Using Blockchain to Improve Security of the Internet of Things

Joshua W. Quist
quist127@morris.umn.edu
Division of Science and Mathematics
University of Minnesota, Morris
Morris, Minnesota, USA

Abstract

The Internet of Things has increased in popularity in recent years, with daily life now being surrounded by "smart devices." This network of smart devices, such as thermostats, refrigerators, and even stationary bikes affords us convenience, but at a cost. Security measures are typically inferior on these devices; considering that they collect our data around the clock, this is a big reason for concern. Recent research shows that blockchain technology may be one way to address these security concerns. This paper discusses the Internet of Things and the current issues with how security is handled, discusses how blockchain can shore up some of these shortcomings, and goes in depth into examples of how blockchain has been implemented to improve the security of the Internet of Things.

Keywords

Blockchain, Internet of Things, Security, Data Privacy

1 Introduction

Imagine you have a job interview scheduled in the morning. You set your alarm for 6 A.M. on your fancy new smart alarm. Your alarm then connects to the internet, which gives it the ability to look at weather and traffic forecasts to adjust the time that it will wake you up. There is one slight problem, however; your device gets compromised by a software attack which wipes all the alarm data from it. This results in your alarm not going off, which leads to you missing your job interview. Even though this is a basic example, it is clear that it is essential for Internet of Things devices to have solid security.

The Internet of Things (IoT) is the name for the network of physical objects, or "things", which are equipped with sensors, software, and other technologies in order to communicate with other devices wirelessly. This gives these devices the ability to connect and share data between each other without requiring human interaction. These devices can range from a simple kitchen appliance that connects to the internet to show you the weather forecast, all the way to a car that is fully autonomous. The IoT is a giant network of these types of connected devices, all of which collect and share data about the way they are used and the environment around them.

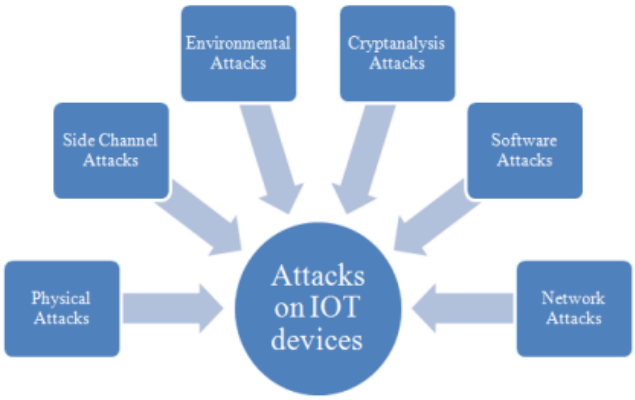

Figure 1. Potential attacks on the Internet of Things.

Although the IoT is very useful to us in everyday life, Harit et al. describe in their paper "Internet of Things Security: Challenges and Perspectives" [2] some of the issues that come with implementing this level of massively connected devices. One of these issues is potential attacks on IoT devices, shown in Figure 1. A few of these attacks which will be addressed in this paper are network attacks, cryptanalysis attacks, and software attacks. A short description of the three is as follows:

- Network attacks aim to collect information on a system in order to exploit vulnerabilities, which may result in unauthorized access to data in the system.
- Software attacks consist of overloading the system with data requests, which slow down the system as a whole and may even shut it down completely.
- Finally, cryptanalysis attacks attempt to break the data protections put in place in order to access data they are unauthorized to access.

An estimated 70% or more of the devices that make up the IoT are vulnerable to attacks such as these. [2]. This means that action must be taken to improve flaws in these devices. This paper describes two systems improving these vulnerabilities, both using blockchain technology.

