CUMUL & Co: High-Impact Artifacts for Website Fingerprinting Research[*]

Jan Pennekamp[×], Martin Henze[§,¶], Andreas Zinnen[†], Fabian Lanze[‡],
Klaus Wehrle[×], and Andriy Panchenko[‖]
[×]*Communication and Distributed Systems*, RWTH Aachen University, Aachen, Germany · [§]*Security and Privacy in Industrial Cooperation*, RWTH Aachen University, Aachen, Germany · [¶]*Cyber Analysis & Defense*, Fraunhofer FKIE Wachtberg, Germany · [†]Hochschule RheinMain, Wiesbaden, Germany · [‡]Huf Group, Velbert, Germany
[‖]*IT Security*, Brandenburg Technical University, Cottbus, Germany
pennekamp@comsys.rwth-aachen.de · henze@cs.rwth-aachen.de · andreas.zinnen@hs-rm.de
fabian@lanze.net · wehrle@comsys.rwth-aachen.de · andriy.panchenko@b-tu.de

ABSTRACT

Anonymous communication on the Internet is about hiding the relationship between communicating parties. At NDSS '16, we presented a new website fingerprinting approach, CUMUL, that utilizes novel features and a simple yet powerful algorithm to attack anonymization networks such as Tor. Based on pattern observation of data flows, this attack aims at identifying the content of encrypted and anonymized connections. Apart from the feature generation and the used classifier, we also provided a large dataset to the research community to study the attack at Internet scale. In this paper, we emphasize the impact of our artifacts by analyzing publications referring to our work with respect to the dataset, feature extraction method, and source code of the implementation. Based on this data, we draw conclusions about the impact of our artifacts on the research field and discuss their influence on related cybersecurity topics. Overall, from 393 unique citations, we discover more than 130 academic references that utilize our artifacts, 61 among them are highly influential (according to SemanticScholar), and at least 35 are from top-ranked security venues. This data underlines the significant relevance and impact of our work as well as of our artifacts in the community and beyond.

CCS CONCEPTS

• **Security and privacy → Pseudonymity, anonymity and untraceability**; • **Networks → Network privacy and anonymity**.

KEYWORDS

Traffic Analysis; Website Fingerprinting; Privacy; Anonymous Communication; Onion Routing; Web Privacy

1 INTRODUCTION: ADDRESSED PROBLEM

Anonymous communication on the Internet is about hiding the relationship between communicating parties. For many people, in particular, for people living in oppressive regimes, the use of anonymization techniques is the only way to exercise their right to freedom of expression and to freely access information, without fearing the consequences. Thus, these techniques are often used to bypass country-level censorship. Hence, users of anonymization techniques strongly rely on the underlying protection as defined in their attacker model. The Tor network [11]—the most popular low-latency anonymous communication system nowadays is used by millions of daily users—promises to hide the relationship between the sender of a message and its destination from a *local observer* (e.g., a local system administrator, an ISP, a Tor node operator, or everyone with the ability to eavesdrop wireless connections).

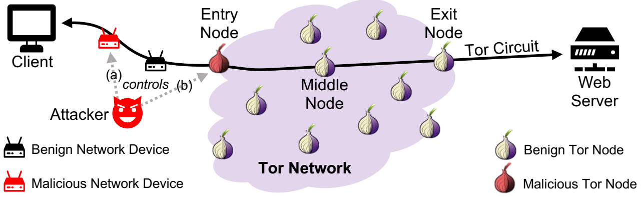

Figure 1: Website fingerprinting (WFP) attack on Tor users.

The website fingerprinting (WFP) attack is a special case of *traffic analysis*. Performed by a local eavesdropper, it aims to infer information about the content (i.e., the website visited) of encrypted and anonymized connections by observing network patterns between the sender and the first anonymizing node (i.e., the entry node). Here, the attacker merely utilizes meta information, such as packet size and its direction, without breaking the encryption, as we illustrate in Figure 1. To (passively) capture the network traffic, the attacker either controls (a) a compromised network device on path or (b) operates a malicious entry node. Before 2011, Tor was considered to be secure against this threat [19, 32]. Since then, WFP has become an active field of research [10, 12, 25, 34, 38, 41, 42].

In our NDSS paper [31], we propose a novel, best-performing WFP attack at the time, based on a subtle method to map network traces into a robust representation of a class (a set of abstracted objects that share a common characteristic; in our case, multiple traces are recorded for the same webpage). We abstract the loading process of a webpage by generating a cumulative behavioral representation of its trace. From this data, we extract features, called CUMUL. These features implicitly cover characteristics of the traffic that other feature sets have to explicitly consider, such as packet ordering or burst behavior. By design, CUMUL is robust against differences in bandwidth, congestion, and the timing of a page load.

