SECURITY AND RELIABILITY OF RATIONAL PLAYERS IN DISTRIBUTED CONSENSUS

Kehelwala Gamaralalage Janani Hansika Kehelwala

189329L

Degree of Master of Science

Department of Computer Science and Engineering

University of Moratuwa
Sri Lanka

May 2021

SECURITY AND RELIABILITY OF RATIONAL PLAYERS IN DISTRIBUTED CONSENSUS

Kehelwala Gamaralalage Janani Hansika Kehelwala

189329L

Dissertation submitted in partial fulfillment of the requirements for the degree of MSc in Computer Science specializing in Security Engineering

Department of Computer Science and Engineering

University of Moratuwa

Sri Lanka

May 2021

DECLARATION

Candidate:

I declare that this is my own work and this dissertation does not incorporate, without acknowledgment, any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief, it does not contain any material previously published or written by another person except where the acknowledgment is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

\-   - - - - - - - - - - - - - - - - -

K. G. J. H. Kehelwala   Date

Supervisor:

The above candidate has carried out research for the Masters Dissertation under my supervision.

\-   - - - - - - - - - - - - - - - - -

Dr. C. D. Gamage   Date

ABSTRACT

Distributed ledgers and their applications in solving centralization problems in both financial and non-financial domains has been in the forefront of information security research since the emergence and the subsequent popularity of Blockchain. While the Proof of Work protocol has been successfully utilized for cryptocurrencies, the requirement for higher throughputs in non-financial domain based distributed ledgers favor alternate protocols whose consensus assumptions usually come with thresholds of Byzantine agents (faulty inputs) the consensus can withstand. Proof of Work is designed so that financial gain from conducting a successful attack is less than what honest participation would provide, eliminating any motivation an adversary might have to attack (within the context of direct gain). This assumption fails for non-financial solutions since resourceful malicious participants may exist where their gain may lie in manipulation of the distributed ledger or the order in which the transactions are recorded. A resourceful attacker could selectively convert rational agents to byzantine agents until the tolerance threshold is exceeded. Therefore, we propose that completeness assurance, and the overall reliability of distributed consensus requires rational and foresighted players to be sufficiently incentivized in affording costs of self-protection. We present a dynamic, complete, and imperfect information game to study the relationships between individual costs and utilities, tolerance threshold of the protocol and environment volatility in terms of exogenous attack probabilities, and observe conditions under which a mixed strategy equilibrium that preserves completeness would be stable. Our research extends existing literature by obtaining realistic resilience measures when considering rational player behavior in volatile environments, and provide a better understanding of mandatory security requirements that need to be implemented by a protocol designer for security in distributed consensus. We evaluate our proposed model using efficiency measurement concepts such as Price of Anarchy and Price of Malice, alongside learning methodologies such as regret matching and bounded rationality for extended insight. Our evaluations follow the theoretical predictions of the proposed model. Our results confirm reputation optimization to be capable of completeness assurance when the benefits are carefully assigned with consideration to tolerance threshold of the network. Our experiments also indicate that reputation optimization has attractive stability and convergence properties that are absent in other learning methodologies considered for evaluation.

Keywords: Incentive Compatibility, Mixed Strategy Equilibria, Social Trust Network, Bounded Rationality, Price of Malice, Game Theory, Distributed Consensus, Mechanism Design

DEDICATION

This dissertation is dedicated to my grandfather, Captain John Jayapala.
I hope there is peace. I hope you have found it.

ACKNOWLEDGMENTS

My gratitude goes to my supervisor, Dr. Chandana Gamage for providing continuous guidance throughout this endeavor. The expertise, resources and supervision provided by him were invaluable to the successful completion of this research. I would also like to thank Dr. Charith Chitraranjan and Dr. Indika Perera for the resources and guidance they provided in assessment and reviewing of literature.

Finally, my gratitude extends to my friends and family, whose support and patience has been a driving force throughout the duration of this effort.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Description
ABS	Agent Based Simulation
BFT	Byzantine fault tolerance
BGP	Byzantine Generals Problem
DES	Discrete Event Simulation
DMS	Dynamic Micro Simulation
DoS	Denial of Service
EFBP	El Farol Bar Problem
IDS	Intrusion Detection Systems
IoT	Internet of Things
IP	Intellectual Property
MABS	Multi Agent Based Simulation
NFT	Non-fungible tokens
OOS	Object Oriented Simulation
PBFT	Practical Byzantine fault tolerance
PoA	Price of Anarchy
PoM	Price of Malice
PoW	Proof of Work
UNL	Unique Node List

LIST OF APPENDICES

CHAPTER 1
INTRODUCTION

1.1 Background

The technical and monetary investment in Blockchain and smart contracts has grown rapidly alongside the success of cryptocurrencies. These disruptive technologies have proven their capability in transforming currently centralized solutions into decentralized solutions with increased security. This has subsequently caused industry encouragement in revamping business solutions to integrate Blockchain in order to stay relevant among the competition willing to cater to differing customer requirements.

More importantly, businesses, academia and government organizations alike have shown interest in alternative use cases of Blockchain. By themselves or alongside smart contracts, properties such as decentralization, traceability, immutability, anonymity and security as a whole or as subsets provide alternative and compelling solutions for many different problems.

Intellectual property protection, supply chain management, medical history management, and reputation management are some of the proposed use cases for utilization of Blockchain in non-financial domains. Problems in domains such as land-registry management and entertainment industry royalty contracts, mainly caused due to their historically centralized mode of operation, are also expected to be solved using Smart Contracts [1]. Even though these solutions would take time to reach the maturity required for wide adoption, related innovations will be frequently introduced in the coming years. This raises the problem of whether Blockchain technology is secure, and scalable enough to support such a wide variety of solutions. Since the technology was designed to enable peer to peer transactions, it has limitations that are only acceptable in the originally proposed context. One such limitation is the low transaction throughputs caused by the size of blocks mined, being crucial in maintaining required security [2], [3] (Refer section 2.3.1). Formerly discussed alternative use cases may not be practical with such latencies nor the accompanying computational burdens.

In order to overcome this problem, multiple strategies have been proposed to facilitate faster throughputs. It should be noted that these protocols are often seen to sacrifice one or more elements of the original Proof of Work, such as decentralized consensus or open participation while retaining most security features. The most transaction-efficient among these protocols are Byzantine Fault Tolerant (BFT) Replication protocols, which are able to withstand $< \frac{1}{3}$ arbitrarily behaving consensus participants. Many variations

of these protocols are being utilized for consensus in Hyperledger Fabric.[1]

Another noteworthy protocol, Ripple, retains the open participation of PoW with the use of Unique Node Lists (UNL) and can withstand $< 20\%$ of malicious or arbitrary consensus inputs. A detailed discussion regarding some varying consensus protocols[2] can be found in section 2.3.

Consensus based protocols are only resilient to a certain fraction of unpredictable agents, called "Byzantine Agents". They are often considered analogous to malicious agents in terms of presenting protocol resilience, but this definition is not optimum for a distributed network context. Byzantine agents are faulty or *individually* malicious processes that provide incorrect results to a process, and whose collective existence below a threshold cannot affect the overall correctness of consensus. These could be transiently failing processes, or intentionally malicious processes, but they cannot affect non-faulty node/agent behavior.

Therefore, we define malicious agents in this discussion as intentionally malicious participants or processes who are resourceful enough to influence external entities. For example, even if a malicious entity is unlikely to obtain the required computing power to attack the system, the entity may be able to influence selfish players in the protocol so that, in order to optimize their utility, their only option would be to cooperate with malicious nodes. This is realized in case of Selfish Mining (See section 2.3.2). Even in cases such as Byzantine Consensus, these selfish players may not even be aware of the existence of malicious players, straining their ability to obtain rewards from consensus participation.

A rational attacker would not invest resources if the potential gain, intrinsic or extrinsic, did not justify the investment. Existing protocols are sufficiently secure against such attacks, given the incentivized honest participation of miners and the decentralization requirement of double-spending attack prevention. Thus, for Blockchain applications in financial domains, subverting consensus in the aforementioned manner would require more expenditure than possible gain, making such attacks irrational.

But given Blockchain's ability to move beyond one domain, and having a multitude of applications that is of practical worth to the society, it has to be questioned whether the existing protocols are capable of defending against a highly resourceful attacker. For example, imagine a timestamped evidence storing Blockchain, and a powerful adversary who wants certain evidence to not be included in the immutable ledger. Expenditure vs financial gain would not be a valid argument here, making the consensus protocol vulnerable to external manipulations of Miners, whether through bribery, or simply denying them network connectivity.

For a distributed ledger that caters to such time-sensitive solutions, apart from the aforementioned decentralization, traceability and immutability, the additional security

[1] An Open Source distributed ledger framework for cross-industry Blockchain technologies. It has pluggable consensus protocols, most prominent being PBFT [4]

[2] It should be noted that when this prose refers to "consensus protocols" it only refers to consensus protocols related to distributed ledgers.

requirement of **completeness** must be considered. This is the property of all authentic and valid blocks presented for consensus being accepted and included in the distributed ledger in their correct chronological order. The order of blocks must be made tamper-proof to powerful adversaries whose gains are not strictly financial.

In considering defense against such an attacker, the autonomy of agents participating in distributed consensus must be ensured, and as such, they should be incentivized to afford the costs of self-protection. This matter is complicated by "free-riding" and "tragedy of commons" scenarios seen to emerge in public goods games in game theory, where only some of the agents bear the entire cost of protection, and are thus discouraged from continuous participation, leading to complete compromise of security in the system. Similarly however, aspects such as the impossibility of prior communication between agents, and the volatility of the network have shown to persuade agents to behave in silent agreement with each other such that everyone is fairly compensated for their collective shared costs.

We postulate that we could identify relationships between the factors influencing the security of distributed consensus, both at an individual level and the network level in such a way that the least amount of cost is exerted while most benefits can be reaped by participants. This would ultimately result in a stable convergence, simultaneously assuring the completeness assurance security requirement of distributed ledgers. We model these factors using game theory in order to provide maximum possible utility to rational agents whilst achieving maximum possible social welfare. We incorporate insights from multiple existing protocol implementations that are hardened against subsets of distributed ledger security requirements in making assumptions for our solution design.

In summary, we analyze research outcomes from multiple domains in order to place our research in a practically applicable context, establish the prominence of our problem, justify our choice of solution methodology, and conclude with conditions under which security and reliability of distributed consensus can be achieved when the participating peers are "selfish" and rational.

1.2 Motivation

The practical applications of distributed ledgers have already been firmly established within the context of cryptocurrencies. While the ability to perform decentralized secure transactions is a noteworthy achievement, their practical use remains questionable. As of now, it seems an excellent investment opportunity, and it further provides the anonymity required for illicit activities among cybercriminals. While the integration of smart contracts can increase Blockchain's financial use cases exponentially (to include market shares, derivatives, bonds and other forms of equity) they are not necessarily within the grasp of the general public. It is truly doubtful whether the general public would abandon using centralized solutions for their transactions due to the insurance and throughput that have already been established. With the current rates of Bitcoin and the expenses that must be afforded for mining, it stands to reason that financial institutions

present the cheaper alternative to them for very little compromise. They do not care for anonymity, and centralized agents are legally bound to implement immutability and fault tolerance, in addition to being well governed by authorities.

These establishments also have heavy influence over how a marketplace operates, and would not stand idly by while they become obsolete. In any case, wide, universal integration of Blockchain in financial domains would take a significant time to solidify.

However, many cases exist where the centralized record keeping and differently interpretable contracts have become a problem for the general public, such as the management of land registries. This mostly concerns public facing governmental agencies who face less scrutiny than financial institutions. Specially in a developing country such as Sri Lanka with bleak hope for heavy re-infrastructure in the near future, further insight into this would be worth its own dedicated research.

Another domain where such bottlenecks exist in distributed supply-chain systems would be the music industry. There are many studies already conducted in this domain with celebrity endorsement, highlighting how a decentralized ledger would help many of the general public in the short term with very little objection for wide adoption.

However, one of the major problems in these alternative applications is securing the chronology of Blocks, where cryptocurrency applications could afford a certain amount of flexibility. For example, if it were to be a bidding application where the participant stakes are high, chain of Blocks must be accurately maintained even though the consensus required is obtained out of selfish participants who may be susceptible to attacks or persuasion.

Therein lies our motivation to understand the fundamental factors that contribute to the security of distributed ledgers, and the limits of its perceived reliability, which so far seems to dismiss completeness assurance in the face of resourceful adversaries. We believe that enforcement of certain environmental conditions could ultimately result in a correlated equilibrium of these seemingly out-of-control factors, allowing completeness alongside security properties such as immutability and transparency of the original protocol. We hope our analysis would facilitate more accurate security assessments of non-financial Blockchain applications, and ultimately support the technology to be utilized in solutions that serve the general public.

1.3 Problem

1.3.1 Definitions

To explain our problem, we first establish the definition disparities between agents in our problem and agents in existing solutions. We do this to highlight the security concerns in applying a solution designed for a different use case with different actors in our specified context of non-financial distributed ledger applications. Through this analysis we decide the scope within which we define our proposed solution.

We base our presentation on the arbitrary choice of using BFT-based consensus in following definitions. Our choice here mainly concerns with the throughput required for transaction order assurance in applications outside financial domains. We refer the reader to section 2.5.3 for a detailed justification of this decision.

1.3.1.1 Byzantine vs Malicious agents

The Byzantine Generals Problem famously solves the problem of n faulty or unpredictable participants, dubbed "byzantine" agents, being unable to thwart a distributed consensus for $3n+1$ participants. The other $2n+1$ agents are considered loyal, and their adherence to the protocol would ensure that the final commitment made is equivalent for all parties. This scenario is demonstrated in figure 1.1.

