June 18, 2019

Distributed Global People Registry

Anonymous

Distributed Global People Registry

ABSTRACT

This disclosure describes a distributed people registry that allows any individual to register for free. A registered user can store personal data, including unique personal identifiers, such as biometric data. The registry includes a personal vault for each user that is directly accessible only to the account owner. The user information is stored in the vault in a semantic triple structure, making it machine-readable. Third-party applications can reliably authenticate individuals using the vault. Transaction data for each application can be stored separate from the vault. The vault supports anonymized access to user's personal information on an as-needed basis for every transaction. The personal vault for each user is stored at zero cost to the user. The personal vault enables anonymity in online transactions, reduces the amount of personal information provided to applications, and enables users to have self-sovereignty over their data.

KEYWORDS

- Biometric data
- User identity
- Distributed hash table
- People registry
- User authentication
- Personal vault
- User profile
- Data privacy
- Semantic triple

BACKGROUND

Reliable identification of an individual is a fundamental building block in offline and online transactions. Most individuals currently rely on identification issued and certified by a third-party authority, such as a government, financial institution, online service provider, etc. However, such identification suffers from various problems, including data inaccuracies and lack of machine readability.

Online identity of an individual is typically stored in the form of a profile. Most people have multiple profiles online. These profiles often contain information that is inaccurate, outdated, and inconsistent with one another. Additionally, these profiles are stored on servers maintained by disparate companies or organizations that can have vastly differing privacy and security practices. Further, in many instances, personal information is not stored in a structure that allows for meaningful machine-readable communication in and across online systems.

One important security concern is that the servers that store online profiles are housed in data centers that are not inherently secure by system design and architecture. The servers are vulnerable to hacking, misuse, and other types of compromise, and are secured by brute-force measures. The data center infrastructure that supports user profiles is also expensive to maintain.

Another concern with third-party storage and verification of user information is that when a person logs into an application using an existing online profile, the application can obtain access to much more of the user's information than is necessary for the application to provide its services and for a much longer time beyond the application session, unless the profile is explicitly configured by the user to prevent such use. Further, once obtained, such information can be shared with others that the user didn't explicitly grant access to. Validation of a person's identity is not always performed using a reliable process and can leave personally identifiable

information vulnerable to unnecessary exposure.

Also, due to poor identity authentication, it is possible to create online identities based on fictitious persons that do not have a corresponding actual person in the physical world. This leaves the possibility for ill-intentioned actors to use such fictitious persons' profiles to masquerade as real people to perform nefarious and even criminal activities online.

DESCRIPTION

Techniques described herein address the many problems of user identification and access to user information. In particular, a personal vault application is provided for use on any type of device such as a phone, tablet, computer, or other computing device. Each instance of the personal vault application stores a portion of a distributed hash table. The distributed hash table includes data vectors that each uniquely represent a particular person. Together, the data stored by each instance of the personal vault application forms a global people registry.

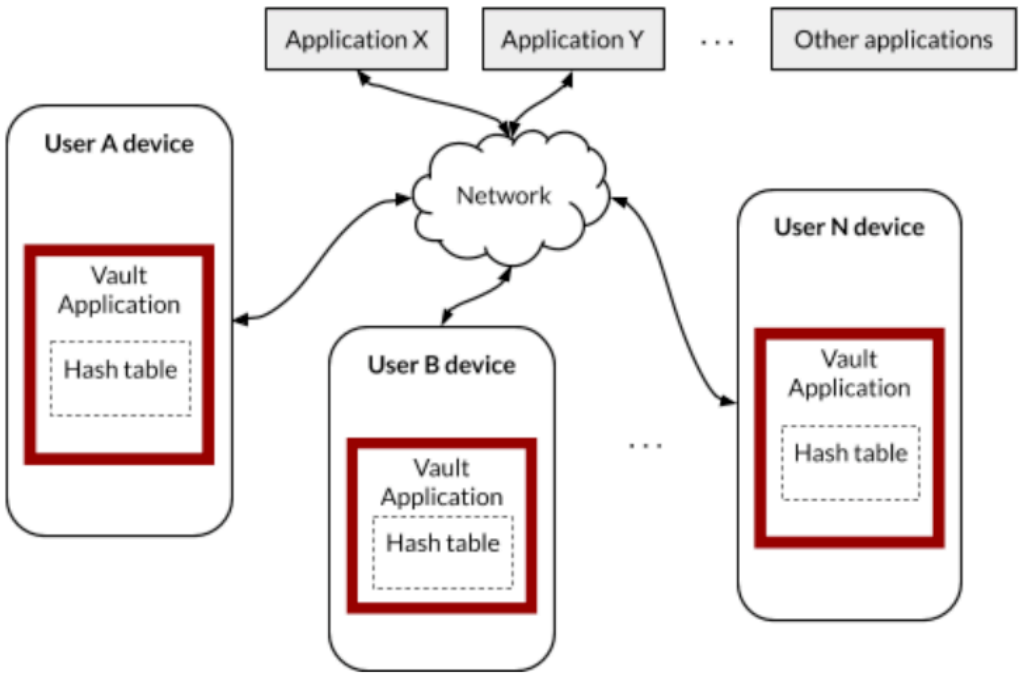

Fig. 1: Global people registry with personal vaults

Fig. 1 illustrates a global people registry with personal vaults. As illustrated, user device of each of N users includes a personal vault application. The personal vault application stores a portion of a distributed hash table. Together, the data stored across the devices forms a global people registry. Applications can access user data from the registry, as permitted by each individual user.