To evaluate the severity of the WFP attack in reality, we built the most representative dataset ever assembled in this domain at the time. The dataset is representative not only because of its size, but also because it is not subject to simplified assumptions made by the related work (e.g., most researchers consider only the index pages, i.e., those pages that web servers provide for a requested domain [12, 19, 42]). Our dataset consists of over 300 000

[*] ACSAC '22: Cybersecurity Artifacts Competition and Impact Award Submission

webpages (which is ten times larger than the biggest existing set at the time [22] and, to the best of our knowledge, still the second biggest as of today [34]). It consists of two parts: a random sample of World Wide Web and URLs from the real Tor traffic. Due to ethical concerns, we could only publish the first part of the dataset.

Using our artifacts (i.e., robust features, a simple yet effective classifier, and representative datasets), we show that our attack outperforms existing state-of-the-art attacks in terms of accuracy while computationally being several orders of magnitude faster [31]. Thanks to the large size of the datasets and our computationally effective classifier, we provide several new insights: We were the first to show that with existing classifiers, *webpage* fingerprinting for *any* webpage is similar to finding a needle in a haystack—in general, it is doomed to fail. However, *website* fingerprinting (detecting complete sites instead of single pages only), despite being a more realistic scenario, is also easier to handle for existing classifiers.

Our artifacts allow researchers to easily benefit from our work and to follow up on our insights. Although WFP is still a major threat to the anonymity of Tor, we initiated a discussion with our artifacts that due to the large number of the webpages in the World Wide Web, the scalability and practical realization of the attack is much harder than thought before and boosted further studies of the limits of webpage and website fingerprinting at Internet scale.

2 OUR ARTIFACTS AND HOW THEY WORK

Subsequently to the publication of our paper [31], we also made our research artifacts publicly and freely available online [30]. Our artifacts cover the complete pipeline to study WFP attacks, starting from datasets and data extraction, through feature generation up to classification and evaluation scripts, as illustrated in Figure 2. In the following, we focus on the three most relevant artifacts: our novel effective features (called CUMUL), an efficient SVM-based classifier, and our representative datasets.

2.1 Our CUMUL Features

Instead of manually identifying characteristics that may contain significant information about the load behavior, we aimed at deriving our features from an abstract representation that implicitly covers all relevant characteristics. As identified before [32, 41], four basic features significantly contribute distinctive information: $\mathcal{N}_{\text{in}}$, the number of incoming packets, $\mathcal{N}_{\text{out}}$, the number of outgoing packets, $\mathcal{S}_{\text{in}}$, the sum of incoming packet sizes, and $\mathcal{S}_{\text{out}}$, the sum of outgoing packet sizes. Additionally, to characterize the progress of the page load, we proposed to use the cumulated sum of packet sizes of a network trace and to sample n additional features.

We define such a network trace, i.e., a sequence of raw TCP packet sizes, as follows. $T = (p_1, \ldots, p_N)$, where $p_i > 0$ indicates an incoming packet and $p_i < 0$ an outgoing packet. The cumulative representation of this trace is then calculated as $C(T) = ((0,0), (a_1, c_1), \ldots, (a_N, c_N))$, where $c_1 = p_1$, $a_1 = |p_1|$, and $c_i = c_{i-1} + p_i$, $a_i = a_{i-1} + |p_i|$ for $i = 2, \ldots, N$. From this representation, we derive n features $C_1, \ldots, C_n$ by sampling the piecewise linear interpolant of C at n equidistant points, as captured in Figure 3.

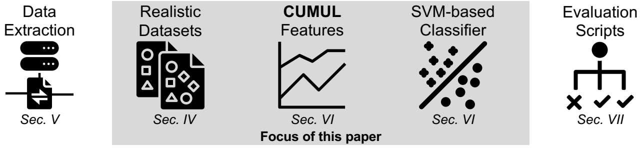

Figure 2: Our provided artifacts cover all required steps.

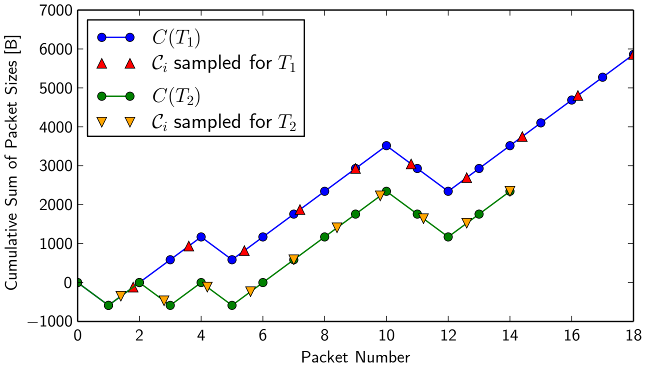

Figure 3: Sampling of two cumulatively represented traces.

