# Poster: Circumventing the GFW with TLS Record Fragmentation

Niklas Niere
Paderborn University
Paderborn, Germany
niklas.niere@upb.de

Sven Hebrok
Paderborn University
Paderborn, Germany
sven.hebrok@upb.de

Juraj Somorovsky
Paderborn University
Paderborn, Germany
juraj.somorovsky@upb.de

Robert Merget
Technology Innovation Institute
Abu Dhabi, United Arab Emirates
robert.merget@tii.ae

## ABSTRACT

State actors around the world censor the HTTPS protocol to block access to certain websites. While many circumvention strategies utilize the TCP layer only little emphasis has been placed on the analysis of TLS—a complex protocol and integral building block of HTTPS. In contrast to the TCP layer, circumvention methods on the TLS layer do not require root privileges since TLS operates on the application layer. With this proposal, we want to motivate a deeper analysis of TLS in regard to censorship circumvention techniques. To prove the existence of such techniques, we present TLS record fragmentation as a novel circumvention technique and circumvent the Great Firewall of China (GFW) using this technique. We hope that our research fosters collaboration between censorship and TLS researchers.

## CCS CONCEPTS

• **Security and privacy** → **Firewalls**; • **Networks** → *Security protocols*; • **Social and professional topics** → **Technology and censorship**; **Censoring filters**.

## KEYWORDS

Censorship, Fragmentation, GFW, TLS

**ACM Reference Format:**
Niklas Niere, Sven Hebrok, Juraj Somorovsky, and Robert Merget. 2023. Poster: Circumventing the GFW with TLS Record Fragmentation. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security (CCS '23), November 26–30, 2023, Copenhagen, Denmark.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3576915.3624372

## 1 MOTIVATION

To restrict the network access of their citizens, censorship is employed by various nation-state actors [1, 3, 19, 21, 27]. A prominent example of such a censor is the Great Firewall of China (GFW). The GFW uses deep packet inspection to censor various websites and protocols: most notably HTTP and HTTPS (HTTP+TLS) [2, 5, 13, 15, 25, 26]. To block websites in the HTTP and HTTPS protocol, the GFW analyzes either the host header in the plain HTTP request or the unencrypted Server Name Indication (SNI) extension in the TLS `ClientHello` message. As more than 99% of the browser traffic issued from China from Mid-July to Mid-August 2023 was encrypted using HTTPS [10], TLS [11, 23] censorship is a central component of the GFW.

Raman et al. [22] change the TLS version, cipher suite, and SNI value of the `ClientHello` message to circumvent TLS censorship. While some of their techniques worked considerably well against some censors, none were conclusive, and all were untested against TLS servers. Chai et al. [8] find that ESNI—the encrypted version of the SNI extension—is unsupported by a large share of TLS servers and Bock et al. [5] found that it is already being censored by the GFW. Thus, the circumventability of plaintext SNI censorship is still essential and has largely been unexplored in the TLS layer. Much research and many circumvention tools attempt to circumvent TLS censorship on the TCP layer [4, 6, 7, 14, 18, 24, 28]. A notable circumvention technique is TCP fragmentation, in which the `ClientHello` message is split over multiple TCP segments to confuse stateless censors that do not reassemble packets.

In this proposal, we transfer the idea behind TCP fragmentation to the TLS layer by fragmenting `ClientHello` messages on the TLS layer alone. This technique, called TLS record fragmentation, does not require elevated privileges to alter TCP traffic and can be combined with existing techniques such as TCP fragmentation. As the GFW shows the first signs of successfully handling TCP fragmentation [4], we deem TLS-layer circumvention techniques integral to the future of censorship circumvention. Overall, we hope to motivate a discussion about and an analysis of other potential TLS-layer circumvention techniques around the world.

We also published our discovery of TLS record fragmentation and analysis results as a blog post [20].

## 2 TLS RECORD FRAGMENTATION

Before being wrapped in a TCP segment, every TLS message is wrapped in a so-called TLS record. As the maximum size of a TLS message ($2^{24}$ bytes) exceeds the maximum size of a TLS record ($2^{16}$ bytes) the TLS specification allows TLS messages to be fragmented over multiple TLS records. The difference between an unfragmented and a TLS record fragmented TLS `ClientHello` message is depicted in Figure 1. In addition to the natural occurrence of TLS fragmentation, it can also be forced by manually wrapping TLS messages in smaller-sized TLS records.


