Summary of Research Achievements

XXX[1,*]

[1] *Computer Science and Technology, Zhejiang University, Hangzhou, China*

**12421094@zju.edu.cn*

Abstract

这篇文章主要用来梳理我的研究成果和思路想法, 这句话为了证明本文档可以支持中文! Your abstract text goes here. The abstract should be between 150 to 300 words and should provide a concise overview of the problem, methodology, main results, and conclusions.

1 to-do

1.1 database

Exploring QUIC Dynamics: A Large-Scale Dataset for Encrypted Traffic Analysis, 待论文录取后数据集公开

1.2 A novel QUIC traffic classifier based on convolutional neural networks

1.based on the convolutional neural network which combines the feature extraction and classification phase into one system.

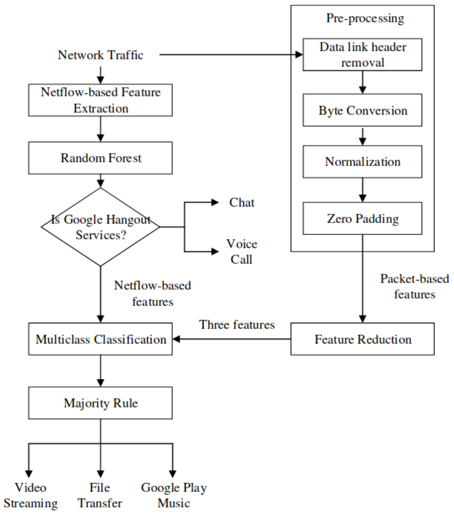


Figure 1: The structure of this database

2 to read

Lee, J., et al. (2023). ”Challenges and Solutions in Identifying QUIC Traffic.” Proceedings of the ACM Conference on Network Traffic Analysis.

Zhang, Y., & Chen, L. (2023). ”Using Machine Learning to Detect QUIC Traffic.” International Journal of Network Security.

Singh, A., et al. (2024). ”Flow-level Identification of QUIC Traffic using Heuristic Methods.” Computer Networks Journal.

Patel, R., & Roy, D. (2023). ”Hybrid DPI and Flow Analysis for QUIC Traffic.” IEEE Transactions on Network and Service Management.

Kim, S., et al. (2023). ”The Privacy Trade-offs in QUIC Traffic Identification.” Journal of Privacy and Security Research.

He, Y., et al. (2024). ”Adapting Traffic Analysis for Evolving QUIC Versions.” Journal of Protocol Engineering.

main points and key words

3 Video Datasets

1、An Encrypted Network Video Stream. Available online 22 June 2023 the local storage path The url that I downloaded the database author：捷克理工大学信息技术学院

the structure is showed in the figure 2.

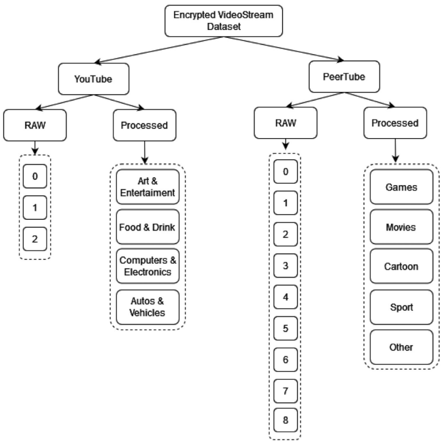


Figure 2: The structure of this database

4 QUIC

4.1 QUIC protocol

One of QUIC′ s main features is a fast connection establishment: In the ideal case, when cached information of a prior connection is available, it does not even take a single round-trip (0-RTT) to send encrypted application data. Yet, in the worst case (without prior connections as in our measurements), QUIC needs at least three round-trips as shown in Figure 3 and explained next.

Clients initiate a connection using a Client Hello (CHLO)(1) including the QUIC version it desires to use. In case the server does not support this version, it may send