Our simplified example shows the cumulative representation of two traces T_1 and T_2, consisting of $|T_1| = 18$ and $|T_2| = 14$ packets (each of size ± 568 bytes) and the corresponding features C_i for $n = 10$. With this method, we are able to extract a fixed number of identifying characteristics from traces with varying lengths. In our paper [31], we show that $n = 100$ yields the best trade-off between classification accuracy and computational efficiency. Given the expressiveness of our features and their small number needed to outperform competitive classifiers, we could study the scalability of real-world WFP attacks by considering huge datasets.

As a beneficial side-effect of our feature set, derived fingerprints can be intuitively visualized and compared. In Figure 4, we exemplarily detail CUMUL fingerprints for two webpages (40 traces each). Distinctive load behaviors characterize the two webpages. Our subtle method to represent this load behavior based on the cumulated packet sizes enables the differentiation of fingerprints of these two pages even by the human eye. Obviously, as the universe size grows (number of considered websites), this differentiation is not that easily possible and requires an efficient classifier.

2.2 Our Classifier

In addition to our novel feature set that reflects a sampled cumulative representation of a trace, we also proposed using an SVM-based classifier as part of our paper [31]. After the collection of relevant network traces and the subsequent feature extraction, a classifier is usually applied to differentiate them, i.e., to study the implications of WFP attacks (their accuracy) for Tor users. Since our extracted fingerprints have a fixed length by definition, we can directly use them as input to train the SVM classifier. In particular, we utilize a

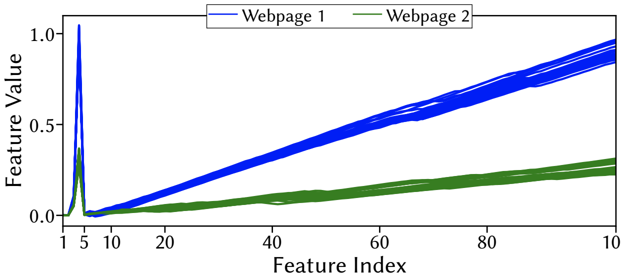

Figure 4: CUMUL's feature visualization of two websites.

Table 1: Paper citations per year based on Google Scholar.

2016*	2017	2018	2019	2020	2021	2022†	**Sum**
13	41	47	55	70	92	62	**382**

*from February 2016 †until September 2022

slightly modified libSVM [6] implementation, which we specially tailored for the needs of the WFP attack evaluation. We provide it as a dedicated artifact to foster additional work in the area.

2.3 Our Dataset for Representative Evaluations

Even though researchers could gather their own datasets, this process requires significant patience and effort to accurately prepare, crawl, record, and post-process representative large-scale datasets. Moreover, fair comparisons of different attacks are only possible if the evaluations are performed on the identical datasets. Hence, when ethically possible (i.e., no real user data that can be misused to harm their privacy is recorded), we published datasets as research artifacts. At that time, we provided the most comprehensive and realistic dataset to evaluate WFP attacks. Instead of limiting our dataset to index pages of popular sites, our artifact contains various webpages that have been actually retrieved. We combined different sources of information (cf. our NDSS paper for more details [31]), such as links distributed via Twitter, Google trends, websites blocked in certain countries, or traces of a Tor exit node, to create a random and representative sample of webpages visited on the Internet (or, over Tor in particular) at the time of evaluation.

We publish two types of datasets: first, we have a *foreground* set that consists of 1125 individual webpages with 40 instances each. This dataset can be applied for closed-world evaluations, and it is by an order of magnitude larger than most datasets at the time (typically, related work only considered closed-world evaluations with 100 webpages). As part of an open-world evaluation, such a set can also be utilized to represent a set of monitored webpages, i.e., those pages that the attacker wants to detect. Second, we provide a *background* set that includes 111 884 individual webpages in a single instance. In this dataset, webpages correspond to unmonitored (unknown) pages. As outlined before (cf. Section 1), our dataset is still the second largest ever collected for WFP. We made our datasets available in both TCP and TLS representation to give researchers as much flexibility as possible when performing their evaluations.

3 EVIDENCE OF THE ARTIFACTS' IMPACT

Following the description of our artifacts, we now highlight their relevance and impact within the research community. According to *Google Scholar*, as of September 2022, our paper has been referenced 382 times. From the statistics per year (Table 1), we observe that the relevance of our paper increases steadily from 41 citations in 2017, over 55 in 2019, to 92 in 2021 (data for 2022 is still incomplete at the time of writing). Moreover, in Google's 2021 Scholar Metrics, our work was ranked as the 11th most cited paper at NDSS [15] within all NDSS papers from the five most recent years (2016–2021).