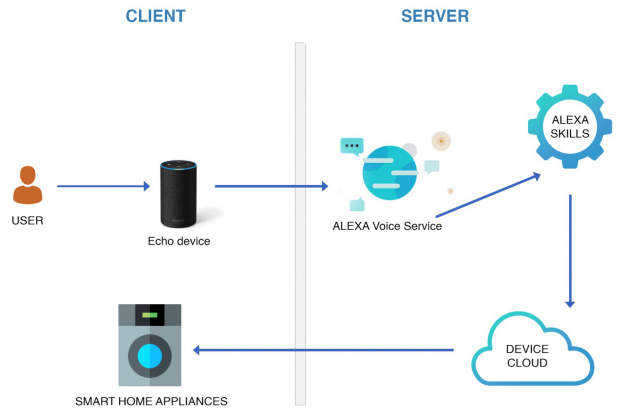

Figure 2. Amazon's Alexa, an example of IoT System design. [4]

2 Background

Before introducing the two examples of implementation, the following section will go more in depth on current IoT system architecture, the basics of blockchain and Bitcoin, and general ways that blockchain improves on current systems.

2.1 IoT System Architecture

The current structures of IoT systems are comprised of four parts: things, gateways, network infrastructure, and cloud infrastructure. These are described as follows:

1. Things are the individual physical objects in the system which communicate between each other.
2. Gateways are the go-between for "Things" and the cloud in order to provide connectivity and security.
3. Network infrastructure is the system of routers and gateways which control data flow.
4. Cloud infrastructure is made up of servers and storage methods which store data produced by the IoT devices.

These four parts work together to securely communicate between each other and store data in order to improve the quality of life of those using the system, such as the smart alarm in the example at the beginning of section 1.

An example of a system like this is used with Amazon's Echo device, as seen in Figure 2. The user speaks a command to the device, which is the Thing in this system. The Alexa device then communicates with the ALEXA voice service in order to process the request, which would be a gateway in this system. The ALEXA voice service then communicates with the Alexa Skills functionality which is connected to the smart home appliance. The Alexa Skills part of the system makes up the network infrastructure in this system. Finally, the data from the request is stored on Amazon's Cloud servers (which makes up the cloud infrastructure) and the request gets sent to the smart home appliance. This whole process would occur when, for example, a user would tell the Echo device "Echo, turn on my washing machine."

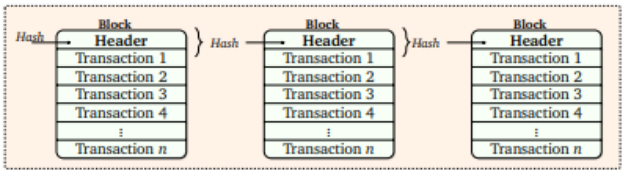

Figure 3. Image of blockchain blocks hashed together. [1]

2.2 Blockchain

A blockchain is a distributed and continuously growing list of records. "Distributed" means that no single entity controls the ledger, but instead that members in the blockchain work together to validate new records. These records are also known as "blocks" which are then chained together using hashing, thus the name blockchain. This system is shown in Figure 3.

2.3 Hashing

A "hash" is a mathematical function that converts an input of any length into an output of a fixed length. This fixed output length means that nothing is able to be determined about the input, such as its size or length. An input will always have the same output when entered into the hashing function. Hashes also can not be used to reverse engineer the original input, giving the hashing algorithm what is known as "One-way functionality."

For a simple example of this one-way functionality, say a friend of yours tells you to pick two random numbers between 1 and 10,000. In this example, assume you pick the numbers 5200 and 300. They then ask you what those two numbers add up to; your answer to them is 5500. The friend now has the task of figuring out what your two original numbers were. Even though they know the input has to be two numbers between 1 and 10,000, your hashing method of adding them together has made that an incredibly difficult task. It then becomes easy to imagine that a hashing algorithm with multiple inputs and much larger numbers can become difficult enough to the point of not being able to be reverse engineered. Because of this difficulty, hashing helps make storing data in a blockchain a secure solution.

For each new record in a blockchain, the hash is calculated by processing data from the previous block's header. This header contains various information such as version numbers, timestamps, the hash of the previous block, and the hash of the root block. Since these unique numbers are used to calculate the hash, no two hashes are the same. The root block is the first entry in the blockchain, which contains the hash of every transaction on the blockchain; this is used to authenticate transactions in the blockchain.