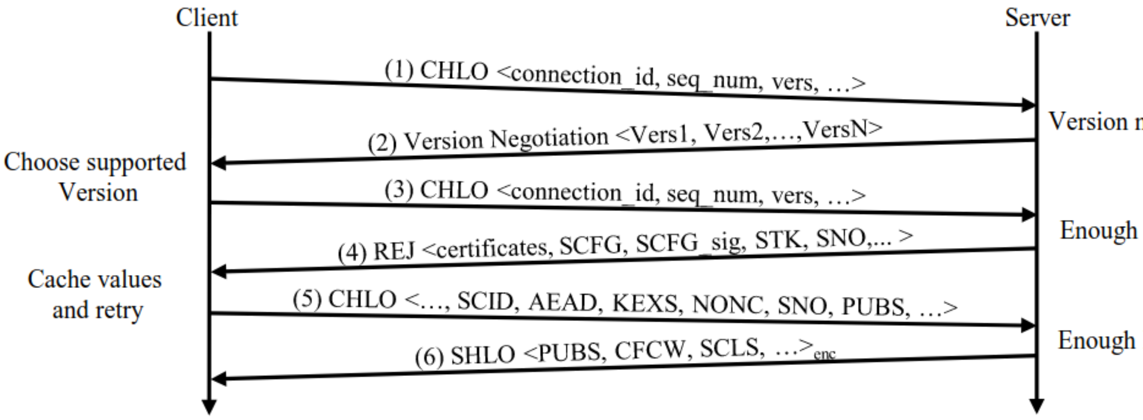

Client
Server
(1) CHLO <connection_id, seq_num, vers, …>
Version r
(2) Version Negotiation <Vers1, Vers2,…,VersN>
Choose supported Version
(3) CHLO <connection_id, seq_num, vers, …>
Enough
(4) REJ <certificates, SCFG, SCFG_sig, STK, SNO,... >
Cache values and retry
(5) CHLO <…, SCID, AEAD, KEXS, NONC, SNO, PUBS, …>
Enough
(6) SHLO <PUBS, CFCW, SCLS, …>enc

Figure 3: The structure of this database

a version negotiation packet (2) enabling the client to choose from a list of supported versions for a second try. We will utilize packet (1) to quickly probe for QUIC-capable hosts with only a single packet exchange and analyze their supported versions provided in (2). Using a supported version, the client may advance in the handshake by sending another CHLO (3), without prior communication, it does not possess enough information about the server to establish a valid connection. The server supplies the necessary information (4), in one or multiple exchanges (i.e., (3) and (4) may be repeated until all required data is available). In these step(s), the client will be given a signed server config (SCFG) including supported ciphers, key exchange algorithms and their public values, and among other things the certificates authenticating the host. We will utilize these information to analyze the server-provided certificates. With this information, the client can issue another CHLO (5) including enough information to establish a connection, the client may even send encrypted data following the CHLO which depicts the optimal case for a 0-RTT connection establishment. Following the CHLO, the server acknowledges (6) the successful connection establishment with a Server Hello (SHLO), containing further key/value-pairs enabling to fully utilize the connection. from [A First Look at QUIC in the Wild]

4.2 特点

□ 加密性：QUIC 使用了 TLS 1.3 进行端到端加密，这意味着所有的数据包都经过加密处理，流量特征在网络中难以直接进行内容分析。

□ 基于 UDP：QUIC 基于 UDP 协议进行传输，绕过了传统的 TCP 流量分析方法。

□ 流量混淆：QUIC 流量在传输过程中具有较高的混淆性，尤其是通过 TLS 1.3 加密的传输数据，使得基于传统的端口号、协议类型等特征进行分类变得更加困难。

4.3 挑战

□ 流量特征不明显：由于 QUIC 使用了加密和混淆技术，传统基于数据包内容（如协议字段、端口号等）的方法无法有效地识别 QUIC 流量。□TLS 加密头的变化：虽然 QUIC 使用了 TLS 1.3 协议进行加密，但其协议头结构与传统的 HTTPS 协议相比有所不同，导致加密流量识别变得更加复杂。□UDP 的多路径特性：QUIC 支持多路径传输（multipath），这意味着同一会话的数据包可以通过不同的路径发送，进一步增加了流量分析的复杂度。

5 Deep Packet Inspection (DPI) Challenges

As QUIC encrypts much of its traffic, it presents a challenge for traditional traffic analysis methods such as DPI. Researchers have explored several techniques to analyze QUIC traffic, such as looking at timing patterns and side-channel information, bypassing encryption restrictions.(Source: Lee et al., 2023), explores the challenges of using traditional DPI with QUIC and presents methods to decode QUIC traffic without needing full decryption, by analyzing the connection establishment phase.

6 Machine Learning for Traffic Classification

Machine learning has been increasingly applied to classify QUIC traffic. By training classifiers on features like packet sizes, timing, and flow duration, researchers have developed systems that can identify QUIC flows without needing to decrypt them. (Source: Zhang & Chen, 2023) discusses applying machine learning for traffic identification and its potential to overcome the encryption challenges imposed by QUIC, even in the presence of packet fragmentation.