*CCS '23, November 26–30, 2023, Copenhagen, Denmark*

ACM ISBN 979-8-4007-0050-7/23/11.
https://doi.org/10.1145/3576915.3624372

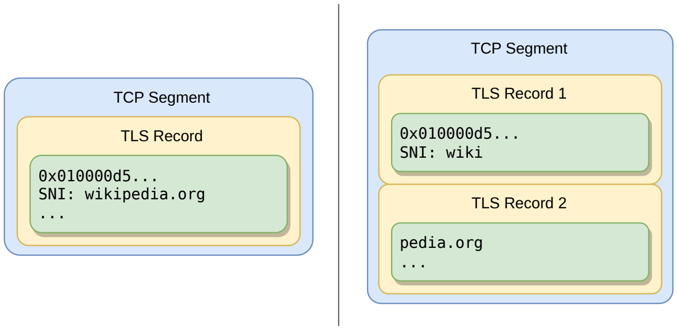


**Figure 1: An unfragmented TLS `ClientHello` (left) and the same message fragmented over two TLS records (right). Note that both fragments are contained in the same TCP segment. Figure taken from [20].**

To confuse censors, a `ClientHello` message can be fragmented so that the SNI extension is not placed in the first TLS record. This forces the censor to allocate memory for the TLS connection state and message reassembly. During our analyses, we found a description of TLS record fragmentation in the context of censorship circumvention by Thomas Pornin in 2014[1]. To the best of our knowledge, TLS record fragmentation has not been implemented in any circumvention tool and has not been subject to practical analysis. We analyze the practicability of TLS record fragmentation as a circumvention technique and provide a circumvention tool that uses TLS record fragmentation.

## 3 CIRCUMVENTING THE GFW

To demonstrate the viability of TLS record fragmentation, we tested it against the world's most sophisticated censor: the GFW. As a stepping stone, we developed DPYProxy: a circumvention tool that implements TCP and TLS record fragmentation. We ran DPYProxy on a vantage point in Mainland China (AS4837) that is subjected to censorship by the GFW. From that vantage point, we queried a censored domain (https://wikipedia.org/wiki/turtle) through DPYProxy using curl[2]. Specifically, DPYProxy fragmented messages on the TCP and TLS layer both before and after the SNI extension. We present the results of our analyses in Table 1.

**Table 1: Circumvention results of the GFW. Any form of TLS record fragmentation circumvents the GFW.**

| Fragmentation | Split | Circumvents Censor |
|---|---|---|
| None | – | No |
| TCP | Before / After SNI | Yes / No |
| TLS | Before / After SNI | Yes / Yes |
| TLS + TCP | Before / After SNI | Yes / Yes |

Our results show that TLS record fragmentation successfully circumvents the GFW. Interestingly, the GFW cannot detect the SNI extension in any TLS record of a TLS-fragmented `ClientHello` message, including the first. The GFW is more successful in detecting TCP fragmentation. Although it is unable to reassemble TCP segments, the GFW still detects the SNI extension when it occurs in the first TCP segment. All of this suggests that the GFW is unaware of TLS fragmentation and cannot analyze TLS records that do not exactly fit into a TCP segment. We conjecture the GFW is similarly unaware of other potential circumvention techniques on the TLS layer.

## 4 TLS SERVER SUPPORT

We also analyzed how many TLS servers on the internet support TLS record fragmentation. Specifically, we analyzed all domains of the Tranco Top 1M list[3] and all `https://` domains from the list of censored domains maintained by CizitenLab[4]. Table 2 shows that around 96% of the censored domains registered by CitizenLab support TLS record fragmentation. 92% of the domains on the Tranco Top 1M list support TLS record fragmentation. TLS servers' support of TLS record fragmentation is slightly biased towards the lower ranks but all ranks support it with over 90%. Overall, TLS record fragmentation is widely usable as a circumvention technique on TLS servers as of today. We hope our research motivates TLS server developers to enable TLS record fragmentation.