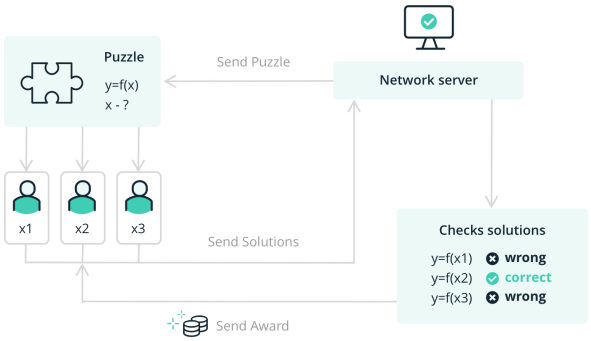

Figure 4. Illustration of Bitcoin's "proof of work." [3]

2.4 Bitcoin

The most successful example of a blockchain implementation is Bitcoin. Bitcoin is a decentralized digital currency built on blockchain. This type of decentralized digital currency is also known as a cryptocurrency. The Bitcoin blockchain records all transactions since the root block. Each Bitcoin transaction contains the sender, receiver, amount of the transferred currency, and the sender's public key, or signature. This public key is unique to each user and is used to anonymously identify the user in their transaction. Since the system has no central authority, each transaction needs to be authenticated by other participants of the blockchain. This process is known as "peer-to-peer authentication." To incentivize participants to authenticate transactions and suggest new blocks, a percentage of a bitcoin is offered to participants called "miners" who solve a puzzle known as a "proof of work."

2.5 Proof of Work

A diagram of proof of work is shown in Figure 4. Proof of work in the Bitcoin system is a system where a the network transmits a mathematical puzzle to network of miners. The answer to this puzzle is the hash of the next block in the blockchain, which ensures that the blockchain will continue to grow to accommodate new transactions. The actual puzzle itself is to find a 64-digit hexadecimal number, which when put through the SHA256 hashing algorithm, is less than or equal to the hash originally produced for the puzzle by the Bitcoin network.

A new proof of work puzzle is transmitted to the miners every ten minutes, and the miner who solves the puzzle is rewarded in a set amount of the bitcoin currency. When a miner finds the solution, it is communicated to the network so the other miners can verify the answer and consider it as the next block in the blockchain. When several verified solutions are suggested simultaneously, miners randomly select the next block.

This system benefits both the miners and the Bitcoin blockchain itself; all transactions are verified without the use of a central entity having control over the system, and miners are rewarded for their work in keeping the system decentralized.

Since there is a large amount of miners working towards solving the proof of work puzzles, this ensures that no single miner will control the entire network. If a single miner did have control over the network, they could potentially manipulate transactions, which would reduce the security and reliability of Bitcoin itself. Proof of work also ensures that each transaction is verified before being added to the blockchain.

Because of Bitcoin's tight security measures, decentralization, and peer-to-peer validation, Shafagh et al. propose an IoT system built on Bitcoin which addresses security concerns from Section 1. This system will be described in section 3.

2.6 Current Methods and Blockchain Potential

Current IoT systems use centralized models, mostly using the system architecture described in the beginning of section 2. Large cloud servers store the data permissions of IoT devices, and handle the data processing and storage side of the system. Communications happen exclusively through the internet, even if devices are only a few feet apart. These systems are expensive due to the cost of server farms, including infrastructure and maintenance expenses. All of these systems outsource data management to a single entity, leaving individual users with no control over how their data is used.

A decentralized system utilizing blockchain would solve many of these issues. Using peer-to-peer authentication similar to the one used in Bitcoin's blockchain would reduce the processing power needed to maintain the system. It would also decrease the need for large server farms to authenticate, process and store data. The decentralization of this system would also prevent a single failure in the network from bringing the entire process to a halt.

Some concerns of implementing systems with blockchain technologies are the processing power of IoT devices, and storage limitations of data. Devices in the IoT are currently not made with blockchain implementations in mind, which could make it difficult for older devices to make the transition over to a blockchain system. While keeping these concerns in mind, multiple IoT systems using blockchain have been implemented and tested.

3 A Blockchain System for Access Control

One example of blockchain being implemented in an IoT system is discussed by Shafagh et. al in "Towards Blockchain-based Auditable Storage and Sharing of Data" [5].

