A Cross-Domain Secure Data Trading Framework Based on Blockchain-Cloud Fusion

Youngho Park[1], Su Jin Shin[2], Ye Hyeon Park[2], and Sang Uk Shin[3]*

[1] Electronics and Information Communications Research Center
Pukyong National University, Republic of Korea
pyhoya@pknu.ac.kr
[2] Department of Information Security, Graduate School
Pukyong National University, Republic of Korea
{inuin9014, yhyeon00612}@gmail.com
[3] Division of Computer Engineering, Pukyong National University
Republic of Korea shinsu@pknu.ac.kr

Abstract

Data has become a valuable asset in the era of the data economy, so data trading enables data owners and data requesters to sell and purchase data. This requires a data marketplace that makes it possible to trade data between owners and requesters in a reliable and efficient manner. In this paper, we propose a blockchain-based secure and fair data trading framework where data owners and data requesters belonging to different domains can trade data. We present a data trading system architecture, in which the blockchain serves as a data trading controller and the cloud acts as an external storage platform. To design a secure data trading protocol that enables access control to the data by the owner and source identification by the requester, we make use of cross-domain matchmaking encryption. We also employ a smart contract to encourage the trading parties to act honestly following the trading rules.

Keywords: Blockchain, Data Marketplace, Fair and Secure Trading, Smart Contract

1 Introduction

The development of the Internet-of-Things (IoT) infrastructure and devices has led to the generation and collection of IoT data at an explosive rate, and these data have become valuable assets in the era of the data economy. Such trend demands a data marketplace that makes it possible to trade the data between owners and requesters in a reliable and efficient manner. Fair data trading requires that a data buyer can obtain the requested data if it pays the corresponding price and a data seller can receive the payment if it sends the data to the data buyer [1]. Unfortunately, there is a mistrust problem in untact data trading between the seller and the buyer. The seller would not provide the data until receiving the payment from the buyer, and the buyer is unwilling to pay first until the seller provides the data.

Blockchain is a distributed ledger with immutable on-chain storage and verifiable state updates [2][3], which can act as a transparent and reliable controller of data trading. Some existing work adopts an on-chain data trading model [4][5][6], where the data is stored and shared via the blockchain. This model can find some practical applications if the data is small.

The 7th International Conference on Mobile Internet Security (MobiSec'23), December 19-21, 2023, Okinawa, JAPAN, Article No.70

*Corresponding author: Division of Computer Engineering, Pukyong National University, Busan, 48513, Republic of Korea, shinsu@pknu.ac.kr, Tel: +82-051-629-6249

However, it is ineffective to manage the data in the on-chain data trading model due to the ever-increasing volume of the IoT data. One solution is for data owners to store and trade their IoT data on a powerful off-chain storage platform such as the cloud [7]. Therefore, a novel paradigm of blockchain and cloud fusion has been widely considered as a promising data trading platform.

IoT is made up of technologies in various domains and application areas, and interested organizations or developers in these domains would want to participate in establishing the IoT data marketplace. For example, cloud providers, technology companies, research institutes, and application developers can form a consortium to manage a blockchain-based data trading platform. Then, the members in these organizations can share and trade their data on the platform. In this system model, it is essential to ensure accountability for fair trading by detecting and penalizing dishonest parties. In addition, it is also required to achieve access control and source identification for the data managed by the external cloud. Data owners may want to specify who can access their data entrusted to the external cloud as a policy. Data requesters may want to specify an attribute for certain data owners from whom they want to purchase data.

In a blockchain-cloud fusion data trading system, the blockchain can be used to keep track of the actions taken by both the seller and the buyer while the cloud can offer a way to store and access the huge volume of IoT data. In addition, with the smart contract technique, the participants can negotiate a trustworthy data trading agreement that enforces consent-based access control over the IoT data and records data trading instances as provenance evidence. In [8], the authors proposed a blockchain-based secure data trading system, which enables access control by the seller and source identification by the buyer. However, their system model is not suitable for dealing with secure data trading between users of different organizations. They need another trusted third party to handle secure data trading when the seller and the buyer belong to different organizations.

