Securing Healthcare IT Infrastructure with Blockchain and Modern Cryptography

Yan Huang
Indiana Univeristy
yh33@indiana.edu

Haixu Tang
Indiana University
hatang@indiana.edu

1. Introduction

We consider novel integration of blockchain technology and several contemporary cryptographic protocols to strengthen the Healthcare IT infrastructure.

The blockchain technology is distinguished by its two interweaving themes *openness* and *decentralization*. Its openness allows anyone in the world to contribute to its execution, hence shoveling the system's operational cost to a swamp of well-motivated volunteering contributors. Meanwhile, its decentralized architecture alleviates many security and reliability issues that have been plaguing the trusted third party paradigm for a long time. These valuable properties would be well-suited for managing health- and medical-data with high security and reliability requirements. We will discuss a few representative applications in § 2.

Patient care is witnessing the coming of its digital era: massive health data are collected routinely using wearable health devices and clinical records are also becoming electronic. Electronic health records (EHRs) come with many applications. For examples, it is much easier for clinical practitioners to learn the complete health history of a patient moving from one city to another or transferring between hospitals to another from the EHRs; it will also be convenient for physicians to devise proper, evidence-based therapeutic approach to specfic patients based on the EHRs of patients with the same diseases; finally, biomedical researchers may exploit EHRs to build predictive models for early diagnosis of diseases and to discover novel methods for effective disease treatment. Despite these obvious advantages, EHR data are still of limited use in practice, largely due to security and privacy concerns. In recent years, cryptogrpahic approaches have been used in biomedical informatics to enable privacy-preserving sharing and analysis of sensitive health data Naveed et al. [10]. But the blockchain technology has yet been considered for health applications. In this whitepaper, we lay out several scinarioes in that blockchain technology can be used alone or combined with other cryptogrpahic protocols to enhance the privacy protection when human health data are used.

Secure Multiparty Computation. Secure multiparty computation enables a set of mutually-distrustful parties to jointly compute any agreed-on target function over their private inputs. It is a prominent enabling-technique behind collaborative computing over encrypted data, promising many important applications such as privacy-preserving biometric identification and personal genome comparison (e.g., to match patients with donors). While the theoretical feasibility results of secure computation was discovered in the 1980s [4, 14], many theoretical [6, 7] as well as implementational [1, 5, 8, 12, 15] advancements have been made throughout the years. A state-of-the-art secure computation implementation can compute the edit distance between whole human-genomes

within 40 seconds over a realistic continental network [13]. We consider using secure multiparty computation protocols in a few application scenarios including auditable and authorized medical record accesses (Section 2.1) and marketing authenticated personal healthcare data (Section 2.3).

Verifiable Computation. The classic goal of verifiable computation is to enable someone to verify the correctness of a computational outcome in a more efficient way than compute the results on his/her own. The state-of-the-art verifiable computation implementations [2, 3, 11] is able to offer *public verifiability* and achieve up to an order of magnitude cost reduction for the verifier. We will speculate combining verifiable computation, secure computation and blockchain technology in marketing authenticated personal records (Section 2.3) and verifiable clinical trials (Section 2.4) to increase the *confidence* on surveys and studies based on those private records.

2. Many Use Cases

It is highly challenging to manage medical records securely and effectively. Below we discuss how today's blockchain technology can be utilized to achieve some of these overarching goals.

2.1. Authorized and Auditable Accesses

Consider realizing a medical record management system to support controlled accesses to medical records. More specifically, it requires data owner authorization before accessing and will automatically log every access (including the access by data owner) that has occurred in the history for tracking purposes. Such a system can be implemented without a trusted third party but utilizing the *consensus* ensured by the blockchain technology. We can imagine such a record retrieval be processed as follows:

1. The requesting party (e.g., a doctor) sends a signed *request* to the record owner.
2. The record owner submits a signed *authorization* to the blockchain.
3. Once the transaction is picked up by the miners to put on the blockchain, the miners will verify both signatures of the accessing party and authorization party and append an *access record* in the access history of the requested record.
4. A secure multiparty computation protocol is launched among a random subset of miners to re-encrypt the requested record with a fresh key K while K itself will be encrypted under the requesting party's public key. Both ciphertexts are delivered to the requesting party.