**Table 2: TLS record fragmentation support of TLS server.**

| List | Scanned Domains[a] | Support TLS Record Fragmentation |
|---|---|---|
| CitizenLab | 1 135 | 1 092 (96.21%) |
| Tranco Top 1M | 830 357 | 766 909 (92.36%) |

[a] We excluded domains when they were unresolvable, they did not handshake TLS, or their owners requested exclusion from our scans.

## 5 DISCUSSION

We were able to circumvent the GFW with TLS record fragmentation. As the GFW is the world's most sophisticated censor[2, 5, 9, 13, 15, 17, 19, 25, 26], we suspect TLS record fragmentation to be similarly successful against other censors. We want to motivate researchers with access to vantage points in other countries to evaluate the viability of TLS record fragmentation as a circumvention technique.

We successfully developed DPYProxy as a censorship circumvention tool that supports TLS record fragmentation. DPYProxy acts as a MITM. It retrieves a TLS record, splits the included TLS message into multiple parts, and places them in different TLS records. As TLS records are not protected during the TLS handshake, DPYProxy does not break the TLS handshake. Overall, TLS record fragmentation can be implemented by circumvention tools that operate as MITMs and TLS client applications such as custom browsers. As TLS records are constructed on the TCP/IP application layer, no elevated privileges need to be given to circumvention tools that

[1] Thomas Pornin, StackOverflow, https://security.stackexchange.com/questions/56338/identifying-ssl-traffic, Accessed: 17.08.2023, 10:16
[2] curl GitHub page, https://github.com/curl/curl, Accessed: 17.08.2023, 12:42
[3] Tranco Top 1M list, https://tranco-list.eu/, Accessed: 17.08.2023, 13:19
[4] Global list of censored domains, CitizenLab, https://github.com/citizenlab/test-lists/blob/master/lists/global.csv, Accessed: 17.08.2023, 13:19

implement TLS record fragmentation. With a poster, we want to motivate the censorship community and circumvention tool developers to integrate TLS record fragmentation into their tools.

With TLS record fragmentation, we propose the first circumvention technique that is TLS-specific. As the TLS protocol is highly complex, we suspect that additional TLS-specific techniques exist. For example, the SNI extension, which contains the domain of the website, has an overly complex definition in the standard [12] as a list that in practice only contains a single element. Additionally, TLS-specific techniques can be combined with TCP-specific techniques such as TCP fragmentation to generate circumvention techniques spanning multiple protocol layers. We suspect that these combined techniques are especially interesting for QUIC [16], a new protocol that combines the functionality of TLS and TCP while being located on the UDP layer. In the end, we suggest a thorough analysis of the TLS protocol for further circumvention techniques and their applicability against real-life censors. We believe that a collaboration between censorship researchers and TLS researchers benefits this process.

## 6 AVAILABILITY

To make our results reproducible and incite further analyses of TLS record fragmentation, we published DPYProxy and the server-specific results of our analyses. The source code of DPYProxy is available under https://github.com/UPB-SysSec/DPYProxy. The results of our analyses are available under https://github.com/UPB-SysSec/TlsRecordFragmentationResults.

## 7 ACKNOWLEDGEMENTS

We would like to thank Jost Rossel for various helpful comments. Sven Hebrok was supported by the research project "North-Rhine Westphalian Experts in Research on Digitalization (NERD II)", sponsored by the state of North Rhine-Westphalia – NERD II 005-2201-0014. Niklas Niere was supported by the German Federal Ministry of Education and Research (BMBF) through the project KoTeBi.

