Research Statement

June 2023

Name	Hyeonmin Lee (Postdoctoral Researcher)
Affiliation	Network Convergence and Security Lab at Seoul National University
Email	min0921110@gmail.com
Webpage	https://hyeonmin-lee.github.io/

1. Research Interests

My research interests lie in the field of **network security**, with a focus on identifying security issues in network systems and designing practical solutions to address them. In my research, I often utilize (large-scale) Internet measurement to identify potential security problems.

My current research focuses on areas including, but not limited to, DNS security, email security, and PKI/TLS, as these play crucial roles in the current Internet. DNS lookups are preceded for almost every activity on the web, and email communication often contains sensitive and private information. TLS provides a foundation for secure communication. Thus, ensuring the security of these systems is essential to protect users' privacy and prevent cyber attacks. My research aims to identify potential security vulnerabilities in these systems and develop practical solutions to enhance their security. Additionally, I am interested in improving existing security protocols from a practical standpoint.

2. Knowledge Background

Throughout my academic career, I have participated in (or led) numerous research projects and collaborated with experts from institutions including Virginia Tech, Rochester Institute of Technology, University of Twente, SIDN Labs, and NLnet Labs. These experiences have provided me with a diverse range of knowledge in various areas, including (but not limited to):

- **Network Security:** I possess an extensive understanding of various networks security techniques including, DNS security (i.e., DNSSEC, DNS over TLS, DNS over HTTPS), email security (i.e., STARTTLS), PKI, TLS, DNS-based Authentication of Named Entities (DANE).
- **Internet Measurement:** I have experience in large-scale Internet measurement (e.g., DNS records and certificate crawling). Also, I am capable of processing and analyzing massive data using frameworks like Spark and Hadoop.
- **Others:** I have conducted research on distributed systems (e.g., edge computing) and the Internet of Things (IoT), with a focus on their security. Additionally, I have knowledge of hardware-supported trusted computing technologies (particularly, Intel SGX) and Software Defined Networking (SDN).

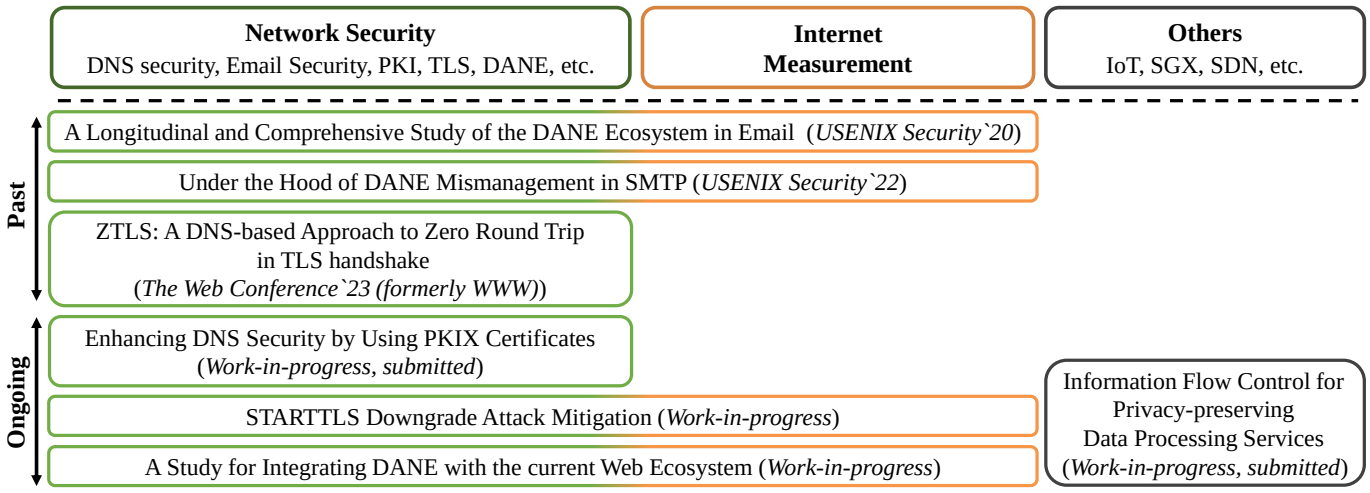

Figure 1: **Research topics and background knowledge.** I have published three papers in top-tier conferences, including USENIX Security and The Web Conference (formerly WWW), and I am actively involved in several ongoing research projects.