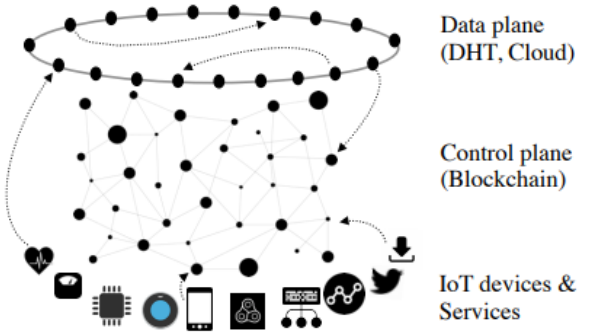

Figure 5. An illustration demonstrating the design of this system. [5]

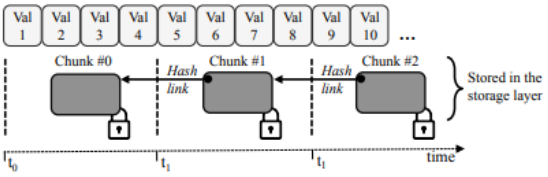

Figure 6. An illustration demonstrating the method of "chunking." [5]

3.1 Purpose

In their paper, the authors talk about the current architecture of Internet of Things data storage. These methods of data storage result in systems "where users have limited control over their data and how it is used" [5]. These systems often end up with companies using third parties to store data, which results in at least two degrees of separation between users and their data. The researchers also mention that current implementations "fall short in addressing security during the life-cycle of data." [5] Users are then forced to blindly trust that the companies with control over their data will responsibly handle and secure their data. With many mentions in the news of data leaks and mismanagement of user data, this results in users being untrusting of IoT devices.

To address these concerns, Shafagh et al. propose in their paper a "blockchain based auditable data-management system for IoT data" , which includes secure data sharing, access revocation, efficient data streams, and a distributed storage layer [5]. This system will provide users more control over their own data.

3.2 System Design

The system proposed by the authors, as illustrated in Figure 5 consists of three main parts: the IoT devices, the control plane and the data plane. The control plane's responsibility is to manage who has access to the system, and the specific permissions of those with access. The control plane manages access by implementing a blockchain without a central trusted entity, which stores access rights. When a data request is received, the storage checks the blockchain for the access rights of the request in the blockchain, and sends the data when the transaction is authenticated.

The data plane's role is to store all data records and their permissions information. This is accomplished in this system by splitting each individual record into multiple sections, or "chunks", which are then cryptographically chained together. They are chained together because each chunk is hashed together as seen in Figure 6. This is effectively the same setup as a blockchain, using hashing to make every chunk point to the next one in the storage system. Each chunk keeps track of who has access to it, which is then checked against the access privileges of any data request. Data storage is able to be managed in multiple ways in this type of a system, but the authors specifically mention that "on-premise storage" (such as a server set up on a user's personal device) and "storage on cloud services" (such as cloud storage on Amazon Web Services) are "compatible with our system" [5].

3.3 Data Encryption

This system also uses techniques to keep data inside the data plane secure. One of these techniques used is AES-GCM encryption during the data chunking process. Encryption is the process of converting data into an irreversible output, similar to how hashing worked as described in section 2.3. However, unlike hashing, this encryption process is able to be reversed by use of a key K. This key is then given to any device authenticated in the IoT system, meaning that only authenticated devices are able to access data in the system.

3.4 Real Life Example

Based on the potential attacks described in section 1, there are two examples of attacks that this system would help prevent: software attacks and cryptanalysis attacks. We can get a good idea of what this system achieves when we look at this through the lens of our smart alarm from earlier in this paper.

As you may recall from earlier in the paper, software attacks are defined as attacks which consist of overloading the system with data requests. If we had our alarm clock set up using the system design, these data requests would be made from a device which was not in the blockchain as an authenticated device. This means that none of the requests would make it through to our data plane, which in theory should not slow down the effectiveness of our alarm; this nullifies the attempted "overload" of the system.

