Papers on Blockchain

Ipek Uyguner

February 23, 2021

1 Owner-Controlled Information [2 hours work]

1.1 Hypotheses

- The centralised-user controlled systems has some drawbacks, most importantly, from security point of view.
- The centralised-user controlled systems cannot be straightforwardly adapted to our current world.
- The current distributed data management system is more reliable than centralised systems.

1.2 Abstract

Information about individuals is currently maintained in many thousands of databases, with much of that information, such as name and address, replicated across multiple databases. However, this proliferation of personal information raises issues of privacy for the individual, as well as maintenance issues in terms of the accuracy of the information. Ideally, each individual would own, maintain and control his personal information, allowing access to those who needed at the time it was needed. Organizations would contact the individual directly to obtain information, therefore being assured of using current and correct information. While research has been performed on users owning and controlling access to their personal information in an electronic commerce environment, we argue that this concept should be extended to all user information including, for example, medical and financial information. The end goal is not for users to simply maintain copies of this information, but to be the source of this information. This paper presents the concept of users owning their personal information and introduces some of the issues involved in users being able to control access to this information. The security requirements, including authentication, access control and audit, as well as user interfaces and trust, for this new paradigm are given particular emphasis.

1.3 Summary

With the advent of the technological developments, the information of the people are mostly stored in digital databases. The online bankings, websites, hospital recordings are such examples that are used widely over the world. Currently, the information system of users are distributed and are controlled by the organisations itself and there is no co-operation between the various entities to share user information. The advantages of these systems include that it is already well adapted and used for many years. The mistakes occur in one system cannot affect to other systems easily since it is based on local frameworks. On the other hand, this kind of systems introduces some disadvantages in terms of privacy, maintenance, availability etc.

The new paradigm is introduced which is shifting to the centralised control of the user information. In this centred method, the user could more easily reach the system, Thus, one pros of it users can manage their information over accessing their data or can update and change. It also ensures that information is accurate and up-to-date and reachable more easily.

However, this gives the user the responsibility to manage their data them self. Moreover, since there is one system, the attacks become more dangerous. Since once the user hacked, the all information could leak. In the paper, there are some state-of-art projects introduced in which the paper is published. This new system also brings some security concerns. Authentication and access control is crucial to ensure the identity of external users who have access, to what extend, to an individual's information. Audit trail is important in the case of a dispute. Moreover, audit tracks are also necessary for backup and recovery issues.

The user interface should be designed according to the different levels of user, such as different levels of expertise in digital world, computer literacy and cognitive ability.

1.4 Open Questions

- How to ensure security for sure?
- Who should be the owner of data of the user?
- How to ensure that the system will work smoothly?
- How to transform the whole decentralised system to the centralised system, what are the transmission processes?

1.5 Papers

- D. Chaum. Security without identification: Transaction systems to make big brother obsolete.

- F. Cranor, J. Reagle, and M. S. Ackerman. Beyond concern: Understanding net users' attitudes about online privacy. Technical Report Tl:t 99.4.3, ATT Labs, Apr. 1999.
- R. S. Sandhu and P. Samarati. Access control: principles and practice. IEEE Communications, 32(9):40 - 48, 1994.

1.6 Key Achievements

It introduces the main idea behind the centralised data management system with its drawbacks and advantages by giving some examples as follows:

- - Microsoft Passport (Microsoft Corporation. Microsoft .NET passport: One easy way to sign in online. http://ww , passport.net/. Last visited: 17 July 2003.)
- - Hailstorm (Microsoft Corporation. Microsoft announces "Hailstorm," a new set of xml web services designed to give users greater control. http ://www .microsoft. com/presspass/f eatures/ 2001/marO1/O3-19hailstorm. aT.sp, 2001. Last visited: 17 July 2003.)
- - Liberty Alliance (Liberty Alliance. Liberty Alliance project. http :// nr . proj ecZliberZy, org/, 2003. Last visited: 17 July 2003.)

2 An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends [3 hours work]

2.1 Hypotheses - Research Questions

- What is Blockchain and how it works?
- What are the obstacles in the Blockchain recently and in future?
- What are the future directions‘in Blockchain?

