Blockchains and Bureaucrats:

Discovering Properties of the Workflow of Elections Based on Blockchain Models

HACS287 Final Research Paper

Students:
Noah Burkhardt
Courtland Climer
Corey Doyen
Jeffrey Wong

Supervisor:
James Purtilo

Abstract

The focus of this research paper is to discover various properties of the workflow of elections based on blockchain models. The motivation for this research arrives from the merger of the relatively new and popular blockchain technology based on Satoshi Nakamoto's "Bitcoin: A Peer-to-Peer Electronic Cash System" and the plethora of security issues seen today in the American election system. The research problems investigated are: What are some of the various security advantages and disadvantages inherent to some models of elections based on their blockchain configurations? What are the basic requirements a researcher should bake into their blockchain-based election when designing such a system for testing? What are the reasonable dimensions for this space to be evaluated? We use our prototypes to elaborate upon some of the issues within blockchain-based voting systems and suggest possibilities for resolving portions of these difficulties. We explore the following issues surrounding the infrastructure of the prototypes: privacy, security, voter convenience, voter confidence, accessibility, verifiability of accurate vote count, and distribution methods for votes.

Through this design, the paper illustrates security trends across multiple blockchain-based election models. Ultimately, we conclude from this work that there are indeed valuable insights to gain from the rapid prototyping of blockchain-based voting systems, but these same prototypes also expose a plethora of security issues and bureaucratic design pitfalls that present themselves with a greater force and vitality than in the election models of the status quo. This paper's simple two-model study shows a number of issues to be addressed as a product of using blockchain-based voting, and future models like this shall reveal even more. While this quickly-popularized technology has taken the stage in many fields of academia and industry, it appears as though more thought and concerted effort must be applied before it is ready to be used as a standard of security in the election market. Our aspiration is that other researchers will be able to utilize the findings from this paper to understand issues around organizational structures for a blockchain-based voting system, so that they may quickly recognize new issues and approaches.

Introduction

-Background-

In the current political climate, with evidence of tampering in our last presidential election by foreign actors, many voters and security experts lack confidence in our current election infrastructure. The current voting infrastructure of the United States differs not only between different states but even between different voting districts (Good). With this decentralization comes a large disparity in security, efficiency and reliability between different districts, with some maintaining effective infrastructure, and others facing vulnerabilities to very real and present security threats. The emerging technology of blockchain has been hailed by some as a possible solution to improve or replace existing election infrastructure. Cited often as a central point in favor of blockchain is its inherent public ledger, which allows users to ensure that their transactions, votes in this case, are cast accurately. This public ledger also uses hashed

identifiers for users to ensure anonymity. With this anonymity of users combined with easy, public verification of results, blockchain is purported by some to open the gateway to secure, mobile voting leveraging the omnipresence of the Internet and the ubiquity of mobile devices in the United States.

Many security experts feel that these features of blockchain, while beneficial, do not ameliorate severe issues involved in actual implementation of an integrated blockchain election infrastructure. According to Timothy Lee of Ars Technica, there are "many ways that foreign governments could compromise an online vote without breaking the core cryptographic algorithms" that comprise blockchain (Lee). A large part of these possible methods exploit the mobile voting platform: the smartphone or personal computer. Through phishing attacks or exploiting vulnerabilities in a user's system, malicious actors can "trick them into revealing their voting credentials -- or simply trick them into thinking they've cast a vote when they haven't" (Lee). There have been several proposed solutions to these perceived vulnerabilities in a blockchain-based voting system, among them the solution of a "post-election paper audit" where paper copies of each vote cast can be stored and counted as needed to ensure an accurate result if the blockchain is somehow compromised and individual votes are either missing or deemed altered (Schwartz). The concerns expressed by many of these experts center around their belief that blockchain does not address the fundamental security issues of elections, or even proves to be harmful, adding additional vulnerability and complexity to a system that is already disorganized and proven to be lacking in necessary security.

-Our Study-

The digital vulnerabilities of America's elections and a lack of standard design patterns for blockchain-based election models led us to conclude that further work is needed to detail some of the properties of different blockchain models. We arrived at these questions from gaps found in literature: What are some of the various security advantages and disadvantages inherent to some models of elections based on their blockchain configurations? What are the basic requirements a researcher should bake into their blockchain-based election when designing such a system for testing? What are the reasonable dimensions for this space to be evaluated? We discuss advantages/disadvantages relative to the following categories: privacy, security, convenience, confidence, accessibility, accuracy of vote count, and distribution of votes. While there are likely more issues to consider, we found these to be the most pertinent issues today, and discuss other potential issues for future researchers to explore with different resources and systems. The contribution of this work is to outline properties of these systems and discuss their implications in the realm of elections. This would be useful for other researchers specializing in studying blockchain-based election models to save time in the design of their models. The structure of this paper is as follows. First, we discuss the methods used in the construction of the general blockchain apparatus and the two prototype models, then mention the various limitations of the project, and afterwards describe the two models in terms of their advantages and disadvantages relative to various design issues. The paper ends with an impact calculus of these advantages and disadvantages and explains directions for future research.

