PhD School: AI and Privacy - Safeguarding IoT Data in the AI Era

Su Wang*
suwang@mines.edu
Department of Computer Science
Colorado School of Mines
Golden, Colorado, USA

Abstract

Internet of Things (IoT) devices have been increasingly deployed in many people's homes for a more convenient life. However, significant research has shown that users' IoT traffic data could be collected by Internet Service Providers (ISPs) and equipment manufacturers, which can lead to information leakage. In other words, on-path hackers and adversaries could analyze users' traffic data to infer their privacy information, which is known as an IoT traffic analytics (TA) attack. Moreover, Artificial Intelligence (AI) development holds great potential for human progress, especially in extracting and analyzing data features. Additionally, AI will bring significant opportunities and challenges to cybersecurity, as it can improve threat detection and defenses while also creating new weaknesses and more advanced ways for attacks. Consequently, integrating AI into cybersecurity has become an urgent research necessity. My Ph.D. research goal is to build a comprehensive framework to help people better manage their IoT data privacy leakage. As we develop systems to handle IoT traffic data, we face challenges such as prior defense or attack methods being not open-source, and most AI models being unstable or too large to run on small devices. Furthermore, IoT devices are updated very quickly, and user behaviors could be diverse. To further address these problems, I will establish a system that uses AI models to defend against TA attacks, which could be widely adopted under various IoT scenarios.

CCS Concepts

• **Computer systems organization** → *Sensor networks*; **Embedded software**; • **Information systems** → *Computing platforms*; • **Security and privacy** → **Vulnerability management**; *Mobile and wireless security.*

Keywords

IoT Privacy, Data Analytics, Inference Attack , AI Powered, Deep learning

1 BACKGROUND AND RELATED WORK

1.1 The Internet of Things (IoT)

The IoT refers to a vast network of devices that are connected to the internet, enabling devices to collect, share, and analyze data. The core of the IoT is the connection between the internet and the smart devices according to user command. Typically, smart homes connect to the internet through WiFi provided by Internet Service Providers (ISPs), using routers or IoT hubs to support multiple devices accessing the internet. Many IoT devices are web and cloud service-enabled, like Google Home and Amazon Alexa, which download traffic data from the manufacturer's server or upload data from user commands.

1.2 IoT traffic data rate

Currently, in modern smart homes, many IoT devices are heavily user-interactive, like voice command or remote control by the phone, so they all generate bidirectional traffic data including incoming and outgoing traffic. Unlike computers and smartphones, many IoT devices transmit small but frequent data packages. While the traffic contains so much information and is changeable, it's still surprisingly easy to infer user privacy information through traffic analytics (TA) attacks. The reason is each IoT device has unique traffic patterns based on its function and user activities. Regarding recent research, it can be clearly identified the IoT device types and user activities by TA [1, 17]. So there is much prior work preventing privacy leakage from IoT traffic data rates [1–8, 10–12, 14–17] .

1.3 The Development of AI

The advent of Artificial Intelligence (AI) has significantly enhanced data processing capabilities, enabling more efficient handling of large-scale and complex datasets. According to the research recently, AI has strong ability to draw the features of data distribution [9, 13].In cybersecurity, AI is ability to quickly analyze and understand patterns has greatly improved the ability of learning features. New advances in machine learning have made intrusion detection systems stronger, helped automate malware analysis, and built better defenses against cyberattacks. By using AI, cybersecurity can not only create stronger protections but also defend itself better from AI-based attacks.

1.4 Privacy Threat Model

As shown in Figure 1 we can see the threat of information leakage in smart home IoT traffic because of external adversaries, especially ISPs, network observers, IoT manufacturers, and other on-path adversaries. Unfortunately, the adversaries can get traffic data details surprising easily like incoming and outgoing traffic data rates, network protocols, source and destination IP addresses. By observing this data from IoT devices, adversaries can apply many approaches to infer user privacy information. For example these types of information, including user occupancy, network traffic patterns, and both short and long-term user activities, may be vulnerable to threats and exploitation by adversaries. Thus, this kind of inference represents a significant privacy concern, as we typically do not wish it to happen in users' smart homes.

2 Challenges

Although recent research proposes significant work to thwart privacy TA attacks on IoT traffic rates, there's currently no systematic method to adaptively help users protect their privacy information.

