JES
Journal of Engineering Sciences

Vol 12, Issue 06, JUNE /2021
ISSN NO:0377-9254

# Securing Data with Block Chain and AI

*CH.JAYA RAMULU, BOREDDY BHAVANI, PALLAPOLU NITHISHA, MADASU HARITHA, PETETI BHAVANI, CHERUKURI MAHESH*

*DEPT OF CSE*

*KRISHNACHAITANYA INSTITUTE OF TECHNOLOGY & SCIENCES*

**ABSTRACT:**

Data is the input for various artificial intelligence(AI) algorithms to mine valuable features, yet data in Internet is scattered everywhere and controlled by different stakeholders who cannot believe in each other, and usage of the data in complex cyberspace is difficult to authorize or to validate. As a result, it is very difficult to enable data sharing in cyberspace for the real big data, as well as a real powerful AI. In this paper, we propose the SecNet, an architecture that can enable secure data storing, computing, and sharing in the large-scale Internet environment, aiming at a more secure cyberspace with real big data and thus enhanced AI with plenty of data source, by integrating three key components: 1) blockchain- based data sharing with ownership guarantee, which enables trusted data sharing in the large-scale environment to form real big data; 2) AI-based secure computing platform to produce more intelligent security rules, which helps to construct a more trustedcyberspace; 3) trusted value-exchange mechanism for purchasing security service, providing a way for participants to gain economic rewards when giving out their data or service, which promotes the data sharing and thus achieves better performance of AI. Moreover, we discuss the typical use scenario of SecNet as well as its potentially alternative way to deploy, as well as analyze its effectiveness from the aspect of network security and economic revenue.

## I. INTRODUCTION

With the development of information technologies, the trend of integrating cyber, physical and social (CPS) systems to a highly unified information society, rather than just a digital Internet, is becoming increasing obvious [1]. In such an information society, data is the asset of its owner, and its usage should be under the full control of its owner, although this is not the common case [2], [3]. Given data is undoubtedly the oil of the information society, almost every big company want to collect data as much as possible, for their future competitiveness [4], [5]. An increasing amount of personal data, including location information, web-searching behavior, user calls, user preference, is being silently collected by the built-in sensors inside the products from those big companies, which brings in huge risk on privacy leakage of data owners [6], [7]. Moreover, the usage of those data is out of control of their owners, since currently there is not a reliable way to record how the data is used and by who, and thus has little methods to trace or punish the violators who abuse those data [8]. That is, lack of ability to effectively manage data makes it very difficult for an individual to control the potential risks associated with the collected data [9]. For example, once the data has been collected by a third party (e.g., a big company), the lack of access to this data hinders an individual to understand or manage the risks related to the collected data from him. Meanwhile, the lack of immutable recording for the usage of data increases the risks to abuse them [10]. If there is an efficient and trusted way to collect and merge the data scattered across the whole CPS to form real big data, the performance of artificial intelligence (AI) will be significantly improved since AI can handle massive amount of data including huge information at the same time, which would bring in great benefits (e.g., achieving enhanced security for data) and even makes AI gaining the ability to exceed human capabilities in more areas [11]. According to the research in [12], if given large amount of data in an orders of magnitude more scale, even the simplest AI algorithm currently (e.g., perceptrons from the 1950s) can achieve fanciest performance to beat many state-of-the-art technologies today. The key lies in how to make data sharing trusted and secured [13]. Fortunately, the blockchain technologies may be the promising way to achieve this goal, via consensus mechanisms throughout the network to guarantee data sharing in a tamper-proof way embedded with economic incentives [14], [15]. Thus, AI can be further empowered by blockchain-protected data sharing [16]–[18]. As a result, enhanced AI can provide better performance and security for data. In this paper, we aim at securing data by combining blockchain and AI together, and design a Secure Networking architecture (termed as SecNet) to

