be successful.

Prevention against attacks

Other related THS works present preventive measures to counteract the previously exposed vulnerabilities. In particular Overlier and Syverson [59], after exposing an attack using the entry nodes, show that it could be counteracted by using what they call "Valet Services" as an extension of the concept of hidden services. The basic idea is to add an additional layer of protection for the entry points, hiding them from users. These valet nodes are now known within the Tor network as guard nodes.

Other ideas for preventing traffic analysis attacks are being put forward by Beitollahi and Deconick [60], and by Yang et al. [61]. Both articles propose a different routing architecture. The first one proposes that the hidden service should also be part of the routers and that a closed circuit should be formed in the form of a ring, where data packets travel along the circuit together with dummy packets or filler packets in both directions of the circuit. In this case all the ORs in the ring see the traffic, but only the hidden server can understand it because only the HS can decode all the encryption layers. The second article presents a cell scheme based on multiple routes that exploits flow mixing and merging to distort or destroy the traffic patterns inserted by an attacker.

Hopper [39] conducted a study on Tor's challenges to protect hidden services from botnet abuse, which can lead to poor network performance due to increased load on nodes. Solutions to this problem include reusing failed partial circuits and isolating circuits for

 doi:10.20944/preprints202107.0600.v1

hidden services. In the latter case, the author states that if a mechanism that allows ORs to recognize that a cell carries hidden service traffic is introduced, a mean can be provided to protect the rest of the system from the effects of this traffic by scheduling priority or simple isolation.

Another interesting direction is the creation of "Honey Onions" (Honeypots + Onion Services) to identify malicious directories. These Honions, as they are called in the literature, are HS whose link (.onion) is not public anywhere, which means that only the service administrator knows about it. Therefore, when there is a possible access it is assumed to be a malicious HSDir connection attempt.

On the other hand, works such as [55] and [56] approach the subject from a forensic perspective. The researchers expose the creation of a persistent fingerprint through a flow of requests that can be recorded on a computer through a hidden service, so that, at a later time, it can be retrieved and used as an evidence that a physical machine hosted a particular content even after this content has been removed.This attack does not actually locate the anonymous server, but authorities can use it as almost foolproof evidence of a criminal provider's guilt.

5.1.2. Content classification

The second topic most often addressed by researchers is the classification of content hosted or offered in hidden services. It is interesting to mention that, as far as we know, the first article that deals with this topic was written in 2013 although Tor was offering hidden services since 2004.

In this first paper on content classification [35], Guitton used three databases that listed an extensive amount of hidden services. They obtained a total of 1,171 individual entries, which they then reviewed and manually classified into 23 categories. They conclude that 45% of all available THS host unethical or illegal content.

Biryukov et al. [34] also did a search for .onion addresses and as a first step found 8,153 addresses, but of these, only 1,813 were functional services and these were sorted into 18 categories. The results are similar Guitton's research we explained above, since they found that 44% are dedicated to illegal services, while the remaining 56% are dedicated to different subjects, but which do not pose a real danger, for example, policy and anonymity forums, information resources, pages similar to WikiLeaks, as well as pages that provide different unknown services such as mail or anonymous hosting.

A different way of classification is presented by different authors [25,29,36]. They apply data mining techniques based on classification models to over a thousand hidden Tor services, to model their thematic organization and linguistic diversity. Their results indicate that most of the hidden services display illegal or controversial content in their dataset. In contrast, Savage and Owen [32] return to manual study and classification, suggesting that given the variety and complex technical nature of some content, automatic classifiers would be insufficient, due to the difficulty of interpreting the context completely. Other documentation of a manual classification reviewing 3480 HS is by Faizan and Khan [28], but they claim that only 38% of the servers found offer illegal services.

On the other hand, Biswas, Fidalgo, and Alegre[24] present ATOL (Automated Tool for Onion Labeling) a classification tool, which includes a new discovery mechanism using keywords and a classification framework using image recognition, which is also used and replicated by Ghosh et al. [26]. Another article showing a classification mechanism using image recognition is by Fidalgo et al [27]. They apply improved techniques, selecting only the regions with the most outstanding information, through what is known, in machine learning, as the Bag of Visual Word. These investigators focus on classifying services that host suspicious or illegal content, in an attempt to collaborate with entities that combat crimes involving human trafficking and child pornography.

Other scientific articles show that the vast majority of hidden services are offered in English, approximately 73.28%, while only 2.14% are offered in Spanish [31].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 July 2021 doi:10.20944/preprints202107.0600.v1

Many more authors mention methods for classifying the content offered in hidden services, but their articles are based on the discovery of these services before classifying them. Therefore, they are mentioned below in the next category.

5.1.3. Discovery and measurement

This section includes articles that retrieve the maximum number of .onion addresses and try measuring the size of the Tor hidden network from different perspectives.

The first document is from 2009 [74] and their authors (Betzwieser2009 et al.) are the first researchers to mention that the dark network is widely interconnected, manually locating 39 hidden services, including directory sites, and from these 39, they find 20,499 more HSs. Later, this concept is taken up again in 2017 by Bernaschi et al. [23] but they show that the THS that offer specific services contain few or no references to other THS, while those that advertise and link to a large number of other pages, are much more likely to be known and accessible from public websites. Later, Sanchez et al. [31] re-analyzed the content of hidden services and found that there is a clear connection between the normal network or surface web and many Hidden Services, for example, claiming that more than 20% of the domains in Tor imported resources from the surface web. Furthermore, they confirm the results obtained by Betzwieser [74] and show that approximately 90% of onion services are interconnected.

