Journal of Engineering and Applied Sciences 13 (3): 691-695, 2018
ISSN: 1816-949X

A Study on Privacy Protection Method Applying Blockchain Technology for Big Data Utilization (Focus on the Blockchain Technology for Monero, Dash and Z-cash to Ensure Anonymity)

[2]Jang Mook Kang, [1]Baehyun Kim and [2]Sung-jun Kim
[1]K-ICT Big Data Center, National Information Society Agency, Seoul, Korea
[2]Department of Big-Data Industrial Security, Namseoul University, 31020 Cheonan City, Korea

Abstract: There are currently about 700 kinds of cryptocurrency. In addition, more than 3,000 cryptocurrencies are being prepared for development or use. Cryptocurrency consists of a blockchain technique. The blockchain technology ensures transparency and objectivity. In other words, it can handle transparent distributed book transactions using cryptocurrency. Already in the world, cryptocurrency is utilized to achieve digital social innovation. We analyzed the possibility of cryptocurrency and blockchain technology such as dash, Z-cash and monero. These new technologies can be applied in a variety of ways depending on the propriety of the personal information they want to deal with business. This study is expected to help experts who develop or utilize the blockchain policy.

Key words: Privacy protection, monero, dash, Z-cash, digital social innovation, cryptocurrency, blockchain, blockchain policy

INTRODUCTION

Big data is a key technology in the hyper-connected society. The discussion of big data technology has been steadily studied over the years.

On the other hand, there is insufficient solution for privacy concern due to use of big data. The World Economic Forum (WEF) which took place from January 25-29, 2012, discussed the balanced development of big data to solve the above research problems.

In particular, the big data ecosystem has begun discussions on "big data, big impact: new opportunities for international development" which was announced in a study related to big data ecosystem (Sasson *et al*., 2014).

New technologies and services are emerging such as Internet, artificial intelligence, wearable service, social network service, cloud infrastructure construction, pin tech and blockchain technology. These various technologies have in common. Service is possible based on big data. Therefore, the data-driven technology policy is in a new phase. Figure 1 shows the data ecosystem.

The data are largely divided into individual, public/development and private sectors. Even with the same data, the level of privacy protection differs for each area. For example, individuals base their opt-out on data processing. On the other hand in the private domain, the privacy policy differs depending on the business model and the ownership relationship of sensitive data. Figure 1 shows the privacy protection required for such data mining and analysis. This study is based on the recognition of the above problems. Big data must be utilized. At the same time, privacy must be protected. To do this, we applied privacy protection to use blockchain technology.

MATERIALS AND METHODS

Big data and concern of privacy

Big data for data-driven society: What is a data driven society? The decision of society members and leaders should be based on data. In the meantime, data from experts is often optimized for limited laboratories. Therefore, it is very difficult to find the data in reality and solve complex social problems based on it.

However, with the advent of big data recently, data-driven decision making has become possible in actual reality.

The data has changed like the oil of the industrial age. The value of data is changing. Although, data has long been valuable, it was either seen as ancillary to the core operations of running a business or limited to relatively narrow categories such as intellectual property or personal information.

Corresponding Author: Sung-Jun Kim, Department of Big-Data Industrial Security, Namseoul University, #13149 Design Information Bldg., 91 Deahakro, Seongjwan-eup, Seobuk-gu, Cheonan City, 31020 Choongnam, Korea