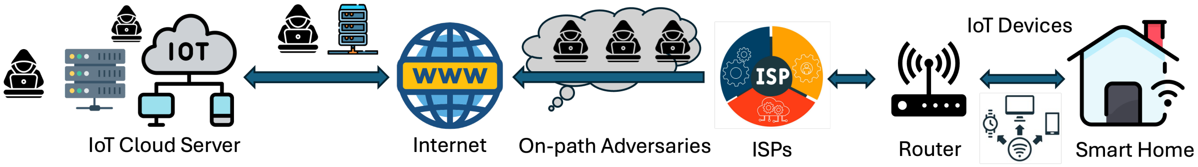

Figure 1: Privacy Threat Model.

In this section, we outlined the key challenges we met when researching, designing, implementing, and evaluating our system.

"Closed" Source Code and Evaluation Dataset Although prior work made a great contribution to IoT cybersecurity, most of the work, both attack and defense approaches, are not open source and the evaluation data are not clarified. Thus, this makes it really hard for me to evaluate each method and calculate the overhead of their presented algorithms on different datasets. To overcome this challenge, we will build an open-source system to assist researchers in further research and help users manage their IoT data privacy leakage.

Large "Volume" and unstable of AI model. Although prior AI models show significant contributions to learning the features from data, unfortunately existing AI models are still facing the large "volume" issue and instability, especially for some small IoT devices. To achieve faster and better performance, AI models tend to rely on advanced computing performance CPUs or even GPUs. However, this is not really friendly to IoT system design. Furthermore, most of the AI models' ability to analyze features is highly sensitive to different datasets, and slight changes in parameters can also cause significant shifts in accuracy and other metrics. To overcome these problems, we plan to build a widely applicable AI model for all IoT traffic data scenarios with robust parameters.

Diversity of IoT Scenarios and Rapid Device Evolution Different users have different IoT usage habits, especially as the diversity of IoT scenarios presents huge differences. Thus, it is a huge challenge to build an algorithm and a system to fit all the scenarios. Additionally, the rapid evolution of IoT devices leads to frequent updates and new functions. This constant change increases the complexity of maintaining robust privacy protection and secure communication across different IoT ecosystems. To overcome this issue, we decided to use large-scale data for learning to build a widely adaptable and robust IoT security system.

Limitation of Computing Resource One major challenge is the limitation of computing resources. While AI has made significant progress, often relying on powerful CPUs or GPUs, when deployed, IoT cybersecurity systems online will lack such high-performance computing hardware, which will impact the real-time processing capabilities. To address this issue, I plan to utilize some lightweight AI tools which will be optimized for smaller devices. We are aiming to reduce the overhead of our system and increase processing speed without reducing performance.

3 Proposed work

Regarding the significant recent research, it shows that launching traffic analytics (TA) attacks is surprisingly easy, since user activities highly correlate with simple time-series data statistical metrics. Thus, IoT device traffic rates alone pose significant user privacy threats. Despite the increasing volume of literature addressing the defense against these malicious IoT traffic analytics, there is currently a lack of a systematic method to compare and assess the comprehensiveness of these existing studies. Thus, my future research will combine AI and cyber security to build a whole framework to help people manage their IoT data privacy leakage and Enhancing IoT cyber security through AI-driven defenses, secure data processing, and advanced privacy protection techniques.

AI-Powered Adversarial Attacks and Defenses. In the future, I will explore the Generative Adversarial Network and Adversarial Machine Learning to simulate potential attacks on IoT systems. By generating and desecrating the traffic data, AI models can better understand the data features. Meanwhile, it can make the system analyze every IoT traffic scenario and implement appropriate protection mechanisms.

Attacks on AI and Deep Learning Models. Investigating the information leakage issues in AI models and deep learning models will be another focus. As AI performance develops, we cannot only focus on improving AI performance, but also need to pay attention to the security issues of AI models and data. Unfortunately, AI models are still facing severe data and model leakage. Thus, my goal is to identify these weaknesses and potential security risks of AI models.

IoT Data Encryption. I plan to explore encryption protocols like zero-knowledge proofs and secure multi-party computation to ensure IoT data remains secure, even when shared across different systems or entities.

Distributed Inference Models. In the future, I will build distributed inference models that process data and models across multiple locations, reducing the risk of data and model exposure. This method will ensure no single point contains enough information to fully infer user behavior and privacy, thus enhancing overall privacy protection in IoT systems.

Benchmarking and Deployment. I will benchmark and deploy protection approaches at various points in the IoT ecosystem, including IoT hubs, gateways, routers, and edge servers. By testing these mechanisms at different levels, I can ensure that they are adaptable and effective in real-world smart home environments. The deployment of protection systems at multiple levels will provide better IoT cybersecurity.