Owen and Savage [32] attempted to discover HS operating 40 ORs over 6 months, each with a bandwidth of approximately 50 kB/s and leaving them active continuously for 25 hours, with the intention that their nodes or one of them would be eligible to obtain the HSDir indicator and be able to recover the maximum amount of .onion addresses possible. In the study period, they were able to observe approximately 80,000 HSs. This is the first article that exposes the short life of the hidden services, many of which were observed to exist at most a few weeks before their closure. In all, only 15% of the MSM they found persisted through their six-month observation period. Liu et al. [68] also take up this technique again, but doing a cost evaluation, they quantify the relationship between resources consumed and hidden services collected.

Taking into account Owen and Savage's work that exposes the short life of the services in Tor, Owenson [75] indicates that the Dark Web is much smaller than many people think, since most Onion Services do tend to be ephemeral, so they do not all coexist at once.

Other research works [33,37] are based on the discovery of Hidden Services through search engines with specific keywords that allow them to extract .onion addresses, for example, from sites on the Dark Web such as wikis or using common Surface Web search engines such as Google. Thus, Bernaschi et al. in [33] emphasize that during their collection process, they only reached 25% to 35% of the total number of hidden services that Tor says are published daily. Furthermore, Li et al. [37] obtained 173,667 .onion addresses but only 4,857 of these were online.

Finally, some articles focus on finding the most popular services on the Tor network. Most of them do so by locating the most referenced .onion addresses, but researchers who had access to HsDir [34] estimate the popularity of hidden services by looking at the request rates of descriptors. The four documents differ among the most popular service, one of which ranks Silk Road [19], which was shut down by the FBI in 2013, as the number one; the same authors also suggest that the top 10 also includes services related to BotNets. Some similar results are shown in the following study, which gives this place to an HS belonging to a zombie network structure called "Goldnet" [34]. More generally, later Savage and Owen [32] found that the first place is given to MSM related to abuse, although for ethical reasons they do not specify what type of abuse or make their .onion address public. The latest study, which is from 2019, shows that the most popular service is dedicated to drugs and narcotics [30]. Although none of them coincide with which is truly the most popular and influential HS, it can be inferred that the best-positioned services are those related to illicit content.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 July 2021 doi:10.20944/preprints202107.0600.v1

To a lesser extent, the research analyzed covers issues related to Hidden Services performance and the analysis of specific parts of the protocol that they propose improvements later. These studies are detailed below.

5.1.4. Performance and Deficiencies

Loesing et al.[67] measured latency during connection to a hidden service with a special focus on general response times and found that connection establishment when using a broadband access network, took an average of 24 seconds. Furthermore, the study revealed that most of the time is spent establishing the connection to the meeting and introduction points. Lenhard et al. [65] make this same measurement but in low-bandwidth environments. They attribute the Tor bottleneck to downloading relay descriptors and building circuits. Their findings suggest an increase in the value of the timeout to avoid repeated retransmissions.

Another article, claiming that because of Tor's unique features, HS can be used to control botnets [66], tests the implementation of new proxy-based botnet architectures that benefit from the anonymity provided by the Tor network to disguise its Command and Control infrastructure. To do this, each bot creates a Tor HS, thus acquiring an .onion address, to allow communication with the rest of the botnet. They mention that while this option can suffer from Tor's high latency and complex administration, it benefits greatly from stealth and anonymity, not to mention that each HS can migrate to another physical location while maintaining the same unique onion address.

5.1.5. Changes in the design

Other authors present improvements or changes in Tor design that might strengthen or end the main shortcomings, for example, point to protocol changes to establish faster connections to hidden services. Their suggestions include reducing the number of nodes involved in the process [45,69,73], which should lead to a decrease in connection establishment times.

This category also includes documents that propose changes to the circuit and data routing, specifically designed to avoid traffic recognition attacks [60,61] (see Section 5.1.1).

Another proposal is to change the .onion address from a Base32 pseudonym to a decentralized DNS that can be secure, searchable, and, above all, readable [71,72].

5.1.6. Others

The two remaining documents have been included in this category. The first work, by He et al. [9], make a study of hidden services from a legal perspective, concluding that while the Tor Project is not passing over any law, HS facilitates the creation and proliferation of services that clearly violate the laws of any country.

At the other extreme, Winter et al.[76] explain how Tor users perceive, use, and administer onion services, present what the main obstacles are for the average user, collect mental models, and obtain real usage patterns through interviews and questionnaires. Among their main findings they mention that many of the people who use the services offered by Tor were not even aware that there is a significant difference to the normal web. Among other findings, they expose the limited ways to discover the existence of onion services and usability problems, which Tor is actually improving with new versions of the design.

5.2. What are the limitations that research work present and in which research lines there is not enough research regarding THS?

Very few of the investigations studied coincide with each other, there are usually many differences, for example, in the number of hidden servers found, while some discover 20,499 [74] others manage to find approximate 80 thousand HSs with 45 thousand of these being persistent [32]. Therefore, there are many inconsistencies between the results obtained.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 July 2021 doi:10.20944/preprints202107.0600.v1

Another deficiency found is the lack of evidence in a realistic environment. Many of the attacks are considered successful when tested in simulated environments and are rarely brought to the real Tor network [19,29,42,53,54,56,61,73,75], among others. Panchenko et al. [50] present a fingerprinting attack that is theoretically fully functional but, when was brought to the real environment, it has a very low recognition rate. Therefore, it cannot be guaranteed that the other approaches will be successful, even if they work under simulated environments.