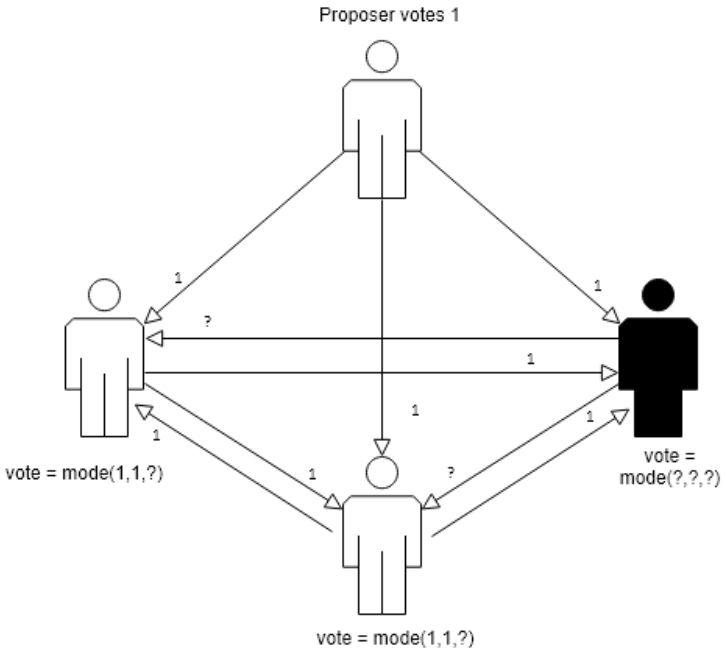


Figure 1.1: Byzantine Fault Tolerant Consensus
$2n+1$ participants reach agreement while n traitors propagate false information

The original Byzantine Generals Problem, as discussed in section 2.4.1 considers the message delivery reliable. For example, in a synchronous scenario, if the messages aren't delivered within the designated time, the solution is to default to a failsafe value upon timeout. The most risk-averse default is not committing to any action, hereon addressed as "committing to inaction".

But imagine a scenario where the traitor is not concerned with subverting the consensus agreement of the network, but rather delaying it until the time is opportune for him, as illustrated through figure 1.2.

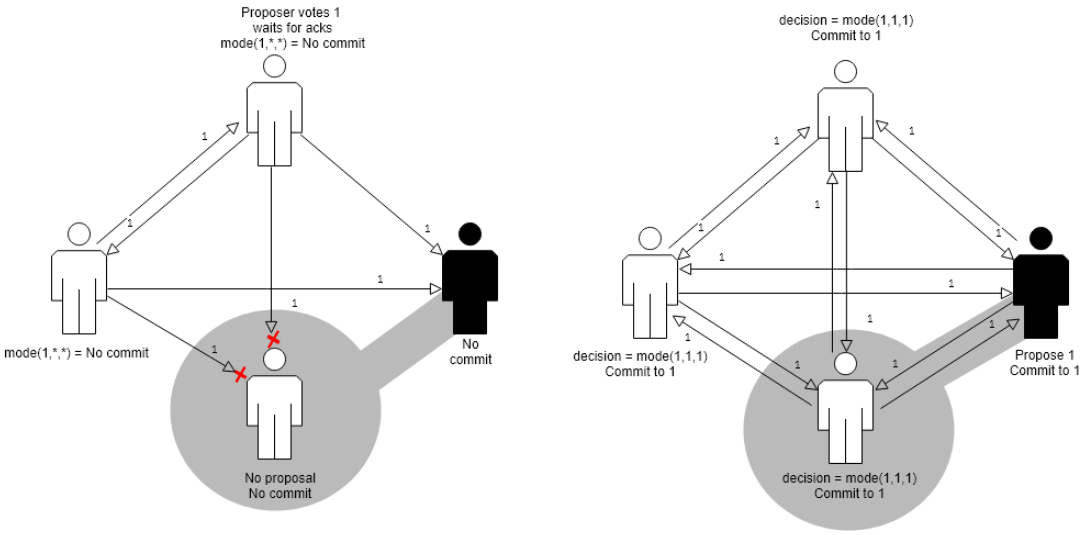


Figure 1.2: Selective interference

The image on the left illustrates where the attacker interferes with one agent so that all agents commit to inaction. The image on the right displays how the attacker can simply use the same network so all agents commit to the action proposed by him.

This scenario does not violate the solution proposed by Byzantine Generals Problem. The solution handles $\frac{n}{3n+1}$ Byzantine nodes, and considering the compromised node and the actually malicious node both as Byzantine, that condition is not fulfilled.

Byzantine Fault Tolerance was proposed to handle abrupt failures of nodes in distributed processing or communication, such as the state machine replication used in database commit synchronization. In the presence of $\frac{n}{3n}$ faults, all nodes commit to inaction. Even in the scenario in figure 1.2, agreement is achieved in all cases, and the nodes retain consistency.

However, in both permissioned and permission-less Blockchain implementations, the number of peers who are voting is not a constant number. Even if we assume that permissioned peers assure 100% availability, in a Blockchain with open participation, the total number of peers would always be an uncertain property. Upon this condition, the absence of one extra "loyal" node may not be detected within the synchronization period.

Furthermore, in the presence of m malicious nodes and $3n$ total nodes, one malicious node would have to selectively attack $\geq \frac{n-m}{m}$ nodes. Given the openly available information between peers, such a denial of service attack can be easily mounted with a high probability of success. They can further avoid disqualification by randomizing the proposer amongst themselves ($m > 1$), combined with randomized victims. Even the compromised nodes would not be disconnected since they would participate in the vote as usual when one of m is the proposer. To summarize, for m nodes to tamper with the order of the distributed ledger, exceeding established resilience is not strictly required. Selfish mining (Refer section 2.3.3) can be seen as a realization of this problem.

Thus, we consider it unsuitable to apply Byzantine Fault Tolerance based solutions to

distributed ledger consensus implementations without any assurance of the security capabilities of its peers. We declare liveness a required property, especially in scenarios where the strict order of Blocks has to be maintained, ruling out solutions based on multiple randomized consensus calls for different blocks. We also consider committing to "inaction" upon failing to reach consensus a vulnerability of directly applying the BFT protocols without preconditions on peer's ability for reliable message delivery and voluntary resignation, given the synchrony and the network reliability requirements.

We conclude that in a security analysis of distributed consensus protocols, peer's protection capability must be evaluated in order to tamper-proof the order of blocks added to the ledger.

1.3.1.2 Loyal Generals vs Selfish Generals

Mounting denial of service attacks is not an exclusive way of compromising $n - m$ followers. Unlike the "loyal" generals in Byzantine Generals Problem who would not be compromised at any cost, there is nothing stopping any of the $(3n - m) - 1$ (given that proposer is loyal) nodes from colluding with the m nodes who may simply offer better financial rewards for their votes. It is a reasonable assumption to think that no peer would be volunteering their resources for the sake of loyalty. This is another scenario where the definitions in Byzantine Generals problem deviate when applied in distributed ledger consensus.

While we declared the vulnerability of defaulting to "inaction", it is the behavior of these selfish agents that makes it a vulnerability. Even without the presence of malicious agents, selfish agents would not have any reason to propagate the messages once they have received acknowledgment from $2n + 1$ nodes, so that the rewards are split amongst the first receivers. The other n nodes would not be aware of any messages, as such their votes would be considered Byzantine and they would not reap any rewards. We consider selfish agents different from malicious agents on the fact that they do not spend resources on denying communication to the n leftover nodes, but only withhold resources spent to maximize the gained utility for the cost afforded.

One solution to this problem would be to increase rewards in message propagation over the rewards in participating in consensus. The discussion comes full circle in considering that the compromised node in figure 1.2 would not be obtaining the message propagation rewards due to malicious node's interference. This would provide incentive for a selfish node to implement the necessary security preconditions since it must always be in the presence of at least one secure node to participate in consensus, and subsequently the message propagation.

While the above analysis uses Byzantine Generals Problem as an example to demonstrate certain security assumptions that are invalid in case of Blockchain, these misconceptions exist in many other consensus protocols, only with differing thresholds. For example, Proof of Work consensus faces the problem of transaction withholding when mining rewards are eliminated in favor of transaction fees [5] (Refer section 2.5.1.2). Therefore an analysis into relative self-protection measures, rates of incentive and the layers of

communication where incentives should be implemented would provide crucial insight in obtaining realistic resilience properties of robust consensus protocols.

1.3.2 Problem Statement

We recognize two areas that haven't been coherently studied in existing literature.

1. The transference of security properties in existing distributed systems consensus to distributed ledger based solutions
2. The transference of security properties in existing distributed ledger based solutions to distributed ledgers in non-financial domains

Therefore, following previous discussions, we establish multiple security parameters that are closely related to each other when considering rational peers participating in consensus.

1. Protocol's tolerance for compromised agents
2. Agent's probability of self protection
3. Agent's cost of self protection
4. Agent's incentive for consensus participation
5. Agent's incentive for message propagation

Using the above factors to establish the scope and domain of our research, we proceed to define our problem.

What are the relative relationships between incentives, costs and peer security requirements a consensus protocol must implement to ensure lasting stability and completeness?

1.4 Research Objective

Our main objective throughout this research is to provide environmental conditions that ensure the autonomy of "selfish" peers in a consensus protocol. It is naïve to develop a distributed ledger for non-financial domains without accounting for the risk of denial of service on participants, and it is difficult to obtain equilibria when the utilities of agents are susceptible to external interferences. An analysis of the volatility that these interrelated parameters cause is required to justify any protocol implementation and we aim to observe and analyze those parameters and their respective limits.

Protection from external interferences in distributed networks is not trivial. Centralized solutions such as dynamic load balancing, or malicious traffic deflecting are not applicable

to peer to peer nodes with limited resources. As such, our secondary objective is to establish the compromises or external considerations the protocol designer should account for in order to provide the benefit-cost ratio that results in a stable network. For example, for a lower tolerance threshold, maybe a better benefit-cost ratio would be feasible, encouraging increased participation.

Following the game theoretical analysis and extraction of required security parameters, we evaluate whether the proposed model is the optimum coincidental collaboration technique that is available to the peers by comparing multiple learning mechanisms. Thus, our tertiary objective is to establish the efficiency of our proposed model alongside other applicable learning mechanisms of equilibria, providing a baseline for future research in evaluating security properties of distributed ledgers.

1.5 Summary

In this chapter, we discuss the possibility of resourceful attackers conducting block history manipulation attacks on non-financial distributed ledgers, and establish the importance of influencing self-protection of financially motivated rational participants of consensus. Since the costs for such self-protection has to be afforded by agents themselves, we consider the reward structure that allows for satisfactory compensation. We recognize obtaining equilibria between incentives, costs and self-security assurances by peers who are using a volatile network as their communication medium as a significant research problem in designing stable consensus protocols. We define our objective as obtaining limits of endurance in equilibria for aforementioned costs, incentives and attack probabilities, maintaining the selfishness of financially motivated agents, resulting in influence-free voting and complete block history maintenance of non-financial distributed ledger applications.

Chapter 2 provides a comprehensive review of current research literature on the core topical areas as well as the papers that introduced seminal concepts relating to Blockchain and Game Theory. The proposed game theoretical model is presented in Chapter 3. We briefly present the implementation of our simulation and the subsequent evaluation of our model using simulation results in Chapter 4. We conclude our discussion in Chapter 5 with an overall summary of our work and a discussion on possible future extensions of our research.

CHAPTER 2
LITERATURE REVIEW

2.1 Introduction

The literature review is structured as follows. A brief justification of our motivation to analyze security properties in Blockchain implementations of non-financial domains is provided in section 2.2. In section 2.3 existing consensus protocol implementations and their respective limitations are discussed as a precursor to determining necessary security parameters that need to be incorporated into our proposed model. The potential of game theory as a model for problem solving is discussed in section 2.6. Game theory applications in general security domains and how their findings relate to our research are discussed in section 2.7. This section further provides insight into how our improvement suggestions can be integrated to a game theoretic model. Security properties of multiplayer public goods games, equilibria observed in distributed system specific games and subsequent efficiency metrics are discussed in section 2.8. In establishing a baseline for evaluating our model, alternative methods of learning which could also converge in equilibria are discussed in section 2.9 with reasons for their selection. Section 2.10 presents structure preservation concerns in implementing our model via simulation and we conclude in section 2.11 with a brief summary of the overall chapter.

2.2 Blockchain Applications Beyond Financial Sector

Given that it is not our technical objective to implement a Blockchain application for non-financial domains itself, we restrict this section of review to briefly present some use cases found in literature, so that our research intent is given sufficient perspective. It should be noted that this is not an exhaustive review, but rather a glimpse of how hopeful multiple sectors are regarding solving their problems through secure, decentralized distributed ledgers.

2.2.1 Intellectual Property

The stakeholders of Intellectual property domain have extensive expectations on Blockchain's use cases. The community is continually frustrated due to problems in centralization, record integrity assurance, and the extensive delays in processing information/transactions through many intermediate parties. And the recent popularity

of Blockchain and the generalization of its security assurances from a specific context to a wider, more generalized context have professionals seeking Blockchain-based solutions to most of their prominent problems.

In [6], the World Intellectual Property Organization itself acknowledges the potential Blockchain has on Bitcoin applications. It credits this feasibility to versatility of data which Blocks can store, along with usual integrity and decentralization assurances.

Out of the many use-cases of Blockchain listed, we pinpoint the following, which requires preservation of order of transactions.

- "Evidence of creatorship and provenance authentication"
- "Registering and clearing IP rights"
- "Providing evidence of genuine and/or first use in trade and/or commerce"

Permissioned Blockchains for Intellectual property registration and copyright protection could provide centralized access to all records for validation purposes, record the steps throughout the IP registration for legal dispute resolution and furthermore, provide evidence of use and frequency of use, where accessing the IP related records are notified back to the original owner.

Genuine product verification by the consumer is near-impossible given the limited information they are provided with, or the purposefully restricted information resources they have regarding products. Customer protection from counterfeits, Standard verification, Warranty validation and Regulatory requirement assurance/evidence maintenance can be facilitated by businesses through Blockchain based traceability maintenance of the supply-chain. Such a blockchain would further aid customs authorities in performing their validations and reduce unnecessary delays in order processing.

The article concludes with the forethought that given the popularity and the wide application possibilities, it is only a matter of time before large-scale Blockchain applications find the laws required introduced so that they can function as legally accepted entities.

[7] introduces Kudos, a currency which represents educational rewards, and commoditizes academia beyond the community of established academics.

The authors suggest a consortium of established educational institutes to be assigned an initial amount of currency based of their reputation (University rankings from an agreed upon entity), which they can transfer to employees whose research interests they want to promote. The employees themselves can associate their Kudos to their publications (not spend but simply associate) to show how prominent the publication is among others. External users can subscribe their respective Kudos to publications they want to advocate or be associated with. Institutions will mine the reputation in a proof of "reputation" stake based consensus.

More traditionally, Exam credentials such as degree certificates, course completion verifications, examination pass/fail verifications are few of the straightforward applications. Existing initiatives such as University of Nicosia Blockchain Based Certificate

authenticity verification and Academic record storage by Sony Global Education attest to the practicality of such implementations. Additionally, such a scheme has unorthodox advantages as well, such as enabling reputation collection for alternative yet related pursuits such as tutoring and volunteering, which may not count as experience in traditional job application contexts but could accumulate to significant amounts over time. (Much like Stackoverflow points, Github commits of Software Engineering professionals.)