## REFERENCES


[1] Chaabane Abdelberi, Terence Chen, Mathieu Cunche, Emiliano De Cristofaro, Arik Friedman, and Mohamed Ali Kâafar. 2014. Censorship in the Wild: Analyzing Internet Filtering in Syria. In *Proceedings of the 2014 Internet Measurement Conference, IMC 2014, Vancouver, BC, Canada, November 5-7, 2014*, Carey Williamson, Aditya Akella, and Nina Taft (Eds.). ACM, 285–298. https://doi.org/10.1145/2663716.2663720

[2] Anonymous. 2014. Towards a Comprehensive Picture of the Great Firewall's DNS Censorship. In *4th USENIX Workshop on Free and Open Communications on the Internet, FOCI '14, San Diego, CA, USA, August 18, 2014*, Jedidiah R. Crandall and Vern Paxson (Eds.). USENIX Association. https://www.usenix.org/conference/foci14/workshop-program/presentation/anonymous

[3] Simurgh Aryan, Homa Aryan, and J. Alex Halderman. 2013. Internet Censorship in Iran: A First Look. In *3rd USENIX Workshop on Free and Open Communications on the Internet, FOCI '13, Washington, D.C., USA, August 13, 2013*, Jedidiah R. Crandall and Joss Wright (Eds.). USENIX Association. https://www.usenix.org/conference/foci13/workshop-program/presentation/aryan

[4] Kevin Bock, George Hughey, Xiao Qiang, and Dave Levin. 2019. Geneva: Evolving Censorship Evasion Strategies. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, CCS 2019, London, UK, November 11-15, 2019*, Lorenzo Cavallaro, Johannes Kinder, XiaoFeng Wang, and Jonathan Katz (Eds.). ACM, 2199–2214. https://doi.org/10.1145/3319535.3363189

[5] Kevin Bock, iyouport, Anonymous, Louis-Henri Merino, David Fifield, Amir Houmansadr, and Dave Levin. 2020. Exposing and Circumventing China's Censorship of ESNI. https://gfw.report/blog/gfw_esni_blocking/en/

[6] Kevin Bock, Gabriel Naval, Kyle Reese, and Dave Levin. 2021. Even Censors Have a Backup: Examining China's Double HTTPS Censorship Middleboxes. In *FOCI '21: Proceedings of the ACM SIGCOMM 2021 Workshop on Free and Open Communications on the Internet, Virtual Event, USA, 27 August 2021*. ACM, 1–7. https://doi.org/10.1145/3473604.3474559

[7] bol van. 2023. zapret. https://github.com/bol-van/zapret

[8] Zimo Chai, Amirhossein Ghafari, and Amir Houmansadr. 2019. On the Importance of Encrypted-SNI (ESNI) to Censorship Circumvention. In *9th USENIX Workshop on Free and Open Communications on the Internet, FOCI 2019, Santa Clara, CA, USA, August 13, 2019*, Susan E. McGregor and Michael Carl Tschantz (Eds.). USENIX Association. https://www.usenix.org/conference/foci19/presentation/chai

[9] Richard Clayton, Steven J. Murdoch, and Robert N. M. Watson. 2006. Ignoring the Great Firewall of China. In *Privacy Enhancing Technologies, 6th International Workshop, PET 2006, Cambridge, UK, June 28-30, 2006, Revised Selected Papers (Lecture Notes in Computer Science, Vol. 4258)*, George Danezis and Philippe Golle (Eds.). Springer, 20–35. https://doi.org/10.1007/11957454_2

[10] Cloudflare. 2023. Internet traffic trends in China. https://radar.cloudflare.com/adoption-and-usage/cn

[11] Tim Dierks and Eric Rescorla. 2008. RFC 5246: The transport layer security (TLS) protocol version 1.2.

[12] D Eastlake 3rd. 2011. RFC 6066: Transport Layer Security (TLS) Extensions: Extension Definitions.

[13] Roya Ensafi, Philipp Winter, Abdullah Mueen, and Jedidiah R. Crandall. 2015. Analyzing the Great Firewall of China Over Space and Time. *Proc. Priv. Enhancing Technol.* 2015, 1 (2015), 61–76. https://doi.org/10.1515/popets-2015-0005

[14] Sadegh Hayeri. 2022. GreenTunnel. https://github.com/SadeghHayeri/GreenTunnel

[15] Nguyen Phong Hoang, Arian Akhavan Niaki, Jakub Dalek, Jeffrey Knockel, Pellaeon Lin, Bill Marczak, Masashi Crete-Nishihata, Phillipa Gill, and Michalis Polychronakis. 2021. How Great is the Great Firewall? Measuring China's DNS Censorship. In *30th USENIX Security Symposium, USENIX Security 2021, August 11-13, 2021*, Michael Bailey and Rachel Greenstadt (Eds.). USENIX Association, 3381–3398. https://www.usenix.org/conference/usenixsecurity21/presentation/hoang