The problem with the research found in front of Tor's network of hidden servers is that it is a very volatile network, and the results of the studies can be confirmed or refuted completely overnight. While many studies claim that the Dark Web is governed by illegal services, others claim that the vast majority of HS host content that may be sensitive in nature but is not unethical.

On the other hand, there is no study so far that validates and compares onion service discovery methods. If we compare the statistics provided by The Tor Project vs. the number of HS discovered, most of the scans only reached between 20% and 35% of the official total presented by Tor, and there is no evidence why the remaining number cannot be reached. On the other hand, until version 2 of Tor Onion Services, there has been no significant change in the protocol that deals with HSs, and no document addresses the evolution of Tor, the increase in users, and the advance in the mechanism that attackers use in the face of an architecture that had evolved very little.

5.3. Are there significant advances in THS during the latest years?

Figure 3 shows the number of research papers published by year, from the first in 2006 to the last in 2019 shortly before the completion of this work. As seen in the Figure 3, it is evident that researchers have shown greater interest as of 2015, with a significant increase in 2017.

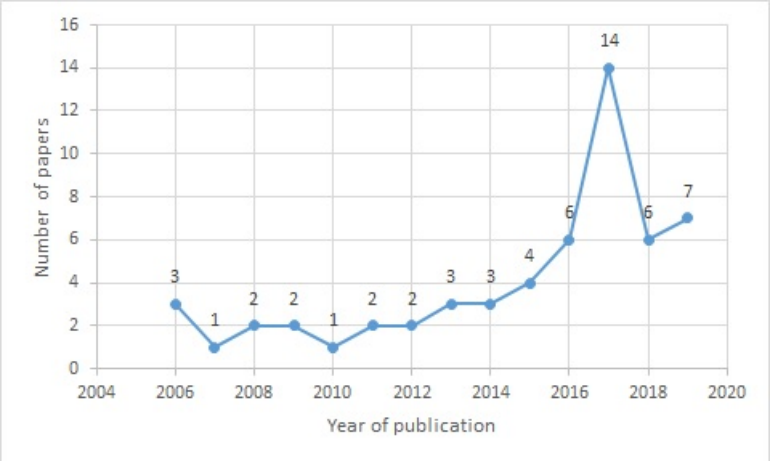

Figure 3. Number of publications by year

The biggest advance in research in the last three years has been in the area of content classification of hidden services, such as the application of improved Machine Learning techniques for image recognition that allow better classification with fewer margins of error and crawling and .onion address discovery techniques, and to a lesser extent, security issues.

Although the techniques used present significant advances over older articles, there is still no way to collect and analyze highly effective content. This issue may be due to the slow access speed of the Tor browser that makes it difficult to correctly observe the dynamics of Hidden Services.

In terms of security, the last published article explains how an eclipse attack can cause the denial of service of an HS [41] but the attack mechanism still involves the control of an HSDir. Therefore, there is nothing new in comparison with the other documents.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 July 2021 doi:10.20944/preprints202107.0600.v1

As for proposed changes to the Tor design, the articles, which can be considered that they cover new issues, are related to changing .onion addresses to more readable domain names. However, the fact that the addresses assigned to each hidden service are a 16-character alphanumeric pseudo-domain has a security objective since it makes them difficult to trace.

5.4. What are the main cited articles in the area of THS?

As mentioned in the methodology, the documents were obtained from three databases: Web of Science, SCOPUS, and Google Scholar. The latter is the search engine that gave the most results, including a large number of repeated results that had already been selected from the other two databases. Therefore, to homogenize the number of citations that each article has received, Google Scholar was used as the source and only in its absence were the other two used.

The three most cited articles are focused on security [19,44,49], more specifically on de-anonymization. Additionally, the third one presents an analysis in terms of HS discovery. Of the articles [19,44,49], the total number of citations is 363, 270, and 169, respectively. It is worth mentioning that the first two were published in 2006, while the last one, which is the least cited, was written in 2013, which could be explained because it is the most recent.

In the latter, Biryukov et al. [19] analyze the weaknesses in hidden services that can be exploited by attackers to detect, measure, and de-anonymize HSs, and describe several relevant attack methods, based on traffic confirmation techniques and the fact that an attacker can easily include malicious nodes in the Tor network. They also claim that large-scale hidden service de-anonymization attacks are practically possible with only a moderate amount of resources, exploit Tor vulnerabilities to collect hidden service descriptors, classify their content, and detect the most popular hidden services. In short, this is one of the most comprehensive documents about Tor and its onion services.

5.5. Is there any relationship between launching OR or updates in Tor project with the research made in THS?

According to the Tor Project [77], in 2008, the idea of making Tor more accessible to people with less technical knowledge was introduced. So they started developing the Tor Browser, a tool beyond the Tor proxy, which from 2010 positioned Tor as an instrumental tool for protecting people's identity online and access to critical resources, social networks, and blocked sites. The versions of the Tor protocol released have been consistent since then and to this day.

In 2008 the most significant updates are to the performance of the Hidden Services descriptors. Then, in 2009, a problem with the descriptors was again corrected that prevented customers from using a cached service descriptor that was more than 15 minutes old and also made it impossible to search for a new one, because there was already one in the cache. Later, in 2011, a change was made to improve bandwidth consumption by having directory servers publishing small summaries of router's information, which customers can use in place of regular HS descriptors. In 2012, a new, fairly comprehensive stable version is released, Tor 0.2.3.25, where they reduce directory overload significantly, implementing micro-descriptors, developing a new TLS link protocol that can better resist fingerprint attacks, and improve scalability for hidden services, among many other stability, security, and privacy fixes.