The effectiveness of cryptanalysis attacks are also reduced in this system due to the chunking method of storing data. Cryptanalysis attacks were defined earlier as attacks which

attempt to break the data protections put in place in order to access or modify the data in the system. Say our alarm clock stores data on when you typically set it for. Even if a bad actor managed to get access to the data plane in this system, they would decipher a small fraction of the data in a single chunk due to the encryption in place in the system. This information would most likely be useless; if the data stored was "User xyz average alarm time is 8:00 AM" then the information in a single chunk could be something like "ge ala."

3.5 System Evaluation

Shafagh et al. performed a security analysis and an initial evaluation of performance numbers. In the security analysis, the "chunking" method of storing the data is talked about, which ensures that even if one chunk is exposed, the whole data entry is not revealed, which was demonstrated in the real life sample above.

The performance evaluation, which utilized Amazon's cloud storage with this system's access control, showed 10% slowdown in data requests. The researchers mentioned that that this number could be improved with more extensive future work. The researchers also mention that they are "currently in the process of finalizing a complete reference implementation of our system and building several IoT applications on top of it." [5]

4 A Blockchain System for Behavior Monitoring

A second example of blockchain being used in an IoT system is talked about in Ali et al.'s "Blockchain-based Smart-IoT Trust Zone Measurement Architecture." [1]

4.1 Purpose

In their paper, the researchers mention that "In IoT the things (devices) communicate and exchange the data without the act of human intervention. Such autonomy and proliferation of IoT ecosystem make the devices more vulnerable to attacks." [1] The researchers then go on to say that this increases the vulnerability to attacks because after a device is authorized within the IoT system, there is no way to detect if it starts behaving maliciously. To help reduce this vulnerability, the authors suggest in their paper a method of monitoring traffic from individual devices in the IoT system in order to detect whether or not devices are behaving maliciously.

4.2 System Design

The proposed architecture of this system "adds a layer of security for behavior monitoring of various IoT-zones in a blockchain setup," with a "zone" in this example meaning a self contained IoT system operating on a local blockchain [1]. A model of a single system using this design is shown in Figure 7. In the system, the device with the most processing

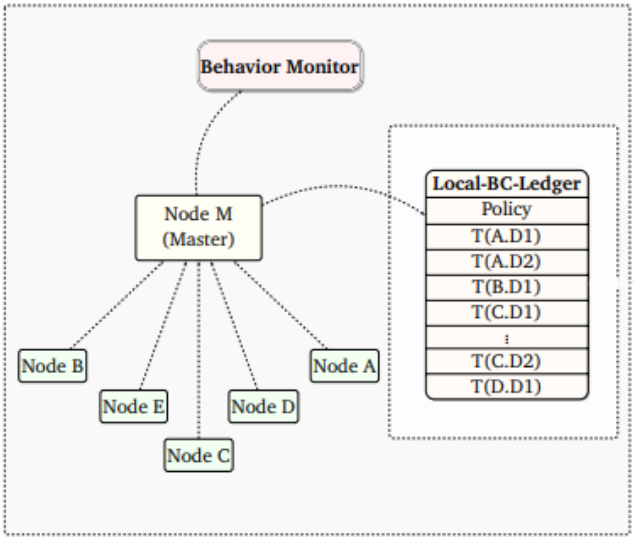

Figure 7. An illustration showing the proposed behavior monitor architecture. [1]

power is designated as the main node. All other devices in that system are designated as a follower. When a device is added to the system, they must send a transaction containing a unique "groupID" of that system to the main node. When the ID of the device is verified, that device's ID is stored in the local blockchain, authenticating it to make transactions in the future. The local blockchain contains the hashes of all transactions generated by the IoT devices in the system. For each communication between IoT devices, a transaction is created and added to the blockchain.

4.3 Behavior Monitor

The researchers then implement a behavior monitor utilizing machine learning in order to monitor devices which have already been authenticated. Machine learning is the process of using computer algorithms to build a model based on data to make predictions. The system proposed by the researchers uses an auto encoder algorithm to detect whether or not data coming into the system is malicious. This auto encoder is trained to detect anomalies, which in this case would be detecting malicious data.

When a device communicates with another part of the system, that data is added to the local blockchain. This data is consistently monitored by the auto encoder, which has been trained on both harmless and malicious data. When it observes data that is potentially malicious, it triggers an alert for the system administrator that a bad actor could potentially be in the system. This then means that the administrator would be able to evaluate the data and decide whether or not to remove that device from the system.