Thus, all "document access" events are logged as a result of step (3) since a missing log disqualifies the transaction to be accepted by the blockchain. Also note that without the data owner's authorization, step (4) won't be accomplished, hence the requesting party cannot retrieve the data. On the other hand, a (potentially) malicious data owner cannot "incriminate" a party by declaring its record was accessed by that party because of step (1) and (3), where the requesting party's signature is needed to enable the transaction. Finally, the requesting party has sufficient information from step (4) to access the record at any time, without additional authentication from data owner.

2.2. Smart Prescription and Insurance

Coupled with blockchain technology comes the capability of enforcing *smart contracts*, which would result in copious opportunities to reduce the administrative effort in carrying out consistent checks on health/medical records. This advantage is especially valuable taking into account the prevalent adoption of intelligent IoT (Internet of Things) devices. Below, we speculate the use of blockchain-enabled smart contracts in implementing smart prescriptions and health insurance.

Suppose an individual obtains from his/her doctor visit a prescription that is electronically certified over the blockchain. That is, a hash of the prescription is signed by the prescribing doctor and timestamped on the blockchain. Later, when this individual shops at any drugstore for the

prescribed medicine, the smart payment device at the store would be able to automatically verify that the sales is indeed backed by a valid prescription. Moreover, it would be even slicker when the payment is required to be in certain form of cryptocurrencies since

1. It will economically deter stores from selling products to customers without an appropriate prescription because such transactions are deemed violating the contract, hence will not be recognized by the majority-honest miners.
2. Some medical insurance articles that can be encoded into smart contracts can be much easier to claim. For example, a smart insurance contract that stipulates full coverage of immunization shots would be uploaded to the blockchain at purchase time such that all transactions of flu shots of the beneficiary will be automatically charged towards the dedicated account of the insurance company.
3. Moreover, using a network of intelligent IoT devices, it would allow insurance companies to implement novel incentive mechanisms (such as reduced premium for customers who work out regularly or watch out their diet) to promote healthy habits.

Since these mechanisms are enforced by decentralized volunteering computers, this approach would reduce the bookkeeping costs meanwhile offering a unique level of guarantee on executing the claims when events happen.

2.3. Marketing (Authenticated) Personal Data

Today abundant personal data such as medical records and personal genomic information are being collected. However, it is generally difficult to fully utilize these data in medical research and practice due to privacy concerns. The blockchain technology along with smart contracts would enable novel mechanisms for individuals to profit from releasing their private data in a controllable manner.

Imagine a person is willing to contribute his/her own private records as long as the uses of the private data satisfy some strong security guarantees, such as that the data is never revealed directly or that it should always be used in surveys above certain size. In addition, it is possible to encourage this kind of personal information sharing with *pay-per-use* monetary rewards. We note the smart contracts are well-suited to implement such security and monetary incentive mechanisms, while secure uses of the personal data can be accomplished using homomorphic encryption and secure multiparty computation protocols.

Further, we remark that the personal medical records marketed this way can enjoy the *authenticity* originally provided by the blockchain mechanism. Moreover, through securely executing (e.g., using a random subset of miners) a verifiable computation's Verify algorithm (based on the access history recorded on blockchain), it would be possible to *publicly verify* the research results over the private data, without actually revealing the private input. This will offer extra credibility to the research results over the private data.

2.4. Verifiable Clinical Trials

Pharmaceutical companies need to undertake serious field trials to demonstrate the effectiveness of their products to obtain drug approvals. Credibility of these field trials relies heavily on the absence of bias in grouping the participating patients. Namely, the participating patients should be randomly picked according to the clinical trial specifications and receive identical treatments throughout the trials. Blockchain would make it easier to digitalize these trial records in a way that allows publicly auditing the fairness of such trial practices.