As can be seen in 3, in 2013 there is a small increase in published articles, including one of the most frequently cited documents (see section 5.4). In contrast, the Tor project did not release any new stable version in that year, so it cannot be said that there is a relationship between the changes in Tor and the written articles, since there is no constant in the four research papers published in 2013 either.

In 2014, Tor is again updated to a stable version that corrects errors that affected the consistency and speed of connection to hidden services. Then, that same year they add a new feature where the positions of the hash table are derived from a random cryptographic key, to correct the vulnerabilities that could cause denial of service attacks due to hash

 doi:10.20944/preprints202107.0600.v1

flooding. Bearing this in mind, of all the documents analysed, none presents an attack of this type.

In 2015, the 0.2.7 series is launched, which again improves the performance of persistent hidden services and makes significant changes to the HS entry nodes, forcing them to use more than one EntryNode to avoid a guard discovery attack.

From this update, a Hidden Service must choose between a maximum of 10 and 3 guard nodes and they eliminate the adaptive algorithm to choose the number of entry points depending on the number of connections the HS sees; with this in mind, it is worth noting that from 2015 onwards, of the remaining 37 documents, only 3 mention attacks where it is necessary to control the entry node.

Tor was updated to 0.2.9 in 2016. This version made several previously optional security features mandatory, and most importantly, includes a single-hop Hidden Service mode to optimise .onion services that really do not want to be hidden, although this option is never considered by researchers.

In 2017 there is an increase of more than 100% in the number of publications, coinciding with this year, Tor Project presents the next generation of onion services at DEF CON 25. Although version 2 is still the most widely used version so far, due to compatibility issues. While a major change was displayed for the Hidden Services, here too there is no recognisable relationship between Tor updates and the amount of research done in 2017; although one recognisable point is that according to Tor metrics on March 18, 2017 Tor experienced the largest peak in the number of .onion addresses with a total of 124,958 unique addresses [8].

Towards the latest stable versions of Tor released to 2018 there are no changes other than the new version of the hidden services, the creators boast a new, much more extensible protocol with a cleaner code base, but the main features of this new generation comprise several aspects: new encryption algorithms, redesign of the directory system to defend against information leaks and reduce the overall attack surface, and most importantly, addresses. In addition, the new generation of fully private addresses with 54 characters instead of 16 [77,78] will largely prevent the malicious discovery of Hidden Services.

Although the deployment of the new generation of the Hidden Services promises an increase in the number of investigations or at least new topics to cover, when we finished our queries, none of the articles studied include this new version, even though they were published some time after the official announcement by the Tor project.

5.6. What are the main future subjects or problems to be investigated regarding THS?

As we have already mentioned, hidden services are often volatile, so it can be said that researchers have an obstacle when it comes to working on discovering and measuring the actual size of the Tor network. This would therefore be one of the points to cover, realistic approaches to how this ecosystem works and whether there is a feasible option for obtaining metrics without violating anonymity beyond those provided by Project Tor itself.

On the other hand, the great revolution of the new generation of hidden services leaves a great horizon open which must be covered from all perspectives, addressing issues such as how this new version really works, whether all previous vulnerabilities has been removed, or whether the fact that new .onion addresses are more private will help the proliferation of increasingly illicit hidden services, among many other points.

Finally, much of the future work should continue to address security issues. While the developers of Tor, and the huge community behind them, work every day to eliminate vulnerabilities and security leaks, there are also an alarming number of malicious agents taking advantage of every update to exploit potential weaknesses to, above all, discourage users and servers.

 doi:10.20944/preprints202107.0600.v1

6. Conclusions

Tor is currently the largest anonymity network, and in addition to this, the Tor protocol lends itself to in-depth research, because all its development is carried out publicly and the source code and specifications of each version are openly available.

In this article, we have presented a systematic review of the literature on Tor's Hidden Services. We found that most of the papers are focused on security, either presenting attacks or proposing protection methods. Secondly, researchers have focused on discovering at the directory level as many .onion addresses as possible and with this to make a classification of the content hosted on these anonymized services, to a lesser extent research formulates changes in the protocol design to achieve faster connections to hidden services or make circuits more secure.

We also find out that many of the researches exposing attacks have not been tested in realistic Tor environments. Therefore, they have wide margins of failure. Another problem encountered is the inconsistency in the number of services reached or discovered using different techniques; the hidden service descriptors are stored in a distributed manner. Thus, there is no central entity storing the complete list of .onion, the number of addresses found may be far from exhaustive and realistic. Therefore, it is not possible to give a size to Tor's hidden server network beyond the metrics themselves.

Finally, emphasis is placed on the lack of research that studies the new generation of Hidden Services that, although, is not the one used by most of the deployed services, according to the Tor Project it is expected that in the following years the HS version 2 will disappear to give way to a much more secure ecosystem.

As future future work, a deeper evaluation of the evolution of the attacks presented and their degree of success or failure can be made, taking into account how the Tor project has closed vulnerabilities that many of these attacks exploited.

In the medium term, this SLR should be taken up again in search of new research that includes the new generation of Onion Services and make a comparison with the research presented here.

