

SECURING DATA WITH BLOCKCHAIN AND AI

^{#1}**N. VENKATESWAR RAO**, *Associate Professor*,

^{#2}**V. LALITHA**, *Assistant Professor*,

Department of CSE,
SAI SPURTHI INSTITUTE OF TECHNOLOGY, SATHUPALLI, KHAMMAM

ABSTRACT- Despite the fact that data is dispersed throughout the Internet and controlled by various stakeholders who do not have faith in one another, it is challenging to authorize or validate the use of data in complex cyberspace. Data is the input that is used by various artificial intelligence (AI) algorithms to mine valuable features. Consequently, it is extremely challenging to enable data sharing in cyberspace for genuine big data and powerful AI. By integrating three key components, we propose the Sec Net architecture in this paper, which aims for a more secure cyberspace with real big data and, consequently, enhanced AI with plenty of data sources. This architecture can enable secure data storing, computing, and sharing in the large-scale Internet environment. 1) Trusted data sharing in a large-scale environment to create genuine big data through block chain-based data sharing with ownership guarantee; 2) An AI-based secure computing platform for the creation of more intelligent security rules that contribute to the creation of a more trustworthy online environment; 3) a dependable value-exchange mechanism for purchasing security services, allowing participants to earn financial rewards for sharing their data or services, encouraging data sharing and improving AI performance. In addition, we examine SecNet's efficacy from the perspectives of network security and revenue generation, as well as its potential alternative deployment methods.

INDEX TERMS: *Data security, data systems, artificial intelligence, cyber space.*

INTRODUCTION

The trend of integrating cyber, physical, and social (CPS) systems into a highly unified information society, rather than just a digital Internet, is becoming increasingly obvious with the advancement of information technologies [1]. Data should be the owner's asset in such an information society, and the owner should have full control over how

it is used, although this is not always the case [2, 3]. Since data is unquestionably the fuel of the information society, almost every major company wants to collect as much data as they can to improve their competitiveness in the future [4, 5]. The built-in sensors in the products of these big companies are silently collecting an increasing amount of personal data, including location information, websearch behavior, user calls, and user preference

[6, 7], which poses a significant risk to the privacy of data owners. In addition, the owners of those data have no control over how they are used because there is currently no reliable method for recording how the data is used and by whom, making it difficult to locate or punish data abusers [8]. That is, it is extremely difficult for a person to control the potential risks associated with the collected data if they lack the ability to effectively manage it [9]. For instance, once the data has been collected by a third party (such as a large corporation), an individual's inability to comprehend or manage the risks associated with the collected data prevents him from doing so. In the meantime, abuse of data is more likely because there is no immutable record of its use [10]. Artificial intelligence (AI) will be able to handle massive amounts of data, including huge amounts of information, at the same time, which would bring about great benefits (such as achieving enhanced data security) and even make AI gaining the ability to surpass human capabilities in more areas if there is an effective and dependable method for collecting and merging the data scattered across the entire CPS into real big data [11].

According to the findings of the study published in [12], even the simplest AI algorithm that is currently in use—such as perceptrons from the 1950s—can achieve the highest level of performance to surpass many of the most recent technologies. How to make data sharing secure and trustworthy is the key [13]. Thanks to consensus mechanisms throughout the network that guarantee data sharing in a tamper-proof manner and are infused with economic incentives [14, 15], the

blockchain technologies may be the promising means of achieving this objective. Subsequently, simulated intelligence can be additionally enabled by blockchain-protected information sharing [16]–[18]. Consequently, enhanced AI can enhance data security and performance. By combining AI and blockchain in this paper, we aim to secure data and design a Secure Networking architecture (SecNet) to significantly enhance the security of data sharing and the entire network, including the CPS. Because users are required to provide their data to service providers in order to use particular applications or services [1, 3], one of the most challenging aspects of data protection in SecNet is determining where and how to store data.

This is because the current service mechanisms' inherent coupling of user data and application significantly impedes the growth of data protection and application innovation. PDC is more suitable for deployment and to deal with this problem because it provides a more secure and intelligent data storage system via physical entities rather than software-based algorithms as in openPDS. SecNet is inspired by the concept of the Personal Data Store (PDS) from openPDS [5] and the Private Data Center (PDC) from HyperNet [1]. Each SecNet user's data is actually stored in a protected and centralized physical space in each PDC. Users would be able to truly control every operation on their own data and achieve fine-grained management of access behaviors for data if PDC were integrated into SecNet. This would enable users to monitor and reason about what and why their data is used as well as by who. In point of fact, depending on a number of requirements, in addition to PDC, other options can also be utilized for the storage of data in SecNet (see Section V).