Adoption of these Blockchains by the wider community is not likely to face heavy public criticism, given that academic reputation is already tracked through measures such as citation counts, author and publication H-indexes.

Blockchain is further suggested to be used as a crowd-sourced method of patenting, providing "Proof of intellectual work". Even if a legal registration has not taken place, malicious parties would be prevented from claiming IP as their own for commercial gains. It can act as a open-source enabler, because users are assured that due credit will be given and due Licensing will be propagated in derived works. An existing implementation of such a scheme would be "Blockai" Copyright infringement protection for creative workers.

The parallels drawn between Kudos and Ted Nelson's Xanadu project, which advocated verifiability of web with unbreakable links, author attributions and consumption based royalty payments, prove to be a highlight of the paper. It is interesting to see that Blockchain enables the realization of a model proposed as early as 1960s, which was also criticized as Vaporware, proving that the research interest in non-financial Blockchain applications is not a misguided byproduct of the Bitcoin bubble.

The authors also claim the exchange rates of actual currency to kudos, and determining financial values of ideas as a potential problem (presumably in a smart contract context), but we believe the essence of it itself to be problematic. Even in current practice, new publications are backed by senior academic professionals with no monetary exchange involved. Changing such a well established norm would have severe repercussions to integrity of academia as a whole, and thus, we propose such conversions to be done within non-monetary, intrinsic currencies. For example, a publication endorsement would increase the author's "reputation" in the system while the user endorsing the publication would increase their "experience credentials". It would be more useful if such endorsements were not expenditures, but simply stakes of "Kudos" which would be lost if the publication was deemed poor by peers, encouraging higher quality publications.

As acknowledged, in utilizing Blockchain for information genesis verification purposes, while the existence at the time is verified, no one accounts for the accuracy of content inside Blocks. While for peer-reviewed publications, organizations may provide the necessary proof of stake, for alternative applications, other Blockchains will have to use the same currency. Authors note that "validity, authenticity or usefulness" may not be guaranteed by a Blockchain, and that a user's claim has the potential to be contested. We partially disagree with this statement, given that any contesting party would also have to prove the existence of records on a verified Blockchain. Revocation in case of

invalid or useless articles could be handled in a way similar to Git version controlling (git revert commits) and validated by checking whether sufficient money exists in an account for it to be spent on a transaction. The authors' solution to this problem is the public scrutiny following openly accessible content, much like the rating systems in eBay, Amazon, Airbnb, and Uber.

The authors conclude with the fact that they are already conducting initial trials of a Ethereum[1] based private blockchain that implements an academic reward system.

Both [6] and [7] highlight the potential use of Blockchain for legal evidence collection, but we point out that without maintaining the correct order in which the transactions are added to the chain, this could never serve as legal evidence. Evidence blocks/transactions have to be added to history immediately after their collection, as seen in general forensic evidence collection procedures. For example, if the binary hash of the source hard disk was delayed until after any decrypted evidence on said hard disk was added to the history, the entire chain of custody could easily be dismissed.

Imagine a scenario where a startup is registering an intellectual property (IP) which is in direct competition with a similar product of a tech-giant. It might be profitable for the more financially capable company to delay their registration, either through obtaining the majority of voting/computational power, or through influencing the network in a timed manner until they have registered their intellectual property. Even if the startup eventually registers their product, it might now be infringing on the earlier IP.

The integrity of voting entities holds no power here, given that they are not aware of any valid transactions with earlier timestamps waiting to be announced to the network. If any automated infringement protection is integrated to the Blockchain, the second, delayed transaction might even get rejected, similar to rejection of transactions when funds are insufficient.

The process could be recorded in multiple steps and thus the final timestamp might not matter, but it still would open the startup to lawsuits which they may not be able to afford. In a decentralized solution, every participant must have equivalent power, a condition that is violated here. Even though the aforementioned situation is more socio-economical, and outside the security assurances of Blockchain itself, it directly affects the wider adoption of Blockchain in different domains. Therefore unless an order can be guaranteed, the security assurances of distributed ledgers have the potential to disintegrate.

We note here that the observation by M. Sharples and J. Domingue [7] of Blockchain enabling "Proof of intellectual work" has also been implemented in the form of non-fungible tokens (NFT). A non-fungible token (NFT) is a certificate of authenticity confirming an asset (digital or otherwise) to be unique, which is stored in a distributed ledger [8]. This contrasts with cryptocurrencies which are fungible, in that they could be exchanged with another currency at an available rate. NFT provide digital assets with the same gravity a physical asset might hold. For example, a renowned painting cannot be

[1] An open-source operating system which implements a generalized platform for Blockchain and smart contracts where users can create their own decentralized applications.

owned by multiple owners due to the simple fact that it can only be available physically at a single location, while digital assets can be copied and transferred in a multitude of ways. Through NFT, the original asset retains its value, while any counterfeits circulating would fail to reap any monetary compensation (unless the original license allows for derivative work).

If the ownership of an NFT is transferred, the chain of custody is recorded alongside the token. Given that the ledger is public, ownership can be verified by any interested parties, making NFT a funding enabler for digital art and creative industries in general, while allowing for a wider array of alternative blockchains wherein ownership could be transferred among content creators, collectors and investors alike.

2.2.1.1 Music, and Creative Industries

While related to intellectual property, and fully harboring the same intentions, creative industries stand to gain extended benefits from Blockchain. Since intellectual property owners, as a majority, are not as underrepresented as indie music artists and the like.

Artists and songwriters are represented, or managed by multiple intermediaries between artists and consumers, such as record labels, performance rights organizations and publishers, all of whom play a part in centralization of content delivery. As [9] clearly points out, Nakamoto's micro-payment concept allows elimination of such intermediaries who obtain the most collective value of creative endeavors, allowing creative content to be provided at a much more affordable fee. It also helps reduce illegal consumption of creative arts, allowing customers to enjoy the creative product legally and support the artists directly, providing them with continuing income.

Of course, not all intermediaries will be eliminated, since administrative functions such as payment negotiations, event organization will require external interference. But they would have much less control regarding how the payment is delivered to the artist, resulting in a healthier ecosystem.

Instantaneous payments could be arranged at the time of download or streaming through smart contracts, making royalty payments fast and effortless, resolving the issue of fund withholding by middle-men for contractual onward payments, which are subject to seizure in case of business failure. Transparency assurances of Blockchain could solve the problems in fund payments to wrong parties (In case of covers of existing songs), and royalty withholding due to owner identification difficulties.

An interesting alternative application purposed is crowd-funding for artists, with their own cryptocurrencies if necessary, in contrast to seed-funding from investors. The function of a record label is replaced with patrons, who are empowered through immutable records in verification of their investments being used for realization of creative endeavors. Such a system would also allow high-level Investors to monitor popularity through investment rates and delivery rates, enabling effective assessment and eventual capitol allocation. Portfolio investors can make use of automated monitoring of artist careers for further analysis and negotiate deals for potential team-ups, or make

investments in marketing mentorship or trainings. Thus, the heart of the initiative would be informed decision making by all parties involved, with or without the help of data analytics.

The authors proceed to say that this adoption is not without its problems. As we observe in section 2.5.3, scalability concerns are a major obstacle for these applications. Furthermore, usability constraints plague naïve end-users and even artists, or their managers themselves. Any application that enables blockchain based music delivery would have to present a version of Blockchain wrapped from both endpoints, but this would not be difficult given that existing cryptocurrency wallets represent these scenarios.

However, a much more critical concern is the centralization of these interfaces, since customers have to get the information regarding new publications somehow. The possibility of these outcomes being manipulated by intermediary parties invalidates most benefits of Blockchain based music delivery, since this can be done through paid advertisements and Search Engine Optimizations. However, this is a concern external to our targeted problem.

Recurrent concerns from intellectual property applications such as the inability of validating inserted data, and the immutability in case of enormous insertions are also noted as potential problems, which we address above. The authors note that distributed file storage could facilitate piracy, but this can be solved through cryptographic means, as seen in the case of consumer specific music delivery by Apple Music.

Practicality of Blockchain based implementations for the Music industry is demonstrated through examples of existing companies such as Ascribe for limited edition image creation, Resonate streaming service, Colu Asset Transference (including Intellectual Property). Higher level artist enabling platforms such as Mycelia[2] which aims for a fair music ecosystem, and DotBlockchain[3] an information Blockchain that protects authorship of creative work, with amendments being inherited forward while the original record is kept intact are also mentioned.

The authors conclude by noting that Blockchain will indeed have extended use cases in creative industries such as fashion, journalism, games, and art in addition to music, given that realistic expectations are kept regarding wide adoption. Disruptive technology itself is insufficient in solving potential social problems, such as the centralization through user interface caused due to usability constraints of Blockchain, calling for a much wider dissemination in future research in enabling alternate Blockchain applications.

2.2.2 Internet of Things

A use-case based discussion regarding the "roles Blockchain could play in strengthening IoT security" is presented in article [10] . The authors discuss the vulnerability of using centralized servers for communication between IoT devices, since this causes individual

[2]http://myceliaformusic.org/

[3]http://dotblockchainmedia.com/

device compromise to propagate through the entire network. This claim is supported by the case of the distributed denial of service attack on "Dyn" DNS Provider caused by Mirai malware infected IoT devices, which themselves were infected after a successful phishing attack.

Blockchain's decentralization, peer to peer communications and strong origin authentication is presented as security features that would solve this situation. However, we add that the sanitizations and lightweight data acceptance required of consensus protocols in general would play a larger role in assuring this, given that this requirement would limit acceptable data types and commands. Decentralization further prevents interferences or data withholding by external parties (See section 2.2.3), which is crucial for IoT services. Forgery resistance of Blockchain is also mentioned as an enabler of potential IoT applications.

IMB's supply chain provenance of high value items and IBM Watson IoT Platform, which supports data storage, authorized sharing, translation and censoring for smart contract APIs are mentioned as existing implementations which have obtained aforementioned benefits of Blockchain. Filament, an IoT service provider who has eliminated the cloud dependency by using Blockchain for authentication, communication and autonomous smart contract execution purposes is also mentioned. Furthermore, Organizations such as "Cisco, Bosch, Bank of New York Mellon, Foxconn Technology, Gemalto" and startups such as "Consensus Systems, BitSE, and Chronicled" have formed a collaboration to standardize Blockchain based IoT applications, showing signs of global adoption in future.

The authors also discuss the immutability of the ledger, and acknowledge that initial accuracy and validation is required before blocks are added, much like all aforementioned applications in different domains. A correlated challenge is the secure authentication and integrity verification of physical devices themselves. A solution is proposed in the form of cryptographically secured hashes of device's initial state which can be used by the permissioned Blockchain nodes to verify IoT device integrity at its current state. Of course, this would ideally require integration of Trusted platform modules, but it is not even remotely scalable or economically feasible for IoT platforms.

Another interesting aspect noted is that Blockchain facilitated fast message propagation among close by devices, wherein cloud based servers require communications to flow through them regardless of physical proximity of endpoints. While Blockchain relieves the unnecessary communication overhead in case of asynchronous protocols, it may not necessarily be true in other cases. Furthermore, if the platform was written to use a gossip protocol, this problem doesn't necessarily require a Blockchain based implementation.

Somewhat related to section 2.2.3, crisis prevention is also expected to benefit from Blockchain based supply chain provenance. Vulnerable products could be immediately tracked down to their current owners and recalled if required. For example, Hangzhou Xiongmai Technologies is mentioned to have recalled the products vulnerable to Mirai malware, while more recently, 2016's Samsung Galaxy Note 7's tendencies in exploding resulted in a similar situation.

In contrast to the conceptual discussion of this article, [11] provides an implementation of a smart home, which uses a private ledger at the local network and a decentralized public ledger with more computationally capable devices as peers for the purpose of fault tolerance and trust maintenance.

2.2.3 Healthcare

[12] makes a strong case for Blockchain based security applications in the medical industry in the interest of gaining benefits of transparency, tamperproof-ness and anonymity. It mentions that Blockchain clearly has many applications beyond the financial sector, such as music rights administration by *Ujo* or *Peertraks* and *stampery* for signed documents, and expresses potential applications in Healthcare such as medical record maintenance for latest and timestamped patient histories, auditable data usage in medical research with rewards given to contributing users, and drug counterfeiting prevention through supply chain provenance up to the ingredients and production conditions that were used.

Patient history maintenance through Blockchain benefits medical professionals by being faster and less intrusive than traditional authentication processes through different security boundaries (i.e. different hospitals for example). Availability of accurate and properly audited information results in less negligence and more accurate diagnosis, benefiting patients. The present success of *Gem*, a startup based on Ethereum Blockchain technology founded in 2011 as a result of Estonia's collaboration with *Guardtime* is mentioned as a practical implementation of such a platform.

Patient health data being produced at exceeding rates through wearables and other smart devices are discussed as a potential for customer revenue. Data will never leave the ownership of the user, but still be available for reuse by medical research professionals for a financial compensation. *Healthbank*, a current Switzerland based health startup aims for such an implementation on their existing personal health data sharing platform. Given [13]'s observation that "personal information is intellectual property that bears negative royalty" we find this particular application highly interesting.

In addition to this, the practical advantages of origin backtracking of data samples can be observed in an incident such as Flint, Michigan, United State's water crisis in 2014, where the officials discarded certain samples to falsify research outcome, distributing water with higher than recommended lead levels [10]. If medical or other standardization research were conducted with immutable records obtained from a publicly available ledger, they become open to scrutiny, creating a more informed and better protected consumer community. One skepticism that remains however, is the if this was implemented in a permissioned blockchain (which would be highly likely), it still becomes susceptible to alteration by powerful parties.

The fight against counterfeit drugs can make significant strides forward through Blockchain based recording of the sensitive production process of their legal counterparts. Liability concerns will have clear evidence to proceed with, and production monitoring by standardizing organizations will also be made much simpler and transpar-

ent. Furthermore, Backtracking the root causes of health issues will be made both fast and cost-efficient. *Hyperledger*'s *counterfeit medicines project* is given as a aspiring implementation which serves this purpose.

As closing arguments, the authors mention that Blockchain would further help drug cost reduction through elimination of middle-men, promising a much more optimized healthcare system.

2.2.4 Governance

[14] introduces smart cities, where internet of things and cloud come together to serve citizens by optimal utilization of available resources which are likely to be limited. Increase of population, and the subsequent public preference of living in cities given the centralization of job opportunities, requires efficient management of resources. Given that such an automated self-sustaining system would require public data on locations, consumption of resources and financial aspects, it is imperative that "privacy, integrity and data confidentiality" of this data is assured. Integrating traditional security into city's infrastructure is impractical, given that physical security cannot be assured with such a high number of devices, which unfortunately are also known to have high vulnerability rates.