To properly assess the impact of our artifacts, we study papers citing our work in more detail. Besides Google Scholar, we also include papers indexed by *Web of Science* and *SemanticScholar*. Overall, we end up with 393 sources that reference our work. According to SemanticScholar [1], our work was highly influential for 74 papers. Next, we look closely at how related work utilizes our artifacts.

Table 2: Relevant works that apply our artifacts per year.

2016*	2017	2018	2019	2020	2021	2022†	**Sum**
10	19	16	23	19	29	18	**134**

*from February 2016 †until September 2022

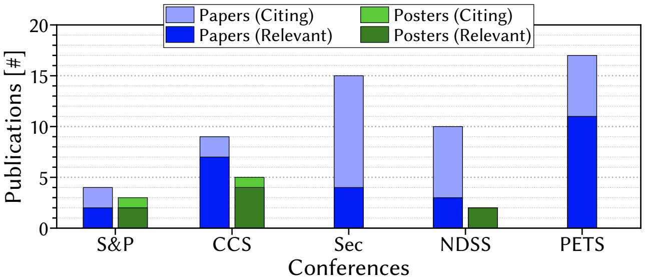

Figure 5: Frequency of references at top-ranked conferences.

3.1 Key Numbers on Artifacts' Use in Research

In addition to frequent citations, we also observe a notable number of papers reusing our artifacts as part of their security research.

Methodology. Out of 393 articles citing our work, we were able to retrieve all but 5 and inspected them all manually. In particular, we investigate the usage of our artifacts: features (cf. Section 2.1), classifier (cf. Section 2.2), and dataset (cf. Section 2.3). With this approach, we highlight relevant papers (those works that actually use our artifacts). We followed a conservative approach, marking articles as irrelevant if we do not find strong evidence of artifacts' usage. Finally, we end up with **134** relevant references, out of which 8 are posters, 26 are theses, and 6 are unreviewed preprints (at the time of writing). Table 2 shows their overall distribution by year.

Top-Ranked Conferences. To assess the venues' reputation of analyzed papers, we refer to the CORE conference [8] and journal [9] rankings. If not listed there, we consult Gu's ranking [17]. Overall, 27 works that use our artifacts were published at A* conferences. Apart from security conferences, two A* papers originate from WWW and one from INFOCOM. We further discovered 42 and 19 papers at A- and B-ranked venues, respectively.

Additionally, we specifically looked into top-ranked conferences in the security field, i.e., the A* conferences IEEE S&P, ACM CCS, USENIX Security (Sec), and NDSS, as well as PETS, a premier venue for research on anonymous communication [33] (supplementary data on PETS' A-ranking with a detailed justification of its quality is available [7]). Figure 5 shows the distribution of relevant papers (and posters) at these top security conferences: most papers (11) were published at PETS, followed by ACM CCS and USENIX Sec. At ACSAC, 2 papers apply our artifacts as part of their research.

Applied Artifact Components. In Figure 6, we grouped the artifacts' use by features (dark blue), classifier (light blue), and dataset (dark green). Additionally, light green indicates how many authors re-published components of our artifacts (including derivations and re-implementations) as part of their own artifacts. We notice that 61 of the 74 highly-influential [1] papers apply our artifacts.

Our proposed features and the provided classifier are frequently applied and lay the foundation for various research efforts, also beyond WFP (cf. Section 3.2). The usage of our artifacts generally

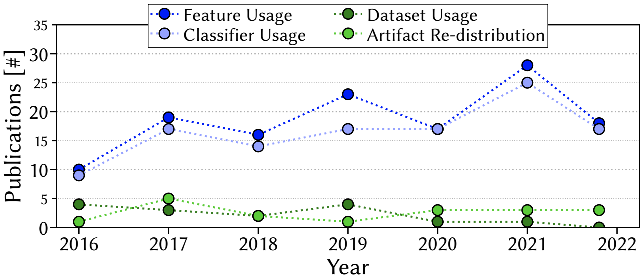

Figure 6: Usage and re-distribution of our artifacts per year.

increases over time, e.g., from 19 citations in 2017 to 29 in 2021. In total, while our features have been applied 131 times, we count 116 uses of our classifier. Compared to the other artifacts, our dataset has the lowest usage, with a total of 15 instances. A possible reason is the emergence of methods using deep learning (e.g., [34, 38, 45]) that require a larger number of instances. Due to its elegant simplicity, we further notice that researchers frequently re-implement CUMUL (≥22 times). Additionally, 15 papers adapt our features, and 6 papers alter the proposed classifier. Within 43 third-party artifacts, we discover a re-distribution of our artifacts in 18 cases. At least 2 follow-up papers even rely directly on such re-distributions.