2.2 Abstract

Blockchain, the foundation of Bitcoin, has received extensive attentions recently. Blockchain serves as an immutable ledger which allows transactions take place in a decentralized manner. Blockchain-based applications are springing up, covering numerous fields including financial services, reputation system and Internet of Things (IoT), and so on. However, there are still many challenges of blockchain technology such as scalability and security problems waiting to be overcome. This paper presents a comprehensive overview on blockchain technology. We provide an overview of blockchain architechture firstly and compare some typical consensus algorithms used in different blockchains. Furthermore, technical challenges and recent advances are briefly listed. We also lay out possible future trends for blockchain.

2.3 Summary

The Blockchain Technology is immutable records of data that is managed without ruling by any central entity. It is mostly used in financial services but the application in other fields, such as public services or Internet of Things, are increasing. It consists of a sequence of blocks, that contains the list of transaction records by using digital signatures. It's main characteristics are decentralisation, persistency, anonymity, auditability. There are three categories, in terms of taxonomy, namely: public, private and consortium blockchain. In public blockchain, all records are visible to the public and everyone could take part in the consensus process whereas only a group of nodes can participate it in consortium blockchain. In private blockchain, only some nodes from specific organization can join the consensus process. The main consensus algorithms are Proof of work(PoW), Proof of Stake(PoS), Practical byzantine fault tolerance(PBFT), Ripple and Tendermint. As an example, in PoW, the miners compete against each other in solving complex computational problem. The winners can be the one who records the transaction. These computations consume too much energy and are time consuming. And in PoS, miners prove the ownership of the amount of currency. One big challenge in the Blockchain technologies are scalabilty. It can only process nearly 7 transactions per second which cannot fulfill the requirement for real time demand. One solution could be storage optimisation by deleting old transactions. The other problem is privacy leakage. Some research show that the transactions can be linked to reveal user's information. The selfish mining, which is strategy for mining bitcoin where groups of miners collude to increase their revenue, is also another problem. This can cause the centralised system again. In the future, the blockchain testing mechanism could be one direction to work on it. There are already over 700 cryptocurrencies created and the mechanism to check if these systems fulfill their requirements could be necessary. Another point to work on in the future is to stop the tendency to centralization. Although the idea behind the blockchain is decentralisation of the system, there is already tendency toward to centralisation in the mining pool. There might some research on this area. Blockchain is also good field to work with big data management, for example, it can be used to store specific data because it is distributed and secure to some extend.

2.4 Open Questions

- How to do blockchain testing?
- What kind of works can be done for preventing tendency toward to centralisation?

2.5 Papers

- S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008. [Online]. Available: https://bitcoin.org/bitcoin.pdf
- G. Foroglou and A.-L. Tsilidou, "Further applications of the blockchain," 2015.
- V. Buterin, "A next-generation smart contract and decentralized application platform," white paper, 2014

2.6 Key Achivements:

This paper introduces the Blockchain technologies with current trends. Few examples are below.

- - J. Kwon, "Tendermint: Consensus without mining," URL http://tendermint.com/docs/tendermint v04. pdf, 2014
- - S. King, "Primecoin: Cryptocurrency with prime number proof-of- work," July 7th, 2013.
- - G. Wood, "Ethereum: A secure decentralised generalised transaction ledger," Ethereum Project Yellow Paper, 2014

- - V. Zamfir, “Introducing casper the friendly ghost,” Ethereum Blog URL: https://blog. ethereum. org/2015/08/01/introducing-casper- friendly-ghost, 2015

3 Decarbonizing Bitcoin: Law and policy choices for reducing the energy consumption of Blockchain technologies and digital currencies [3.5 hours work]

3.1 Hypothesis

- The Blockchain technologies threaten our future because it is environmentally not viable with the current trends.
- The fiscal tools such as taxation on Blockchain can be an option to convince people to work on environment friendly solutions in Blockchain.
- Recently, there is no big effort to change the current energy intensive trends in Blockchain. Moreover, it will be harder to change current trends once they are used more widely.

3.2 Abstract

The vast transactional, trust and security advantages of Bitcoin are dwarfed by the intentionally resource-in- tensive design in its transaction verification process which now threatens the climate we depend upon for survival. Indeed Bitcoin mining and transactions are an application of Blockchain technology employing an inefficient use of scarce energy resources for a financial activity at a point in human development where world governments are scrambling to reduce energy consumption through their Paris Agreement climate change commitments and beyond to mitigate future climate change implications. Without encouraging more sustainable development of the potential applications of Blockchain technologies which can have significant social and economic benefits, their resource-intensive design combined now pose a serious threat to the global commitment to mitigate greenhouse gas emissions. The article examines government intervention choices to desocialise negative environmental externalities caused by high-energy consuming Blockchain technology designs. The research question explores how to promote the environmentally sustainable development of applications of Blockchain without damaging this valuable sector. It studies existing regulatory and fiscal policy approaches towards digital currencies in order to provide a basis for further legal and policy tools targeted at mitigating energy consumption of Blockchain technologies. The article concludes by identifying appropriate fiscal policy options for this purpose, as well as further considerations on the potential for Blockchain technology in climate change mitigation.