Given the security assurance of proof of work, (resilience to $< 50\%$ computing power), the authors explore the idea of using Blockchain to assure the security requirements of smart cities. They propose a high-level security framework with the following layers.

- **Physical layer** consisting of strong encryption and access control mechanisms for sensory devices and a shared communication standard for cross-functional usage
- **Communication layer** where data are broadcasted using Ethereum Telehash over BitTorrent to multiple Blockchains
- **Database layer** which is the distributed ledger of a private Blockchain, invulnerable to Sybil attacks
- **Interface layer**, where smart applications correlate their date to provide enhanced services to citizens

Fault tolerance, reliability and scalability are highlighted as benefits obtained by blockchain. Efficiency, a critical requirement for a smart city, is also mentioned as an advantage, but they acknowledge the limitation in proof of work's consensus finality and suggest private Blockchains as a solution. They further endorse Blockchain for allowing direct communication between citizens and governments.

2.2.4.1 Enabling disadvantaged parties

A recent social phenomenon observed in United States is past-abuse related accusations towards public figures. Victims claimed that the delay in accusations were due to their

sense of responsibility towards the society, in preventing unsuitable officials' rise to influential positions, as opposed to accusing at the time of abuse for their personal justice. Legal authorities were not equipped to handle, or deliver justice in these situations, where false victims could attack arbitrary public figures, while real victims remained void of justice.

The internet has a reputation of not losing anything that was once posted, but the data are stored in centralized, inaccessible locations where its origin, time of creation will never be verifiable enough to use as evidence. But a public Blockchain would be highly suitable for storing such abuse related information, specially for impoverished, or socially discriminated individuals who may want to seek justice in future. Collected evidence and independent witness statements (including any medical examination results) could be stored in encrypted format, and by simply securing the decryption key, existence of evidence at the time of incident could be proved if required. Even anonymous evidence of corruption can be provided to the legal authorities by simply supplying them with the decryption keys.

2.2.5 Influence of Resourceful Adversaries

In extension to our discussion regarding alternative blockchain applications, the existence of attacks by resourceful adversaries must also be considered in justifying our problem statement. A prominent example from financial blockchains for manifestation of such attacks would be "Goldfinger attacks" and "altcoin infanticide" [15], conducted by powerful adversaries exclusively to destabilize rival cryptocurrencies irrespective of any personal financial loss. In case of non-financial blockchains, manifestation of such attacks are hard to pinpoint given that our strict requirement of completeness might be an accepted limitation of the non-financial blockchain solution by design. However, this does not mean that the potential for such attacks are not present. For example, we recall the water crisis incident of Flint, Michigan where research outcomes were falsified by officials in power [10]. Abuse of power by officials can be expected in any location. As such, if alternative blockchain solutions are going to be utilized, strengthening its underlying infrastructure to withstand such interference would best serve the public whilst holding responsible officials accountable. Similarly, evidence collection blockchains would be highly attractive to resourceful adversaries who plan to escape legal repercussions. We note here that this is where the completeness requirement must be ensured with absolute certainty, considering that chronology assurances would be crucial in legal arguments.

2.3 Consensus Algorithms In Blockchain

While a comparison of Blockchain consensus protocols in existence is critical to our research, we first focus on Proof of Work to gain a realistic insight of the purpose of Blockchain. Due to the protocol's early introduction, research has already been conducted on modeling behavior of selfish players, and how they can be incentivized to

act honestly in reaching consensus. Thus, the protocol review follows an analysis of these limitations and the solutions in existence that overcome them. Secondly, our focus moves to Byzantine fault tolerant protocols due to the higher throughputs they facilitate, making them a great candidate for non-financial Blockchain applications. Thirdly, we proceed to describe protocols that slightly deviate from both aforementioned protocols, and the novel solutions they provide for limitations in the original protocols they are based on. Finally, we compare the discussed protocols and the existing problems in Blockchain protocols in general to finalize the parameters that should be represented in an accurate game theoretical model.

2.3.1 Proof of Work

Conversation of consensus protocols in distributed ledgers started with Nakamoto's celebrated white paper on Bitcoin [2]. Despite being the epicenter of decentralized ledger's rise to popularity, the original paper simply focuses on solving the double-spending problem in the absence of centralized parties. The function of a third party relies on identity verification and mediating disputes, and the paper argues that in context of anonymous irreversible transactions, the need for trusting a centralized third party can be eliminated. While generating transaction records can be accomplished through public key infrastructure, recognizing whether a digital coin was previously spent requires a record of all coin transfers. The authors solve this problem through a ledger of chained blocks recording batches of transactions. The chain of blocks, "the Blockchain" need to be immutable and chronological. Immutability is ensured with peer distributed storage, and the chronology is ensured through including the hash of preceding block in any new block's header. The anonymity lies in using anonymous public keys or using new key pairs per transaction.

For the block to be accepted to the chain, its hash require fulfilling some mathematical parameters, such as having a defined number of preceding zeros. Proof of work is the random value that generates such a hash when considering the time-stamp, new set of chosen transactions, and previous block's hash, while the computationally intensive task of finding it is referred to as "mining". This works efficiently, since once the random value is found, hash can be instantaneously verified by any node. Longest chain, the chain that most proof of work is invested in, is universally accepted as the correct history. Therefore, consensus agreement is reached when the block broadcast to the network is accepted by the peers to the longest chain by starting working on a new block. Action of adding the block to individually maintained ledgers represents voting, giving each CPU exactly one vote. Initially, coins come into existence through generation of proof of works by the expenditure of processing and electrical power by miners, and incentive is the coins themselves. But with time, the aggregated transaction fees in a block are given to the miner who found the block's proof of work as incentive to keep mining.

Altering the ledger would require an attacker to find such values for the entire chain of altered blocks to the point that it is longer than the existing longest chain. This requires 51% of the computing power of the entire network in the least. To ensure security is achieved this way, difficulty of proof of work, which is exponential to the number of

leading zeros, is contained by limiting the number of blocks added per hour. This also aids in adjusting the rate of new coin generation.

The security assurance through proof of work is substantiated through two ways. Firstly, an argument is established that participant would always try to maximize their utility/ personal gain, in which case the attack scenario is made less rewarding. If an attacker were to manipulate the network by obtaining 51% of computation power, honest participation would generate him more coins while dishonest participation would undermine his own wealth in the system. Secondly, calculations are provided for the exponential probability reduction of succeeding at a double spending attack within the parameters and the context of the proof of work protocol. Due to the immutable storage, the system remains closed to deceptive blocks with false history, limiting the attacker's possibility of double-spending coins by reversing his own transactions that are currently in the history. This requires the attacker to maintain a secret fork from point of transaction, trying to make the version with false history the longest chain of the network. However without 51% of computation power, success probability drops as honest participants keeps adding new blocks which the attacker would then have to overtake.

If 51% attack were to happen, one solution would be to simply wait out until the attacker depletes all his resources, but the "Snowball effect" can come into effect, as confidence in the network is lost, causing honest miners to leave and the malicious miner to obtain majority of the computing power [5] before resource depletion.

The limitation on transaction throughput remains integral to the security of Blockchain while being the major obstacle in its scalability. The time taken to mine a block, dubbed "block frequency", increases with the size of the block, and thus transaction throughput becomes limited. If block size is reduced (i.e. Lesser transactions per block), then the possibility of multiple forks becomes higher resulting in wasted computational power. Furthermore, for a transaction to be considered immutable, it require to be followed by several more blocks, which in turn demands consensus per each block, accumulating in latency. This property translates to how proof of work governs consensus finality, in that while the existence of forks indicate that consensus finality is not achieved, the acceptance of longest chain ensures that the outcome of consensus finality is eventually obtained nonetheless. This however is not an acceptable condition for time-sensitive distributed ledgers.

An interesting point to note from the nature of digital coin transactions is how it resembles "transactions" in certain alternative domains. In a transaction, multiple transactions previously paid to payer are combined to be equal to or greater than the amount payable (and the transaction fees), with the output directed to payee and any change directed back to payer. In domains where the end result need to be mapped to different origins, such as land registry, digital rights, intellectual property and even academic credit management, this concept proves to be naturally adaptable.

While the success of Blockchain applications provides validity of the claims the paper has presented, the reliance provided by incentive in assuring honesty of participants has to be discussed. As elaborated in section 2.3.2, due to this very reliance on incentive

and the rarity of being able to mine a block individually, mining pools have come to existence where a group of participants pool resources together so that rewards can be split amongst themselves. This reduces the variance of their income and presents itself as a worthy alternative participation strategy for mining blocks. With emphasis on such incentive, given alternative reward maximizing strategies (Refer section 2.3.2), the pool would gain majority control, losing system's decentralization, violating the purpose of the protocol.

In the context of our research, exclusive dependence on such monetary reward based security assumptions presents further problems. It cannot be assumed that the system only contains selfish players who seek to maximize monetary rewards. Alternative applications may exist where certain participants are willing to sacrifice resources for different gains. The motive could be preventing or delaying a block being added to the chain for taking advantage in time sensitive developments. In a different perspective, if a different coin presents a market challenge to a more powerful coin, the existing party could launch an attack to destabilize the newer coin. Attacks where powerful opponents destabilize Blockchains are called "Goldfinger attacks", while this happening at new coin launches is dubbed "altcoin infanticide" [15]. Therefore we can conclude that presenting upper limits of security by considering only selfish participants who may or may not be susceptible to attack as an insufficient assurance, which if naively taken at its word could make a distributed ledger open to some disastrous attacks.

In conclusion, proof of work presents some fundamental security properties which any distributed ledger should aspire to retain. We identify these properties to be Open/ Permission-less participation (Ability for nodes to join/leave at will), decentralization based immutability assurance, chronology maintenance, sybil-proof ness (One vote for one CPU), and extensive scalability. However, given that security assurance of proof of work relies on the throughput limitations of blocks per hour, alternative applications call for more efficient approaches.

2.3.2 Incentive Incompatibility

Public participation is imperative for the proposed decentralization in Blockchain. It further accounts for the security of the protocol since to prevent a 51% attack the committed computing power must be distributed. Subsequently, rewards are necessary to justify the extensive CPU and electricity costs afforded by the public in generating proof of work. In fact, the initial coin mining process is compared to "gold miners expending resources to add gold to circulation" [2]. However, it is unlikely that any such participants would invest in obtaining the necessary resources, and committing them to the protocol for extended periods of time without expecting consistent rewards, the likes of which Proof of Work (PoW) does not promise.

In fact, the computations put into mining a block by all except one would be wasted. With the hardness of proof of work adjusting with more computing power being committed (due to coin generation rate adjustment) and the consequential limit on throughput, it takes over a year [16] before a single miner could mine one block that gets added to the

Blockchain. This has resulted in formation of mining pools (Refer section 2.3.1). Since self-interested agents would leave individual participation or any protocol-adhering pool in favor of maximizing their rewards, if a given mining pool is capable of generating blocks above average throughput, it can easily aggregate the required majority of processing power.

[17] argues that in such a situation, since the quantity of proofs produced is analogous with rewards obtained, PoW becomes incentive incompatible. They present an attack which could be utilized by a mining pool to maximize rewards, making a 51% attack very much a reality. For example, once the 51% limit is reached, the pool could close itself to new joiners in favor of retaining centralization, and the ability to keep generating more rewards for its participants.

2.3.3 Selfish-mining

[17] presents the "Selfish-mining" attack, where a mining pool keeps the blocks mined by its participants private until the moment is opportune to publish them. Upon finding a new block, it is added to the private chain. When the public chain broadcasts a block, the pool considers the difference between two chains.

- If public chain is ahead, pool synchronizes with it
- If both chains are at same level, pool publishes hidden private block to compete with public block.
- If private chain is ahead by 1, publish all blocks so that private chain becomes the longest.
- If private chain is ahead by more, keep mining at the head and keep publishing to the private chain.

It is evident that this causes massive profit and resource loss for the non-pool participants since they have invested almost as much in generating proofs of work for the existing public chain. Since the internal peers would decidedly admit the pool's blocks, and the external peers would continue to propagate information of the first valid next block found in the network and discard the others [15], the number of internal peers and their communication capabilities also contribute heavily to the success of this attack.

They present simulation results to evaluate the accuracy of their analysis. As depicted from figure 2.1, even if the block propagation was only within the pool ($\gamma = 0$, no honest miners are mining on top of pool's branch), if the pool size exceeds 33% mining power, then the protocol becomes vulnerable to this attack.

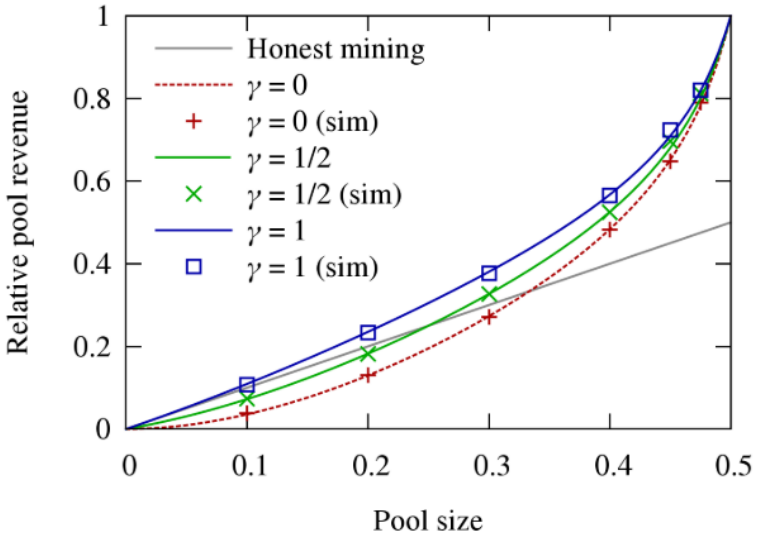


Figure 2.1: Selfish Mining Rewards

Graph from [17] depicting the number of honest participants that must mine on top of pool's published branch (γ) for a pool of a given size to obtain more rewards by selfish mining over honest mining

γ is likely to be a higher value, and Bitcoin protocol provides no threshold limit on this value, allowing even the smallest selfish mining pool to monopolize on rewards. Thus, they propose a solution where honest miners broadcast messages on all new found blocks, rather than the first one, and starts to mine on one of the broadcast blocks at random. This makes $\gamma = 0.5$, raising the threshold of tolerable computing power of a pool from $> 0\%$ to 25%.