Influential Citing Authors. When studying the application of our artifacts by well-established researchers, we discover at least 36 professors (14 full professors) who publish at top venues in the domain (A^* and PETS). In total, they account for contributions in 54 papers and 14 posters. When looking at their h-indexes (according to Google Scholar), we count 10 professors with an h-index over 40 and 22 professors with an h-index over 20. If we also consider A-ranked venues, these numbers change to 21 and 41, respectively. Since not all authors use Google Scholar, our results are conservative (the actual number is expected to be higher). The most influential author has an impressive h-index of 101.

3.2 Associated Discoveries and Trivia

After our quantitative analysis, we now focus on qualitative aspects that underscore the impact of our artifacts for security research.

Quotes. During our manual study, we came across several quotes that underline the performance of CUMUL and our proposed feature set. Several researchers argue that our features are very effective, especially given their manual selection: ► "CUMUL is the best performing manual feature extraction attack in vanilla WF settings" [3], ► "The CUMUL algorithm [...] is one of the most accurate in the literature." [5], and ► "The cumulative sum features are very effective" [46]. Moreover, they praise the computational efficiency when applying our artifacts: ► "By sampling features from the cumulative representation of a trace, CUMUL outperforms previous attacks while staying computationally efficient." [18] and ► "CUMUL performed the best [...], and proved to be more practically feasible" [34]. Thus, in addition to quantitative facts, we also observe a qualitative distinction of our work by security researchers in the area.

Best Practices Survey. Arp et al. [2] conducted a study on the use of machine learning in security research and contacted us to include our work "because of [our] outstanding contributions". With the authors, we thoroughly discussed their derived pitfalls in the context of our work and WFP in general. Thereby, we contributed to (i) raising the awareness of methodological pitfalls and (ii) supporting a collection of best practices for future research. As confirmed by Arp et al., back in 2016, we already initially discussed possible pitfalls (e.g., sampling bias and simplified lab-only evaluation) that are relevant to security research, our domain, and our artifacts.

Website Fingerprinting Research. Our paper is part of the renowned list of selected papers on anonymity [13]. Moreover, we are confident to have motivated further research on the fingerprinting and fingerprintability of complete websites, an overlooked direction. Jiang et al. stress: "From [[31]], it can be seen that the existing studies on [website fingerprinting] are unrealistic and have an impact on the collection of the dataset because they only consider the homepage of the website and not the subpages." [21, translated].

Concerning the prevalence of WFP at ACSAC, we discover 6 research papers on website fingerprinting or with reference to it.

Other Impacted Research Areas. Apart from the publications at A^* conferences in other domains, e.g., at WWW (web) and INFOCOM (networking), we came across various relevant papers that do not focus on WFP. We discovered several references to our work in the context of DNS privacy [4, 37], user profiling [14], or Tor measurements [20]. Additionally, CUMUL was applied for various fingerprinting tasks, e.g., apps [39, 40], search keywords [26–28], or hidden services [43, 44]. Likewise, at least 10 papers utilize our artifacts to classify web traffic on a general level. Our work is further used to generate packet traces [16], optimize features [36], predict the fingerprintability [29], estimate information leakage [23], and study WFP metrics [24]. Consequently, the impact of our work and artifacts goes beyond (intended) applications for WFP. We even discovered a patent that cites our work [35].

4 CONCLUSION: OUR SECURITY ARTIFACTS

In this paper, after introducing the addressed problem and how our artifacts work, including our intuitive CUMUL feature set, we have highlighted the impact of our approach and published artifacts. Our approach utilizes a novel feature set and a simple yet powerful algorithm when performing website fingerprinting attacks against anonymization networks such as Tor. By analyzing publications that refer to our work w.r.t. our artifacts (features, classifier, and dataset), we demonstrated the enormous impact our artifacts had and have on the research field and beyond. Out of 393 papers citing our work, more than 134 references rely on our security artifacts, 61 among them are highly influential (according to SemanticScholar), and at least 35 are from top-ranked venues. Thus, by making our easy-to-use, easy-to-adapt, and re-distributable artifacts available to the research community, we facilitated significant progress on the general challenge of understanding limits and protecting against traffic analysis attacks. Finally, our presentation of key facts and influence on other areas greatly underlines the impact and relevance of our security artifacts, even beyond the domain of WFP.

ACKNOWLEDGMENTS

We kindly thank Thomas Engel for his contributions to the original NDSS paper [31]. Additionally, we are grateful to Asya Mitseva for her effort in maintaining our artifacts and for supporting this submission. Finally, Ike Kunze deserves credit for plotting our figures.