Therefore, in this paper, we propose a secure data trading framework based on blockchain that enables access control and source identification between data owners and data requesters in multiple domains. To achieve our design goals, we consider the use of the cross-domain identity-based matchmaking encryption (cd-IBME) scheme [9]. The rest of this paper is organized as follows: Section 2 introduces related work on blockchain-based data trading systems. Section 3 presents the proposed system architecture and data trading protocols. Finally, Section 4 concludes this paper.

2 Related Work

With the emergence of blockchain technology, accompanied with cryptocurrecies and smart contracts, it is regarded that fair data trading can be achieved on the blockchain. Research on blockchain-based data trading models has received a great deal of attention, and several system models have been introduced. The existing data trading system models can be roughly classified into on-chain model and on/off-chain model. The data is directly shared or traded via the blockchain in an on-chain model [6][4] while an off-chain storage service (Cloud) is employed to host a huge volume of data in an on/off-chain model [10][8][7]. The data is usually encrypted before being outsourced to the blockchain or the cloud storage. For secure sharing/trading of the data, the data owner can specify access policy and manage the decryption rights to the data.

With regard to fair data trading, Li et al. introduced a decentralized fair data trading system based on blockchain to guarantee fair data transactions with authentication [11]. Dixit et al.

proposed a decentralized platform of a digital data marketplace enabled by blockchain which hosts IoT data in a reliable and fault-tolerant manner [12]. Chen et al. proposed a blockchain-based IoT data trading system [13], where the trading behaviors of the seller and the buyer are recorded on the blockchain to facilitate on-chain and off-chain arbitration. However, Chen et al. do not address secure data trading with access control for data confidentiality.

For secure data trading, Alsharif et al. proposed a blockchain-based medical data marketplace model [14], in which sellers can enforce access control policies on the encrypted records by using the attribute-based encryption scheme [15]. In [33], Li et al. proposed a blockchain-based secure data trading platform by using the plaintext checkable encryption scheme [16]. However, these systems only focused on the access control of the seller for data confidentiality and burdened the on-chain procedures with complex cryptographic operations.

3 Proposed System

3.1 System Architecture

We consider the data marketplace architecture as shown in Figure 1 which consists of Supervising Authority, Data Owner, Data Requester, Blockchain Layer, and Cloud Layer.

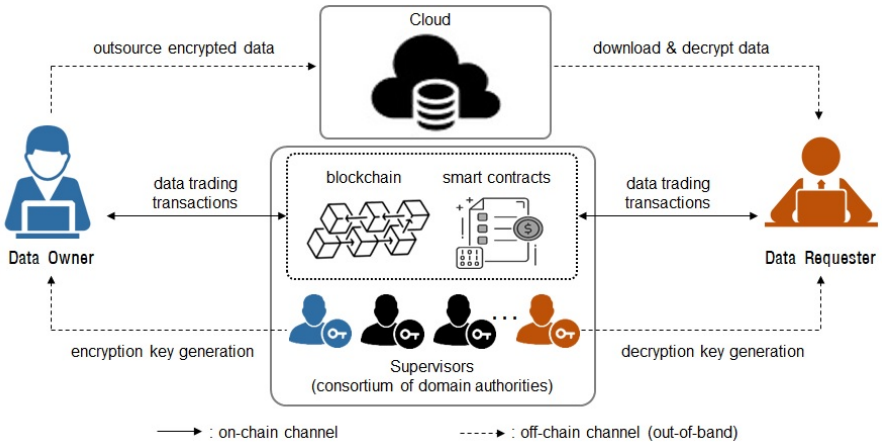

Figure 1: Overview of the proposed data trading system architecture.