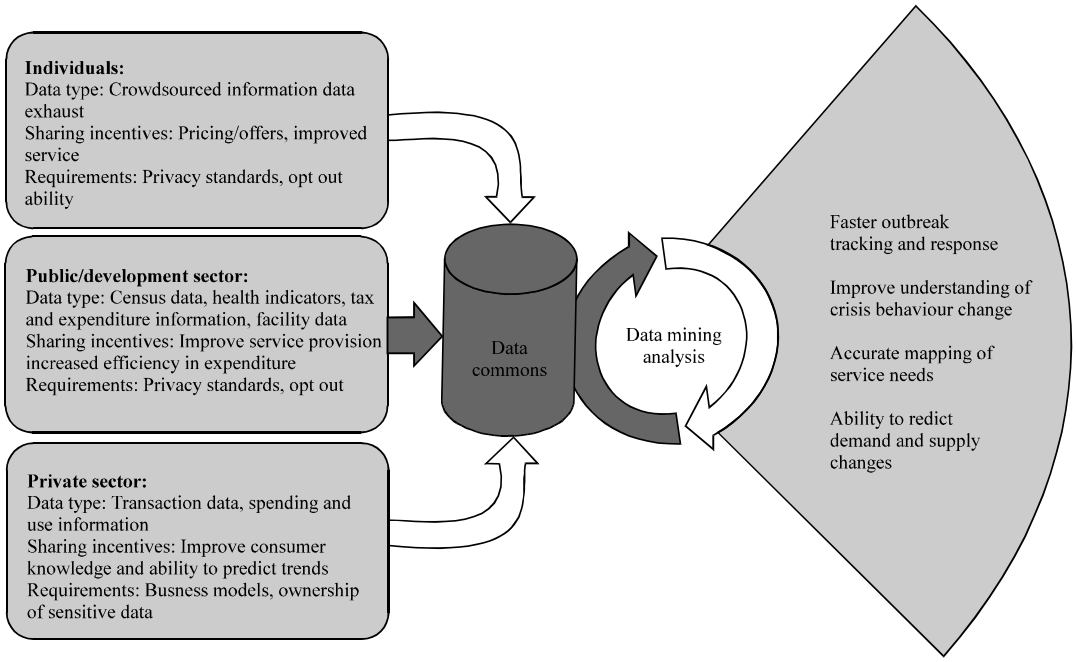

Fig. 1: The dynamics of data ecocyctem (Anonymous, 2015a, b)

In contrast in the age of big data, all data will be regarded as valuable in and of itself. Data is seen as a "non-rivalrous" good because it can be used many times for the same purpose, harnessed for multiple purposes and one person's use of it does not impede researcher's (Mayer-Schonberger and Cukier, 2013).

In order to use the data freely, it is necessary to reduce the concern about privacy. In essence, it seems difficult to eliminate the threat to privacy. However, it should prevent it and reduce the damage. Vulnerabilities in privacy are often the result of business activity or commerce behavior.

Much of the technology required for big-data computing is developing at a satisfactory rate due to market forces and technological evolution (Anonymous, 2014). The following section looks at ways to reduce privacy concerns while adjusting market and technology development.

A concern for privacy using big-data: Big data is dominated by unstructured data. Unstructured data refers to consumer behavior data, emotional data of voter and so on. In the past, this data was not collected. In the big data era, however, data such as emotion, movement and body temperature are important. The above-mentioned unstructured data threatens the privacy of individuals. From a data security perspective, there are some important challenges with the protection of big data-most distributed systems have only a single level of protection which isn't ideal,' says Sijbrandij. Non-relational databases (NoSQL) are actively evolving, making it difficult for security solutions to keep up with the demand (Rossi, 2016).

In other words, unstructured data increases security vulnerability. Therefore, the concept of traditional privacy must change. Of course, how you protect your privacy must also change. Let us consider the definition of initial privacy for this purpose.

In a legal context, privacy has been considered to be largely synonymous with a 'right to be let alone' sources (Buchanan *et al*., 2007). Privacy which was initially subjective, changed with the development of information and communication technology. Recent privacy includes the meaning of personal information that an individual is identified. To be able to identify an individual, it was possible with a unique personal identification code. For example, the internet currently used by researchers has a unique address. At this address, the behavior of the author can be identified. However, with the emergence of a hyperlinked society, existing personal information also changes.

For much of that information relates to not just things but to people. Information about us is accessed, stored, manipulated, data mined, shared, bought and sold, analyzed and potentially lost, stolen or misused by countless government, corporate, public and private agencies, often without our knowledge or consent (Buchanan *et al*., 2007). Therefore, the ecosystem must be designed to protect privacy while utilizing big data.