The lack of trust between various data

stakeholders significantly impedes the sharing of data across the Internet; consequently, the data utilized for AI training or analysis is limited in quantity and partial in variety. Thanks to the real big data that is being collected from more places on the Internet, the rise of block chain technologies offers a hopeful, efficient, and effective method for enabling trust data sharing in a trustless environment. SecNet makes use of the upcoming blockchain technologies to stop data abuse and make trusted data sharing possible in environments where trust is low or even nonexistent. For instance, it may make it possible for various edge computing paradigms to collaborate in order to enhance the overall system performance of edge networks [19]. Blockchain can provide a transparent, tamper-proof metadata infrastructure to seriously recode all data usage, enabling trusted mechanisms [17]. As a result, SecNet introduces ownership-guaranteed blockchain-based data sharing mechanisms in which any data that is ready for sharing must be recorded in a Data Recording Blockchain (DRB) to indicate its availability for sharing. This chain should also validate and record every data access behavior by parties other than the data owner. Additionally, only DRB can validate the authenticity and integrity of data. Moreover, SecNet empowers monetary motivation between various substances assuming they share information or trade security administration, by implanting shrewd agreement on information to set off programmed and sealed esteem trade.

SecNet encourages data sharing throughout the CPS and ensures data security in this way. In addition, data is the engine of AI [11], and if it can be effectively networked and fused, it can significantly enhance the performance of AI algorithms. A method for maximizing the

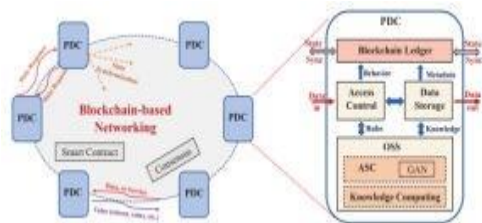
utilization of scattered data in separate entities with potential conflicts of interest can be to enable a more powerful AI by enabling data sharing among multiple service providers. Given an adequate number of information and blockchainbased shrewd agreement [20] on secure information sharing, it isn't shocked that simulated intelligence can become quite possibly of the most remarkable innovation and instruments to further develop network protection, since it can check colossal measure of information all the more rapidly to save time, and distinguish and moderate dangers all the more quickly, and in the mean time give more precise expectation and choice help on security decides that a PDC ought to convey. In addition, AI can constantly learn patterns by applying existing data or artificial data generated by GAN [22] to improve its strategies over time and strengthen its ability to identify any deviation in data or behaviors 24 hours a day, 7 days a week. This is possible because AI is embedded with Machine Learning [21]. These cutting-edge AI technologies can be incorporated into SecNet's Operation Support System (OSS) to adaptively identify more unusual data-related behaviors. Swarm intelligence can also be used in SecNet to improve data security by collecting security information from a large number of intelligent agents scattered throughout the CPS using trusted incentive token exchange mechanisms.

SYSTEM ANALYSIS

EXISTING SYSTEM

Since data is unquestionably the fuel of the information society, almost every major company wants to collect as much data as they can to improve their competitiveness in the future [4, 5]. The built-in sensors in the products of these big companies are silently collecting an increasing amount of personal data, including location information, web-search behavior, user calls, and user preference [6, 7], which poses a significant

risk to the privacy of data owners. In addition, the owners of those data have no control over how they are used, as Chi-Yuan Chen is currently the associate editor in charge of coordinating the review of this manuscript and authorizing its publication. There isn't a reliable way to keep track of who uses the data and how, so it's hard to find or punish those who misuse it [8]. That is, it is extremely difficult for a person to control the potential risks associated with the collected data if they lack the ability to effectively manage it [9]. For instance, once the data has been collected by a third party (such as a



large corporation), an individual's inability to comprehend or manage the risks associated with the collected data prevents him from doing so. In the meantime, abuse of data is more likely because there is no immutable record of its use [10].

System Concept:

we target getting information by joining blockchain and computer based intelligence together, and plan a Safe

Organizing engineering (named as SecNet) to essentially work on the security of information sharing, and afterward the security of the entire organization, even the entire CPS. Because users are required to provide their data to service providers in order to use particular applications or services [1, 3], one of the most challenging aspects of data protection in SecNet is determining where and how to store data. This is because the current service mechanisms' inherent coupling of user data and application significantly impedes the growth of data protection and application innovation. PDC is more suitable for deployment and to deal with this problem because it provides more secure and

intelligent data storagesystem via physical entities instead of software-based algorithms as in openPDS. SecNet is inspired by the concept of the Personal Data Store (PDS) from openPDS [5] and the Private Data Center (PDC) from HyperNet [1]. Each SecNet user's data is actually stored in a protected and centralized physical space in each PDC. Users would be able to truly control every operation on their own data and achieve fine-grained management of access behaviors for data if PDC were integrated into SecNet. This would enable users to monitor and reason about what and why their data is used as well as by who. In fact, depending on a particular requirement, other options for data storage in SecNet can be utilized in addition to PDC.