- Supervising Authority (SA) consists of a set of domain authorities which cooperate to manage the data marketplace. For instance, a collection of organizations that want to participate in the data marketplace may form a consortium and collaborate in operating the data trading platform based on the blockchain. SA is responsible for setting up the public system parameters, and each domain authority acts as a key generator that issues encryption/decryption keys for owners/requesters belonging to it. In addition, if a dispute occurs, SA is involved in dispute resolution.

- Data Owner (O) offers his/her data through the data marketplace. To allow only a valid data requester authorized by the owner to access the data, O sets the requester policy which specifies the allowed type of requester. O needs to obtain the identity-based

encryption key for cd-IBME from the domain authority to which O belongs. Then, O encrypts the data by using the cd-IBME scheme under the access policy.

- Data Requester (R) initiates the data purchase when R finds the interesting data provided by the desirable data owner O who satisfies the owner policy required by R. To access the data of O, R is issued the decryption key under its identity from the domain authority to which R belongs. R can obtain the data if and only if the encrypted data is associated with the identity of R which corresponds to the requester policy of O.

- Blockchain Layer provides a decentralized trusted platform which enforces data trading rules, coordinates the data trading state, and deals with payment. Blockchain is a consortium ledger which records data trading state between the owner and the requester. Smart contract on the blockchain implements the business logic of data trading and specifies the terms and actions to allow involved parties to execute certain aspects of data trading transactions.

- Cloud acts as an efficient storage unit to host a huge volume of data in trading. The cloud is an external storage, so data owners outsource their data in the form of an encrypted package for data confidentiality.

Table 1 describes the notations used in the proposed system. For the detailed algorithm descriptions of the cd-IBME, we can refer to [9].

Table 1: Notations and descriptions.

notation	description
$\mathcal{SA}$	supervising authorities, $\mathcal{SA} = \{SA_1, SA_2, ..., SA_n\}$
SA_O, $SA_R \in \mathcal{SA}$	owner-authority and requester-authority
mpk_O, mpk_R	master public keys of SA_O and SA_R for cd-IBME
$H(x)$	cryptographic hash for the input message x
F	data file of the owner for marketing
SC	smart contract for the data trading
id_O, id_R	identities for the owner and the requester
ek_O	owner's encryption key for cd-IBME under id_O
dk_R	requester's decryption key for cd-IBME under id_R
$cd\text{-}ibme.Enc(M, mpk_O, mpk_R, ek_O, id_R)$	encryption of the message M under ek_O and id_R
$cd\text{-}ibme.Dec(C, mpk_O, dk_R, id_O)$	decryption of the cyphertext C under dk_R and id_O
$Enc_K(X)$	symmetric encryption of the input X under the key K
$Dec_K(Y)$	symmetric decryption of the input Y under the key K

3.2 Setup

The consortium of supervising authorities $\mathcal{SA}$ establishes an efficient blockchain as a transparent and reliable controller for data trading. Supervising authorities agree on public system parameters, and each $SA_i \in \mathcal{SA}$ generates its master secret key and public key pair (msk_i, mpk_i) for supporting cd-IBME. At this phase, we assume that the public parameters are published to the system and the smart contract to control the data trading is deployed on the blockchain.

3.3 Data Trading

Suppose that the owner O wants to sell a data F and the requester R wants to purchase the data through the data trading system. Assuming that O and R are respectively under the authorization of SA_O and SA_R, they perform the data trading protocol as follows.

1. To announce data selling, O submits the advertisement transaction $adv := [desc_F, rcv, price_F]$ to the system, where $desc_F$ is the description about the data F, rcv is the access policy that the requester must satisfy, and $price_F$ is the price for F. If R finds an interesting data description, R submits the request transaction $req := [ind_{adv}, snd]$, where ind_{adv} is the index of the referenced advertisement and snd is the policy specifying the attributes required to the owner.