3.3 Summary

The Blockchain Technologies has already huge impact on a range of sectors. However, beside all the multi-sectoral benefits it brings, there are also environmental concerns of it. Because the current trends on Blockchain in transaction verification, which is a big part of the system, is intentionally resource-intensive design. There are some works on how to encourage a change to less energy intensive Blockchain technology. One option is government intervention by taxes. However, to evaluate tax on Blockchain trends, one should first clarify to how digital currencies are legally classified. Whether a digital currency can be considered as money is an open question worldwide. The Bank for International Settlements do no provide definition of "money" or "currency". Some believe that it can be treated as money because it can be exchanged for goods, services etc. However, another view is since there are lack of government control and liability from central bank, it cannot be actual currency. So far the regulators' focus for Blockchain technology were on financial technologies, such as preventing money laundering, whereas there is no record of charging on digital currencies based on their carbon footprint. It is also necessary to differentiate between energy intensive and less intensive versions of the technologies in order to regulate fiscal tools on them. It could be based on their electricity consumption. For example, "Proof of Stake," would be a less energy intensive process than "Proof of Work". The other issue needs to be worked on is that the determination of which person or entity to target for policymakers. One option could be the developers of the technology. But the taxation on developers might be impractical but the alternative fiscal tools such as governmental grants, academic research funds could be nice alternatives. The other option is about that energy intensive machine requiring significant computational power. So, the taxes on mining businesses could be applied or manufacturers can be targeted. The next one is the users of Blockchain technology. For instance, people doing an electronic transaction would be discouraged to utilise a certain technology, in this case high energy intensive technologies, if a tax were payable.

3.4 Open Questions

- The limits to the universal solution
- The most efficient technological design options

3.5 Paper

- Bitcoin Energy Consumption Index. https://digiconomist.net/bitcoin-energy- consumption. (Accessed 23 May 2018)
- Is Bitcoin Worth the Energy? Financial Times' The Week, 2017 11 Dec http://theweek.com/articles/742253/bitcoin-worth-energy (Accessed 24 May 2018)

- T. Brosens, Why Bitcoin Transactions are More Expensive Than You Think, (2017) (Accessed 23 May 2018), https://think.ing.com/opinions/why-bitcoin- transactions-are-more-expensive-than-you-think/.

3.6 Key Achievements

It introduces the possible solutions to how to promote the environmentally sustainable development of Blockchain Technologies for policymakers.

4 Provenance-Aware Storage Systems [3 hours work]

4.1 Hypothesis

- The adaptation of provenance-aware storage system (PASS) is not straightforward.
- The current provided prototypes has many advantages but it also has some missing functionalities. It needs more work to do.
- The provided PASS brings new overhead that can be still negligible.

4.2 Abstract

A Provenance-Aware Storage System (PASS) is a storage system that automatically collects and maintains provenance or lineage, the complete history or ancestry of an item. We discuss the advantages of treating provenance as meta-data collected and maintained by the storage system, rather than as manual annotations stored in a separately administered database. We describe a PASS implementation, discussing the challenges it presents, performance cost it incurs, and the new functionality it enables. We show that with reasonable overhead, we can provide useful functionality not available in today's file systems or provenance management systems.

