This abstract was submitted to and presented at CCRC International Conference 2022.

Dependency among Data, Code, Governance, and Operation in Trust

Shigeya Suzuki*1, Tatsuya Kurosaka*1, Jun Murai*2

*1 Graduate School of Media and Governance, Keio University
*2 Keio University

In this extended abstract, we discuss the chain of the trust model in terms of Data, Code, Governance, and Operation to reveal the issues around the scheme to establish the chain of trusts in a holistic view.

Online Services, Digital Identity, and Privacy

People capable of working with information technology typically have multiple Internet service accounts. In the beginning, Internet service providers just allowed users to create their identities within their system. This identity information typically includes login identifier and password. But nowadays, the service providers often require users to bind their real identities with the account for the system. These "Digital Identities" consist of the Person Identifiable Information such as name, day of birth, postal addresses, mobile phone numbers, etc. Technically, these pieces of information are called "Attributes of Identity [REF-Attribute]."

Some services require our Digital Identity for our sake. Banking services require our detailed information because they maintain our essential assets, identify us for later necessary communication, and for regulations. In other words, the reason why they need our identity is to correctly bind our information with the assets they're maintaining for us. Banking services don't need to use the customer's identity outside their services.

On the contrary, some businesses want our Digital Identity for their own business merits. Social networking service providers often want detailed information to provide better results. The results may or may not provide good values for the users. In other words, they need our identity because they want to profile us from our activities on the Internet as much as possible. Thus, to achieve their purpose, they want to track and consolidate our activities on the Internet.

Internet services are evolving quickly, and sometimes, we need to reveal our Digital Identities in specific contexts — at least partially. Since we live in a digitalized world, once we reveal part of the attribute, it is usually tough to revert the disclosure of the attribute later. Due to this fact, it is always safer to provide a minimum amount of information. For example, if we want to buy alcohol, we may need to prove that we can purchase alcohol or above. We will use identification cards with photos in the Physical World, such as driver's licenses.

While you only need to prove that you are at a certain age, we reveal the actual age unnecessarily. From this point of view, current physical identification card-based proof is not ideal. Suppose we need to reveal a digital identification card to somebody since the person who checks the age can easily copy the entire information. In that case, it might be more severe than the Physical World case. Thus, there is an intense desire for attribute-based proof with minimum disclosure of the person's attributes.

Trust and Authentication work

Trust is a complex concept. Let us introduce a definition of "Trust" in this paper: When we trust a subject, we do not need to rely on other entities to judge the trustworthiness — assumption that the subject acts as expected. But, before taking action based on trust, you need to be sure the entity you're acting on really is the subject you believe to be with; Before the action, you need to judge the entity's authenticity. You may need to rely on other entities with trust to verify authenticity.

The following section discusses the authentication model and the relationship to Trust in each Physical World and the Cyber World.

Authenticating Subjects in Physical World

Trust in the physical world is a blurred concept. Most of us request banks to keep our assets safely. One of the possible reasons we opt to do this is that keeping money at a banking service is secure than keeping money in our residence. Why do we trust these banks? Maybe, it is a major bank, has a long business history, owns multiple bank locations, or possibly your parents using the service.

Besides trusting a subject, you also need to be sure about the subject's authenticity. Some documents help you judge the subject's authenticity in Physical World. You can use certificates (including identification cards), physical confirmation of existence, or other attribute-like information as a basis for making judgments on the authenticity of the subject.

Unfortunately, the authenticities of these certificates are weak. The certificates may have physically unique characteristics, but it is hard to judge whether the certificate is valid without the help of devices or other schemes that can judge the certificates' authenticity. Such devices or schemes are not common; thus, people need to decide only with a visual inspection which might be not good enough or inefficient.

Authenticating subjects in Cyber World

Trust in the Cyber World is modeled based on the Physical World; due to that, the model has both strengths and weaknesses. When creating a user service on the Internet, the system architect must choose an authentication mechanism that resolves the user-provided information into the user's Digital Identity. Depending on the authentication mechanism, the mechanism itself requires the trust of the entities involved. For example, suppose the system requires a user-specified delivery-verifiable email address which the user owns. In that case, the system relies on the email delivery mechanism, which depends on multiple components,

including mail servers, domain name system, etc. The fundamental component of authenticating a subject in the Cyber World depends on Public Key Cryptography. With Public Key Cryptography, the owner of a particular key can prove the ownership of the private key by providing the signature of some information to the verifier by using its private key. Assuming each entity in the system owns Public Key Cryptography key pairs, the entity holds complete control of the private key. One entity — issuer — can describe another entity — subject — as relationship data with a signature using the issuer's key. The relationship data consists of information of the subject, including the subject's public key, related metadata, information of the issuer, which includes the public key of the issuer, signed with the issuer's private key. The relationship data is often called the "Public Key certificate of the subject." X.509 Public Key Infrastructure (PKI)[PKI] is the standard widely deployed on the Internet.