Funding: This work was funded by the European Commission's H2020 Programme under the Grant Agreement Number 830929 and the Spanish Ministry of Science, Innovation and Universities, FEDER funds, under grant numbers RTI2018-095855-B-I00 and TIN2017-86885-R.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 27 July 2021 doi:10.20944/preprints202107.0600.v1

Table A1. List of all papers analysed

Number	Ref	Authors	Year
1	[45]	Overlier and Syverson	2006
2	[49]	Murdoch	2006
3	[59]	Øverlier and Syverson	2006
4	[67]	Loesing *et al.*	2008
5	[51]	Zander and Murdoch	2008
6	[65]	Lenhard *et al.*	2009
7	[55]	Shebaro *et al.*	2010
8	[56]	Elices *et al.*	2011
9	[40]	Lu Zhang *et al.*	2011
10	[42]	Elices and Perez-Gonzalez	2012
11	[60]	Beitollahi and Deconinck	2012
12	[43]	Elices and Perez-Gonzalez	2013
13	[44]	Ling *et al.*	2013
14	[35]	Guitton	2013
15	[19]	Biryukov *et al.*	2013
16	[34]	Biryukov *et al.*	2014
17	[36]	Spitters *et al.*	2014
18	[39]	Hopper	2014
19	[52]	Matic *et al.*	2015
20	[21]	Nepal *et al.*	2015
21	[61]	Yang and Li	2015
22	[63]	Kwon *et al.*	2015
23	[64]	Nurmi *et al.*	2016
24	[53]	Wang *et al.*	2016
25	[54]	Wang *et al.*	2016
26	[32]	Savage and Owen	2016
27	[37]	Li *et al.*	2016
28	[62]	Sanatinia and Noubir	2016
29	[23]	Bernaschi *et al.*	2017
30	[24]	Biswas *et al.*	2017
31	[25]	Nabki *et al.*	2017
32	[26]	Ghosh *et al.*	2017
33	[31]	Sanchez-Rola *et al.*	2017
34	[47]	Ma and Xu	2017
35	[48]	Overdorf *et al.*	2017
36	[50]	Panchenko *et al.*	2017
37	[57]	Sanatinia and Noubir	2017
38	[69]	Meng *et al.*	2017
39	[66]	Anagnostopoulos *et al.*	2017
40	[46]	Iacovazzi *et al.*	2018
41	[68]	Liu *et al.*	2018
42	[71]	Meng *et al.*	2018
43	[76]	Winter *et al.*	2018
44	[9]	He *et al.*	2019
45	[27]	Fidalgo *et al.*	2019
46	[28]	Faizan and Khan	2019
47	[29]	Takaaki and Atsuo	2019
48	[30]	Al-Nabki *et al.*	2019
49	[33]	Bernaschi *et al.*	2019
50	[38]	Park *et al.*	2019
51	[41]	Tan *et al.*	2019
52	[70]	Øverlier and Syverson	2007
53	[72]	Victors *et al.*	2016
54	[73]	Liang and Liu	2018
55	[74]	Betzwieser *et al.*	2009
56	[75]	Owenson *et al.*	2018
57	[33]	Bernaschi *et al.*	2019