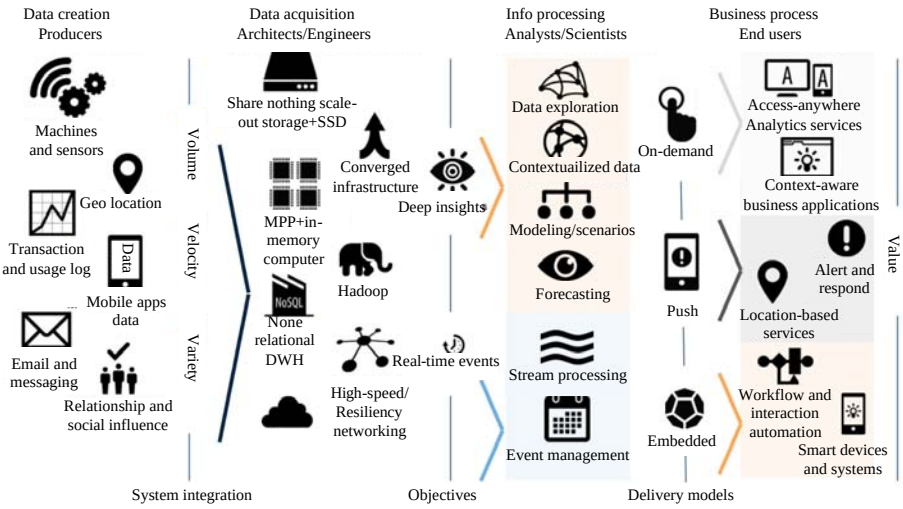

Fig. 2: Echo system for processing big data (Anonymous, 2017)

A big-data ecosystem: Infrastructural technologies are the core of the big data ecosystem. However, the volume, velocity and variety of data mean that relational databases often cannot deliver the performance and latency required to handle large, complex data. The rise of unstructured data in particular meant that data capture had to move beyond merely rows and tables (McNulty, 2014).

Existing formal data can be analyzed in a relational database. However, unstructured data is often raw-data. In other words, raw-data is complex such as data entry, preprocessing, standardization, storage, analysis, post-processing and sharing. Therefore, an echo system for processing big data is required at each step.

Figure 2 shows an ecosystem for activating big data. In each of these processes, privacy protection must be considered to maximize the effect.

RESULTS AND DISCUSSION

Blockchain technology for privacy protection: Blockchain technology supports distributed environments. If you include sensitive information among your personal information, it is safe to distribute it. However, in this case, it is necessary to confirm transparent information about who saw it who deleted it immediately and so on. For this, we can explain the protection example based on blockchain as shown in Fig. 3. Figure 3 shows a computational process similar to the technology of homogeneous or threshold ciphers. The top left block (e-Vote procedure) is the unsecure code where the arguments marked in (*) are private and stored as shares on the DHT. The network selects a subset of nodes at random to compute a secure version of e-Vote and

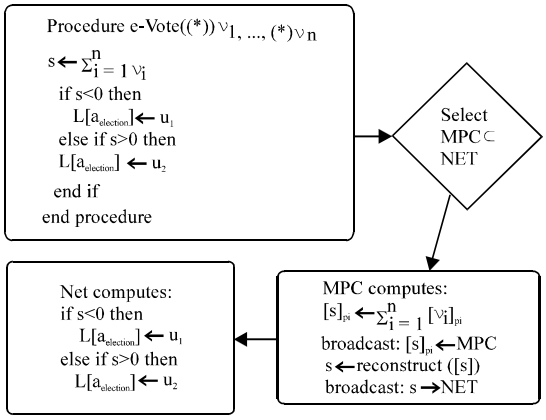

Fig. 3: Example of a flow of secure computation in a blockchain network (Zyskind and Nathan, 2015)

broadcasts the results back to the entire network, that stores it on the ledger (Zyskind and Nathan, 2015).

Figure 3 shows privacy protection in a distributed environment that considers the way Bitcoin operate with basic cryptocurrency. In the following paragraphs, we analyze Z-cash, dash, etc. which are cryptocurrency with anonymity.

Monero, dash and Z-cash to ensure anonymity: When raw data and unstructured data are processed as virtuous processors for Big-data, personal information becomes vulnerable. Companies, individuals and countries that use the big-data want to see more personal information as they use it. Until now, privacy has been protected by a

trusted third party. However, the cryptocurrency introduced below can effectively protect privacy in a distributed environment.

Firstly, monero uses a ring signature system to provide anonymous transactions. In a ring signature setting, a group of users have a set of keys that can confirm a transaction without revealing which user made it (Anonymous, 2015a, b). Users can set the how many mixins they want for their transaction which makes it easier or harder to trace. The problem withthe setting is that the majority of transactions occur with only one mixin, meaning that mixed transactions can still be traced (Warren and Brandeis, 1890).

Secondly, Darkcoin was born later to be rebranded as DASH (Digital Cash). Darkcoin was created a privacy centric cryptocurrency that allows users to pre-mix their coins by combining identical inputes from multiple senders into a single transaction that has serveral outputs. This system was known as Darksend and was recently rebranded to PrivateSend, it is currently a built-in feature in the DASH wallet. This is a step-up from third party coin-mixing services but it still allows us to determine the receivers of the transactions which leaves room for expert analysis.