significantly improve the security of data sharing, and then the security of the whole network, even the whole CPS. In SecNet, to protect data, one of the biggest challenges is where and how to store data, because users have to give their data to service providers if they want to use certain services or applications [1], [3]. This is caused by the inherent coupling of user data and application in current service mechanisms, which significantly hinders the development of data protection and application innovation. Inspired by the concept of Personal Data Store (PDS) from openPDS [5] and the Private Data Center (PDC) from HyperNet [1], SecNet finally inherits and adopts PDC instead of PDS, as PDC is more suitable to deploy and to deal with this problem, since it provides more secure and intelligent data storage system via physical entities instead of software-based algorithms as in openPDS. Each PDC actually serves as a secured as well as centralized physical space for each SecNet user where his/her data lives in. Embedding PDC into SecNet would allow users to monitor and reason about what and why their data is used as well as by who, meaning the users can truly control every operation on their own data and achieve fine-grained management on access behaviors for data. Actually, besides PDC, other choices can also be applied for the data storing in SecNet according to certain requirements (see Section V). The trust-less relationship between different data stakeholders significantly thwarts the data sharing in the whole Internet, thus the data used for AI training or analyzing is limited in amount as well as partial in variety. Fortunately, the rise of Blockchain technologies bring in a hopeful, efficient and effective way to enable trust data sharing in trustless environment, which can help AI make more accurate decisions due to the real big data collected from more places in the Internet. SecNet leverages the emerging blockchain technologies to prevent the abuse of data, and to enable trusted data sharing in trust-less or even untrusted environment. For instance, it can enable cooperations between different edge computing paradigms to work together to improve the whole system performance of edge networks [19]. The reason why blockchain can enable trusted mechanisms is that it can provide a transparent, tamper-proof metadata infrastructure to seriously recode all the usage of data [17]. Thus, SecNet introduces blockchain-based data sharing mechanisms with ownership guarantee, where any data ready for sharing should be registered into a blockchain, named Data Recording Blockchain (DRB), to announce its availability for sharing. Each access behavior on data by other parties (not the data owner) should also be validated and recorded in this chain. In addition, the authenticity and integrity of data can only be validated by DRB as well. Besides, SecNet enables economic incentive between different entities if they share data or exchange security service, by embedding smart contract on data to trigger automatic and tamper-proof value exchange. In this way, SecNet guarantees the data security and encourages data sharing throughout the CPS. Furthermore, data is the fuel of AI [11], and it can greatly help to improve the performance of AI algorithms if data can be efficiently networked and properly fused. Enabling data sharing across multiple service providers can be a way to maximize the utilization of scattered data in separate entities with potential conflicts of interest, which can enables a more powerful AI. Given enough data and blockchainbased smart contract [20] on secure data sharing, it is not surprised that AI can become one of the most powerful technologies and tools to improve cybersecurity, since it can check huge amount of data more quickly to save time, and identify and mitigate threats more rapidly, and meanwhile give more accurate prediction and decision support on security rules that a PDC should deploy. Besides, embedded with Machine Learning [21] inside, AI can constantly learn patterns by applying existing data or artificial data generated by GAN [22] to improve its strategies over time, to strengthen its ability on identifying any deviation on data or behaviors on a 24/7/365 basis. SecNet can apply these advanced AI technologies into its Operation Support System (OSS) to adaptively identify more suspicious data-related behaviors, even they are never seen before. In addition, swarm intelligence can be used in SecNet to further improve the data security, by collecting different security knowledge from huge amount of intelligent agents scattered everywhere in the CPS, with the help of trusted exchange mechanisms for incentive tokens.

## II. EXISTING SYSTEM:

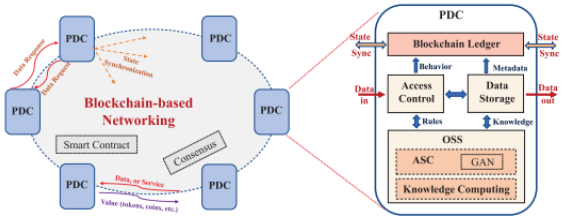
Given data is undoubtedly the oil of the information society, almost every big company want to collect data as much as possible, for their futurecompetitiveness [4], [5]. An increasing amount of personal data, including location information, web- searching behavior, user calls, user preference, is being silently collected by the built-in sensors inside the products from those big companies, which brings in huge risk on privacy leakage of data owners [6], [7]. Moreover, the usage of those data is out of control of their owners, since currently The associate editor coordinating the review of this manuscript and

approving it for publication was Chi-Yuan Chen. there is not a reliable way to record how the data is used and by who, and thus has little methods to trace or punish the violators who abuse those data [8]. That is, lack of ability to effectively manage data makes it very difficult for an individual to control the potential risks associated with the collected data [9]. For example, once the data has been collected by a third party (e.g., a big company), the lack of access to this data hinders an individual to understand or manage the risks related to the collected data from him. Meanwhile, the lack of immutable recording for the usage of data increases the risks to abuse them [10].