2. When O finds the request for its advertised data, O chooses a random secret data key K to encrypt the data file as $Y \leftarrow Enc_K(F)$ and outsources Y to the cloud storage. O computes $\Delta_F = H(F)$, $\Delta_Y = H(Y)$, $\Delta_K = H(K)$, then invokes SC to register the proof of the data by submitting the transaction $reg := [ind_{req}, uri, \Delta_F, \Delta_Y, \Delta_K, dep_O]$, where uri is the link for downloading the encrypted data from the cloud storage and dep_O is a guarantee deposit that will be confiscated if O acts dishonestly during the data trading.

3. R can download the encrypted data Y from the cloud storage linked by the uri, however, cannot recover the original data F yet. R computes $\Delta'_Y = H(Y)$ and, if $\Delta'_Y \stackrel{?}{=} \Delta_Y$ holds, submits the order transaction $ord := [ind_{reg}, pay_R, exp_{ord}]$, where pay_B is the payment for the data and exp_{ord} is the expiration time. After this order, if the data key K is not offered by O within exp_{ord} then this trading will be canceled by SC and pay_R be returned back to R.

4. Upon receiving the order, O provides the data key K in a secure manner by using the cd-IBME scheme. To do that, O is required to obtain the encryption key ek_O under the $id_O = snd$ from the owner-authority SA_O, then encrypts the data key as $C \leftarrow$ *cd-ibme.Enc*(K, mpk_O, mpk_R, ek_O, rcv). O offers C to R through the blockchain by submitting the transaction $offer := [ind_{ord}, C]$.

5. To get the secret key offered by O, R requests the decryption dk_R under $id_R = rcv$ from the requester-authority SA_R, then retrieves the key as $K' \leftarrow$ *cd-ibme.Dec*(C, mpk_O, dk_R, snd). If $H(K') \stackrel{?}{=} \Delta_K$ holds, R decrypts the Y to recover the data as $F' \leftarrow Dec_{K'}(Y)$. R verifies the authenticity of F' by computing $\Delta_{F'} = H(F')$ and checking if $\Delta_{F'} \stackrel{?}{=} \Delta_F$ is valid. If so, R invokes the confirmation procedure. Then, SC transfers pay_R to O's account and returns dep_O to O. The data trading between O and R is normally completed.

Note that the policies (rcv and snd) required to the requester and the owner are associated with the encryption and decryption of the data key K. Hence, the data encrypted by the owner of $id_O = snd$ can only be decrypted by the requester of $id_R = rcv$.

3.4 Dispute Resolution

When the owner and the requester honestly behave, following the data trading protocol, they can respectively get the payment and the data. However, one party may repudiate its trading behavior. R would be at a disadvantage because R first pays the price prior to O's offer of the data. Hence, a dispute resolution to such a problematic situation is needed.

At the confirmation phase of the above data trading, if R finds that the authenticity of the data F' is violated, i.e., $\Delta_{F'} \neq \Delta_F$, R claims resolution for this problematic situation. $\mathcal{SA}$ forms a committee to examine the correctness of the traded data and the proofs recorded on the blockchain. When SA judges that O behaved dishonestly by giving wrong data, SC refunds the pre-paid pay_R to R and confiscates dep_O as a penalty for the fraudulent behavior of O.

4 Conclusion

With the emergence of blockchain technology, research on blockchain-based data trading models has received a great deal of attention. Therefore, in this paper, we proposed a secure data trading framework based on a blockchain-cloud fusion model, which enables access control and source identification between data owners and data requesters in multiple domains. We utilized cd-IBME to guarantee the security of the traded data and smart contract to achieve fair data trading. The elaborate design and evaluation of the data trading and dispute resolution protocols remain as future work.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (No. 2022R1I1A3063257), and supported by Electronics and Telecommunications Research Institute (ETRI) grant funded by the Korean government [22ZR1300, Research on Intelligent Cyber Security and Trust Infra].

References

[1] L. Xue, J. Ni, D. Liu, X. Lin, and X. S. Shen. Blockchain-based Fair and Fine-Grained Data Trading with Privacy Preservation. *IEEE Transactions on Computers*, 72(9):2440–2453, March 2023.