REFERENCES

[1] Allen Institute for AI. 2018 (accessed September 20, 2022). Semantic Scholar – Frequently Asked Questions. https://www.semanticscholar.org/faq#influential-citations.

[2] Daniel Arp, Erwin Quiring, Feargus Pendlebury, Alexander Warnecke, Fabio Pierazzi, Christian Wressnegger, Lorenzo Cavallaro, and Konrad Rieck. 2022. Dos and Don'ts of Machine Learning in Computer Security. In *Proceedings of the 31st USENIX Security Symposium (SEC '22)*. USENIX Association, 3971–3988.

[3] Sanjit Bhat, David Lu, Albert Kwon, and Srinivas Devadas. 2019. Var-CNN: A Data-Efficient Website Fingerprinting Attack Based on Deep Learning. *Proceedings on Privacy Enhancing Technologies* 4 (2019), 292–310. https://doi.org/10.2478/popets-2019-0070

[4] Jonas Bushart and Christian Rossow. 2020. Padding Ain't Enough: Assessing the Privacy Guarantees of Encrypted DNS. In *Proceedings of the 10th USENIX Workshop on Free and Open Communications on the Internet (FOCI '20)*. USENIX Association.

[5] Eric Chan-Tin, Taejoon Kim, and Jinoh Kim. 2018. Website Fingerprinting Attack Mitigation Using Traffic Morphing. In *Proceedings of the 2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS '18)*. IEEE, 1575–1578. https://doi.org/10.1109/ICDCS.2018.00174

[6] Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2, 3 (2011). https://doi.org/10.1145/1961189.1961199

[7] Shaanan Cohney, Matthew Wright, Aaron Johnson, and Veelasha Moonsamy. 2020. Submission Data for 2020-2021 CORE conference Ranking process Privacy Enhancing Technologies Symposium (was International Workshop of Privacy Enhancing Technologies). http://portal.core.edu.au/core/media/conference_submission_2020/Data_Higherrank_1035.pdf.

[8] Computing Research & Education. 2013. CORE Conference Portal. http://portal.core.edu.au/conf-ranks/.

[9] Computing Research & Education. 2013. CORE Journal Portal. http://portal.core.edu.au/jnl-ranks/. This ranking has been discontinued in February 2022 and has not been updated since..

[10] Wladimir De la Cadena, Asya Mitseva, Jens Hiller, Jan Pennekamp, Sebastian Reuter, Julian Filter, Klaus Wehrle, Thomas Engel, and Andriy Panchenko. 2020. TrafficSliver: Fighting Website Fingerprinting Attacks with Traffic Splitting. In *Proceedings of the 27th ACM SIGSAC Conference on Computer and Communications Security (CCS '20)*. ACM, 1971–1985. https://doi.org/10.1145/3372297.3423351

[11] Roger Dingledine, Nick Mathewson, and Paul Syverson. 2004. Tor: The Second-Generation Onion Router. In *Proceedings of the 13th USENIX Security Symposium (SEC '04)*. USENIX Association.

[12] Kevin P. Dyer, Scott E. Coull, Thomas Ristenpart, and Thomas Shrimpton. 2012. Peek-a-Boo, I Still See You: Why Efficient Traffic Analysis Countermeasures Fail. In *Proceedings of the 2012 IEEE Symposium on Security and Privacy (S&P '12)*. IEEE, 332–346. https://doi.org/10.1109/SP.2012.28

[13] Freehaven. 2003. Selected Papers in Anonymity. https://www.freehaven.net/anonbib/.

[14] Roberto Gonzalez, Claudio Soriente, and Nikolaos Laoutaris. 2016. User Profiling in the Time of HTTPS. In *Proceedings of the 2016 Internet Measurement Conference (IMC '16)*. ACM, 373–379. https://doi.org/10.1145/2987443.2987451

[15] Google Scholar. 2021. Network and Distributed System Security Symposium (NDSS) - Google Scholar Metrics. https://scholar.google.com/citations?hl=en&alert_preview_top_rm=2&vq=eng_computersecuritycryptography&view_op=list_hcore&venue=q2FcImd5qbgJ.2021. This metric covers articles published in 2016–2020 and includes citations from all articles that were indexed in Google Scholar as of July 2021..

[16] Alexander Griessel, Maximilian Stephan, Martin Mieth, Wolfgang Kellerer, and Patrick Krämer. 2022. RLBrowse: Generating Realistic Packet Traces with Reinforcement Learning. In *Proceedings of the 2022 IEEE/IFIP Network Operations and Management Symposium (NOMS '22)*. IEEE. https://doi.org/10.1109/NOMS54207.2022.9789851

[17] Guofei Gu. 2020. Computer Security Conference Ranking and Statistic. https://people.engr.tamu.edu/guofei/sec_conf_stat.htm.