## III.PROPOSED SYSTEM:

we aim at securing data by combining blockchain and AI together, and design a Secure Networking architecture (termed as SecNet) to significantly improve the security of data sharing, and then the security of the whole network, even the whole CPS. In SecNet, to protect data, one of the biggest challenges is where and how to store data, because users have to give their data to service providers if they want to use certain services or applications [1], [3]. This is caused by the inherent coupling of user data and application in current service mechanisms, which significantly hinders the development of data protection and application innovation. Inspired by the concept of Personal Data Store (PDS) from openPDS [5] and the Private Data Center (PDC) from HyperNet [1], SecNet finally inherits and adopts PDC instead of PDS, as PDC is more suitable to deploy and to deal with this problem, since it provides more secure and intelligent data toragesystem via physical entities instead of software-basedalgorithms as in openPDS. Each PDC actually serves as a secured as well as centralized physical space for each SecNet user where his/her data lives in. Embedding PDC into SecNet would allow users to monitor and reason about what and why their data is used as well as by who, meaning the users can truly control every operation on their own data and achievefine-grained management on access behaviors for data. Actually, besides PDC, other choices can also be applied for the data storing in SecNet according to certain requirement

## IV.SYSTEM ARCHITECTURE




## V.MODULES:

Modules Information:

This project consists of two modules

1) Patients: Patients first create his profile with all disease details and then select desired hospital with whom he wishes to share/subscribe data. While creating profile application will create Blockchain object with allowable permission and it will allow only those hospitals to access data.

   Patient Login: Patient can login to application with his profile id and check total rewards he earned from sharing data.

2) Hospital: Hospital1 and Hospital2 are using in this application as two organizations with whom patient can share data. At a time any hospital can login to application and then enter search string as disease name.

AI algorithm will take input disease string and then perform search operation on all patients to get similar disease patients and then check whether this hospital has permission to access that patient data or not, if hospital has access permission then it will display those patients records to that hospital.

First create database in MYSQL by copying content from 'DB.txt' file and paste in MYSQL.

In settings file change port no from 3308 to 3306 and in 'views.py' file also change port no to 3306

Deploy code on DJANOGO and start server and run in browser to get below screen

## VI. SYSTEM SPECIFICATION:

**HARDWARE REQUIREMENTS:**

- **System** : Pentium IV 2.4 GHz.
- **Hard Disk** : 40 GB.
- **Floppy Drive** : 1.44 Mb.
- **Monitor** : 14' Colour Monitor.
- **Mouse** : Optical Mouse.
- **Ram** : 512 Mb.

**SOFTWARE REQUIREMENTS:**

- **Operating system** : Windows 7Ultimate.
- **Coding Language** : Python. anaconda

**Front-End** : html

## VII.CONCLUSION:

In order to leverage AI and blockchain to fit the problem of abusing data, as well as empower AI with the help of blockchain for trusted data management in trust-less environment, we propose the SecNet, which is a new networking paradigm focusing on secure data storing, sharing and computing instead of communicating. SecNet provides data ownership guaranteeing with the help of blockchain technologies, and AI-based secure computing platform as well as blockchain-based incentive mechanism, offering paradigm and incentives for data merging and more powerful AI to finally achieve better network security. Moreover, we discuss the typical use scenario of SecNet in medical care system, and gives alternative ways for employing the storage function of SecNet. Furthermore, we evaluate its improvement on network vulnerability when countering DDoS attacks, and analyze the inventive aspect on encouraging users to share security rules for a more secure network. In future work, we will explore how to leverage blockchain for the access authorization on data requests, and design secure and detailed smart contracts for data sharing and AI-based computing service in SecNet. In addition, we will model SecNet and analyze its performance through extensive experiments based on advanced platforms (e.g., integrating IPFS [27] and Ethereum [28] to form a SecNet-like architecture).

## VIII.REFERENCES

[1] H. Yin, D. Guo, K. Wang, Z. Jiang, Y. Lyu, and J. Xing, ''Hyperconnected network: A decentralized trusted computing and networking paradigm,'' IEEE Netw., vol. 32, no. 1, pp. 112–117, Jan./Feb. 2018.

[2] K. Fan, W. Jiang, H. Li, and Y. Yang, ''Lightweight RFID protocol for medical privacy protection in IoT,'' IEEE Trans Ind. Informat., vol. 14, no. 4, pp. 1656–1665, Apr. 2018.