Data vector for identity expression

With permission from participating users, different features of user identity, e.g., based on personally identifiable attributes such as fingerprints, face, voice, three-dimensional shape of body, gait, genome sequence, etc. are obtained and are combined to form a data vector. This data vector uniquely expresses the identity of each person. The framework is extensible and permits any number of attributes to be stored or added to the data vector corresponding to a person. A higher number of attributes used in the vector results in a greater degree of confidence with which an individual is uniquely represented and accurately identified via the stored data vector. Each user is provided with complete control over their personal data vector, e.g., what parameters are stored, which parameters are used for authentication, etc.

Personal vault application

Once the data vector for identity expression of a person is generated, it is stored in a personal vault application. The personal vault application is free for all users to use. Authentication to the personal vault application that stores a user's profile is based on biometric authentication. Personally identifiable data in the vault is hashed such that application developers cannot access it directly. When a pre-authorized application requests data, such data is programmatically unhashed within the pre-authorized app session, as permitted by the user.

The various distributed instances of the personal vault application across multiple user

devices are utilized to provide a global people registry. The registry is implemented as an encrypted, indestructible distributed hash table, with data being stored on each device that has the personal vault application.

The global people registry is implemented such that applications can authenticate people based on the registry, without having access to their identity. It allows an application to authenticate an individual and request access for purposes such as identity verification, person-to-business transactions, storage and access to personal data such as health data, education data, genetics/medical data, legal data, etc.

Storing data as semantic triples

The personal vault application stores data in the form of semantic triples. These are sets of three atomic data entities (subject-predicate-object) in the Resource Description Framework (RDF) data model. This format enables knowledge to be represented in a machine-readable way. Given this precise representation, semantic data can be unambiguously queried and reasoned about. Once a personal vault for a particular user has been created, the vault owner can connect various applications to their vault. The vault owner can provide explicitly approvals of the specific semantic triples from the vault that a particular application can access.

The table below shows examples of semantic triples. Any number of triples can be linked together and enable inference of new facts. The semantic triples are machine-readable.

Subject	Predicate	Object
John Doe	lives in	City A
City A	is in	Michigan
Michigan	is in	United States

Table 1: Example semantic triples

Ontologies

Ontologies can be stored on public blockchains. An ontology is a formalized description of taxonomies and classification networks, defining the structure of knowledge for various domains: the nouns representing classes of objects and the verbs representing relations between the objects. Qualified contributors, who demonstrate their domain expertise through continuous testing and endorsement, collectively maintain ontologies. Edits are peer-reviewed and contributors' and reviewers' identities are masked, for privacy and to eliminate bias. Ontologies can be utilized, e.g., for programmatic interpretation of semantic triples.

Distributed hash table

The personal vault for each user is encrypted and hosted in a distributed hash table spread across all participating computing devices, e.g., devices that have the personal vault application installed. Each device acts as both a client as well as a server. By hosting data in each person's vault entirely in a distributed manner, using other available devices, there is no storage or hosting cost to maintain the data. Further, the data is distributed in a resilient manner, such that the unavailability of a subset of devices has no impact on availability of the vault of any particular individual. Distribution of the data can be performed in such a manner that a larger proportion of an individual's vault is stored on devices owned or controlled by the individual. Each individual has exclusive control over their personal vault such that no one other than the vault owner is able to edit their profile or have agency over their data.

Advantages of global people registry and personal vault

Autonomous Operation

The main resources needed to host a personal vault - computing power, memory,

bandwidth, and energy - are supplied by the devices that support the vault application. Every personal computing device in the world contains these resources. With scale, as more devices are configured with the vault application, the proportion of available excess capacity increases. The vault includes only master data (related to user identity/profile) and no transaction data or content. Transaction data can be stored by each application separately, e.g., on a separate compliant blockchain.

Anonymity

There are relatively limited situations in which an application has a genuine need to access the identity of the person it is serving. The vast majority of applications are able to provide the exact same services when the identity of the people using the application is masked. When an application requests data from the global people registry, such data is provided after removing personally identifiable information of the person.

Zero-knowledge Proof

With the availability of the global registry, individual users do not need to provide any information directly to any application. When an application needs access to an attribute of a person, the application accesses the vault for that person. The person using the application approves the application to have access to only that specific bit of information via the vault. For example, if a person is making a purchase that requires verification that the person is at least 21 years old via the application, the application accesses the vault to obtain a Yes/No answer to the question “is this person 21 years of age or older?” The vault application provides this answer, rather than actual information of the person’s date of birth.

Self-sovereignty

Provision of the vault and the global people registry in this manner allows users self-sovereignty. Control over a user's data is entirely in the user's hands and no third-party such as a government or corporation can control, influence, or manipulate the Global People Registry and the personal vault. Storing data in a distributed manner and providing applications access to the database based on user-specific permissions removes the need for people to depend on government agencies or other parties to provide them with documentation certifying their identity.

CONCLUSION

This disclosure describes a distributed people registry that allows any individual to register for free. A registered user can store personal data, including unique personal identifiers, such as biometric data. The registry includes a personal vault for each user that is directly accessible only to the account owner. The user information is stored in the vault in a semantic triple structure, making it machine-readable. Third-party applications can reliably authenticate individuals using the vault. Transaction data for each application can be stored separate from the vault. The vault supports anonymized access to user's personal information on an as-needed basis for every transaction. The personal vault for each user is stored at zero cost to the user. The personal vault enables anonymity in online transactions, reduces the amount of personal information provided to applications, and enables users to have self-sovereignty over their data.