4.4 Real Life Example

The final type of attack mentioned in section 1 that has not been addressed yet in this paper is a network attack. Network attacks were defined as attacks that aim to collect information on a system in order to exploit vulnerabilities, which may result in unauthorized access to data in the system. If someone was using devices to try and collect information on the system described in this section, that is a type of malicious behavior that would be flagged by the behavior monitor.

4.5 System Evaluation

The researchers performed an evaluation of this system, training their algorithm using data from three different IoT devices: a thermostat, a webcam, and a security camera. They then used a "mirai" attack as an evaluation of their system. A mirai attack is an attack where default username and password combinations are used to attempt to log in to IoT devices, and when accessed, direct their traffic to a single source in order to take down servers in a DDoS (Distributed Denial of Service) type attack. This type of attack falls under the category of a software attack, which is one of the attacks described in section 1. For the evaluation, their system using the auto encoder algorithm was compared to results from three other types of algorithms commonly used for anomaly detection: Support vector machine, isolation forest, and local outlier factor.

1. Auto encoder: The algorithm used in the researchers' system. This model uses compression of data to reduce "noise" in the data, and use the result for anomaly detection. In this case, an anomaly would be malicious data.
2. Support vector machine: Learning model used to separate data using a linear divider. In two dimensional data like this data (Either malicious or harmless) this would be used to detect when a data point would fall on the "malicious" side of the data.
3. Isolation forest: Model which randomly selects a data point, then randomly selects a value between the maximum and minimum value of that data's category in order to detect outliers. This would be used to detect malicious data, which would categorize as an outlier.
4. Local outlier factor: Model which measures the standard deviation of a data point with that of its neighbors in order to detect outliers. Like isolation forest, this would be used to detect potential malicious data.

In their tests, the researchers' proposed model had a faster detection time than the other three models, while also having a similar true positive/false positive rate.

The researchers described the work in their paper as a "preliminary step" towards classifying IoT devices as malicious or benign. More work is needed to be done analyzing different methods of attack, and implementing their design in a more large-scale way.

5 Conclusion

This paper discusses how blockchain technology can be used to improve the security of IoT systems. The current shortcomings of IoT systems are discussed with blockchain being discussed as a solution to those shortcomings. Two different systems are explored; one with the purpose of improving storage methods and access control to IoT data, and the other with improving monitoring of current members of the system in order to discover devices with bad intentions. More real life implementations and tests are needed to improve these systems using blockchain, which would potentially surpass current systems in terms of data ownership, security, speed, and decentralization.

Acknowledgements

Special thanks to Kristin Lamberty and Elena Machkasova for the feedback and advising throughout this research and writing process. I would also like to thank Morris alum Melissa Helgeson for their helpful feedback on the paper.

References

[1] J. Ali, T. Ali, Y. Alsaawy, A. S. Khalid, and S. Musa. Blockchain-based smart-iot trust zone measurement architecture. In *Proceedings of the International Conference on Omni-Layer Intelligent Systems*, COINS '19, page 152–157, New York, NY, USA, 2019. Association for Computing Machinery.

[2] A. Harit, A. Ezzati, and R. Elharti. Internet of things security: Challenges and perspectives. In *Proceedings of the Second International Conference on Internet of Things, Data and Cloud Computing*, ICC '17, New York, NY, USA, 2017. Association for Computing Machinery.

[3] Ledger. What is proof-of-work, 2019. https://www.ledger.com/academy/blockchain/what-is-proof-of-work, Accessed on 2022-4-14.

[4] QSS Technosoft. How does amazon alexa work?, 2019. https://www.qsstechnosoft.com/how-does-amazon-alexa-work, Accessed on 2022-4-10.

[5] H. Shafagh, L. Burkhalter, A. Hithnawi, and S. Duquennoy. Towards blockchain-based auditable storage and sharing of iot data. In *Proceedings of the 2017 on Cloud Computing Security Workshop*, CCSW '17, page 45–50, New York, NY, USA, 2017. Association for Computing Machinery.