[18] Sébastien Henri, Ginés García, Pablo Serrano, Albert Banchs, Patrick Thiran, et al. 2020. Protecting against Website Fingerprinting with Multihoming. *Proceedings on Privacy Enhancing Technologies* 2020, 2 (2020), 89–110. https://doi.org/10.2478/popets-2020-0019

[19] Dominik Herrmann, Rolf Wendolsky, and Hannes Federrath. 2009. Website Fingerprinting: Attacking Popular Privacy Enhancing Technologies with the Multinomial Naïve-Bayes Classifier. In *Proceedings of the 2009 ACM Workshop on Cloud Computing Security (CCSW '09)*. ACM, 31–42. https://doi.org/10.1145/1655008.1655013

[20] Rob Jansen, Marc Juarez, Rafa Galvez, Tariq Elahi, and Claudia Diaz. 2018. Inside Job: Applying Traffic Analysis to Measure Tor from Within. In *Proceedings of the 25th Annual Network and Distributed System Security Symposium (NDSS '18)*. Internet Society. https://doi.org/10.14722/ndss.2018.23261

[21] Chen Jing, Wang Wendu, Du Ruiying, and Zeng Cheng. 2017. A Randomized Website Fingerprint Defense Method. *Journal of Wuhan University (Natural Science Edition)* 63, 5 (2017), 397–402. https://doi.org/10.14188/j.1671-8836.2017.05.003

[22] Marc Juarez, Sadia Afroz, Gunes Acar, Claudia Diaz, and Rachel Greenstadt. 2014. A Critical Evaluation of Website Fingerprinting Attacks. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS '14)*. ACM, 263–274. https://doi.org/10.1145/2660267.2660368

[23] Shuai Li, Huajun Guo, and Nicholas Hopper. 2018. Measuring Information Leakage in Website Fingerprinting Attacks and Defenses. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (CCS '18)*. ACM, 1977–1992. https://doi.org/10.1145/3243734.3243832

[24] Nate Mathews, Mohammad Saidur Rahman, and Matthew Wright. 2019. Poster: Evaluating Security Metrics for Website Fingerprinting. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS '19)*. ACM, 2625–2627. https://doi.org/10.1145/3319535.3363272

[25] Asya Mitseva, Jan Pennekamp, Johannes Lohmöller, Torsten Ziemann, Carl Hoerchner, Klaus Wehrle, and Andriy Panchenko. 2021. POSTER: How Dangerous is My Click? Boosting Website Fingerprinting By Considering Sequences of Webpages. In *Proceedings of the 28th ACM SIGSAC Conference on Computer and Communications Security (CCS '21)*. ACM, 2411–2413. https://doi.org/10.1145/3460120.3485347

[26] Se Eun Oh. 2021. *Towards More Effective Traffic Analysis in the Tor Network*. Ph. D. Dissertation. University of Minnesota.

[27] Se Eun Oh and Nicholas Hopper. 2017. Poster: Fingerprinting Past the Front Page: Identifying Keywords in Search Queries over Tor. 24th Annual Network and Distributed System Security Symposium (NDSS '17).

[28] Se Eun Oh, Shuai Li, and Nicholas Hopper. 2017. Fingerprinting Keywords in Search Queries over Tor. *Proceedings on Privacy Enhancing Technologies* 2017, 4 (2017), 251–270. https://doi.org/10.1515/popets-2017-0033

[29] Se Eun Oh, Saikrishna Sunkam, and Nicholas Hopper. 2019. \$p\$^{1}-FP: Extraction, Classification, and Prediction of Website Fingerprints with Deep Learning. *Proceedings on Privacy Enhancing Technologies* 2019, 3 (2019), 191–209. https://doi.org/10.2478/popets-2019-0043

[30] Andriy Panchenko. 2016. zwiebelfreunde. https://www.informatik.tu-cottbus.de/~andriy/zwiebelfreunde/.

[31] Andriy Panchenko, Fabian Lanze, Andreas Zinnen, Martin Henze, Jan Pennekamp, Klaus Wehrle, and Thomas Engel. 2016. Website Fingerprinting at Internet Scale. In *Proceedings of the 23rd Annual Network and Distributed System Security Symposium (NDSS '16)*. Internet Society. https://doi.org/10.14722/ndss.2016.23477

[32] Andriy Panchenko, Lukas Niessen, Andreas Zinnen, and Thomas Engel. 2011. Website Fingerprinting in Onion Routing Based Anonymization Networks. In *Proceedings of 10th Annual ACM Workshop on Privacy in the Electronic Society (WPES '11)*. ACM, 103–114. https://doi.org/10.1145/2046556.2046570

[33] Privacy Enhancing Technologies Board. 2008. PoPETs/PETS. https://petsymposium.org/.