References

1. Dingledine, R.; Mathewson, N.; Syverson, P. Tor: The Second-Generation Onion Router. Technical report, Naval Research Lab, Washington DC., 2004.
2. Alidoost Nia, M.; Ruiz-Martínez, A. Systematic literature review on the state of the art and future research work in anonymous communications systems. *Computers & Electrical Engineering* **2018**, *69*, 497–520. doi:10.1016/j.compeleceng.2017.11.027.
3. Basyoni, L.; Fetais, N.; Erbad, A.; Mohamed, A.; Guizani, M. Traffic Analysis Attacks on Tor: A Survey. 2020, p. 183–188. doi:10.1109/ICIoT48696.2020.9089497.
4. Cambiaso, E.; Vaccari, I.; Patti, L.; Aiello, M. Darknet Security: A Categorization of Attacks to the Tor Network. ITASEC, 2019.
5. Saleh, S.; Qadir, J.; Ilyas, M.U. Shedding Light on the Dark Corners of the Internet: A Survey of Tor Research. *Journal of Network and Computer Applications* **2018**, *114*, 1–28. doi:10.1016/j.jnca.2018.04.002.
6. De la Cadena, W.; Mitseva, A.; Hiller, J.; Pennekamp, J.; Reuter, S.; Filter, J.; Engel, T.; Wehrle, K.; Panchenko, A. TrafficSliver: Fighting Website Fingerprinting Attacks with Traffic Splitting. Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security, 2020, pp. 1971–1985.
7. Pulls, T. Towards Effective and Efficient Padding Machines for Tor. *arXiv:2011.13471 [cs]* **2020**. arXiv: 2011.13471.
8. Tor Project. Tor Metrics, 2021.
9. He, S.; He, Y.; Li, M. Classification of Illegal Activities on the Dark Web. ACM International Conference Proceeding Series, 2019, pp. 73–78.
10. Perry, M. This is What a Tor Supporter Looks Like: Edward Snowden | Tor Blog, 2015.
11. Alharbi, A.; Faizan, M.; Alosaimi, W.; Alyami, H.; Agrawal, A.; Kumar, R.; Khan, R.A. Exploring the Topological Properties of the Tor Dark Web. *IEEE Access* **2021**, *9*, 21746–21758. doi:10.1109/ACCESS.2021.3055532.
12. Ball, M.; Broadhurst, R. *Data Capture and Analysis of Darknet Markets*; Number ID 3344936, 2021. doi:10.2139/ssrn.3344936.
13. Platzer, F.; Schäfer, M.; Steinebach, M. Critical Traffic Analysis on the Tor Network. *Journal of Cyber Security and Mobility* **2021**, pp. 133–160–133–160. doi:10.13052/jcsm2245-1439.1015.
14. David, H.; Sebastian, P.; Kai, R. Explaining the Technology Use Behavior of Privacy-Enhancing Technologies: The Case of Tor and JonDonym. *Proceedings on Privacy Enhancing Technologies* **2020**, *2020*, 111–128. doi:10.2478/popets-2020-0020.
15. Kitchenham, B.; Charters, S. *Guidelines for performing Systematic Literature Reviews in Software Engineering*; Number EBSE-2007-01, 2007.
16. Tor Project. Tor Design Documents, 2021.
17. Ren, J.; Wu, J. Survey on anonymous communications in computer networks. *Computer Communications* **2010**, *33*, 420–431. doi:10.1016/j.comcom.2009.11.009.
18. Alsabah, M.; Goldberg, I. Performance and Security Improvements for Tor. *ACM Computing Surveys* **2016**, *49*, 1–36. doi:10.1145/2946802.
19. Biryukov, A.; Pustogarov, I.; Weinmann, R.P. Trawling for tor hidden services: Detection, measurement, deanonymization. *Proceedings - IEEE Symposium on Security and Privacy* **2013**, pp. 80–94. doi:10.1109/SP.2013.15.
20. Basyoni, L.; Fetais, N.; Erbad, A.; Mohamed, A.; Guizani, M. Traffic Analysis Attacks on Tor: A Survey. 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), 2020, pp. 183–188. doi:10.1109/ICIoT48696.2020.9089497.
21. Nepal, S.; Dahal, S.; Shin, S. Deanonymizing schemes of hidden services in tor network: A survey. 2015 International Conference on Information Networking (ICOIN). IEEE, 2015, pp. 468–473. doi:10.1109/ICOIN.2015.7057949.
22. Kitchenham, B. Procedures for Performing Systematic Reviews. Technical report, 2004.
23. Bernaschi, M.; Celestini, A.; Guarino, S.; Lombardi, F. Exploring and Analyzing the Tor Hidden Services Graph. *ACM Transactions on the Web* **2017**, *11*, 1–26. doi:10.1145/3008662.
24. Biswas, R.; Fidalgo, E.; Alegre, E. Recognition of service domains on TOR dark net using perceptual hashing and image classification techniques. 8th International Conference on Imaging for Crime Detection and Prevention (ICDP 2017). Institution of Engineering and Technology, 2017, pp. 7–12. doi:10.1049/ic.2017.0041.
25. Nabki, M.W.A.; Fidalgo, E.; Alegre, E.; De Paz, I. Classifying illegal activities on tor network based on web textual contents. *15th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2017* - **2017**, *1*, 35–43.
26. Ghosh, S.; Das, A.; Porras, P.; Yegneswaran, V.; Gehani, A. Automated Categorization of Onion Sites for Analyzing the Darkweb Ecosystem. 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '17; ACM Press: New York, New York, USA, 2017; pp. 1793–1802. doi:10.1145/3097983.3098193.
27. Fidalgo, E.; Alegre, E.; Fernández-robles, L. Classifying Suspicious Content in Tor Darknet through Semantic Attention Keypoint Filtering $. *Digital Investigation* **2019**.
28. Faizan, M.; Khan, R.A. Exploring and analyzing the dark Web: A new alchemy. *First Monday* **2019**, *24*. doi:https://doi.org/10.5210/fm.v24i5
29. Takaaki, S.; Atsuo, I. Dark web content analysis and visualization. IWSPA 2019 - Proceedings of the ACM International Workshop on Security and Privacy Analytics, co-located with CODASPY 2019, 2019, Vol. March 2019, pp. 53–59. doi:10.1145/3309182.3309189.
30. Al-Nabki, M.W.; Fidalgo, E.; Alegre, E.; Fernández-Robles, L. ToRank: Identifying the most influential suspicious domains in the Tor network. *Expert Systems with Applications* **2019**, *123*, 212–226. doi:10.1016/j.eswa.2019.01.029.