Methodology

-Methods-

Our primary method of research for this experiment was to design a prototype that would allow us to determine the advantages and disadvantages of a blockchain-based voting system. The blockchain itself was baked into the backend server. This server was written in Python, running on a Flask backend. The server frontend was written primarily in Javascript, specifically running on an NodeJS server running VueJS. This allowed us to serve up the voting system as a website to users and testers. Our intention was that we would build out this tool as a means for testing its viability in a real life voting setup. Our blockchain implementation itself was not a typical one. In a standard blockchain model, there would be several different blockchain implementations, all simultaneously running on their own servers. As one blockchain has a transaction added to it, it would then propagate that information to the other blockchains. This is where some of the security and risk comes into play in dealing with this data structure. However, our blockchain was not designed in this way. Our blockchain only ran a single instance on the backend server. Using this design, many of the security aspects of the blockchain data structure remained. Our ultimate goal for this prototype was to be used in a mock election. Users (mock voters in this scenario) logged in with provided credentials on 'election day'. They were then given a representative coin by the voting administrator (a member of our team). This coin stood in place for a single transaction in a block on the chain, and allowed the voter to cast a single vote into the system. The chain then validated that the user indeed possessed the coin needed to vote before it added the transaction to the chain. Once the transaction was then added to the chain, the system mined for another coin. Each block on the chain held a vote transaction and a mining transaction. The model that is described above serves as the foundation for the two sub-models described below.

We used two different models of the product in determining conclusions for this experiment. We shall call them Model A and Model B. We shall now briefly describe the two models and their capabilities. Model A is a model that runs offline. It runs a server locally, and serves up the frontend website locally via a localhost server. It is not connected to the Internet; it is strictly local. The blockchain itself is stored locally on the system, and can be accessed via API calls to the backend. This model functions closer to the current system of voting. Registration is performed through the voting office, and at the time of registration account confirmation, the coin is distributed to the voter. This is to protect the coin and minimize the time that it is in the voter's possession. Model A performs better at a smaller scale, likely to be used at a single polling location instead of nationally. For this model, we implemented the ability to view the blockchain as it stands, so that the voter may confirm that their vote has been successfully cast. It shows their blockchain address (a hash of the user's ID that has been converted to Base58) and the blockchain address of the candidate that they cast their vote for. The advantage here is an increase in voter confidence, as the voter is able to not only see that their vote has been cast successfully, but also that the hashes do indeed match. Possible disadvantages in this scenario include a small reduction in security and user privacy.

Model B runs online. The primary difference with respect to Model A is that this model can be accessed anywhere via the Internet. Model B runs on a server, and serves

up a website via the Internet. Again, the blockchain is stored locally on the server, and can only be changed via API calls to the backend. You have the ability to register to vote, and then subsequently vote, entirely via the online portal. Via this model, the coin mentioned above would be distributed to the account at the time of account creation, instead of account activation. Model B also has the advantage of being very easy to scale. A disadvantage to this model, one that we found was necessary to implement, is the removal of the ability to view the chain and verify your vote. Leaving the chain open to the public runs the risk of being able to brute-force the hashes and expose private voter information to the public. The decision to leave this feature out increases both the security and the privacy of this model. However, as a direct result, voter confidence is decreased significantly.

-Limitations-

As to be expected, our methodology has several limitations. These are the qualitative approach of the paper, scalability of the project, time, and the choice to use a customized blockchain over that of Ethereum or otherwise. Since many issues we observed were not practically measurable in the scope of our study, especially those of privacy and convenience, we opted for a qualitative approach for the paper. This limits us from providing results with a particular confidence interval and adds a layer of subjectivity to the study. It also creates a challenge for data collection and survey design. We worked around this limitation by limiting our results to clearly observable occurrences in the testing of our models. Our prototypes were both deployed at a very small scale and only handled a single vote for each voter in one location. In a real election, there would be far more votes for far more positions and the voting would take place across a large number of machines across an even larger number of polling places. We accept this as a limitation, but we believe that we found a sufficient number of issues to discuss even with our limited model.