[3] T. Chajed, J. Gjengset, J. Van Den Hooff, M. F. Kaashoek, J. Mickens, R. Morris, and N. Zeldovich, ''Amber: Decoupling user data from Web applications,'' in Proc. 15th Workshop Hot Topics Oper. Syst. (HotOS XV), Warth-Weiningen, Switzerland, 2015, pp. 1–6.

[4] M. Lecuyer, R. Spahn, R. Geambasu, T.-K. Huang, and S. Sen, ''Enhancing selectivity in big data,'' IEEE Security Privacy, vol. 16, no. 1, pp. 34– 42, Jan./Feb. 2018.

[5] Y.-A. de Montjoye, E. Shmueli, S. S. Wang, and A. S. Pentland, ''openPDS: Protecting the privacy of metadata through SafeAnswers,'' PLoS ONE, vol. 9, no. 7, 2014, Art. no. e98790.

[6] C. Perera, R. Ranjan, and L. Wang, ''End-to-end privacy for open big data markets,'' IEEE Cloud Comput., vol. 2, no. 4, pp. 44–53, Apr. 2015.

[7] X. Zheng, Z. Cai, and Y. Li, ''Data linkage in smart Internet of Things systems: A consideration from a privacy perspective,'' IEEE Commun. Mag., vol. 56, no. 9, pp. 55–61, Sep. 2018. [8] Q. Lu and X. Xu, ''Adaptable blockchain-based systems: A case study for product traceability,'' IEEE Softw., vol. 34,no. 6, pp. 21–27, Nov./Dec. 2017.

[9] Y. Liang, Z. Cai, J. Yu, Q. Han, and Y. Li, ''Deep learning based inference of private information using embedded sensors in smart devices'' IEEE Netw. Mag., vol. 32, no. 4, pp. 8–14, Jul./Aug. 2018.

[10] Q. Xia, E. B. Sifah, K. O. Asamoah, J. Gao, X. Du, and M. Guizani, ''MeDShare: Trust-less medical data sharing among cloud service providers via blockchain,'' IEEE Access, vol. 5, pp. 14757–14767, 2017.

[11] D. E. O'Leary, ''Artificial intelligence and big data,'' IEEE Intell. Syst., vol. 28, no. 2, pp. 96–99, Mar. 2013.

[12] A. Halevy, P. Norvig, and F. Pereira, ''The unreasonable effectiveness of data,'' IEEE Intell. Syst., vol. 24, no. 2, pp. 8–12, Mar. 2009.

[13] Z. Cai and X. Zheng, ''A private and efficient

mechanism for data upload- ing in smart cyber-physical systems,'' IEEE Trans. Netw. Sci. Eng., to be published. doi: 10.1109/TNSE.2018.2830307.

[14] A. Dorri, M. Steger, S. S. Kanhere, and R. Jurdak, ''BlockChain: A dis- tributed solution to automotive security and privacy,'' IEEE Commun. Mag., vol. 55, no. 12, pp. 119–125, Dec. 2017.

[15] J. Wang, M. Li, Y. He, H. Li, K. Xiao, and C. Wang, ''A blockchain based privacy-preserving incentive mechanism in crowdsensing applications,'' IEEE Access, vol. 6, pp. 17545–17556, 2018.

[16] C. Sun, A. Shrivastava, S. Singh, and A. Gupta, ''Revisiting unreasonable effectiveness of data in deep learning era,'' in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 843–852.

[17] W. Meng, E. W. Tischhauser, Q. Wang, Y. Wang, and J. Han, ''When intrusion detection meets blockchain technology: A review,'' IEEE Access, vol. 6, pp. 10179–10188, 2018.

[18] J.-H. Lee, ''BIDaaS: Blockchain based ID as a service,'' IEEE Access, vol. 6, pp. 2274–2278, 2017.

[19] K. Wang, H. Yin, W. Quan, and G. Min, ''Enabling collaborative edge computing for softwaredefined vehicular networks,'' IEEE Netw., vol. 32, no. 5, pp. 112–117, Sep./Oct. 2018.

[20] A. B. Kurtulmus and K. Daniel, ''Trustless machine learning con- tracts; evaluating and exchanging machine learning models on the ethereum blockchain,'' 2018, arXiv:1802.10185. Available: https://arxiv.org/abs/1802.10185