Block Chain for privacy protection of medical data

Degree programme: BSc in Computer Science | Specialisation: IT-Security
Thesis advisor: Prof. Emmanuel Benoist
Expert: Jean-Marie Leclerc (Sword Group)

In this project, we have analyzed the process of sharing data and the authorization algorithms of two service providers Pryv and MIDATA. Our analysis showed that a user may repudiate the access, thus we implemented a prototype API using the Blockchain database that works as a ledger for every sharing/revocation request, therefore protecting the identity of a user.

Blockchain in sharing of medical-grade data

Over the past few years, we see the natural market evolution, where medical data collected from different sources became important and valuable for medical studies, it helps doctors to be more efficient in predicting and preventing diseases as well as developing new cures and treatments. Collecting and sharing of the medical-grade data imply a high level of privacy protection which is not easy to achieve. In this project, we implement a prototype with the use of a Blockchain database that adds an additional protection layer to the data-sharing process between two medical-data collectors Pryv and MIDATA.

Scope

The following goals are defined. Find a way to protect the privacy of a user (usually a patient), who is sharing the data between Pryv and MIDATA: it should not be possible to find out if the same user has shared data before. Make it impossible for a user to repudiate the access (to deny that he/she has shared the data between partners). Both Pryv and MIDATA have different authorization mechanisms (e.g. OAuth), thus the solution has to be authorization-agnostic. The blockchain database has to be searchable. Implement the feature of looking up for a shared stream (identified dataset) in order to be able to check the validity of a shared access.

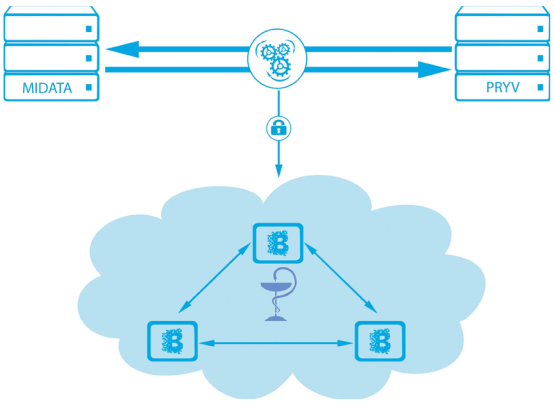

Figure 1 - Data Flows

Implementation

The prototype consists of two nodes with data, representing MIDATA and Pryv (Figure 1), and a network of 3 peer-to-peer Blockchain nodes with an extended API. Data nodes have a simplified authorization API based on the API description of the corresponding partner. Data flows during the authorization process are shown in Figure 1. Figure 2 illustrates the JSON-structured payload formed by MIDATA and Pryv during the process of data sharing (revocation process is similar). The payload will be embedded into a new transaction and after a while, the transaction will be validated by the Blockchain nodes. If the transaction is validated the data stream may be accessed by the requesting partner.

A transaction may be retrieved using a unique transaction ID, or it may be searched by the stream ID within the payload. In the latter case, the latest transaction with the corresponding stream ID will be returned.

Using the implemented prototype we showcase the idea of using the Blockchain for protecting user privacy when their data is shared between Pryv and MIDATA.

```
"payload": {
    "process": "<share|revoke>",
    "requesting": {
        "user_id": "<Protected Requesting User ID>",
        "app_id": "<Requesting Application ID>"
    },
    "sharing": {
        "user_id": "<Protected Sharing User ID>",
        "app_id": "<Sharing Application ID>"
    },
    "request": [
        {
            "stream_id": "<ID of a data Stream (defined by sharing app)>",
            "level": "read"
        },
        {
            "stream_id": "<ID of a data Stream (defined by sharing app)>",
            "level": "read"
        }...
    ]
}
```

Figure 2 - Structure of a Blockchain Transaction

Sergii Bilousov
sergii.bilousov@gmail.com