Another limitation was that we only had a semester to complete the project. This meant that we lacked the time necessary to build out further models for testing and working with a great number of possible configurations. Fortunately, this limitation allowed us to hone in on our discussion of two models and explore each to a greater depth with regard to a greater number of design and implementation issues. Since we utilized a more customized blockchain, there may be a significant amount of ignored security issues present in both models. This means that when building a practical application, and not building a system for testing purposes like our apparatus, one should be very careful to follow common cybersecurity practices like using an appropriately secure hashing function and ensuring the security of their networks. Future studies with more resources could easily expand on this research by focusing on the scalability of the project and observing the impacts of a more scalable design on security and vote distribution. Similarly, even though we chose to limit our tool suite to that described above, we find it important for future researchers to test models using more developed and professional technologies such as Hyperledger Composer. Not every blockchain model should be built from scratch every time, for both efficiency and developer sanity, and perhaps with other interests and insights, more standardized systems would be easier to implement while achieving similar desired effects.

Results

Due to the small-scale nature of this project, most of the results are qualitative rather than quantitative. In order to obtain more quantitative results, a larger scale edition of this research paper would be required. Our first discovery of interest is in the blockchain itself. Initially, as we set up the development environments for the team members, we discovered that many current blockchain implementations, such as Hyperledger Composer, did not fit into our model for this blockchain voting system. Specifically, we ran into several issues related to software compatibility with the operating system. After discussing this issue for several days as a team, we concluded that we would design our own blockchain implementation. This allowed us to design the blockchain down to the very minute details, so that our prototype would run more efficiently.

A primary example of these changes comes in the form of the level of centralization of our blockchain. Traditionally, blockchains are incredibly decentralized. Our application is far more centralized, seeing as it only runs a single instance of the blockchain. This is due to our blockchain being designed for election services. The elections covered by the scope of our project require a central authority. This is necessary to acknowledge voter registration. Additionally, this assists in designating vote administration, meaning the determination of the appropriate period during which the electorate can cast their votes. These requirements for a designated central authority to ultimately determine an election's validity are not optimal with more traditional blockchain implementations.

Going further, by comparing our two models, we detail more specifics involving the trade-offs of varying blockchain implementations. The most obvious difference between the two models is that Model A is offline whereas Model B is online. This means that to vote in an election using Model A, voters would have to travel to a polling station, similar to conventional voting techniques. By comparison, Model B would allow voters to vote by computer, from another location, such as their homes. There is research that "voters tend to be more confident when [...] they vote in person rather than by mail" ("Voter Confidence"). This would indicate that voters would also prefer voting in person to voting online, and as such Model A's in-person requirement would increase voter confidence. Furthermore, since Model A's implementation more closely matches conventional elections, the infrastructure for its use in federal and state elections is better established when compared to the cost of implementing Model B. On the other hand, Model B is far more optimized for use on a large scale, such as a national election. As it is online, Model B provides a standardized system that could easily be implemented to provide identical voting experiences to voters across the country.

However, the online system could be a double-edged sword. Many election security experts believe that the current, decentralized election system provides an increase in security over a centralized system (Good). Steve Povolny, the head of advanced threat research at McAfee, is quoted saying that the "complexity [of having different systems] fundamentally makes it more challenging to use a single vulnerability or chain of vulnerabilities to exploit a large number of systems simultaneously" (Good).

This raises valid concerns about using a single, online system across the country. In the case that a vulnerability happens to be found in the blockchain, it puts the entire country's election at risk. There is a similar risk for Model A since it is the same implementation at each polling station and a vulnerability found in the system could put numerous locations at risk. However, since there are multiple instances, it would require attackers at each location in person to compromise the system. Overall, this could result in a decrease in both security and voter confidence in Model B compared to Model A.

With regard to the accessibility and convenience of the different models, there are also trade-offs. Model A's voting procedure is most similar to the current election process. This means that the model and the current process share the same issues relating to accessibility and convenience. These include the requirement of traveling to a polling station and the difficulties this entails for some voters, particularly those who have economic constraints (Jacobs). In these cases, Model B is potentially a more convenient system. As voting occurs online, the ability to vote without needing to travel to a polling station would be preferable to many. The online system would be faster, provided voters own or have easy access to a computer. Since Model B requires access to a computer to register and cast a vote, people lacking computer access would have greater difficulty in casting their vote.

Related to the accessibility of voting is the registration process of our models. In Model B, a voter would create their own account online prior to voting. This would have to be reviewed by an official, who would determine that the account belongs to a valid, registered voter and give the account a coin to cast a vote. In contrast, Model A requires voters to register to vote similarly to the current voting system. After they do so, an account in the system would be created for them with a single coin. The credentials to access this account would not be available to the voter until they arrive at the polling place on election day. This raises an important issue; Model B requires voters to be able to access a computer before they cast their vote, so that they can make their account. This would have to occur at a different time than voting as they would need to allow for a verification period. This only compounds the necessity of voters to have regular access to computers during the period surrounding voting. Model A does not have this issue as the account creation is not on the voter end, meaning compared to conventional elections, there are no extra steps for the voter in regard to registration.