Single Public Key Certificate only describes a single issuer-to-subject relationship. When connecting a client to a server using PKI certificates, it is possible to verify the public key in use if the verifying entity owns the peer's PKI certificate before connecting. But the most cases, the client does not know the server's certificate before the connection. Also, since a self-issued certificate is not trustworthy, a certificate issued by a trusted third party is necessary. Thus, PKI provides the model to create a hierarchy of relationships of certificates. Entities named "Certificate Authorities" (CAs) issue a certificate for a subject. By hierarchically organizing CAs, when verifying the peer's public key, the verifier only needs to know "root" CA, then traversing tree hierarchy towards the peer's certificate to verify. This traversed path creates a chain of certificates starting from the "root" CA towards the peer's (leaf) certificate. We call this "chain of trust." The chain of trust provides the way to prove a public key in the certificate is owned by the subject described in the certificate. The certificate only verifies the integrity of the public key certificate. That means only the subject's relationship with the public key and nothing beyond that. The PKI does not provide the "Trust" similar to the "Trust in the Physical World" described in the prior section. It merely provides the information that the peer is the peer the program wants to connect.

The Root of Trust and its variation

The PKI system provides a distributed trust architecture. Each piece of the "chain" of trust can be separately issued and verified. The chains of trust are helpful if and only if the participating parties share the list of roots. Thus, how these root CAs operates is extremely important. The operation of the CAs, especially CAs in closer to the root, needs special care. Also, it is essential to understand how the software system trusts and refers to the Root CAs. Each of the major operating systems and some Web Browsers maintains Trust Stores for the software. Each software vendor maintains each of the Trust Store according to its policies. Their policies, including Certificate Authority Policy, are documented [MS][APPLE][CAB] publicly.

Other than relying on Trust Stores, a number of the Root of Trust discovery mechanisms have been proposed and deployed recently. One way is to specify Root of Trust by a well-known URL scheme to retrieve JWKS[JWKS]. The Smart Health Card

specification [SHC] uses this scheme. The Smart Health Card is one of the formats selected for the Digital Agency of Japanese Government [JVC] vaccine certificate. The issuer of the scheme is an URL. There is a standardized step to translate the URL into a JSON file with the help of DNS, HTTP, and TLS protocols.

Governance of the Elements of Trust

The establishment of the Chain of Trust is related to managing two data objects. The first data objects are the Public Key or the Public Key relationship data (a piece or a part of a chain). The second data objects are the data related to the discovery mechanism of the first object.

For a typical X.509 based PKI deployment, CAs handle the first data objects. All data objects are ready when the verification process starts without relying on anything. On building the chain of trust, usually, the software retrieves the subject's certificates alongside the certificates of intermediate CAs. Then, find the root of trust matches with the just retrieved chain of trust in the Trust Store.

For the Smart Health Card (SHC) style JWKS chain of trust deployment, multiple configurations among various parties are necessary: configuration of a domain name, a server that holds the JWKS, and the generation of the public key X.509 chain of trust is necessary. For the retrieval, the verifier synthesizes the URL for the JWKS JSON from the issuer URL value. Then verifier retrieves the file using HTTP with the help of DNS, possibly with TLS. After parsing the JWKS JSON object, the verifier finds either public keys or a set of X.509 certificates. Some small but influential data is involved in resolving the process of the chain of trust. The entity that controls the data manages the data according to a predetermined policy, and governance is in place.

What is important here is whether the subject certificate owner influences these influential data. For the X.509 case, each Trust Store for the software is such data controlled by the software owner; the subject certificate owner, alongside with intermediate CA owner, needs to follow the Trust Store's policy. For the SHC case, the subject certificate owner can fully control the certificate's existence and resolve as expected, as far as dependency on DNS and the server's ownership stays intact.

Conclusion

Verification of the authenticity of Digital Identity is an essential mechanism for providing services in Cyber World. As we discussed, Public Key Cryptography is the key to establishing the chain of trusts. While X.509 PKI-based mechanism with Trust Store in the operating system is the most prevalent way to establish the chain of the trust, the ownership and control of the Trust Store are sometimes unfavorable.

The mechanisms introduced in Smart Health Cards and related standards such as Verifiable Credentials [VC] and Decentralized Identifier allow establishing the chain of trusts. Since these new alternative schemes possibly provide less restriction of establishing a chain of trust, we may apply the scheme with a detailed study on the different, possibly complex governance model.

References

[REF-Attribute] International Organization for Standardization. (2019). IT Security and Privacy — A framework for identity management — Part 1: Terminology and concepts (ISO Standard No. 24760-1:2019). https://www.iso.org/standard/77582.html

[PKI] Internet Engineering Task Force. (2008). Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile (Internet RFC No. 5280). https://www.rfc-editor.org/rfc/rfc5280

[MS] Program Requirements - Microsoft Trusted Root Program. https://docs.microsoft.com/en-us/security/trusted-root/program-requirements

[APPLE] Apple PKI. https://www.apple.com/certificateauthority/

[CAB] CA/Browser Forum. https://cabforum.org/

[JWKS] Internet Engineering Task Force. (2008). JSON Web Key (JWKS) (Internet RFC 7517). https://datatracker.ietf.org/doc/html/rfc7517

[SHC] Smart Health Cards. https://smarthealth.cards

[JVC] Ministry of Health, Labor and Welfare, Japan. (2021). New Coronavirus Infection Vaccination Certificate (Vaccination Certificate). https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/vaccine_certificate.html

[VC] World Wide Web Consortium. (2021). Verifiable Credentials Data Model 1.1. W3C Recommendation. https://www.w3.org/TR/vc-data-model/

[DID] World Wide Web Consortium. (2021). Decentralized Identifier (DIDs) v1.0, Proposed Recommendation. https://www.w3.org/TR/2021/PR-did-core-20210803/