3. Past and Ongoing Work

3.a. Past Work

A Longitudinal and Comprehensive Study of the DANE Ecosystem in Email (USENIX Security '20 - First author) [9]

Motivation. DNS-based Authentication of Named Entities (DANE) [4] was proposed to enable authentication of communication peers without relying on Certificate Authorities (CAs). At present, the DANE protocol is primarily used in mail transfer where it is used to authenticate SMTP servers. However, the operation of DANE involves complex procedures, including the cooperation of DNS servers and SMTP servers, which can be prone to mismanagement by server operators. In light of this, we conducted research to investigate how DANE is managed in the email ecosystem in practice.

Our approach. We carried out a large-scale network measurement that spanned two years. We scanned DNS records related to DANE (`TLSA` and DNSSEC records) for all second-level domains within three generic top-level domains (`.com`, `.net`, and `.org`) and two country code top-level domains (`.nl` and `.se`). Furthermore, we collected certificates from SMTP servers operating under those domains. Our dataset revealed that DANE deployment is still in its early stages, with less than 1% of the domains in `.com`, `.net`, and `.org` supporting DANE. Additionally, we found that a significant number of SMTP servers were not managing DANE properly. Finally, to provide a comprehensive analysis of DANE support from the client-side, we conducted a survey of 29 popular email providers, 4 widely-used mail programs, and 11 DNS software.

Notes. This paper was the result of international joint research conducted with researchers from Rochester Institute of Technology, Amrita Vishwa Vidyapeetham, University of Twente, and NLnet Labs. I presented the paper at USENIX Security 2020.

- I have made all datasets and measurement codes publicly available at `https://dane-study.github.io/`

Under the Hood of DANE Mismanagement in SMTP (USENIX Security '22 - First author) [8]

Motivation. In a previous study on DANE deployment [9], we discovered that numerous DANE-supporting SMTP servers experienced mismanagement issues, including missing essential DNSSEC records or unmatching between `TLSA` records and (SMTP servers') certificates. Our objective in this study was to identify the root causes of these issues.

Our approach. For the DANE protocol, domain owners have the option to either self-manage or outsource the management of their DNS and SMTP servers. We conducted research to determine whether the managing entity of these servers affects the quality of DANE management. To accomplish this, we developed a methodology that utilizes `NS` and `MX` records to identify the managing entity of DNS and SMTP servers. Our findings indicate that self-management of DNS or SMTP servers leads to more errors compared to outsourcing. Additionally, we found that a significant portion of SMTP servers failed to correctly rollover their keys (in certificates). Furthermore, we found that a significant number of SMTP servers failed to rollover their keys (in certificates) correctly. This was primarily due to many servers using automated certificate renewal (such as *Let's Encrypt*) without updating their corresponding `TLSA` records. To address this issue, we provided an automation tool that supports DANE management.

Notes. Notes: This paper was the result of international joint research conducted with researchers from Virginia Tech, University of Twente, NLnet Labs, and SIDN Labs. I presented the paper at USENIX Security 2022.

- I have made all datasets and measurement codes publicly available at `https://dane-study.github.io/`

ZTLS: A DNS-based Approach to Zero Round Trip in TLS handshake (The Web Conference '23 - Second author) [11]

Motivation. Transport Layer Security (TLS) is the de facto standard for a secure connection. However, the TLS handshake process requires one additional round trip time (RTT) to negotiate a session key, which could degrade user satisfaction. We focus on the fact that (1) establishing a TLS session usually requires a DNS lookup (e.g., the `A` record lookup to fetch the IP address of a server) and (2) the DNS infrastructure can disseminate server-related data at some points close to clients.

Our approach. We leverage the DNS to distribute TLS handshake-related data (i.e., Diffie-Hellman elements) in advance as a DNS record called Z-data. This record, published by a server, includes information such as a signature and a certificate that provides authentication for its issuer (i.e., domain) and the integrity of itself. After fetching a server's IP address (`A` record) and Z-data simultaneously, a client can generate a session key and send encrypted data with a 0-RTT delay. ZTLS is the first approach that exploits DNS to send encrypted data with 0-RTT delay.