31. Sanchez-Rola, I.; Balzarotti, D.; Santos, I. The Onions Have Eyes: A Comprehensive Structure and Privacy Analysis of Tor Hidden Services. Proceedings of the 26th International Conference on World Wide Web - WWW '17; ACM Press: New York, New York, USA, 2017; pp. 1251–1260. doi:10.1145/3038912.3052657.
32. Savage, N.; Owen, G. Empirical analysis of Tor Hidden Services. *IET Information Security* **2016**, *10*, 113–118. doi:10.1049/iet-ifs.2015.0121.
33. Bernaschi, M.; Celestini, A.; Guarino, S.; Lombardi, F.; Mastrostefano, E. Spiders like Onions: on the Network of Tor Hidden Services. The World Wide Web Conference on - WWW '19; ACM Press: New York, New York, USA, 2019; pp. 105–115. doi:10.1145/3308558.3313687.
34. Biryukov, A.; Pustogarov, I.; Thill, F.; Weinmann, R.P. Content and Popularity Analysis of Tor Hidden Services. 2014 IEEE 34th International Conference on Distributed Computing Systems Workshops. IEEE, 2014, pp. 188–193. doi:10.1109/ICDCSW.2014.20.
35. Guitton, C. A review of the available content on Tor hidden services: The case against further development. *Computers in Human Behavior* **2013**, *29*, 2805–2815. doi:10.1016/J.CHB.2013.07.031.
36. Spitters, M.; Verbruggen, S.; van Staalduinen, M. Towards a Comprehensive Insight into the Thematic Organization of the Tor Hidden Services. 2014 IEEE Joint Intelligence and Security Informatics Conference. IEEE, 2014, pp. 220–223. doi:10.1109/JISIC.2014.40.
37. Li, K.; Liu, P.; Tan, Q.; Shi, J.; Gao, Y.; Wang, X. Out-of-band discovery and evaluation for tor hidden services. Proceedings of the 31st Annual ACM Symposium on Applied Computing - SAC '16; ACM Press: New York, New York, USA, 2016; pp. 2057–2062. doi:10.1145/2851613.2851798.
38. Park, J.; Mun, H.; Lee, Y. Improving Tor Hidden Service Crawler Performance. *DSC 2018 - 2018 IEEE Conference on Dependable and Secure Computing* **2019**, pp. 1–8. doi:10.1109/DESEC.2018.8625103.
39. Hopper, N. Challenges in Protecting Tor Hidden Services from Botnet Abuse. International Conference on Financial Cryptography and Data Security. Springer, Berlin, Heidelberg, 2014, pp. 316–325. doi:10.1007/978-3-662-45472-5_21.
40. Lu Zhang.; Junzhou Luo.; Ming Yang.; Gaofeng He. Application-level attack against Tor's hidden service. 2011 6th International Conference on Pervasive Computing and Applications. IEEE, 2011, pp. 509–516. doi:10.1109/ICPCA.2011.6106555.
41. Tan, Q.; Gao, Y.; Shi, J.; Wang, X.; Fang, B.; Tian, Z. Toward a Comprehensive Insight Into the Eclipse Attacks of Tor Hidden Services. *IEEE Internet of Things Journal* **2019**, *6*, 1584–1593. doi:10.1109/JIOT.2018.2846624.
42. Elices, J.A.; Perez-Gonzalez, F. Fingerprinting a flow of messages to an anonymous server. 2012 IEEE International Workshop on Information Forensics and Security (WIFS). IEEE, 2012, pp. 97–102. doi:10.1109/WIFS.2012.6412632.
43. Elices, J.A.; Perez-Gonzalez, F. Locating Tor hidden services through an interval-based traffic-correlation attack. 2013 IEEE Conference on Communications and Network Security (CNS). IEEE, 2013, pp. 385–386. doi:10.1109/CNS.2013.6682740.
44. Ling, Z.; Luo, J.; Wu, K.; Fu, X. Protocol-level hidden server discovery. 2013 Proceedings IEEE INFOCOM. IEEE, 2013, pp. 1043–1051. doi:10.1109/INFCOM.2013.6566894.
45. Overlier, L.; Syverson, P. Locating hidden servers. 2006 IEEE Symposium on Security and Privacy (S&P'06). IEEE, 2006, pp. 15 pp.–114. doi:10.1109/SP.2006.24.
46. Iacovazzi, A.; Sarda, S.; Elovici, Y. Inflow: Inverse Network Flow Watermarking for Detecting Hidden Servers. IEEE INFOCOM 2018 - IEEE Conference on Computer Communications. IEEE, 2018, pp. 747–755. doi:10.1109/INFOCOM.2018.8486375.
47. Ma, Y.; Xu, X. Locating tor's hidden service clients based on protocol feature. 2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). IEEE, 2017, pp. 282–285. doi:10.1109/ITNEC.2017.8284989.
48. Overdorf, R.; Juarez, M.; Acar, G.; Greenstadt, R.; Diaz, C. How Unique is Your .onion? An Analysis of the Fingerprintability of Tor Onion Services. 2017 ACM SIGSAC Conference on Computer and Communications Security - CCS '17; ACM Press: New York, New York, USA, 2017; pp. 2021–2036. doi:10.1145/3133956.3134005.
49. Murdoch, S.J. Hot or Not : Revealing Hidden Services by their Clock Skew Categories and Subject Descriptors. 13th ACM conference on Computer and communications security, 2006, pp. 27–36.
50. Panchenko, A.; Mitseva, A.; Henze, M.; Lanze, F.; Wehrle, K.; Engel, T. Analysis of Fingerprinting Techniques for Tor Hidden Services. Proceedings of the 2017 on Workshop on Privacy in the Electronic Society - WPES '17; ACM Press: New York, New York, USA, 2017; pp. 165–175. doi:10.1145/3139550.3139564.
51. Zander, S.; Murdoch, S. An Improved Clock-skew Measurement Technique for Revealing Hidden Services. *USENIX Security Symposium* **2008**, pp. 211 – 225.
52. Matic, S.; Kotzias, P.; Caballero, J. CARONTE: Detecting Location Leaks for Deanonymizing Tor Hidden Services. 22nd ACM SIGSAC Conference on Computer and Communications Security - CCS '15; ACM Press: New York, New York, USA, 2015; pp. 1455–1466. doi:10.1145/2810103.2813667.
53. Wang, R.; Wen, Q.; Zhang, H.; Qin, S.; LI, W. Transparent Discovery of Hidden Service. *IEICE Transactions on Information and Systems* **2016**, *E99.D*, 2817–2820. doi:10.1587/transinf.2016EDL8100.
54. Wang, R.; Wen, Q.; Zhang, H.; LI, X. A Novel Protocol-Feature Attack against Tor's Hidden Service. *IEICE Transactions on Information and Systems* **2016**, *E99.D*, 839–849. doi:10.1587/transinf.2015ICP0001.
55. Shebaro, B.; Perez-Gonzalez, F.; Crandall, J.R. Leaving timing-channel fingerprints in hidden service log files. DFRWS 2010 Annual Conference. Elsevier, 2010, Vol. 7, pp. S104–S113. doi:10.1016/J.DIIN.2010.05.013.
56. Elices, J.A.; Perez-Gonzalez, F.; Troncoso, C. Fingerprinting Tor's hidden service log files using a timing channel. 2011 IEEE International Workshop on Information Forensics and Security. IEEE, 2011, pp. 1–6. doi:10.1109/WIFS.2011.6123154.