The basic idea is to randomly mix the medicine and placebo and let the blockchain miners tag them with random labels created by miners using secure multi-party computation. Then both the manufacturer and the miners commit their random permutation and upload the commitments to the blockchain. Later on, information on how the medicine/placebo are actually mapped to

patients will be recorded with the blockchain as regular medical records. Now no one can find out the medicine/placebo-to-patient mapping given that the manufacturer cannot corrupt all the miners. Finally, after all treatments are done, the manufacturer and blockchain miners will release the random mappings that they committed earlier, thus revealing the *case* and *control* grouping. Since the full medical history of participating patients before and after taking the medicine/placebo has been timestamped on the blockchain, it is straightforward to evaluate and verify the effectiveness of the medicine. Alternatively, the certificate of effectiveness can be implemented as smart contracts to leverage the blockchain's automated consensus reaching mechanism.

2.5. Securely Outsourcing Medical Dataset

Due to the advancement in data collecting methods (e.g., through wearable IoT devices), the size of medical records nowadays can be inundating. This can raise many challenges to securely store, update and retrieve these sensitive datasets.

It is promising to tackle these challenges with novel combination of blockchain and certain *proof-of-storage* techniques. The basic idea is to incentivising every miner to store a random subset of the encrypted dataset locally in order to keep mining. Here special mechanisms are needed to translate the proof-of-work puzzles into ones that also require proofs of storage. Comparing to existing commercial cloud storage services, this approach is distinguished by its cost-effectiveness and high availability thanks to a free market of decentralized volunteers. As a side benefit, access pattern privacy would be less of a concern as it were with centralized commercial storage providers.

Permacoin [9] has explored a few first steps in this direction. However, a number of details still need to be fleshed out. For example, while Permacoin focused on outsourcing static dataset such as the library of congress archives, it remains open to support dynamic dataset updates effectively. Moreover, the impacts on the incentives needs to be carefully evaluated. Most cryptocurrency miners contribute to the network with some expectation of monetary rewards. The extra investment of storage would influence the current dynamics of cryptocurrency mining activities. Would the extra burden discourage the miners to join the network (and by how much)? Additionally, existing proposals did not prevent selfish miners from ignoring data retrieval requests (even if they locally stored the corresponding data chunks) as the rewards are not associated with answering such utility queries.

3. Challenges

We anticipate several outstanding challenges towards the bright future of blockchain- and modern-cryptography-enabled secure infrastructure to support better Healthcare.

1. **Efficient, Trustworthy Implementations.** First of all, any endeavor in this trade relies heavily on slick software implementations of the novel ideas to summon up adoption and volunteering. In the administrator's perspective, since the whole system is designed to work autonomously given the initial implementation, it would bring up big concern on the trustworthiness of the implementations. Moreover, due to the use of some modern cryptographic constructions, extra care is needed to make sure the throughput of the implementation match with the practical requirements.

2. **Hassel-free System Updates.** The ever-changing requirements on complex real world applications stipulate its system design to accommodating future upgrades. However, this would be a largely overlooked design consideration in existing distributed cryptocurrency implementations. Novel system upgrading mechanisms are needed to enable a deployed blockchain network to evolve without abrupt discontinuation of their critical services.

3. **Administration and Legislation.** Educational efforts are essential to disseminate the notion of security along with its underlying assumptions to future adopters. We can imagine that a successful adoption of such system may require adequate initial investment to get the system rolling. Therefore, some administrative promotion will be tremendously helpful to actually deploy these systems.

4. Conclusion

We overviewed the opportunities and challenges in re-targeting blockchain technology to building trustworthy IT infrastructure and applications for better Healthcare. We discussed in more detail five prospective directions and anticipated three major obstacles along this line of work. Generally speaking, we believe this is a promising but underexplored area of research.