The simplicity of the solution and its intuitiveness in being effective even through partial adoption must be appreciated. As pointed out by [15], it is not a trivial task to enforce a change in the protocol without forking the network and thus altering the history, facing much backslash from users who have already committed significant amount of work to the current longest chain. However, it should be considered whether peers would be willing to work on a chain head that may or may not be included in the chain. It seems likely that peers will cooperate given the alternative of receding all possibility of reward collection.

Another important point highlighted by the paper is that the attacking pool can protect itself from penalties and detection, since any detection algorithm deployed on the honest protocol would be public information. We consider this to be untrue, as explained in the following section (2.3.4). They further criticize how the protocol is incentive incompatible while its security remains highly dependent on it. Since pools with $> 25\%$ collective power already exist, they discuss how even their solution would not fully solve this problem.

The presentation by [17] discards block propagation time entirely due to it being negligible compared to proof of work's mining time. Per the ensued discussion, the

crucial γ variable, the choice of honest participants who randomly choose to mine on pool's head depends entirely on the ability of propagation. This is further supported by their discussion on how a pool could conduct a Sybil attack (Refer section 2.3.4.1) on honest miners in order to better disseminate selective information, raising the value of γ to their advantage. While it is true that the connectivity of nodes both internal and external to the mining pool would be roughly similar, we have no assumptions regarding the physical proximity of pool participants. We consider this a minor point overlooked by this research.

In conclusion, we find the insights presented by [17] on importance of data dissemination and incentive compatibility truly useful and acknowledge their heavy influence in leading us to our research problem.

2.3.4 Fair Mining

[16] discusses difficulty in ensuring fairness given how miners would be tempted to

(a) Adopt selfish mining
(b) Include transactions with greater transaction fees, forking the chain trying to obtain more rewards
(c) Form pools with majority power for reduced profit variance, harming decentralization

They propose a solution to ensure that the amount of work put in by a miner is proportional to the rewards they receive. An attacker not being able to significantly exceed their share of rewards by any of the above strategies is defined as "Fairness".

For transaction fee regarding deviations, the solution entails splitting a block reward within a sequence of blocks preceding it. It seems to be an intuitive notion given that blocks following a transaction help establish its permanence in the chain, and distributing rewards this way would compensate for it in an inverse manner. They eliminate the variance of rewards through reducing the mining difficulty, thus eliminating the need for pool formation and limiting the potential of selfish-mining.

The notion of "fruits" is introduced, which are mined simultaneously with blocks. For example, the hardness of finding a suitable fruit hash is defined from the length of the prefix while hardness of block hash is defined by the length of the suffix. Which ever is found first can be published. Fruits contain the actual transactions but bear no incentive to be mined. Unconfirmed fruits are placed in a queue, sets of which are to be included in Blocks. Rewards of such fruits in Blocks can be obtained by miners. Fruits must point to a previous block in the recent sequence of blocks (i.e. the hash of block at head when fruit mining started). If multiple fruits are confirmed at once, the fruit pointing to the farthest block within the accepting range is included first.

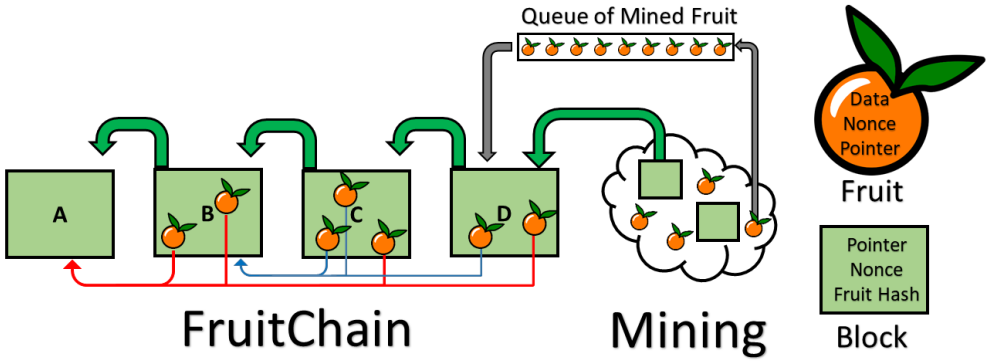


Figure 2.2: Fruitchains
Relationship between fruits and blocks [4]

As explained above, deincentivizing the formation of mining pools is achieved by transferring the hardness of block mining to fruit mining. Since the block a fruit points to must be recent, and this must be true for all fruits in a block, miners must aspire to mine fruits despite them having no incentive. Due to this recency parameter, fruit or block withholding is unlikely to give high proportion of rewards to "selfish-miners". And since the block mining hardness and the number of fruits included in a block is adjusted to reduce reward variance per individual miner (From 2 to 5 years per block to half a day per block), pool formation becomes obsolete.

This paper presents useful insight into the utilization of game theory in modelling consensus protocols, and how it can yield useful security relationships. In their discussion on related work, they dispute on a Nash equilibrium achieved for pools controlling less than $\frac{1}{3}$ resources due to the assumption on 0 communication latency and adversaries not being able to perform network level attacks. As we highlighted in our problem statement, it seems to be an important parameter excluded from many analysis on reaching equilibria. A reason for this might be that including such minute details would further complicate the game theoretical analysis. Therefore we use the concept of self-protection to account for this uncertainty in our proposed solution and the ensure liveness through vote-based consensus.

Apart from the informal proof presented, the authors validate their insights on incentive compatibility by discussing their applications in existing proof of stake solutions, claiming that they are applicable to non PoW Blockchains as well. We approach this level of generalization with caution, given that by their own admission, the protocol relies on monetary rewards. While freshness is accounted for by this protocol, when the block gets added during the allowed period still remains unaudited. Specially since our

[4]Fruitchains Image: `https://flintbox.com/public/project/51008`

focus is on alternative applications of Blockchain, we would like stronger chronology assumptions to be included in a protocol that solves incentive incompatibility.

2.3.4.1 Sybil Attack

As discussed multiple times in the preceding discussion, Sybil attack represents the ability of a malicious node to present itself as multiple identities in order to thwart any known limits of fault tolerance in a distributed system. As the attack's introductory paper [18] claims, it is impossible to defend against these attacks unless a centralized identity verification authority is put in place. They further present that only under extreme circumstances or somewhat unrealistic assumptions can such an attack be prevented.

Many implementations of Blockchain protocols such as Proof of Work, and Proof of stake overcome this attack by expecting the trust to be from a resource expenditure or a given commitment, as opposed to a claimed identity. However, moving to BFT replication protocols, where multiple resource expenditures required can be tolerated by the same malicious object, this attack becomes one of increasing possibility.

While the attack bares a certain resemblance to the problem we discuss, where a malicious node can control an outcome by "speaking for others" in a distributed system, the importance of implementing resilience for the attack comes in our requirement for providing incentive for message propagation. It is superfluous to say that employing a central verification authority as solution would severely limit the scalability and stunt the decentralization requirements of the protocol, and thus not suitable for a Blockchain consensus protocol.

However, a game theoretical analysis is presented by [19] which provides valuable intuition towards a solution that could be applicable to our scenario.

2.3.5 Propagation Incentive

While the requirement for propagation incentive in competitive reward schemes sets the stage for [19]'s problem definition, their solution intertwines heavily with Sybil-proofing reward schemes used for incentive, and with good reason.

The paper relates the example of 2009 DARPA Network Challenge and the concept of raffle draws, where the organizer wants to increase the number of participants while the participants would rather keep the pool smaller to increase chances of winning. For a reward scheme to succeed, a balance must be kept between rewards offered for winning and recruiting.

Elaborating on the 2009 DARPA Network Challenge, where teams had to find a given number of Red weather balloons set loose across the country, they discuss the winning team's strategy of outsourcing the sightings to general public with a per-balloon reward. Upon facing difficultly in dispersing the information regarding the rewards across country, an information dissemination incentive was offered for recruiting balloon

finders, and a second but lesser reward for the recruiter's recruiter. However, this strategy was vulnerable to Sybil attacks, where both recruitment rewards can be obtained if a recruiter pretended to be both first and intermediary recruiter.

[19] rightly claims that introducing propagation incentive would have to be resilient to Sybil attacks. While any such virtual node can not perform heavy computations (i.e. find the balloons, calculate the hash), they would be able to reap message broadcast rewards dishonestly.

The model considers the "distribution" and "computation" phases in a network of nodes structured as d-ary directed trees of a fixed height H. Root nodes get the messages which are propagated downwards to all children. If a given node decides to not propagate the message, it loses the "distribution" rewards it may acquire if a child succeeds in the computation phase. The more children the node has, the higher the probability of rewards it can reap. However, this depends on the competition provided by other siblings and their children sub-trees. If the competition is less, the node is more likely to succeed at computation phase. If the node does replicate itself, it increases the "distribution" rewards that can be collected if a descendent were to succeed in computation phase. However, since the height of the tree is fixed, the node loses the d children at the bottom most level who might have succeeded. This situation is depicted in figure 2.3.

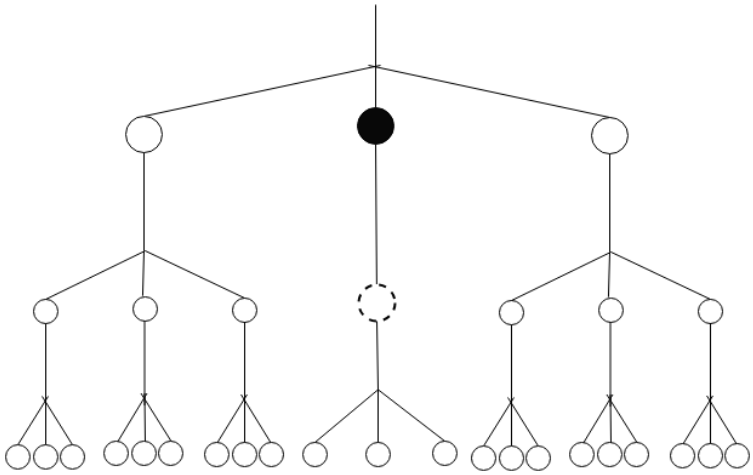


Figure 2.3: A d-ary tree with a duplicating node

A 3-ary tree with a duplicating node in the middle (Black circle), and a second level Sybil node (Dashed circle). Honest nodes (Solid circles) at the first level has more children, while Sybil node allows the duplicating node to collect both rewards if a child succeeds. Due to fixed height of tree, leaf nodes cannot propagate further

General properties of d-ary directed trees provide two potential values for incentive reward which depend on the height of the tree and the constancy of number of top-level parents. A hybrid scheme is designed using these values such that top-level parents are kept limited, and the expected payment is normalized by the tree height, reducing overhead.

The analysis presented utilizes the game theoretic concept of iterated removal of

dominated strategies. If a player has strategies that yield higher utility irrespective of other player's strategy, those strategies are kept while others are removed from consideration. The remaining strategies are considered for other player. If they are dominated by yet another higher utility yielding strategy of that respective player, they are also removed. The process is continued until a remainder of strategies for both players are obtained where their utilities cannot be further maximized.

Using this concept, [19] models an equilibrium for a Sybil-proof Hybrid scheme discussed above. The authors prove that all surviving strategies are non-duplicating, non-blocking for all players for a given order of elimination, and that non-duplicating, non-blocking strategies that are independent of the order of removal do survive, if not exclusively.

As acknowledged by authors in suggested future work, one limitation of the solution is its dependency on a network representing a d-ary trees of a fixed height, which is not an accurate representation of a peer to peer network, nor how a Byzantine Fault Tolerance protocol would have to broadcast messages (Refer section 2.4.1). They do suggest an analysis for the p-normal graphs as future research, used in analyzing oral message solution of Byzantine Generals' Problem in [20].

While the paper defines its problem upon the Proof of Work protocol, it does not consider how the mining delay of a node that intends to keep the mining rewards to itself, or the extensive resource expenditure would serve as incentive for message propagation. They acknowledge these duration differences but elude it in their problem definition. We consider this consideration out of scope ourselves, but only given that distributed ledger applications we consider require higher throughputs where this consideration is void.

We borrow intuition from this paper heavily for our proposed model, since the idea they present regarding balancing the incentive offered depending on the security requirement is a key idea that applies to our research problem as well. Their application of game theory in solving their problem also supports the direction of our research. However, we believe our approach to be more inclusive in its analysis due to increased number of variables considered, such as players having homogenous or heterogenous costs and the protocol having differing tolerance thresholds.

2.4 Byzantine Fault Tolerance

Byzantine Fault Tolerance is another leading area of distributed consensus which spans over several subtopics. Herein we discuss the origin of this problem, practical solutions proposed, and existing applications of these solutions in the context of distributed ledgers.

2.4.1 Byzantine Generals Problem

Byzantine Fault Tolerance derives from the seminal work of [20], "The Byzantine Generals Problem", which introduces an algorithm that allow $> \frac{2}{3}$ of accurately

functioning processes to reach agreement with unsigned messages. They extend their research into another solution under message unforgeability assumptions and proceed to a discussion on communication path reliability between nodes, and the practicality of implementing their solution given the probability of assuring said assumptions. Throughout majority of the work reviewed for our research, findings of this paper prove to be fundamental, given its applications in distributed information processing in computer systems, such as state machine replication with limited replicas and data collection or coordination of sensor networks.

The authors present their problem in an analogy including a number of loyal generals ($> \frac{2}{3}$) trying to reach agreement while the remaining generals are traitors. An enemy city is surrounded by these loyal generals who must decide whether to attack or not, while they can only communicate through oral messages. They must reach the same agreement in order to mount a successful attack, while a small number of traitors deviating from the plan would have no effect. In order to reach agreement, the generals all agree to take the majority vote as their plan of action. A commanding general issues the first vote, and the lieutenants each mutually assure the command they got from the general. In taking the majority vote, lieutenants withstand both traitorous lieutenants and generals as long as they do not exceed the $\frac{1}{3}$ threshold (Refer figure 1.1). Any two loyal generals obtain the same output value whether it means "attack" or "retreat". In an application, this analogy is interpreted as the general being equivalent to a party providing an input over unreliable channels, and the lieutenants being the processes which must reach a common decision regarding the output.

The authors prove that no solution exists for 3 generals with 1 traitor for both exact and approximate agreement through impossibility results, and proceed to present an algorithm as a solution for $3m + 1$ generals with m traitors. The solution assumes correct delivery of messages, protection of message authenticity, and detection on message absence.

Every command received by a lieutenant is multicasted to all other generals. This is a major drawback of their solution since it degrades the performance exponentially as the number of replicas increase. This causes the direct application of the solution to be highly unsuitable for distributed ledgers in Blockchain, but the alternative solutions with additional connectivity assumptions explained later prove to be relatively more efficient. Another point presented in this paper is that if a majority vote cannot be obtained, a default action of "Retreat" is provided in the algorithm. Unlike a computer system, in a distributed Blockchain consensus "lieutenants" themselves (or their enrolled representatives) can generate the inputs. This points to our originally discussed research problem of consensus and chronology manipulation (Refer figure 1.2).