[34] Vera Rimmer, Davy Preuveneers, Marc Juarez, Tom Van Goethem, and Wouter Joosen. 2018. Automated Website Fingerprinting through Deep Learning. In *Proceedings of the 25th Annual Network and Distributed System Security Symposium (NDSS '18)*. Internet Society. https://doi.org/10.14722/ndss.2018.23105

[35] Roberto Gonzalez Sanchez, Claudio Soriente, and Nikolaos Laoutaris. 2021. Method for performing user profiling from encrypted network traffic flows. US10885466B2.

[36] Meng Shen, Yiting Liu, Liehuang Zhu, Ke Xu, Xiaojiang Du, and Nadra Guizani. 2020. Optimizing Feature Selection for Efficient Encrypted Traffic Classification: A Systematic Approach. *IEEE Network* 34, 4 (2020), 20–27. https://doi.org/10.1109/MNET.011.1900366

[37] Sandra Siby, Marc Juarez, Claudia Diaz, Narseo Vallina-Rodriguez, and Carmela Troncoso. 2020. Encrypted DNS –> Privacy? A Traffic Analysis Perspective. In *Proceedings of the 27th Annual Network and Distributed System Security Symposium (NDSS '20)*. Internet Society. https://doi.org/10.14722/ndss.2020.24301

[38] Payap Sirinam, Mohsen Imani, Marc Juarez, and Matthew Wright. 2018. Deep Fingerprinting: Undermining Website Fingerprinting Defenses with Deep Learning. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (CCS '18)*. ACM, 1928–1943. https://doi.org/10.1145/3243734.3243768

[39] Riccardo Spolaor. 2017. *Security and Privacy Threats on Mobile Devices through Side-Channels Analysis*. Ph. D. Dissertation. University of Padua.

[40] Vincent F. Taylor, Riccardo Spolaor, Mauro Conti, and Ivan Martinovic. 2017. Robust Smartphone App Identification via Encrypted Network Traffic Analysis. *IEEE Transactions on Information Forensics and Security* 13, 1 (2017), 63–78. https://doi.org/10.1109/TIFS.2017.2737970

[41] Tao Wang, Xiang Cai, Rishab Nithyanand, Rob Johnson, and Ian Goldberg. 2014. Effective Attacks and Provable Defenses for Website Fingerprinting. In *Proceedings of the 23rd USENIX Security Symposium (SEC '14)*. USENIX Association, 143–157.

[42] Tao Wang and Ian Goldberg. 2013. Improved Website Fingerprinting on Tor. In *Proceedings of 12th ACM Workshop on Privacy in the Electronic Society (WPES '13)*. ACM, 201–212. https://doi.org/10.1145/2517840.2517851

[43] Xuebin Wang, Zhipeng Chen, Zeyu Li, Wentao Huang, Meiqi Wang, Shengli Pan, and Jinqiao Shi. 2022. Identification of MEEK-Based TOR Hidden Service Access Using the Key Packet Sequence. In *Proceedings of the 22nd International Conference on Computational Science (ICCS '22)*, Vol. 13350. Springer, 554–568. https://doi.org/10.1007/978-3-031-08751-6_40

[44] Xuebin Wang, Zeyu Li, Wentao Huang, Meiqi Wang, Jinqiao Shi, and Yanyan Yang. 2021. Towards Comprehensive Analysis of Tor Hidden Service Access Behavior Identification Under Obfs4 Scenario. In *Proceedings of the 2021 ACM International Conference on Intelligent Computing and its Emerging Applications (ICEA '21)*. ACM, 205–210. https://doi.org/10.1145/3491396.3506532

[45] Yanbin Wang, Haitao Xu, Zhenhao Guo, Zhan Qin, and Kui Ren. 2022. snWF: Website Fingerprinting Attack by Ensembling the Snapshot of Deep Learning. *IEEE Transactions on Information Forensics and Security* 17 (2022), 1214–1226. https://doi.org/10.1109/TIFS.2022.3158086

[46] Shi-Jie Xu, Guang-Gang Geng, Xiao-Bo Jin, Dong-Jie Liu, and Jian Weng. 2022. Seeing Traffic Paths: Encrypted Traffic Classification with Path Signature Features. *IEEE Transactions on Information Forensics and Security* 17 (2022), 2166–2181. https://doi.org/10.1109/TIFS.2022.3179955