4 Summary

In conclusion, my research focuses on addressing privacy issues in IoT systems by developing a comprehensive framework to manage IoT data privacy leakage. With the increasing deployment of IoT devices, user privacy becomes vulnerable to TA attacks, which can infer sensitive user activities by analyzing traffic rate patterns. unfortunately, Traditional AI models, the rapid evolution of IoT devices, diverse user habits and the closure of the code challenge our research. To overcome these challenges, my research will build a system to help researches to manage the IoT traffic and I will focus on proposing an approach that includes lightweight AI-powered tools optimized for IoT device systems. Also, we will focus on advanced encryption methods like zero-knowledge proofs. In addition, by utilizing distributed inference models, the system will reduce the risk of data leakage by spreading data processing across multiple locations. Meanwhile, we aim to benchmark and deploy these protection mechanisms across various levels of the IoT ecosystem. This framework will give users greater control over their privacy and protect their data from leakage.

References

[1] 2023. PAROS. https://github.com/cyber-physical-systems/paros.

[2] Noah Apthorpe, Danny Yuxing Huang, Dillon Reisman, Arvind Narayanan, and Nick Feamster. 2019. Keeping the smart home private with smart (er) iot traffic shaping. *Proceedings on Privacy Enhancing Technologies* 2019, 3 (2019), 128–148.

[3] Phuthipong Bovornkeeratiroj, Srinivasan Iyengar, Stephen Lee, David Irwin, and Prashant Shenoy. [n. d.]. RepEL: A Utility-preserving Privacy System for IoT-based Energy Meters. In *2020 IEEE/ACM Fifth International Conference on Internet-of-Things Design and Implementation (IoTDI'20)*.

[4] Xiang Cai, Rishab Nithyanand, Tao Wang, Rob Johnson, and Ian Goldberg. 2014. A systematic approach to developing and evaluating website fingerprinting defenses. In *Proceedings of ACM SIGSAC Conference on Computer and Communications Security*. ACM, 227–238.

[5] Dong Chen, David Irwin, Prashant Shenoy, and Jeannie Albrecht. 2014. Combined heat and privacy: Preventing occupancy detection from smart meters. In *IEEE International Conference on Pervasive Computing and Communications*. 208–215.

[6] Roger Dingledine, Nick Mathewson, Paul F Syverson, et al. 2004. Tor: The second-generation onion router.. In *USENIX security symposium*, Vol. 4. 303–320.

[7] 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 *2012 IEEE symposium on security and privacy*. IEEE, 332–346.

[8] Md Kamrul Hasan, Husne Ara Rubaiyeat, Yong-Koo Lee, and Sungyoung Lee. 2008. A reconfigurable HMM for activity recognition. In *2008 10th International Conference on Advanced Communication Technology*, Vol. 1. IEEE, 843–846.

[9] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long Short-Term Memory. *Neural Comput.* 9, 8 (Nov. 1997), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

[10] Homin Park, Can Basaran, Taejoon Park, and Sang Hyuk Son. 2014. Energy-efficient privacy protection for smart home environments using behavioral semantics. *Sensors* 14, 9 (2014), 16235–16257.

[11] Vasanthan Raghavan, Greg Ver Steeg, Aram Galstyan, and Alexander G Tartakovsky. 2014. Modeling temporal activity patterns in dynamic social networks. *IEEE Transactions on Computational Social Systems* (2014).

[12] Karsten Rothmeier, Nicolas Pflanzl, Joschka Hüllmann, and Mike Preuss. 2020. Prediction of Player Churn and Disengagement Based on User Activity Data of a Freemium Online Strategy Game. *IEEE Transactions on Games* (2020).

[13] Alex Sherstinsky. 2020. Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network. *Physica D: Nonlinear Phenomena* 404 (March 2020), 132306. https://doi.org/10.1016/j.physd.2019.132306

[14] Vitaly Shmatikov and Ming-Hsiu Wang. 2006. Timing analysis in low-latency mix networks: Attacks and defenses. In *European Symposium on Research in Computer Security*. Springer, 18–33.

[15] Tao Wang and Ian Goldberg. 2017. Walkie-talkie: An efficient defense against passive website fingerprinting attacks. In *26th USENIX Security Symposium*.

[16] Wei Wang, Mehul Motani, and Vikram Srinivasan. 2008. Dependent link padding algorithms for low latency anonymity systems. In *Proc. of 15th ACM conference on Computer and communications security*.

[17] Keyang Yu, Qi Li, Dong Chen, Mohammad Rahman, and Shiqiang Wang. 2021. PrivacyGuard: Enhancing Smart Home User Privacy. In *Proceedings of the 20th International Conference on Information Processing in Sensor Networks (IPSN'21)* (Nashville, TN, USA) *(IPSN '21)*. 62–76.