The solution with signed messages eliminates the possibility that a traitor could forge a message sent by a loyal general, and allows the total number of participants to be $m + 2$ at minimum in the presence of m traitors. In the original 3 general problem, a loyal lieutenant could not distinguish between a lying general and a lying lieutenant. But since messages are unforgeable, and contains the "stamp" of each recipient since its inception, lieutenant can identify when conflicting messages are received under

general's stamp. Both loyal lieutenants can use a common function such as median on an ordered list of received orders to reach consensus. Each received message is signed by the receiving node as acknowledgment and then broadcasted to nodes whose signature is not present in the message. For termination of agreement, either a time-out function or a acknowledgment by the last recipient is suggested. Still, in case of indecision the nodes proceed to "retreat", or commit to inaction.

2.4.1.1 Connectivity Assumptions

Authors proceed to evaluate the assumption of direct links of communication between participants for both the oral message and signed message algorithms. While the oral message algorithm requires $3m + 1$ participants to tolerate m byzantine nodes, with nodes placed in a $3m$ regular graph (i.e. complete connectivity), signed message algorithm is proved to withstand a weakest possible connectivity hypothesis. If the loyal lieutenants are connected in a subgraph with d diameter, signed message algorithm can withstand m Byzantine nodes with $m + d - 1$ total participants, derived through following reasoning. Since at least d participants are loyal, traitors must be $m < n - d$, thus, maximum traitors possible becomes $m = n - d - 1$.

This result is advantageous in a scalability perspective, given that in the oral message algorithm, 1 new malicious agent requires 2 more loyal agents for reliable operation, while in the signed message algorithm with subgraph connectivity, no additional loyal agents are required.

It should be noted that this subgraph connectivity problem is encapsulated in our problem definition in the context of switching networks, given that any routing optimization is not straightforward where the subgraph would not remain connected. The authors themselves acknowledge that if a switching network is used rather than fixed lines of communication, byzantine general's problem reappears for faulty network nodes. Therefore our analysis inadvertently includes studying the probability of connected subgraph formation by required number of selfish participants who are also satisfactorily incentivized to invest in self-protection.

Even though the subgraph connectivity of loyal nodes solves the problem of lack of direct communication paths, it does not reduce the communication overhead of messages. All peers must be aware of what each of the other peers have voted, and a simple delivery of the message to all participants will be cumbersome and thus cause throttling. We give ourselves the liberty of excluding communication overhead as a concern, and leave the implementation of optimized message propagation in BFT consensus to future research.

While the number of required honest participation is reduced to $d - 1$ from $2m$ for the signed message algorithm, the unforgeable signature concerns are followed by increased synchrony assumptions. We postulate that they must be facilitated in any case since a relative identity, if not absolute, must be maintained to distribute the rewards. The difficulty of signature replication is also considered a solved problem in the present for network communications. The paper further considers clock synchronization when accounting for synchrony in a processor based application of the algorithm, which we

consider to be superfluous for distributed ledger consensus.

In conclusion, the Byzantine generals problem in a synchronous context is thoroughly studied in this paper, and the presented solutions include many interesting conclusions one could use in designing fault tolerant systems. Their application in Blockchain consensus however intensifies the assumptions that must be enforced to assure correctness, while the assumptions themselves become harder to justify. Therefore, while the intuitions are undoubtedly useful for solution modelling, it remains doubtful whether the expense, connectivity assurance in a global scale, and loyalty assumptions on selfish nodes will allow a consensus protocol to be modeled through direct application of concepts provided by the paper.

2.4.2 Practical Byzantine Fault Tolerance

Software fault tolerance in the face of frequent transient errors caused by ever-growing complexity of current applications, or simply malicious attacks has become an exceedingly difficult problem to solve. Synchronous consensus protocols may not be suitable for these applications, most of which are deployed over the Internet. As such, Practical Byzantine Fault Tolerance (PBFT) [21] is proposed as an asynchronous sate machine replication algorithm which tolerates m Byzantine faults in a network of $3m + 1$ peers.

The authors describe existing solutions for asynchronous consensus to be too slow for practical use or bound by assumptions on known delays of message delivery and processing steps. But given that asynchronous consensus is significantly slower than synchronous consensus, they aim to reduce response times through different optimizations. One such improvement is using public key authentication measures only in the presence of faults, while resorting to message authentication codes in normal operation.

The fault tolerance of the model is based on the assumptions that network could delay, duplicate, drop or reorder messages, and that node failures are independent, but the malicious influence from a strong adversary who has control over all faulty nodes $< \frac{n-1}{3}$ cannot indefinitely interfere with node operations.

The **safety** requirement is defined as assurance of linearizability, which is all non-faulty replicas maintaining the same state at a given time, similar to a centralized service. This property is designed to be agnostic of any synchrony requirements, and remains analogous to consensus finality obtained in synchronous protocols. **Liveness** requirement is defined as eventual response to all client requests.

A weak synchrony assumption of retransmission until receipt is used to ensure liveness, given that [22] proves distributed consensus with even one faulty process in an asynchronous environment to be impossible. Replicas are required to be deterministic in state transitions and to start off in the same state, as per usual state machine replication requirements.

The protocol's interaction sequences between clients, primary and replicas are summarized in figure 2.4.

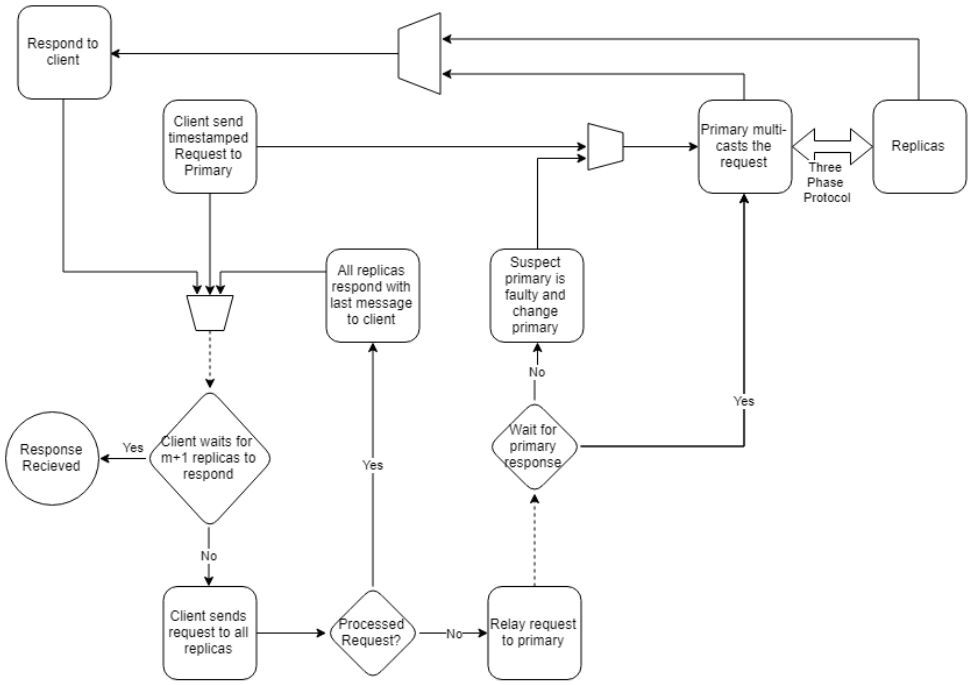


Figure 2.4: PFBT Client, Primary and Replica interactions
Dashed lines indicate timed operations, and sinks indicate that the interaction sequence resumes from there onwards.

The three phase protocol is a sophisticated interaction sequence between primary and the replicas, which consist of pre-prepare, prepare and commit messages. At the completion of the protocol, a "view" is considered as completed and its number incremented. Primary of the succeeding view is selected through the modular operation on view number and number of replicas.

Primary broadcasts a pre-prepare message with necessary security parameters (such as request sequence number for ordering and message digest) and the current view. Replicas accept the message after cryptographic and sequence number validations, and multi-cast a prepare message. All replicas accept the prepare message after validating security parameters, and once $2m$ prepares are received ($2m + 1$ including self), it is validated against the previous pre-prepare message. When these conditions are met, each replica multi-casts commits to other replicas with a commit message. If $2m + 1$ messages are received, the request is performed and client is responded to. Once $m + 1$ responses are obtained by client, the request is completed. This multi-level agreement in prepare and commit phases ensure that requests are committed in order (per timestamp) across all replicas in the same view without the need for synchronous communication.

Since ordered message delivery is not guaranteed, requests can be committed out of order, but all requests are ultimately responded to by the majority, and therefore state transitions remain constant. All three phase messages are logged for this purpose, and checkpointing is conducted only on requests which have response logs in at least $m + 1$ replicas. Checkpointing itself requires $2m + 1$ sequence number validated messages to

ensure its proof of correctness.

In case of Primary failure, (Refer figure 2.4), View change messages are generated by replicas whose timers (started after receiving client request) have expired. View change message increments the current view, and thus a new primary is selected, who waits until $2m + 1$ View Change messages are received to broadcast a "New View" message, upon which the system is considered ready to accept new requests. Requests since last checkpoint up to maximum sequence numbered request are redone at the new view, and missing information at one replica are obtained from another if required. Therefore, this "View change" mechanism ensures the order of requests at different views, and since the order of requests at the same view is ensured by prepare and commit phases, the safety requirement of the protocol is fulfilled.

Liveness is facilitated through dynamic adjustment of waiting times for request execution. If a new view isn't initiated during the designated time period, view number is again incremented but the waiting time is increased by twofold. To ensure that view-change occurs before above occurrence, replicas only have to wait for $m + 1$ view change messages to generate their own, irrespective of timeout occurrence. The time adjustment is protected by faulty replicas' attempts of illegal view changes through requirement of $m + 1$ requests. Here the assumption of not being able to indefinitely delays messages becomes significant, since the timeout period grows while the delays are unlikely to do so.

The ambiguity induced by certain service specific property values are also discussed, and an additional step of byzantine vote is suggested, such that the median value of replica suggestions could be accepted as the universal value to be used as the configuration property under question. Performance improvements gained from modes of operation dependent cryptographic parameter complexity are also discussed.

The paper is concluded with a comprehensive evaluation of a fault tolerant file system implemented using the PBFT protocol, whose performance is compared with other state of the art sate machine replication or fault tolerant algorithms. This section is not reviewed specifically given the obvious difference in use cases for Blockchain vs File Systems.

An important point highlighted by the paper is that synchrony assumptions always pose a danger of attackers who may delay non-faulty nodes until they are excluded from the consensus. This can be easily done as opposed to fully compromising a non-faulty node, causing the attack much more probable than a coordinated attack from $m + 1$ Byzantine nodes. In state machine replication, the risk is heightened due to reduced number of nodes. But even in a Blockchain peer to peer network, such an attack is not far fetched, as explained in [5] and [23], reviewed respectively in sections 2.5.1.2 and 2.5.2.1.

In multiple round based protocols, selective interference problem becomes extended due to widened attack surface. But one advantage of asynchronous protocols is that nodes are not excluded upon failure to respond. The interference by an attacker is only effective if it is indefinite. If this was a concern then it could be possible to blacklist the respective attacker and continue the protocol as usual. But it is worthwhile to

acknowledge that in the interest of latency, a blockchain protocol might not be able to afford this level of communication overhead, nor the step-wise synchronization given the larger scale of the peer to peer network. A similar protocol Tendermint, overcomes the communication overhead issue through use of gossip protocols, as well as the liveness property, as discussed in section 2.5.2.2.

Retaining all replicas despite failures is important when nodes can either be misclassified as faulty and eliminated from the network under an attacker's influence. Thus, replica set being immutable is an interesting security property implemented by the protocol. But for networks with fluctuating number of peers, assurance of this property would be of high importance for security assurance. The significance of this paper to our research lies in the above requirement extraction. While PBFT provides an intriguing asynchronous alternative for distributed ledger consensus, we still find synchronous consensus more applicable in the context of our research due to aforementioned scalability and throughput concerns.

2.5 Derived Consensus Protocols

Proceeding from a discussion on PoW, and BFT, we summarize few other protocols in existence which claim either energy efficiency or much less latency or both as candidates for alternative Blockchain applications in non-financial domains. Following each summary, we analyze their suitability and their ability to retain the important properties we identified from previous protocols, and extract other features that prove insightful in modelling our problem.

2.5.1 Mining Based Protocols

2.5.1.1 Proof of Stake

PPCoin [24], introduces the concept of proof of stake as a favorable alternative to the resource intensive PoW protocol. They eliminate the need for extensive energy consumption in establishing security by replacing it with the concept of coin-age instead. Coin-age is determined per each user through timestamped transactions, and is calculated by multiplying the coin value with the period it was owned by a given user. As transactions are validated, coin-age is consumed, making it a resource as scarce and unforgeable as the computational power required for PoW mining.

A miner, generates a proof of stake block by paying himself and consuming his own coin age for the opportunity of generating a new block. If the block is added to the chain, the mining rewards are obtained by the miner. The mining is done over a hash space limited by the stake committed, providing an upper limit for the possible energy consumption, in direct contrast with PoWs unlimited search space. The protocol is further tuned for hash generation efficiency, where the difficulty of meeting hash target is reduced as the consumed coin-age becomes greater.

The blocks are scored by the sum of consumed coin age of each included transaction, and the block with maximum value extends the chain. The translates to the PoW's longest chain principle, allowing similar agreement to be reached regarding chain forks. The blocks mined include owner's signature and the respective proof of stake reference, such that if many blocks were created using same proof of stake, nodes have sufficient information to discard them until a valid block with a new stake reference is broadcasted.

Competing miners are prevented from hording transactions for their transaction fees through elimination of paying transaction fees to the block owner. Instead, they are integrated at the protocol level in defense of Block bloating attacks, and can be aggregated and distributed as block rewards. Cost of 51% attack is also raised alongside the replacement of computing power with stake of coin age, since even if such majority stake could be obtained, the coin-age is fully consumed during multiple block additions.

However, such stake allows an attacker to reorder the block history or perform double-spending attacks, to which the authors propose centrally broadcasted frequent check-pointing as a solution. The authors rightly note here that this property disrupts full decentralization, but justifies it given that a similar mechanism was introduced for Bitcoin and PoW as well.