4.3 Summary

The Digital Provenance is the history of the object. It includes different information regarding of how the object was derived, changed. We can split the current provenance systems into below groups in terms of their architectures. The first, File, File System and Database Approaches, puts the provenance in the corresponding data such as header of the data or a file system that automatically tracks provenance at the file system level. The second, the Service-oriented Architectures, is mostly used in the computational sciences and designed for grid environments. The third, scripting architectures; is for source code control and build systems. The last one is the environment architectures where users track provenance in a unified environment. They are typically stored in separate database systems. But, this brings some problems such that lack of consistency between the provenance and the data, enforcing provenance maintenance, and preserving provenance during backup, maintenance. Thus, A provenance-aware storage system (PASS) is introduced as a storage system that automatically collects, stores, manages, and provides search for provenance. The PASS has some advantages such that the detection of system changes, intrusion detection, build debugging, understanding system dependencies are easier. The PASS is the collection of the provenance for all the objects stored in it. The PASS should fulfill some requirements as follows. It should support application-generated provenance. It should provide security for provenance and support for queries on provenance. The duplicate elimination, cycles and versions should be carefully handled. For instance, the cycles are problematic, because it means violations of causality such as an object depending on the existence of its children. So, when adding a provenance record, the system checks the ancestry graph for cycles. If the new record would create a cycle, the cycle-breaking algorithm is invoked. The Berkeley DB embedded database library is used for storage in this prototype. The query system for provenances is also built an easy-to-use query tool in the databases. As an query example, DUMP ALL, reveals the all provenance for a selected file. However, this current prototype do not meet all the requirement that is mentioned above yet. It doesn't create provenance for files that are transmitted over the net. Moreover, the security part is also not implemented yet. Some future researches on this work could be some provenance pruning because the current version is append-only and it is not viable in the long run. Cycle-free provenance collection is another purpose for the future improvements.

4.4 Papers

- MUNISWAMY-REDDY, K.-K. Deciding How to Store Provenance. Technical Report TR-03-06, Harvard University, Jan. 2006
- YOGESH SIMMHAN, BETH PLALE, D. G. A Survey of Data Provenance Techniques. Technical Report IUB-CS-TR618, Indiana University, Bloomington, 2005
- GROTH, P., MOREAU, L., AND LUCK, M. Formalising a protocol for recording provenance in grids. In Proceedings of the UK OST e-Science Third All Hands Meeting (Nottingham, UK, Sept. 2004)

4.5 Open Questions

- Which approaches will be followed for building provenance-aware applications?
- Which approaches will be followed for Cycle-free provenance?

4.6 Key Achievements

- The introduction of the prototype for PASS with evaluations.

5 A privacy threat analysis framework: supporting the elicitation and fulfillment of privacy requirements [4.5 hours]

5.1 Abstract

Ready or not, the digitalization of information has come, and privacy is standing out there, possibly at stake. Although digital privacy is an identified priority in our society, few systematic, effective methodologies exist that deal with privacy threats thoroughly. This paper pre- sents a comprehensive framework to model privacy threats in software-based systems. First, this work provides a systematic methodology to model privacy-specific threats. Analogous to STRIDE, an information flow–oriented model of the system is leveraged to guide the analysis and to provide broad coverage. The methodology instructs the analyst on what issues should be investigated, and where in the model those issues could emerge. This is achieved by (i) defining a list of privacy threat types and (ii) providing the mappings between threat types and the elements in the system model. Second, this work provides an extensive catalog of privacy-specific threat tree patterns that can be used to detail the threat analysis outlined above. Finally, this work provides the means to map the existing privacy- enhancing technologies (PETs) to the identified privacy threats. Therefore, the selection of sound privacy counter- measures is simplified.

5.2 Hypothesis

- The security threats are focused on more than privacy threats in terms of methodological approaches to threat modeling.
- The STRIDE model lacks the privacy threats support.
- This paper presents a systematic model for privacy threats inspired by security threats models.

5.3 Summary

The security and privacy analysis is studying of what can possibly go wrong in a system in terms of security and privacy. Threat modeling is one popular approach to this end. For example, Microsoft's STRIDE is an methodology to revealing threat scenarios. The STRIDE threat modeling process involves nine steps as follows shortly:

- Define use scenarios

- Gather a list of external dependencies
- Define security assumptions
- Create external security notes
- Create one or more data flow diagram of the application being analyzed
- Determine threat types
- Identify the threats to the system
- Determine risk
- Plan mitigation

However, these steps are more security related and lacks privacy perspective. In this paper, the LINDDUN methodology is introduced for covering privacy threat modelling. In order to grasp the LINDDUN better, the privacy properties are studied first. Some example from these properties are examples below:

- **Unlinkability**: Hiding the link between pieces of information.
- **Anonymity**: Hiding the link between an identity and an action or a piece of information.
- **Pseudonymity**: An identification of a subject other than user's real name.
- **Plausible deniability**: To be able to deny having an action that others can neither confirm nor contradict.
- **Undetectability and unobservability**: Hiding the user's activities.
- **Confidentiality**: Hiding the data content or controlled release of data content.
- **Content awareness**: The user's consciousness regarding his own data.