57. Sanatinia, A.; Noubir, G. Off-path man-in-the-middle attack on tor hidden services. *New England Security Day, NESD* **2017**.
58. Tan, Q.; Gao, Y.; Shi, J.; Wang, X.; Fang, B. A closer look at Eclipse attacks against Tor hidden services. 2017 IEEE International Conference on Communications (ICC). IEEE, 2017, pp. 1–6. doi:10.1109/ICC.2017.7996832.
59. Øverlier, L.; Syverson, P. Valet Services: Improving Hidden Servers with a Personal Touch. International Symposium on Privacy Enhancing Technologies Symposium. Springer, Berlin, Heidelberg, 2006, pp. 223–244. doi:10.1007/11957454_13.
60. Beitollahi, H.; Deconinck, G. Ferris wheel: A ring based onion circuit for hidden services. *Computer Communications* **2012**, *35*, 829–841. doi:10.1016/J.COMCOM.2012.01.008.
61. Yang, L.; Li, F. Enhancing traffic analysis resistance for Tor hidden services with multipath routing. 2015 IEEE Conference on Communications and NetworkSecurity, CNS 2015. Springer, Cham, 2015, pp. 745–746. doi:10.1109/CNS.2015.7346915.
62. Sanatinia, A.; Noubir, G. Honey Onions: A framework for characterizing and identifying misbehaving Tor HSDirs. 2016 IEEE Conference on Communications and Network Security (CNS). IEEE, 2016, pp. 127–135. doi:10.1109/CNS.2016.7860478.
63. Kwon, A.; AlSabah, M.; Lazar, D.; Dacier, M.; Devadas, S. Circuit Fingerprinting Attacks: Passive Deanonymization of Tor Hidden Services. 24th USENIX Security Symposium, 2015.
64. Nurmi, J.; Kannisto, J.; Vajaranta, M. Observing Hidden Service Directory Spying with a Private Hidden Service Honeynet. 2016 11th Asia Joint Conference on Information Security (AsiaJCIS). IEEE, 2016, pp. 55–59. doi:10.1109/AsiaJCIS.2016.31.
65. Lenhard, J.; Loesing, K.; Wirtz, G. Performance Measurements of Tor Hidden Services in Low-Bandwidth Access Networks. International Conference on Applied Cryptography and Network Security. Springer, Berlin, Heidelberg, 2009. doi:10.1007/978-3-642-01957-9_20.
66. Anagnostopoulos, M.; Kambourakis, G.; Drakatos, P.; Karavolos, M.; Kotsilitis, S.; Yau, D.K.Y. Botnet Command and Control Architectures Revisited: Tor Hidden Services and Fluxing. Web Information Systems Engineering - WISE 2017, 2017, pp. 517–527. doi:10.1007/978-3-319-68786-5_41.
67. Loesing, K.; Sandmann, W.; Wilms, C.; Wirtz, G. Performance Measurements and Statistics of Tor Hidden Services. 2008 International Symposium on Applications and the Internet. IEEE, 2008, pp. 1–7. doi:10.1109/SAINT.2008.69.
68. Liu, P.; Wang, X.; He, X.; Li, C.; Cao, S.; He, L.; Zhu, J. A quantitative model for analysis and evaluation of tor hidden service discovery. Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST, 2018, Vol. 219, pp. 70–77. doi:10.1007/978-3-319-73317-3_10.
69. Meng, Y.; Zhao, X.; Fei, J.; Zhu, Y. A hidden service model based on HS-HS anonymous network. AIP Conference Proceedings 1890, 2017, p. 020006. doi:10.1063/1.5005184.
70. Øverlier, L.; Syverson, P. Improving Efficiency and Simplicity of Tor Circuit Establishment and Hidden Services. In *Privacy Enhancing Technologies*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2007; pp. 134–152. doi:10.1007/978-3-540-75551-7_9.
71. Meng, Y.; Fei, J.; Chen, Y.; Zhu, Y. A Domain Name Model of Anonymous Network Hidden Service. International Conference on Cloud Computing and Security; Springer, Cham: Haikou; China, 2018. doi:10.1007/978-3-030-00015-8_10.
72. Victors, J.; Li, M.; Fu, X. The Onion Name System. *Proceedings on Privacy Enhancing Technologies* **2016**, *2017*, 21–41. doi:10.1515/popets-2017-0003.
73. Liang, J.; Liu, Y. An Improved Method to Build a Circuit of Tor Hidden Service. International Conference on Automation, Mechanical Control and Computational Engineering, 2018, Vol. 166, pp. 120–126.
74. Betzwieser, J.D.; Mason, W.R.; Redmann, R.F.; Taylor, Z.S.; Tsao, S.H.; Brown, D.E.; Conklin, J.H. Systems methodology to characterizing the threat posed by anonymous systems on the internet. 2009 Systems and Information Engineering Design Symposium. IEEE, 2009, pp. 159–164. doi:10.1109/SIEDS.2009.5166173.
75. Owenson, G.; Cortes, S.; Lewman, A. The darknet's smaller than we thought: The life cycle of Tor Hidden Services. *Digital Investigation* **2018**, *27*, 17–22. doi:10.1016/j.diin.2018.09.005.
76. Winter, P.; Edmundson, A.; Roberts, L.M.; Dutkowska, A.; Chetty, M.; Feamster, N. How Do Tor Users Interact With Onion Services ? USENIX Security Symposium, 2018.
77. Tor Project. Tor Rendezvous Specification - Version 3, 2019.
78. Tor Project. Onion Services traffic - Tor Metrics, 2021.