Another major consideration we had in our design of our models was privacy versus confidence. The visual representation of the chain feature that was part of Model A directly addresses the confidence of voters. The ability to verify one's own vote in the blockchain increases confidence in the security of the election model. Alternatively, having a method of viewing how someone voted, even only one's own vote, creates a privacy risk. It increases the potential that if the system is compromised, individuals could learn how others voted and use said knowledge for malicious actions. Model B does not have this feature and, as such, it cannot be used to create a vulnerability in the system. Voters also lose the ability to see their cast ballot, thereby decreasing confidence in the election model. This potential risk demonstrates the trade-off between privacy and confidence in a blockchain-based voting system.

The qualitative results from our project serve as an examination of the various trade-offs between different implementations of a blockchain-based election application.

Some of these trade-offs include features of an online system versus an offline one, accessibility and convenience, and privacy versus confidence. These discussions provide an overview of how differing systems and features can affect the traits and effectiveness of an election.

Conclusion

When organizing an election, one must consider how the specific circumstances impact how the election should be designed. The two different models we developed provide some insight into the options that one would have to consider.

One of the first considerations for running an election should be the scale. The number of individuals who form the electorate and how spread out they are over an area can greatly affect the optimal system for the election. Model B is the more scalable of the two models. As it is run online, it can be handle any number of expected voters without requiring additional resources beyond computer servers. Compared to Model A, Model B has no concern about wait times. Furthermore, Model B can be implemented across any distance whereas Model A requires physical polling locations and computers to vote with. This means that if an election is being held with a small population, Model A may be a better system, but with a large electorate, Model B has the distinct advantage. Also, in the case the electorate is spread out over a large area, Model B has the advantage that it would not require a large number of polling locations or force people to travel far distances to vote.

Similar to scale is accessibility. Both models present some hindrances to voting for individuals with economic constraints as mentioned earlier, so it is valuable to consider which constraints are more important in a given election. Model A is similar to conventional systems, so similar solutions that are currently being discussed to existing accessibility issues could be baked into Model A. On the other hand, Model B generates a voter dependency upon having reliable access to a computer both before and during the voting period. These potential obstacles for voters casting their votes are important considerations. For example, if a number of the voters do not have access to computers, then Model B would make it incredibly difficult for those voters to cast their votes and one should favor Model A.

Another major deliberation is how confident voters are that the votes they cast in an election are fairly and accurately counted. In regard to this, both models have benefits and costs. As such, more research into how different variables affect voter confidence is required in order to draw further conclusions. Expanding upon confidence in voting systems, is the potential risk of administrative incompetence and intentional misuse. Further experimentation is necessary to be able to address the risks involved with the blockchain from the administrative side of the models. In our models, privacy is directly tied to confidence. Going forward, if one is trying to maximize user privacy at all costs, they should opt for using Model B. On the contrary, if one is trying to maximize voter confidence, they should favor Model A.

Even with these conclusions, a significant amount of work remains for a glut of issues in the designs. These include, but are not limited to, efficiency of consensus algorithms used in the implementation, choice of consensus algorithm used in the

implementation, voter confidence based on extreme edge cases such as absentee ballots for an implementation like Model A, administrative incompetence, and administrative intentional misuse. Our models were both based on the infamous Proof-of-Work consensus algorithm, which has proven to carry several vulnerabilities including the dangerous and costly 51% attack. Other researchers would likely benefit from testing a variety of consensus algorithms, perhaps with a particular focus to more centralized algorithms like Delegated Proof-of-Stake. Measurements of the efficiencies of these algorithms could follow. As well, little work was done on analyzing the administrative end of the project. There are certainly methods of testing these issues using third parties, which should be deeply considered by future researchers. Finally, many news articles detail countless edge cases in our current election system and a deeper study into these peripheral problems would likely provide further insight into how to increase voter confidence.

We provide these conclusions with the hope that future academics will use them to determine the feasibility of a blockchain-based voting system for our modern society, and do so in a more organized and standardized way.

Bibliography

Bagley, Judd. "What Is Blockchain Technology? A Step-by-Step Guide For Beginners." *Blockgeeks*, 1 Jan. 1968, blockgeeks.com/guides/what-is-blockchain-technology/.