SYSTEMARCHITECTURE

I. IMPLEMENTATION

ModulesInformation:

Thisproject consistsoftwomodules

- 1) **Patients:** Patients first create his profilewithalldiseasedetailsandthenselect desired hospital with whom he wishes toshare/subscribedata.Whilecreatingprofile application will create Blockchainobject with allowable permission and itwill allow only those hospitals to accessdata.

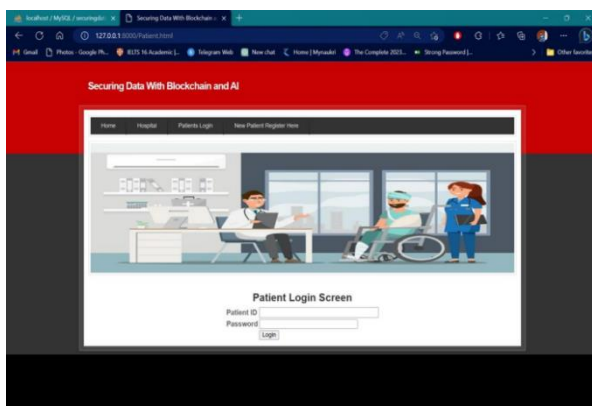
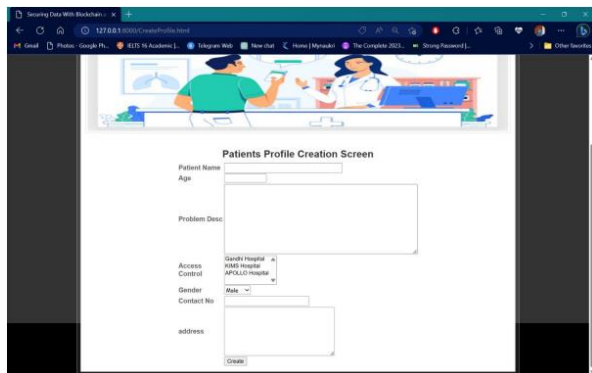
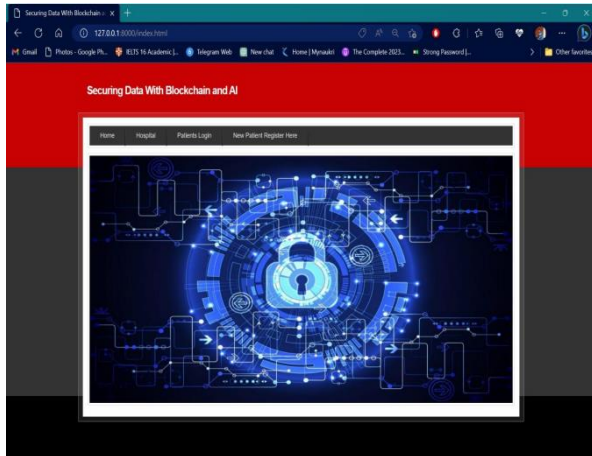
PatientLogin:Patientcanloginintoapplication with his profile id and checktotalrewardsheearnedfromsharingdata.

- 2) **Hospital:**Hospital1andHospital2areusing inthisapplicationastwoorganizationswith whompatientcansharedata.Atatimeanyhospitalcanlogin to application and then enter searchstringas disease name.

AIalgorithmwilltakeinputdiseasestring and then perform search operationonallpatientstogetsimilardiseasepatientsandthencheckwhetherthishospitalhaspermissiontoaccessthatpatient data

or not, if hospital has access permission it will display those patients records to that hospital.

II. RESULTS



III. CONCLUSION

In order to leverage AI and block chain to fit the problem of abusing data, as well as empower AI with the help of block chain for trusted data management in trust-less environment, we propose the Sec Net, which is a new

networking paradigm focusing on secured 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 give alternative ways for employing the storage function of SecNet. Furthermore, we evaluate its improvement on network vulnerability when countering DDoS attacks, and analyze the incentive 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 Sec Net. In addition, we will model Sec Net 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).

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. (HotOSXV)*, 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 Safe Answers,” *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: Trustless 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 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 Jurdak, “BlockChain: A distributed 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.