In PoW design, the long term sustainability of miner incentive is excepted to be borne by transaction fees while block rewards themselves become reduced [2]. After eliminating transaction fee based miner revenue entirely, proof of stake uses reduction of participation costs in ensuring Blockchain's continuity. Peers are no longer required to make up for the energy expenditure through their mining revenue, nor are they required to waste as much energy, and thus, they have no motivation to withhold transactions which declare higher transaction fees. This is a noteworthy optimization of the proof of work protocol, which further serves as a desirable security property given that transaction fees are a known cause of instability once block rewards have been sufficiently subsided. (As illustrated in section 2.5.1.2 Proof of Activity).

However, while its goals in energy efficiency and increased tolerance of 51% attack are fulfilled, it does not have significant throughput improvements to be applicable for non-financial Blockchain applications. In fact, the throughput could be lower than proof of work itself, given that nodes will have to wait for days to mine new coins. This creates income variations which may cause pool formation, making the subsequent problems of selfish-mining and others much more severe given that hash generation complexity is also reduced. The protocol further suffers from potential centralization, given that less resourceful miners would be starved due to lack of coins. In any case, we consider the protocol noteworthy in its unique insight of stake based mining, and how a scarcer stake than computing power can be defined in a creative manner to solve extensive energy wastage of proof of work.

2.5.1.2 Proof of Activity

[5] specifically focuses on the how stable the security of a proof of work based Blockchain would be once the block rewards are subsided in favor of transaction fees. Miners will

be in competition to mine any available transaction into their blocks, given that they cannot afford to wait for transactions with higher fees. Similarly, the transaction issuers won't be offering higher fees since they have no incentive to do so. Ultimately, the costs of proof of work would not be met through transaction fees and the network will collapse. Thus, given the insufficiency of proof of work in maintaining the network health, and proof of activity is proposed to solve this problem.

The article highlights the "Tragedy of the Commons" phenomenon from game theory and economics, which is when each participant of a system assumes that others would follow the protocol and therefore he or she can maximize their utility without any repercussions to the social welfare. Price of Anarchy (Refer section 2.8.3) is a game theoretical concept which measures this social welfare degradation. The authors continue to explain how "Tragedy of the Commons" occurs multiple times in a proof of work network exclusively dependent on Transaction fees.

Security (i.e. Health and stability) of the network depends on miners being provided sufficient funds to maintain decentralization and immutability of the ledger, but the rational users would prefer not to pay assuming that others would sufficiently fulfill this requirement. Rational miners could choose to reject transactions with insufficient fees, since the alternative would be freefall of transaction fees allocated, diminishing their return of investment. But similar to rational user behavior, some miners would prefer to obtain higher rewards by including lower fee transactions while delegating the network health maintenance to others, ending up back at the original problem.

A transaction fee limit per block enforced on miners by the protocol is proposed as a potential solution, where miners will be forced to hold agreement, causing reduction of Price of Anarchy and increase of collective rewards (i.e. Social Welfare). Block space becomes a resource competed against by users, and miners are forced to select a set of transactions per block which collectively exceed the fee limit requirement. However, it is doubtful whether this is sustainable since transaction fee limit grows with volume of transactions (per user competition), becoming a more attractive target to attackers. Therefore the ultimate cost of security maintenance increases and the solution becomes yet again, unsustainable.

Proof of Activity, a proof of stake based protocol, which has low overall costs due to stake-based miner selection, is proposed as a solution which makes the transaction fee limit solution sustainable.

"Tragedy of commons" recurs again at transaction propagation level, (as discussed in section 1.3), and limits on data size and CPU cycles are proposed as a solution. This would cause unpropagated transactions to expire, rendering them useless to the miner.

Proof of work miners also have the capability to alter their hardware resource commitment between most profitable cryptocurrencies, and might further be unaware or indifferent to the honesty or malicious nature of their commitment. Such volatile behavior makes them unsuitable as exclusive authority on network security maintenance, justifying the partial delegation of responsibility to stakeholders. Authors present the rationale that dependence on honesty of online stakeholder majority is a more sound decision than

depending on honesty of majority of computing power.

A subroutine for choosing the stakeholder is as follows. A random coin unit is selected from all mined coins, and its transaction history is followed up to its current owner. Thus, a stakeholder with more coins hold more chances of being selected. N stakeholders are derived within the protocol, where $N - 1$ stakeholders validate an empty block as the next block candidate, and the Nth node includes the transactions and finalizes the empty block header. Fees are distributed among the N stakeholders.

If certain stakeholders are offline, the selection process is repeated with adjustments to the mining difficulty. The calculations presented discuss how a potential attack would need to generate the hash for the empty block, while correctly predicting and influencing the N stakeholder nodes. The probability of succeeding at both being quite low, demonstrates the high resilience against both computing power based and stake-based attacks. The dependence on stake is sufficiently decentralized through the N stakeholder selection.

We agree with the authors' analysis regarding transaction fee induced instability. An additional concern we point out here is that if transaction fee limit is exceeded beyond a reasonable amount, given the "tragedy of commons" scenario described above, it fails to enable micro-payments, which was one of the original aspirations for seeking a decentralized market. Thus, a monotonically increasing transaction fee limit would render the protocol pointless if the limits were not kept at a reasonable level.

Incentivizing stakeholders' online presence through randomized stakeholder selection and mining difficulty reduction, is another aspect that can be used to justify self-protection costs of nodes, which is a requirement in our proposed model for defense against selective interference. We note here that this research supports our attempt in ensuring constant availability of miners, making it a reasonable expectation to be borne by peers.

The protocol establishes enhanced network security through availability assurances, and causes less overall energy consumption. While the limit enforcement on data size and CPU cycles presents an interesting alternative to incentivizing transaction propagation, a question to answer would be how this enforcement is ensured in a protocol open to manipulation by end-users. It further proves to be a security problem for the end-users privacy, but this concern can be considered negligible in case of mining with dedicated resources.

In conclusion, the paper addresses certain problems in both proof of work and proof of stake protocols. It solves proof of work's transaction fee based instability problems through stakeholders, and Proof of stake's centralization issues through randomized stake owner selection and subsequent reward distribution. It further sheds light on the selfish behavior of an often overlooked party in consensus calculations, the transaction issuers.

2.5.2 Voting Based Protocols

2.5.2.1 Ripple

Ripple [23] tries to solve the latency problem in Blockchain while assuring fault tolerance against Byzantine Failures through trusted subnetworks centralized around a chosen node. Collectively, this allows 80% of the network to be trusted, providing a 20% Byzantine node tolerance.

Ripple's servers (i.e. miners) maintain a Unique Node List (UNL) of peer servers whose consensus determines the agreement, as opposed to the entire network. In each consensus round, servers propose client transactions to be included in the ledger and the collective agreement of this subnetwork allows the transaction to be added to the ledger. To ensure that only one consensus is reached amongst all subnetworks, (preventing forks), intersection of servers (which are cryptographically identifiable) in any two UNLs is required to be greater than 20% of the number of nodes in the larger UNL. Thus no two subnetworks can reach differing consensus, accomplishing the consensus finality requirement.

The correctness requirement is considered to be satisfied with sufficient assurance of UNL node diversity, which in turn minimizes the probability of a node's decision to collude. Given the unlikelihood of union of powerful colluding cartels assured by diversity factors, the paper claims that this approach yields much lower collusion probabilities in reality with $< 20\%$ being the highest possible tolerance. The probability of consensus thwarting cartel existence is analyzed through simulations and is shown to drastically drop with the increase in size of UNLs when the user probability of collusion remains below 0.15.

Given that the steps of transaction acceptance, transaction set proposal, accuracy voting and agreement are fixed time operations, already latency-optimized by design of the protocol, the objective of preserving/maximizing utility falls primarily on ensuring low network latencies. A bound is obtained for the maximum tolerant latency, and nodes that exceed it are excluded from all UNLs, guaranteeing convergence within a fixed time. In such a scenario, the remaining nodes must hold the aforementioned correctness and agreement guarantees for protocol's continued resilience.

Protocol further integrates the capability for finding and flagging identifiable maliciously behaving nodes, and a split detection algorithm for nodes facing temporary network latencies allowing them to acknowledge their presence, avoiding network latency induced fork possibilities. To avoid the latency overhead by slower nodes, a multiple round based structure can also be utilized where the consensus requirement is increased at each round, concluding in 80% requirement. This allows identification of slower nodes, assuring eventual latency optimization. A default list of nodes is provided optionally for all nodes to choose from, allowing even new users to directly participate in consensus, and for any two node lists to obtain required intersection.

The authors acknowledge that while their security properties are not optimum, they are well understood and suited for rapid convergence, node count fluctuations, and varying

network topologies, justifying the lower tolerance thresholds. Since the original goal of the protocol is utility maximization these shortcomings can be seen as acceptable. Note that this is not a reference to game theoretic notion of utility, but the utility definition of the protocol which aims to optimize convergence latency and computing power usage.

While the transactions are applied deterministically, and thus double spending fails once one of the transactions are included in the ledger, the chronology can still be compromised through network based attacks. This occurs due to the problem of "committing to inaction", since the protocol either discards or pushes transactions back to next validation queue when they fail to receive required votes. This can be easily used by an attacker to their advantage, despite the diversity requirement's assurance of correctness.

Given the condition that the remaining network must hold the UNL intersection properties mentioned, we cannot definitively declare node exclusion scenarios proposed for optimized latencies as a security weakness. But again, in a volatile network susceptible to denial of service attacks, attackers can target non malicious nodes randomly in shorter time spans, causing their exclusion at many rounds of consensus, reducing the final number of servers in the UNL up to a point where cartel formation becomes probable. However, such a sophisticated attack would be highly unlikely.

The correctness assurance through diversity itself has to guarantee the integrity of the server diversity profiles. But it can be assumed with good authority that strict UNL intersection requirements would sufficiently thwart any impersonation attacks, unless they are conducted in massive scale exceeding the Byzantine agent tolerance.

The UNL intersection property also resembles the randomized distribution mechanism we later highlight in the literature review, which has game theoretically proved optimization properties (Refer section 2.8.2). However, our model does not depend on claimed diversity properties but the property of security investment commitment, and optionally the presumed heterogeneity of peers and the volatility of the network, and therefore is more robust than aforementioned mechanism.

The eventual asynchronous consensus through multiple rounds of non convergence, while latency optimizing, might not be security optimum as authors have proposed due to possibility of "committing to inaction" at each of such rounds. While the high throughput provided by Ripple through tradeoff between latency and centralization may be optimum in the cryptocurrency context, we observe it to be not suitable for chronology ensuring distributed ledger maintenance.

2.5.2.2 Tendermint

Motivation for the Tendermint protocol [25] stems from PoW's expensive resource consumption, and the subsequent proposals which solve this problem being dependent on trust assumptions (i.e. Proof of stake, Proof of activity). The author argues that for PoW, security analysis is complicated by self-interested participants while these existing solution security analysis is complicated by trust dependencies. They propose a

BGP based solution with quantifiable security properties. The original BGP tolerance threshold of $< \frac{1}{3}$ is maintained and the solution simply requires partial synchrony and peer ability to keep time.

Paper criticizes usage of extrinsic factors such as computing power for security assurances in Blockchain, since they cannot be measured (users can buy more resources). Direct usage of intrinsic measures such as stake is also criticized for the *nothing at stake* problem they impose. Users could collect enough stake and reorganize block history, successfully conducting double spending attacks with no repercussions. This is avoided in proof of stake [24] through checkpointing and stake age prerequisites. But users only lose coin-age, and therefore the author's argument that existing protocols do not solve the nothing-at-stake problem stands. Furthermore, proof of stake transfers the security model from a computational impossibility bound one to a trust bound one, again making security analysis unquantifiable.

Therefore in order for a proposed solution to make honest participation of self-interested agents quantifiable, extrinsic dependencies should be eliminated by transferring security insurance to a intrinsic stake-based mining mechanism, while simultaneously augmenting it with an intrinsic penalty to replace trust requirement with a rationality based model. If the penalty is high enough, the users would not attempt to attack the network.

Tendermint validators stake their coins in the form of a bond deposit, as opposed to [24]'s coin-age expenditure. Each validator is given voting power equivalent to the amount of bonded coins, and consensus is reached when $< \frac{2}{3}$ of votes (in the form of validator signatures) are received. Coins can later be unlocked for expenditure when a predesignated time has passed after an "unbonding " transaction is performed.

Tendermint block structure contains a height parameter containing the proposed block's location in the chain. If two blocks have been validated with the same height, the intersecting validators have signed deceitfully and therefore double-spending attack is detected. This information is committed as an "evidence" transaction, upon which the guilty players' bonded coins are destroyed. Evidence transaction has to be committed before attacker has time to unlock their bonds, and as such, unbonding period, and network synchronization are imperative for resilience against double-spending attacks.

For such two blocks to exist, greater than $\frac{1}{3}$ of validators who has signed duplicitously for both Blocks must also exist, causing considerable amount of coins to be destroyed in attack attempts. Bonded coins to voting power ratio can be adjusted so that minimum cost of attack is more expensive than possible gains of successful attacks.

The protocol poses many similarities to the PBFT protocol [21] reviewed in section 2.4.2, both in its round-based protocol and synchrony assumptions for liveness and safety, and therefore we restrict its analysis to pointing out their similarities and main differences.

PBFT's pre-prepare, prepare and commit messages and phases resemble Tendermint's propose, pre-vote and pre-commit messages, while Client response and New View messages resembles the commit and New Height messages. Somewhat similar to primary selection, each phase has a proposer selected similar to the primary in PBFT.

Since proposer varies, pre-vote commits are collected by nodes as a "proof of lock" in order to submit with pre-commit if chosen. One difference is the locking at pre-commit phase until enough agreement or disagreement votes are received, which somewhat resembles PBFT's message digest and view number validation mechanisms.

View change message is abstracted into a step where a CommitTime is set before transitioning to a New Height. This isn't a significant implementation difference given that views serve multiple requests and require history synchronization while Height only concerns itself with one block at a time. Certain properties of views can also be seen in alternative steps. View change timeout increment in case of primary failure is implemented for the voting rounds themselves to ensure eventual consensus. View synchronization can also be compared to Tendermint's property of lock release upon receiving a proof of lock with greater height or later round than that of the locked block.

The lock mechanism at pre-commit step ensures safety of the protocol, and deadlock possibility is removed by lock release caused upon receiving a "proof of locks" proposal at a later round of consensus, ensuring liveness. These assurances also bear similarity to approaches taken by PBFT in each aspect.

As a Blockchain protocol Tendermint is required to implement certain features unrelated to state machine replication, incentivizing cooperation being the first. Tendermint briefly analyzes this by deriving the benefit function of two validators who exclude each other in retaliation, and concludes that it simply causes other validators better benefit while they each reap lesser rewards, preventing any motivation to do so.