7 Flow-Level Identification Techniques

Several techniques aim at distinguishing QUIC flows from traditional TCP/UDP flows based on unique characteristics like flow-level behavior or handshakes. One significant development is analyzing the QUIC handshake phase, which is distinct and can often be used for identification. (Source: Singh et al., 2024), in their work on flow identification, introduces new heuristics that focus on QUIC's handshake pattern, which could be leveraged by network operators to detect QUIC traffic early.

8 Enhanced DPI Approaches with QUIC-specific Features

Despite QUIC's encryption, there are still several identifiable features within the packet headers that can be exploited for traffic classification. Researchers are developing DPI tools that combine these features with machine learning to detect QUIC traffic more effectively. (Source: Patel & Roy, 2023), introduces a hybrid approach combining DPI and flow analysis, which can identify QUIC traffic by leveraging non-encrypted portions of the QUIC handshake.

9 Privacy and Anonymity Concerns

As QUIC's design emphasizes privacy and low latency, there is ongoing research into whether traffic classification efforts compromise these principles. Researchers argue that excessive reliance on identification techniques might undermine the core advantages of QUIC. (Source: Kim et al., 2023) discusses the trade-off between privacy, security, and traffic identification in QUIC networks, suggesting potential solutions for maintaining anonymity without compromising network analysis.

10 Protocol Analysis and Evolution

There have also been studies focusing on the evolving nature of QUIC itself. As new versions of QUIC are introduced (like QUICv2), there are concerns about the adaptability of existing traffic identification techniques. (Source: He et al., 2024) investigates the evolution of QUIC and provides a comprehensive overview of how traffic analysis must adapt to the upcoming versions of the protocol.

11 Hybrid Approaches for Identification

Hybrid methods combining DPI, machine learning, and statistical analysis are being developed to identify QUIC traffic. These methods aim to strike a balance between accuracy and minimal interference with encryption mechanisms. (Source: Gupta & Zhao, 2023), presents a hybrid methodology that uses both DPI and timing analysis to improve identification accuracy for QUIC traffic, even when encrypted.

12 measurement

The F1-score-

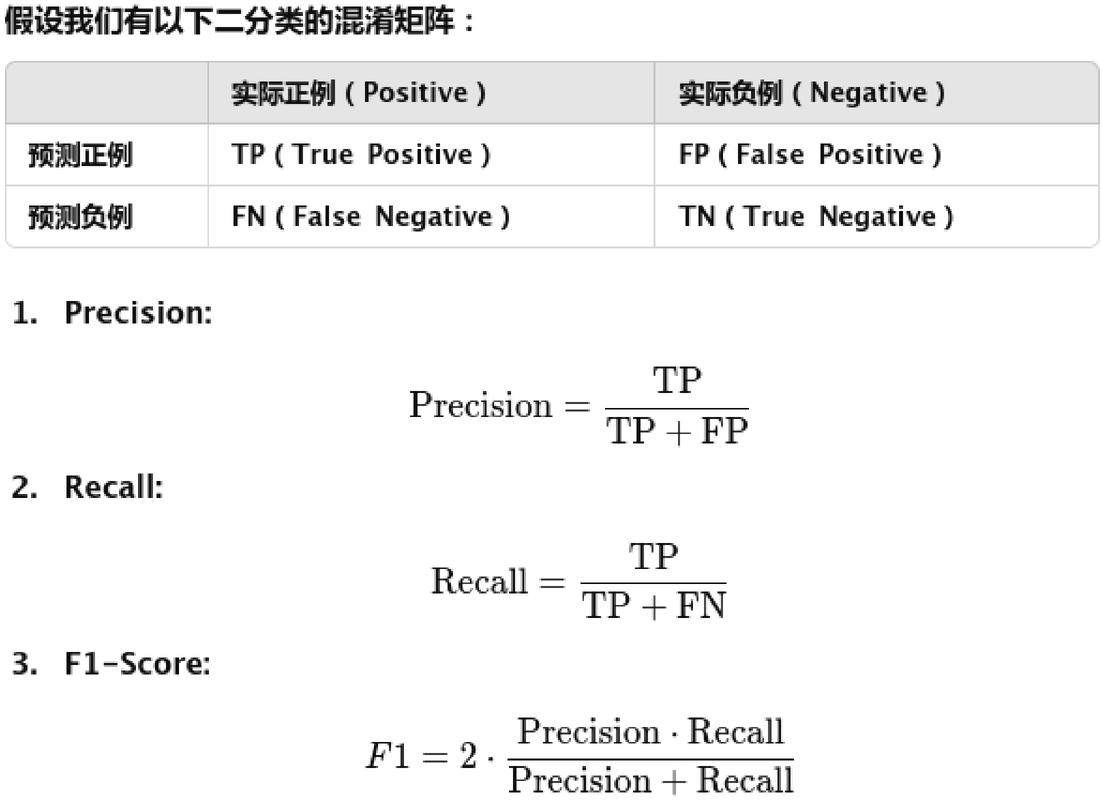
	实际正例（Positive）	实际负例（Negative）
预测正例	TP	FP
预测负例	FN	TN