Thirdly, Z-cash is a cryptocurrency which provides a mechanism to obscure the source, destination and amounts of transactions. These transactions are known as shielded transactions. The privacy of shielded transactions is achieved by using zk-SNARKs which are a type of zero-knowledge pro by (Sasson *et al*., 2014). Transactions can be "transparent" and similar to bitcoin transactions in which case they are controlled by a t-addr or can be a type of zero-knowledge proof called zk-SNARKs (Quesnelle, 2017).

These three cryptocurrencies are expected to be used safely in the future for big data. With this reference, the block chain technique can be applied to privacy protection.

Especially, it is effective for companies to utilize public cryptography for detailed technology. Big data can be utilized while protecting private information in the private sector.

CONCLUSION

Big data utilization and privacy are trade-offs. Therefore, the more you use big data, the more vulnerable or compromised your personal information.

Policy authorities should adopt advanced technology that activates for big data using while protecting personal information. Actual industrial sites utilize big data and these days are limited by the Personal Information Protection Act. In other words, it slows industrial development.

This study explores the protection of personal information with advanced technology of block chain technology. Blockchain technology deals with personal information in a distributed environment.

It is essentially different from the way traditional government or third-party institutions process personal information. What is the most effective way for company and person, individuals and governments, person and person to deal with sensitive personal information in a distributed environment?

In this study, we analyzed the possibility of cryptocurrency and block chain technology such as dash, Z-cash and monero. These new technologies can be applied in a variety of ways depending on the propriety of the personal information they want to deal with business. Therefore, the blockchain technique can replace the de-identification technique or the anonymization technique that has been studied so far. This study helps to analyze the possibility for blockchain and grasp the current situation for cryptocurrency.

ACKNOWLEDGEMENTS

This research is supported by Barun ICT Research Center (Research topics: Diffusion and Sharing of Good Experience: Model Development of the best case and Construction Method Development of Archive of participation, Study period: 2016-2017).

REFERENCES

Anonymous, 2014. IDC's big data ecosystem. International Data Corporation, Framingham, Massachusetts, USA. http://3.bp.blogspot.com/-kbSxczFbqfQ/VMZjXMln8XI/AAAAAAAAAEs/cgsocye-Tt0/s1600/IDC.png.

Anonymous, 2015a. A lesson in anonymity: Bitcoin, Dash, Monero and Zcash. Steemit, https://steemit.com/money/@thecryptodrive/a-lesson-in-anonymity-bitcoin-dash-monero-and-zcash.

Anonymous, 2015b. Monero services. Monero, Wilmington, Delaware. http://monero.org/.

Anonymous, 2017. Big data, big impact: New possibilities for international development. World Economic Forum, Cologny, Switzerland. http://www3.weforum.org/docs/WEF_TC_MFS_BigDataBigImpact_Briefing_2012.pdf.

Buchanan, T., C. Paine, A.N. Joinson and U.D. Reips, 2007. Development of measures of online privacy concern and protection for use on the internet. J. Assoc. Inf. Sci. Technol., 58: 157-165.

Mayer-Schonberger, V. and K. Cukier, 2013. Big Data: A Revolution that will Transform How We Live, Work and Think. Houghton Mifflin Harcourt, Boston, Massachusetts, USA., ISBN:978-0-544-00269-2, Pages: 245.

McNulty, E., 2014. Data science understanding big data understanding big data: The ecosystem. Dataconomy Media, Berlin, Germany. http://dataconomy.com/2014/06/understanding-big-data-ecosystem/.

Quesnelle, J., 2017. On the linkability of Zcash transactions. Cryptography Secur., 1: 1-5.

Rossi, B., 2016. Big data vs privacy: The big balancing act. Vitesse Media, London, England, UK. http://www.information-age.com/big-data-vs-privacy-big-balancing-act-123461795/s.

Sasson, E.B., A. Chiesa, C. Garman, M. Green and I. Miers *et al*., 2014. Zerocash: Decentralized anonymous payments from bitcoin. Proceedings of the 2014 IEEE Symposium on Security and Privacy (SP'14), May 18-21, 2014, IEEE, San Jose, California, ISBN:978-1-4799-4686-0, pp: 459-474.

Warren, S.D. and L.D. Brandeis, 1890. The right to privacy. Harvard Law Rev., 4: 193-220.

Zyskind, G. and O. Nathan, 2015. Decentralizing privacy: Using blockchain to protect personal data. Proceedings of the 2015 IEEE Workshops on Security and Privacy (SPW'15), May 21-22, 2015, IEEE, San Jose, California, USA., ISBN:978-1-4799-9933-0, pp: 180-184.