[16] Jana Iyengar and Martin Thomson. 2021. QUIC: A UDP-Based Multiplexed and Secure Transport. RFC 9000. https://doi.org/10.17487/RFC9000

[17] Lin Jin, Shuai Hao, Haining Wang, and Chase Cotton. 2022. Understanding the Practices of Global Censorship through Accurate, End-to-End Measurements. In *SIGMETRICS/PERFORMANCE '22: ACM SIGMETRICS/IFIP PERFORMANCE Joint International Conference on Measurement and Modeling of Computer Systems, Mumbai, India, June 6 - 10, 2022*, D. Manjunath, Jayakrishnan Nair, Niklas Carlsson, Edith Cohen, and Philippe Robert (Eds.). ACM, 17–18. https://doi.org/10.1145/3489048.3522640

[18] krlvm. 2022. PowerTunnel. https://github.com/krlvm/PowerTunnel

[19] Alexander Master and Christina Garman. 2023. A Worldwide View of Nation-state Internet Censorship. *Free and Open Communications on the Internet* (2023).

[20] Niklas Niere. 2023. *Circumventing the GFW with TLS Record Fragmentation*. https://upb-syssec.github.io/blog/2023/record-fragmentation/

[21] Ram Sundara Raman, Leonid Evdokimov, Eric Wustrow, J. Alex Halderman, and Roya Ensafi. 2020. Investigating Large Scale HTTPS Interception in Kazakhstan. In *IMC '20: ACM Internet Measurement Conference, Virtual Event, USA, October 27-29, 2020*. ACM, 125–132. https://doi.org/10.1145/3419394.3423665

[22] Ram Sundara Raman, Mona Wang, Jakub Dalek, Jonathan Mayer, and Roya Ensafi. 2022. Network measurement methods for locating and examining censorship devices. In *Proceedings of the 18th International Conference on emerging Networking EXperiments and Technologies, CoNEXT 2022, Roma, Italy, December 6-9, 2022*, Giuseppe Bianchi and Alessandro Mei (Eds.). ACM, 18–34. https://doi.org/10.1145/3555050.3569133

[23] Eric Rescorla. 2018. Rfc 8446: The transport layer security (tls) protocol version 1.3.

[24] ValdikSS. 2022. GoodbyeDPI — Deep Packet Inspection circumvention utility. https://github.com/ValdikSS/GoodbyeDPI

[25] Philipp Winter and Stefan Lindskog. 2012. How the Great Firewall of China is Blocking Tor. In *2nd USENIX Workshop on Free and Open Communications on the Internet, FOCI '12, Bellevue, WA, USA, August 6, 2012*, Roger Dingledine and Joss Wright (Eds.). USENIX Association. https://www.usenix.org/conference/foci12/workshop-program/presentation/winter

[26] Xueyang Xu, Zhuoqing Morley Mao, and J. Alex Halderman. 2011. Internet Censorship in China: Where Does the Filtering Occur?. In *Passive and Active Measurement - 12th International Conference, PAM 2011, Atlanta, GA, USA, March 20-22, 2011. Proceedings (Lecture Notes in Computer Science, Vol. 6579)*, Neil Spring and George F. Riley (Eds.). Springer, 133–142. https://doi.org/10.1007/978-3-642-19260-9_14

[27] Diwen Xue, Reethika Ramesh, Valdik S. S, Leonid Evdokimov, Andrey Viktorov, Arham Jain, Eric Wustrow, Simone Basso, and Roya Ensafi. 2021. Throttling Twitter: an emerging censorship technique in Russia. In *IMC '21: ACM Internet Measurement Conference, Virtual Event, USA, November 2-4, 2021*, Dave Levin, Alan Mislove, Johanna Amann, and Matthew Luckie (Eds.). ACM, 435–443. https://doi.org/10.1145/3487552.3487858

[28] xvzc. 2023. SpoofDPI. https://github.com/xvzc/SpoofDPI