References

[1] A. Ben-Efraim, Y. Lindell, and E. Omri. "Optimizing Semi-Honest Secure Multiparty Computation for the Internet." In *Proceedings of the ACM SIGSAC Conference on Computer and Communications Security*. ACM. 2016.

[2] Eli Ben-Sasson, Alessandro Chiesa, Daniel Genkin, Eran Tromer, and Madars Virza. "SNARKs for C: Verifying Program Executions Succinctly and in Zero Knowledge." In *Advances in Cryptology–CRYPTO 2013*, 90–108. Springer. 2013.

[3] Eli Ben-Sasson, Alessandro Chiesa, Eran Tromer, and Madars Virza. "Succinct Non-Interactive Zero Knowledge for a von Neumann Architecture." In *23rd USENIX Security Symposium (USENIX Security 14)*, 781–796. 2014.

[4] Oded Goldreich, Silvio Micali, and Avi Wigderson. "How to Play Any Mental Game." In *Proceedings of the Nineteenth Annual ACM Symposium on Theory of Computing*, 218–229. ACM. 1987.

[5] Yan Huang, David Evans, Jonathan Katz, and Lior Malka. "Faster Secure Two-Party Computation Using Garbled Circuits." In *USENIX Security Symposium*, volume 201. 1. 2011.

[6] Yuval Ishai, Ranjit Kumaresan, Eyal Kushilevitz, and Anat Paskin-Cherniavsky. "Secure Computation with Minimal Interaction, Revisited." In *Annual Cryptology Conference*, 359–378. Springer. 2015.

[7] Yehuda Lindell, Benny Pinkas, Nigel P Smart, and Avishay Yanai. "Efficient Constant Round Multi-Party Computation Combining BMR and SPDZ." In *Annual Cryptology Conference*, 319–338. Springer. 2015.

[8] Chang Liu, Xiao Shaun Wang, Kartik Nayak, Yan Huang, and Elaine Shi. "Oblivm: A Programming Framework for Secure Computation." In *2015 IEEE Symposium on Security and Privacy*, 359–376. IEEE. 2015.

[9] Andrew Miller, Ari Juels, Elaine Shi, Bryan Parno, and Jonathan Katz. "Permacoin: Repurposing Bitcoin Work for Data Preservation." In *2014 IEEE Symposium on Security and Privacy*, 475–490. IEEE. 2014.

[10] Muhammad Naveed, Erman Ayday, Ellen W Clayton, Jacques Fellay, Carl A Gunter, Jean-Pierre Hubaux, Bradley A Malin, and XiaoFeng Wang. "Privacy in the Genomic Era." *ACM Computing Surveys (CSUR)* 48 (1). ACM: 6. 2015.

[11] Bryan Parno, Jon Howell, Craig Gentry, and Mariana Raykova. "Pinocchio: Nearly Practical Verifiable Computation." In *Security and Privacy (SP), 2013 IEEE Symposium on*, 238–252. IEEE. 2013.

[12] Ebrahim M Songhori, Siam U Hussain, Ahmad-Reza Sadeghi, Thomas Schneider, and Farinaz Koushanfar. "TinyGarble: Highly Compressed and Scalable Sequential Garbled Circuits." In *2015 IEEE Symposium on Security and Privacy*, 411–428. IEEE. 2015.

[13] Xiao Shaun Wang, Yan Huang, Yongan Zhao, Haixu Tang, XiaoFeng Wang, and Diyue Bu. "Efficient Genome-Wide, Privacy-Preserving Similar Patient Query Based on Private Edit Distance." In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, 492–503. ACM. 2015.

[14] Andrew Chi-Chih Yao. "How to Generate and Exchange Secrets." In *Foundations of Computer Science, 1986., 27th Annual Symposium on*, 162–167. IEEE. 1986.

[15] Samee Zahur, and David Evans. "Obliv-c: A Lightweight Compiler for Data-Oblivious Computation." In *Workshop on Applied Multi-Party Computation. Microsoft Research, Redmond*. 2014.