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

其中:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$



假设我们有以下二分类的混淆矩阵：

	实际正例 (Positive)	实际负例 (Negative)
预测正例	TP (True Positive)	FP (False Positive)
预测负例	FN (False Negative)	TN (True Negative)

1. **Precision:**

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

2. **Recall:**

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

3. **F1-Score:**

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

Figure 4: The structure of this database

13 old cotent,just for template

Start by giving some motivations, then describe your work and your results. Accents can be written directly without any trouble as for instance the village of Vogüé in Ardèche. For units, siunitx is loaded, so do not hesitate to write about your high magnetic field of 9 T or your spin waves with frequencies up to 250 GHz, or even your spatial resolution in the nm range. You can add equations like this, and refer to them (see Eq. 1):

$$Q_{\text{topo}} = \frac{1}{4\pi}\int \vec{m}(\vec{r})\cdot\left(\frac{\partial \vec{m}}{\partial x}\times\frac{\partial \vec{m}}{\partial y}\right)\mathrm{d}\vec{r} \tag{1}$$

For references, that you may cite like this[**wolfSpintronicsSpinBasedElectronics2001**], use the Bibtex format in a file named `biblio.bib` and use biber as a compiler. You can get the Bibtex description from Zotero or from a DOI on this website: `https://www.doi2bib.org/`. The references section will then be filled automatically. We also suggest to add a nice figure with some caption to catch the attention of your audience (see Fig. 5), following this example:

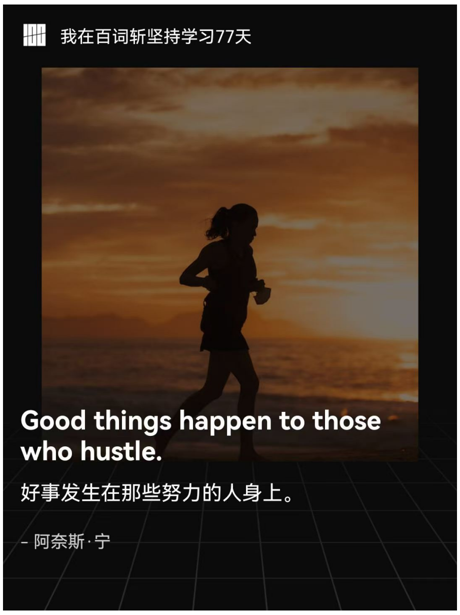

我在百词斩坚持学习77天
Good things happen to those who hustle.
好事发生在那些努力的人身上。
- 阿奈斯·宁

Figure 5: This is a spin, or maybe a chestnut.

Note that you can use figures in any usual format (tif, png, jpg... and even pdf), you simply need to upload the corresponding file in your Overleaf project. The size of the figure can easily be chosen with the “width” parameter.

Once you are ready, go to the conference website to upload your pdf abstract as the main file. This pdf file can be downloaded from Overleaf by clicking on the “Download PDF” icon in the right pannel (next to the “Recompile” button). **As a supplementary file, we ask you to upload a .zip file containing your .tex file, together with your biblio.bib file and your figure(s) (if any)**. Such a zip can be downloaded directly from Overleaf: in the (top left) overleaf Menu, click to download the Source; or from the project list that you see on your Overleaf home page (when clicking on the Home icon), you can click on the download icon which indicates “Download .zip file”. **This is mandatory, we need these files to create the final booklet.**

Table 1: xxxxxx.

me	nihao		
	xx	xx	xxe
xx	73%	66%	68%
xxxx	9%	73%	91%
xxxx	9%	75%	1%
xxx	91%	89.78%	92.28%
xxx	**9**%	**7**%	**9**%

Acknowledgments

If you wish to acknowledge people or funding, do it there.

Text of the abstract (one page maximum, including references).

References

[1] template, *teest*

[2] Codd, E. F. (1983). A relational model of data for large shared data banks. *Communications of the ACM*, 26(1), 64–69.