Notes. This paper was published at The Web Conference (formerly known as WWW) in 2023.

3.b. Ongoing Work

Enhancing DNS Security by Using PKIX Certificates (Submitted to a conference - Second author / IETF Internet-Draft - First author) [10]

Motivation. DNSSEC was proposed 20 years ago to guarantee the integrity of DNS records, but only around 7% of second-level domains have deployed it [6]. To deploy DNSSEC, a domain has to upload its DS record (hash of a public key used to sign the domain's DNS records) to its parent zone, which is a manual process and is often prone to misconfiguration. Due to this complex procedure, it is reported that many domains suffer from prevalent misconfiguration; 30% of domains that support DNSSEC do not have DS records in their parent zones [1]. Thus, we need a more practical method that does not require the cooperation of other entities (i.e., parent zones) in the DNS infrastructure.

Our approach. We leverage the fact that the vast majority of domains already use certificates to support HTTPS/TLS (e.g., 95% of Web traffic to Google is HTTPS [2]). This suggests that domains are already familiar with how to use and manage certificates. Thus, we propose using certificates (issued by CAs) to sign their DNS records, ensuring the integrity of their DNS records with a low technical barrier. In our design, a server signs its DNS records using a public key (in a certificate) and uploads the signature, public key, and certificate (chain) as DNS records. A client can check the integrity of a DNS record by (1) fetching the corresponding signature, public key, and certificate (chain), (2) validating the certificate chain, and (3) verifying the signature.

Notes. This research has been documented as an Internet-Draft at IETF.

STARTTLS Downgrade Attack Mitigation (Work-in-progress)

Motivation. STARTTLS [3] is commonly used to encrypt traffic between mail servers. However, it is well-known that STARTTLS is vulnerable to downgrade attacks. Although there are mechanisms to mitigate these attacks, such as MTA-STS [12] and DANE, they each have limitations. MTA-STS uses a Trust-On-First-Use (TOFU) mechanism that is susceptible to preventing policy discovery attacks [12]. Also, DANE has a very low deployment rate.

Our approach. One option to prevent STARTTLS downgrade attacks is to use a Certificate Transparency (CT) log [7]. A client-side SMTP server can check whether a server-side SMTP server has been issued a certificate by a CA using the CT log. However, the limitation of this approach is that not all CAs append their logs to CT. Therefore, we are exploring more viable ways to prevent STARTTLS downgrade attacks. Additionally, we are examining a design that can detect and warn users about STARTTLS downgrade attacks.

Notes. I am currently funded as a postdoctoral researcher by the BK21 FOUR Intelligence Computing group at Seoul National University for this project.

A Study for Integrating DANE with the current Web Ecosystem (Work-in-progress)

Motivation. Currently, the DANE protocol is mainly used for SMTP server authentication in mail transfers. There have been attempts to adapt DANE for the web, but these attempts have not been successful. One of the reasons is that middleboxes often discard DNS records, including DNSSEC records, which hinders clients such as browsers from performing DANE validation [5]. However, recent proposals such as DNS over TLS (DoT) and DNS over HTTPS (DoH) have the potential to facilitate the adoption of DANE to the Web. Also, research in this area has been limited even though it can solve significant problems that exist in the current CA-based PKI.

Our approach. In this study, we provide a blueprint of how the Web ecosystem could be transformed by the adoption of DANE. We examine the issues that currently exist in the (CA-based) Web ecosystem, such as private key delegation to CDN servers or problems arising from Certificate Transparency (CT), and explore how these problems can be resolved through the adoption of DANE to the Web. Additionally, we will conduct experimental tests to evaluate the overheads associated with DANE adoption, such as network delay, payload overhead, and management overhead.

Notes. This research project is supported by the National Research Foundation of Korea (NRF), and I am the primary investigator of the project.

Information Flow Control for Privacy-preserving Data Processing Services (Submitted to a top-tier security conference - Third author)

Motivation. When using online data-processing services such as Pixlr (photo editing service) or ezyZip (file compression service), users often upload their data to the service provider's server, which can result in a loss of control over their private data. Users are unable to determine how their data is being processed and whether the service provider is storing their data for purposes beyond the stated privacy policy. Therefore, a mechanism is required to ensure that users' data is not leaked to the service provider or cloud provider if the service is deployed on a cloud.