[2] S. Underwood. Blockchain beyond Bitcoin. *Communications of the ACM*, 59(11):15–17, November 2016.

[3] S. Ruoti, B. Kaiser, A. Yerukhimovich, J. Clark, and R. Cunningham. Blockchain Technology: What is it Good for? *Communications of the ACM*, 63(1):46–53, January 2020.

[4] E. Kokoris-Kogias, E. C. Alp, L. Gasser, P. Jovanovic, E. Syta, and B. Ford. Calypso: Private Data Management for Decentralized Ledgers. In *Proceedings of the VLDB Endowment*, volume 14, pages 586–599. ACM, December 2020.

[5] H. Gunasinghe, A. Kundu, E. Bertino, H. Krawczyk, S. Chari, K. Singh, and D. Su. Prividex: Privacy Preserving and Secure Exchange of Digital Identity Assets. In *Proc. of The World Wide Web Conference*, pages 594–604. ACM, May 2019.

[6] K. Bhaskaran, P. Ilfrich, D. Liffman, C. Vecchiola, P. Jayachandran, A. Kumar, F. Lim, K. Nandakumar, Z. Qin, V. Ramakrishna, E. GS Teo, and C. H. Suen. Double-Blind Consent-Driven Data Sharing on Blockchain. In *Proc. of the 2018 IEEE International Conference on Cloud Engineering (IC2E)*, pages 385–391. IEEE, April 2018.

[7] D. Liu, C. Huang, J. Ni, X. Lin, and X. S. Shen. Blockchain-Cloud Transparent Data Marketing: Consortium Management and Fairness. *IEEE Transactions on Computers*, 71(12):3322–3335, February 2022.

[8] Y. Park, M. H. Jeon, and S. U. Shin. A Blockchain-based Secure and Fair Decentralized Data Trading System. In *Proc. of the 6th International Symposium on Mobile Internet Security (MobiSec'22)*, December 2022.

[9] W. Luo A. Yang J.-N. Liu A. Wu, J. Weng and Zike Jiang. Cross-Domain Identity-based Matchmaking Encryption. https://eprint.iacr.org/2022/085, 2022.

[10] L. Zhu, Y. Wu, K. Gai, and K.-K. R. Choo. Controllable and Brustworthy Blockchain-Based Cloud Data Management. *Future Generation Computer Systems*, 91:527–535, February 2019.

[11] Y. Li, L. Li, Y. Zhao, N. Guizani, Y. Yu, and X. Du. Toward Decentralized Fair Data Trading Based on Blockchain. *IEEE Network*, 35(1):304–310, September 2020.

[12] A. Dixit, A. Singh, Y. Rahulamathavan, and M. Rajarajan. FAST DATA: A Fair, Secure and Trusted Decentralized IIoT Data Marketplace Enabled by Blockchain. *IEEE Internet of Things Journal*, Early Access, October 2021.

[13] F. Chen, J. Wang, C. Jiang, T. Xiang, and Y. Yang. Blockchain Based Non-repudiable IoT Data Trading: Simpler, Faster, and Cheaper. In *Proc. of the IEEE INFOCOM 2022 - IEEE Conference on Computer Communications*, pages 1958–1967. IEEE, May 2022.

[14] A. Alsharif and M. Nabil. A Blockchain-based Medical Data Marketplace with Trustless Fair Exchange and Access Control. In *Proceedings of the GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, pages 1–6. IEEE, December 2020.

[15] J. Bethencourt, A. Sahai, and B. Waters. Ciphertext-Policy Attribute-Based Encryption. In *Proc. of the 2007 IEEE Symposium on Security and Privacy (SP'07)*, pages 321–334. IEEE, 2007.

[16] S. Ma, Y. Mu, and W. Susilo. A Generic Scheme of Plaintext-Checkable Database Encryption. *Information Sciences*, 429:88–101, March 2018.