Bashir, Imran. *Mastering Blockchain: Distributed Ledger Technology, Decentralization, and Smart Contracts Explained*. Packt, 2018.

"Blockchain Voting and Its Effects on Election Transparency and Voter Confidence." *Communications of the ACM*, ACM, dl.acm.org/citation.cfm?id=3085263.

Drescher, Daniel. *Blockchain Basics: a Non-Technical Introduction in 25 Steps*. Apress, 2017.

Good, Chris. "When It Comes to Election Cybersecurity, Decentralized System Is Viewed as Both Blessing and Curse." *ABC News*, ABC News Network, 31 Oct. 2018, abcnews.go.com/Politics/election-cybersecurity-decentralized-system-viewed-blessing-curse/story?id=58877082.

Heemaiah, Kariappa. *The Blockchain Alternative: Rethinking Macroeconomic Policy and Economic Theory*. Apress, 2017.

Hillar, Gaston C. *Building RESTFUL Python Web Services*. PACKT Publishing Limited, 2016.

Jacobs, Tom. "How Polling Places Can Affect Your Vote." *Pacific Standard*, Pacific Standard, 19 Aug. 2010, psmag.com/news/how-polling-places-can-affect-your-vote-20318.

Kelly, Makena. "Nearly 150 West Virginians Voted with a Mobile Blockchain App." *The Verge*, The Verge, 10 Nov. 2018, www.theverge.com/2018/11/10/18080518/blockchain-voting-mobile-app-west-virginia-voatz.

Lee, Timothy B., et al. "Blockchain-Based Elections Would Be a Disaster for Democracy." *Ars Technica*, Ars Technica, 6 Nov. 2018, arstechnica.com/tech-policy/2018/11/blockchain-based-elections-would-be-a-disaster-for-democracy/.

Malone, Clare. "What We Know And Don't Know About Election Hacking." *FiveThirtyEight*, FiveThirtyEight, 10 Apr. 2018, fivethirtyeight.com/features/what-we-know-and-dont-know-about-election-hacking/.

Matishak, Martin, et al. "What We Know about Russia's Election Hacking." *About Us*, POLITICO, 19 July 2018, www.politico.com/story/2018/07/18/russia-election-hacking-trump-putin-698087.

Marvin, Rob. "Under Attack: How Election Hacking Threatens the Midterms." *PCMAG*, PCMAG.COM, 29 Oct. 2018, www.pcmag.com/feature/364358/under-attack-how-election-hacking-threatens-the-midterms.

Osgood, Ryan. *The Future of Democracy: Blockchain Voting*. 14 Dec. 2016, www.cs.tufts.edu/comp/116/archive/fall2016/rosgood.pdf.

Prusty, Narayan. *Building Blockchain Projects: Develop Real-Time Practical DApps Using Ethereum and JavaScript*. Packt Publishing, 2017.

Qureshi, Haseeb. "The Authoritative Guide to Blockchain Development – FreeCodeCamp." *FreeCodeCamp*, FreeCodeCamp, 28 Jan. 2018, medium.freecodecamp.org/the-authoritative-guide-to-blockchain-development-855ab65b58bc.

Regan, Michael D. "An 11-Year-Old Changed Election Results on a Replica Florida State Website in under 10 Minutes." *PBS*, Public Broadcasting Service, 12 Aug. 2018, www.pbs.org/newshour/nation/an-11-year-old-changed-election-results-on-a-replica-florida-state-website-in-under-10-minutes.

"Russian Hacks on U.S. Voting System Wider Than Previously Known." *Bloomberg.com*, Bloomberg, www.bloomberg.com/news/articles/2017-06-13/russian-breach-of-39-states-threatens-future-u-s-elections.

Saltman, Roy G.. "Accuracy, Integrity and Security in Computerized Vote-Tallying." *Commun. ACM* 31 (1988): 1184-1191, 1218.

Schwartz, Jen. "The Vulnerabilities of Our Voting Machines." *Scientific American*, 1 Nov. 2018, www.scientificamerican.com/article/the-vulnerabilities-of-our-voting-machines/.

"Voter Confidence." MIT Election Lab, MIT, electionlab.mit.edu/research/voter-confidence.

Zaninotto, François. "The Blockchain Explained to Web Developers, Part 1: The Theory." *Marmelab.com*, 28 Apr. 2016, marmelab.com/blog/2016/04/28/blockchain-for-web-developers-the-theory.html.

Zetter, Kim. "We're Still Way Too Vulnerable to Election Hacking – Member Feature Stories – Medium." *Medium.com*, Medium, 1 Nov. 2018, medium.com/s/story/were-still-way-too-vulnerable-to-election-hacking-63817b9ff211.