Our approach. We have developed an Information Flow Control (IFC) framework that provides a sandboxed execution environment for arbitrary data-processing programs. The framework includes a blackbox monitor that intercepts every data flow into and out of the sandbox and blocks any data flow that could potentially leak user data. Our implementation is based on Intel Software Guard eXtensions (SGX) and extends NativeClient (NaCl) to enforce IFC rules.

4. Conclusion and Future Direction

My past and ongoing work have focused on identifying and addressing security issues in networks. I have attempted to deeply understand the mechanisms and ecosystems of DNS, email, and PKI/TLS, and often employ large-scale Internet measurement to gain insights into these systems. As a solution, I have designed several practical security protocols.

Currently, I am finalizing some of my ongoing projects. My next plan is to advance my other ongoing projects, including mitigating STARTTLS downgrade attacks. Additionally, I am interested in measuring the deployment and management of newer security protocols, such as MTA-STS or TLS-RPT, and analyzing the impact of middleboxes on DNS security.

References

[1] Taejoong Chung, Roland van Rijswijk-Deij, Balakrishnan Chandrasekaran, David Choffnes, Dave Levin, Bruce M Maggs, Alan Mislove, and Christo Wilson. A Longitudinal, End-to-End View of the DNSSEC Ecosystem. In *Proceedings of the 26th USENIX Security Symposium (USENIX Security 17)*, pages 1307–1322, 2017. https://www.usenix.org/system/files/conference/usenixsecurity17/sec17-chung.pdf.

[2] Google. HTTPS encryption on the web. https://transparencyreport.google.com/https/overview.

[3] Paul Hoffman. SMTP Service Extension for Secure SMTP over Transport Layer Security. RFC 3207, 2002. https://tools.ietf.org/html/rfc3207.

[4] Paul Hoffman and Jakob Schlyter. The DNS-Based Authentication of Named Entities (DANE) Transport Layer Security (TLS) Protocol: TLSA. RFC 6698, 2012. https://tools.ietf.org/html/rfc6698.

[5] Austin Hounsel. Measuring Middlebox Interference with DNS Records. https://blog.mozilla.org/security/2020/11/17/measuring-middlebox-interference-with-dns-records/.

[6] Richard Lamb. DNSSEC Deployment Report. http://rick.eng.br/dnssecstat/.

[7] Ben Laurie, Eran Messeri, and Rob Stradling. Certificate Transparency Version 2.0. RFC 9162, 2021. https://tools.ietf.org/html/rfc9162.

[8] Hyeonmin Lee, Md. Ishtiaq Ashiq, Moritz Müller, Roland van Rijswijk-Deij, Taekyoung "Ted" Kwon, and Taejoong Chung. Under the Hood of DANE Mismanagement in SMTP. In *Proceedings of the 31st USENIX Security Symposium (USENIX Security 22)*, 2022. https://www.usenix.org/system/files/sec22-lee.pdf.

[9] Hyeonmin Lee, Aniketh Gireesh, Roland van Rijswijk-Deij, Taekyoung "Ted" Kwon, and Taejoong Chung. A Longitudinal and Comprehensive Study of the DANE Ecosystem in Email. In *Proceedings of the 29th USENIX Security Symposium (USENIX Security 20)*, 2020. https://hyeonmin-lee.github.io/publication/paper/lee-2020-dane.pdf.

[10] Hyeonmin Lee and Taekyoung Kwon. DNSSEC Extension by Using PKIX Certificates. Internet-Draft, 2023. https://datatracker.ietf.org/doc/draft-dnsop-dnssec-extension-pkix/.

[11] Sangwon Lim, Hyeonmin Lee, Hyunsoo Kim, Hyunwoo Lee, and Taekyoung Kwon. ZTLS: A DNS-based Approach to Zero Round Trip in TLS handshake. In *Proceedings of the ACM Web Conference 2023 (TheWebConf 23, formerly WWW)*, 2023. https://doi.acm.org?doi=3543507.3583516.

[12] Daniel Margolis, Mark Risher, Binu Ramakrishnan, Alexander Brotman, and Janet Jones. SMTP MTA Strict Transport Security (MTA-STS). RFC 8461, 2018. https://tools.ietf.org/html/rfc8461.