The proposed framework considers 7 types of threats as described below:

- **Linkability:** Attacker can distinguish whether th subjects are related or not in the system.
- **Identifiability:** Attacker can identify the subject associated to an items of interest.
- **Detectability:** Attacker can distinguish whether such an item exists or not.
- **Information Disclosure:** Expose personal information to individuals who do not have access to it.

- **Content Unawareness:** Related to the user is unaware of the information disclosed to the system.
- **Policy and consent Non-compliance:** Although the system reveals its privacy policies, there is no guarantee that the system actually complies with the advertised policies.

In this framework, data flow diagrams(DFD) are used for the application's use case scenarios. The privacy threats are affecting different parts of the systems. The main threads and the part of the system which are affected from those threads are shown in the figure and some examples for the preconditions for these threads are exampled below.

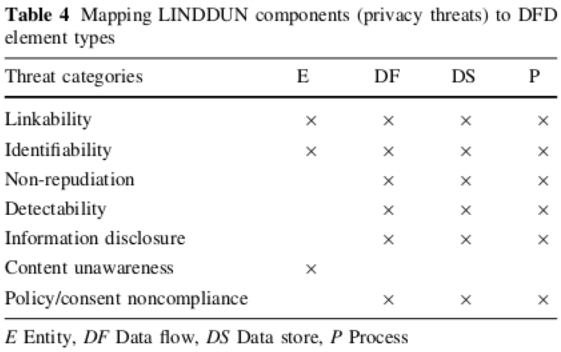

Table 4 Mapping LINDDUN components (privacy threats) to DFD element types

Threat categories	E	DF	DS	P
Linkability	×	×	×	×
Identifiability	×	×	×	×
Non-repudiation		×	×	×
Detectability		×	×	×
Information disclosure		×	×	×
Content unawareness	×			
Policy/consent noncompliance		×	×	×

E Entity, *DF* Data flow, *DS* Data store, *P* Process

- **Linkability of entity:** Attackers can distinguish whether some entities are related or not in the system. The preconditions for this might be data flow or data store not fully protected or personal identifiable information can be linked based on some personal information.

- **Linkability of data flow:** The preconditions for this are that data flows are not fully protected and insufficient anonymous communication are linked.

- **Linkability of data store:** The preconditions for this are insufficient access control of the data store resulting to the information disclosure threat and weak data anonymization.

- **Identifiability of entity:** The main preconditions here are user's actual identity is used, secret (e.g. a password) is used as log-in,use of weak token as login or biometrics is used as log-in.

- **Identifiability of data flow:** The preconditions are similar with linkability of data flow as above.

- **Identifiability of data store:** The preconditions are similar with linkability of data store as above.

- **Non-repudiation of data flow:** One precondition is insufficient obfuscation for data sources or data flows. Another is weak deniable encryption technique is used to protect data flow. Or, weak message authentication codes are used to ensure integrity of data flow content.

- **Detectability of data flow:** The preconditions here are lack of covert channel, side channel attacks, weak information hiding techniques, insufficient dummy traffic etc.

The documentations of the threat scenarios misuses cases are used. In a misuse case, the summary of the threat, primary misactor, the basic flow of threat, the initiation of the misuse case and the system preconditions should be defined.

Afterwards, the risk analysis should be done which reveals the prioritisation of the threats. The next step is,the extraction of the system's requirements from the misuse cases. The solutions for the privacy requirements could be warning the users, removing those features or countering threats with privacy-enhancing technology.

5.4 Key Achievements

- KAOS, a goal-oriented requirements analysis framework, Lamsweerde AV, Brohez S, Landtsheer RD, Janssens D, Infor- matique DD (2003) From system goals to intruder anti-goals: attack generation and resolution for security requirements engi- neering. In: Proceedings of the RE03 workshop on requirements for high assurance systems (RHAS03), pp 49–56
- Distinguishing between hard and soft privacy. Danezis G (2007) Talk: introduction to privacy technology. http://research.microsoft.com/en-us/um/people/gdane/talks/Privacy_Technology_cosic.pdf

5.5 Papers

- Howard M, Lipner S (2006) The security development lifecycle. Microsoft Press, Redmond, WA
- Lamsweerde AV, Brohez S, Landtsheer RD, Janssens D, Infor- matique DD (2003) From system goals to intruder anti-goals: attack generation and resolution for security requirements engi- neering. In: Proceedings of the RE03 workshop on requirements for high assurance systems (RHAS03), pp 49–56
- Mcgraw G (2006) Software security: building security. Addison- Wesley Professional, Boston, NY (book)