Secondly, PBFT's point to point communication is obviously too expensive for massive peer to peer networks, and therefore peer to peer gossiping is used by Tendermint for message propagation purposes. This is a major difference between the two protocols.

Another concern of Tendermint is offline validators, whose coins are automatically unbonded after a time-out. But since this changes active validator count, incentivizing explicit unbounding transactions is proposed in the form of a penalty. If the transaction is explicit, then active users can adjust immediately rather than waiting for timeouts, which opens an avenue of attack.

We agree with the concerns presented by this paper regarding trust dependencies. Such as in proof of work, they must be assured through solid incentive principles such that the security properties are quantifiable, an aspect that makes game theoretic modelling suitable for our problem.

A noteworthy feature of Tendermint is the disagreement votes casted by nodes who obtain inconclusive verification results, called "nil" votes. It is an intuitive inclusion that allows detection of active notes in the network, providing a solution to the problem of "committing to inaction". Agreement only need to exceed $\frac{2}{3}$ of all received votes, including "nil" votes.

Incentivizing users to explicitly commit unbonding transactions through penalty enforcement, is another property of Tendermint which validates node level availability assurance we propose for our protocol as a reasonable requirement. Notion of penalty can also be

introduced in non-financial blockchains as well, but given the non-financial rewards the miners are likely to receive, such penalties may make the network unsustainable and unattractive to miners.

In conclusion, Tendermint is a unique protocol which assures asynchronous consensus, along with higher throughputs and consensus finality in an energy efficient implementation. Even though it has a scalability of supporting about hundred nodes at maximum, it is significantly higher than traditional BFT replication protocols, proving it to be a practically beneficial implementation.

2.5.3 Scalability of Consensus Protocols

While section 2.2 provides a summary of possible applications of Blockchain technology outside financial domains, a problem either casually overlooked or simply dismissed in favor of permissioned Blockchains in this section, is the scalability.

[3] provides a comprehensive analysis regarding this problem. They categorize the consensus protocols to two groups; proof of work based and BFT replication based, and discuss the throughput and scalability tradeoff in choosing between these protocols. It establishes that most proposals for Blockchain based applications require much higher performance than what PoW can provide, given that Bitcoin facilitates about 7 transactions/second[5] wherein throughputs of 20000 transactions/second or higher are required for applications such as credit card payment processing. The interests of domains external to financial markets in seeking distributed ledger based solutions are mostly rooted in the business enrichment nature of Smart Contracts. For example, as observed in the case of [9], musicians advocate for smart contract enabled instantaneous royalty payments over payments delayed by intermediary parties. Therefore going beyond PoW and looking for ways to ensure higher throughputs has become a time-sensitive topic of research.

Recalling from section 2.3.1, PoW dynamically adjusts mining difficulty so that block frequency is kept at a constant rate. Reducing block size allow faster block propagation through the network and thus faster insertions causes forks to occur at faster rates, increasing success probability of double spending attacks. Therefore, these 3 PoW parameters are in a state of convergence where any tradeoff attempts will weaken the protocol's security assurances. On the other hand, PoW cannot assure consensus finality for transactions. Confirmation of a block requires multiple following blocks to be added, and given the usual 10 minutes per block frequency, PoW becomes a highly unsuitable choice for latency critical or insertion order conscious applications. Lack of consensus finality is the cause for existence of forks themselves, resolution of which would also increase transaction latency, or even cause transactions to be cancelled.

It was sufficient for State Machine Replication protocols to have limited scalability of a

[5]It should be noted here 7 transactions/second is the transaction throughput for Bitcoin network, and different cryptocurrency networks and protocols they may utilize would have differing transaction throughputs depending on the hash function, size of the blocks mined, frequency of forks and the manner they are handled by the protocol.

handful of replicas, considering that such systems only had to ensure fault tolerance and linearizability. For reasons such as PoW's throughput limitations, and the industrial preference for permissioned Blockchains which does not require network to scale indefinitely, the BFT variants of State Machine Replication protocols with 33.33% arbitrary input tolerance has become favorable over PoW, given its ability to support a wider range of applications.

BFT variants quite amicably satisfies the throughput requirements of Blockchain applications outside financial domains. Through node availability assurances (Protection against "selective interference") obtained by "nil" votes of Tendermint or Liveness properties in PBFT, BFT can also provide complete defense against "committing to inaction", ensuring insertion order. BFT operates at network speed latencies, provides consensus finality, and given above defensive strategy integrations, could also facilitate security for versatile Blockchain applications in a much larger domain space. However, the protocol does have a high communication overhead. For n participants $O(n^2)$ messages are needed per round of consensus. Given the requirement for synchrony, scaling to a higher number of votes could throttle the network, diminishing the throughput advantage entirely. Even asynchronous consensus protocols such as PBFT and Tendermint have weak synchrony assumptions to ensure liveness, which is another must have requirement. We believe this extensive communication overhead can be reduced through methods suggested in [20], where subgraph connectivity of non-malicious nodes could enable less intensive propagation requirements, which we can actually provide through node availability assurances.

The authors further mention that these BFT replication protocols to not have been tested for more than 20 nodes. With such reduced node scalability, the protocol becomes centralized and immutability protection of the ledger falls on a small number of nodes, making it vulnerable to modification. However, a similar centralization based concern is PoW's mining pools, concocted by miners who want a more stable income for their expenditure commitment. Given that this phenomenon results in attacks such as selfish mining, which causes PoW tolerance threshold to drop to as low as 25%, a percentage exceeded by some mining pools in existence, centralization seems to be a common problem for both protocols with BFT providing better resilience (33%).

Another limitation of BFT protocols is the permissioned node participation, which they must maintain in order to prevent Sybil attacks. However, these limitations are seen to have overcome (at least partially) by protocols such as Ripple, which use diversity and intersection based UNLs ensure scalability and open participation. Even though these protocols have their own limitations, their practical success suggest that the idea of a scalable BFT replication protocol is not far-fetched.

The authors further highlight that alternative Blockchain applications would not require to be permission-less anyway, given that the associated legal compliance would require authenticity assurances of content origin, a concern not present in their financial transaction processing counterparts (As long as you possess the required amount for payment, your identity doesn't matter). This is a claim we feel inclined to agree with, considering most alternative use cases we discussed in section 2.2. But we recall here

of Flint's water crisis [10] where officials were involved in foul play, and point that in terms of security, it would be better to aspire for permission-less implementations of BFT protocols.

We present a compact extension to [3]'s descriptive comparison of PoW and BFT protocols in table 2.1. We extend their comparison horizontally through showing how other protocols discussed previously fare against PoW and BFT, while extending it vertically to include certain security and operational properties we identified as desirable in the literature review.

The comparison follows a discussion regarding ongoing research efforts that improve on the limitations of both protocols. The conflict resolution algorithm GHOST (used by Ethereum platform), and the leader election based BitCoin-NG are mentioned as performance gaining improvements on PoW, both of which however fall short in providing BFT-like throughput, and continues to suffer from lack of consensus finality. A scaling strategy is also mentioned, where blocks are stored in a directed acyclic graph as opposed to the "chain of blocks", so that non-conflicting transactions can be mined together and added as a new block. Optimistic BFT protocols, where byzantine vote is considered only in case of unstable networks (reducing the $O(n^2)$ message requirement when network is stable) and Randomized BFT, where correctness is non-deterministic but assured with high probability are mentioned as variants of BFT which trades off certain security properties in favor of performance. Parallelization, where independent requests are processed at different replicas, and delegating cryptographic operations to hardware are also suggested as avenues for performance improvements in BFT protocols.

In conclusion, alongside the analysis of [3], we have presented the shortcomings of PoW which are somewhat closed to modification, and thus halts us from achieving our desired security and operational properties. In contrast, we mention the advantages of BFT, and how various other protocols have overcome its limitations in a satisfactory manner.

Thus, we choose to base our proposed model on a hypothetical BFT protocol and attempt to conceptually integrate the desirable security properties of protocols discussed in preceding sections. One misgiving in this approach is the lack of correctness proof in many of these protocols which demonstrate their security through strategic reasoning or their practical implementation's success and adoption. This is where we enter into the domain of game theory, which allows us to quantify our security assurances and obtain equilibria and respective environmental conditions within which our proposed model would remain stable.

2.6 Game Theory As A Solution Concept

In choosing a suitable theoretical approach for solving a problem, the pros and cons of the approach must be carefully assessed. With this interest in mind, we review two complementary papers in establishing game theory as a field that yields solutions with practical applications. In the first paper [26], Goeree and Holt address some of the criticism directed towards the "rational" player plausibility, and analyzes certain

Table 2.1: Desirable Security and Operational Properties of Blockchain Protocols

Protocol Property	PoW	PoS	PoA	BFT	PBFT	Ripple	Tendermint
Decentralization	High	Medium	Medium	Low	Low	High	Medium
Miner scalability	High	High	High	Low	Low	High	Medium
Miner identity	Open	Open	Open	Permissioned	Permissioned	Open	Permissioned
Synchrony required	No	No	Partially	Yes	For liveness	No	For liveness
Liveness	No	No	No	Yes	Partial	Yes	Yes
Throughput	Low	Low	Low	High	High	High	High
Latency	High	Medium	Medium	Low	Low	Low	Low
Throttle	Medium	Medium	Medium	High	High	High	Medium
Correctness of consensus	No	No	No	Yes	Yes	Yes	Yes
Committing to inaction	No	No	No	Yes	No	Yes	No
Consensus Finality/ Linearity/ Safety/ No forking / Deterministic insertion	No	No	No	Yes	Yes	Yes	Yes
Security Reliance (Quantifiable Measure)	Hardware	Trust	Trust	Trust	Permission	Diversity	Monetary
Reward structure stability (Secure distribution of transaction fees)	No	Yes	Yes	Yes	Yes	Yes	Yes
Violation Consequences (Nothing at stake)	None	None	None	None	None	Exclusion	High
Reward variance / Incentive Incompatibility	Yes	Yes	Yes	No	No	No	No
Propagation Incentive	No	No	No	No	No	No	No
Tolerance to node count fluctuations	Yes	Yes	Yes	No	Yes	Yes	No
Power Consumption	High	Medium	Low	Low	Low	Low	Low

equilibria alongside the noise derived by diverse sets of players. In the second paper [13], Papadimitriou discusses how suited game theory's mathematical abstractions are in modelling the massive network of Internet.

For readers unfamiliar with game theory, we suggest referring to [27], [28] and [29] for ease of comprehension in following sections. Optionally, we repeat some of the game theoretic definitions we have mentioned throughout the prose in Appendix A as a centralized source of reference.

2.6.1 Noise and Game Theory

[26] addresses the criticism directed at the "perfect rationality and foresight" required of players in games, and the subsequent argument that this makes game theory incapable of interpreting real-world data. It analyzes this phenomenon with stochastic game theory under 3 aspects; evolutionary dynamics; learning dynamics and noisy introspection.

Evolutionary dynamics can be described as players as a community restricting themselves to actions that provided them best payoffs in the past, so that the future iteration payoffs keep improving.

In a game of choosing "effort levels", where the payoff function is the minimum reduced by a proportion of each player's original estimation ($u_i = min(g_1, g_2) - cg_i$), Nash equilibrium predicts best response to be a common effort estimation by the two players. But the real-world experiment obtains different results for different values of c. For a lower cs, effort estimations linearly rise while for higher c, effort estimates linearly decline. The latter is not observed by Nash equilibrium in any form. The authors proceed to demonstrate how player decision evolution over time, proportional to the derivative of payoff, can be captured through *Fokker-Planck equation*, and present it as a model which interpret these observed real-world results with moderate accuracy.

Learning dynamics describe the "reinforced learning" aspect of each individual player, who takes an action based on some probability and then assign it an increased weight for the next iteration if it provided a better payoff. Thus, given the intuition that "decisions with higher expected payoffs" are chosen more frequently, "logit probabilistic rule" is used to interpret behavior in reinforced learning. The rule states that "probability of a decision to be proportional to an exponential function of its expected payoff". Given that the expected payoffs are normalized by respective noise, equation allows all choices to be equally probable when noise goes to infinity, but simultaneously, choice probabilities become highly varying as noise reaches zero.

This behavior properly explains the data authors present from a social dilemma game, where two participants are expected to guess two numbers within a range, and the players are rewarded with $u = min(g_1, g_2) \pm R$ in favor of the minimum guess. Nash equilibrium of the game predicts that all players will guess the minimum, as per the perfect rationality and foresight assumption. In the actual results, the variability of choice probabilities are observed with respective slopes for each potential loss R, where highest loss causes them to quickly iterate up to Nash equilibrium (high variability),

while lower losses vary closer to their original, higher predictions (less variability).

Static games, occurring only once are devoid of learning opportunities. Therefore they are naturally expected to converge at Nash equilibrium. But the authors present the concept of iterated introspection of other player's noisy decisions, and apply "logit probabilistic rule" in interpreting lack of Nash equilibria in certain static games. The illustration that follows shows that resulting choice probabilities to be independent of belief probabilities (Originally, the learned probabilities in iterative learning dynamics), and the authors express the need of relaxing perfect correlation assumptions of beliefs and choices in correctly interpreting such naturally observed static game outcomes.

The paper highlights the effects of stochastic manifestations such as errors and unforeseeable human preferences under iterative conditions. It further proves game theory to be capable of producing practically applicable results, given the accurate representation of noise, as opposed to a theoretical model with no practical applications.

The paper reasserts some basic human behavior, such as overestimation when payoff is higher, in gambling or in time estimations, and the eventual adjustment which finally reaches close to Nash equilibrium. It also provides the intuition that any potential simulation for a game design involving human behavior will have to have some noise introduced for it to represent realistic results, if it doesn't already show properties of learned behavior.

This type of noisy observations were also observed in the Stylized version of Keynes Beauty Contest Game[6] presented in [30] . The game play is as follows. Players name an integer within a known range, and the player who names $\frac{2}{3}$ of the average integer wins the game. If multiple players tie, someone is chosen randomly as the winner.

The rational player first considers the natural average of the range μ, and considers $\frac{2}{3} \times \mu$ as the winning value. But given the assumption that other players are rational, and therefore are likely to name this value, the next iteration suggests naming $(\frac{2}{3})^2 \times \mu$. All players follow this reasoning in many iterations until they cannot go beyond the minimum value. This is the Nash equilibrium of the game. Since all rational players propose the minimum value, a random winner is selected.

The figure 2.5 depicts two histograms with values collected from participants of an online course in two iterations of the game.

[6]The stylized version is a combination of the original Keynes beauty contest game and the "guess 2/3 of the average game"

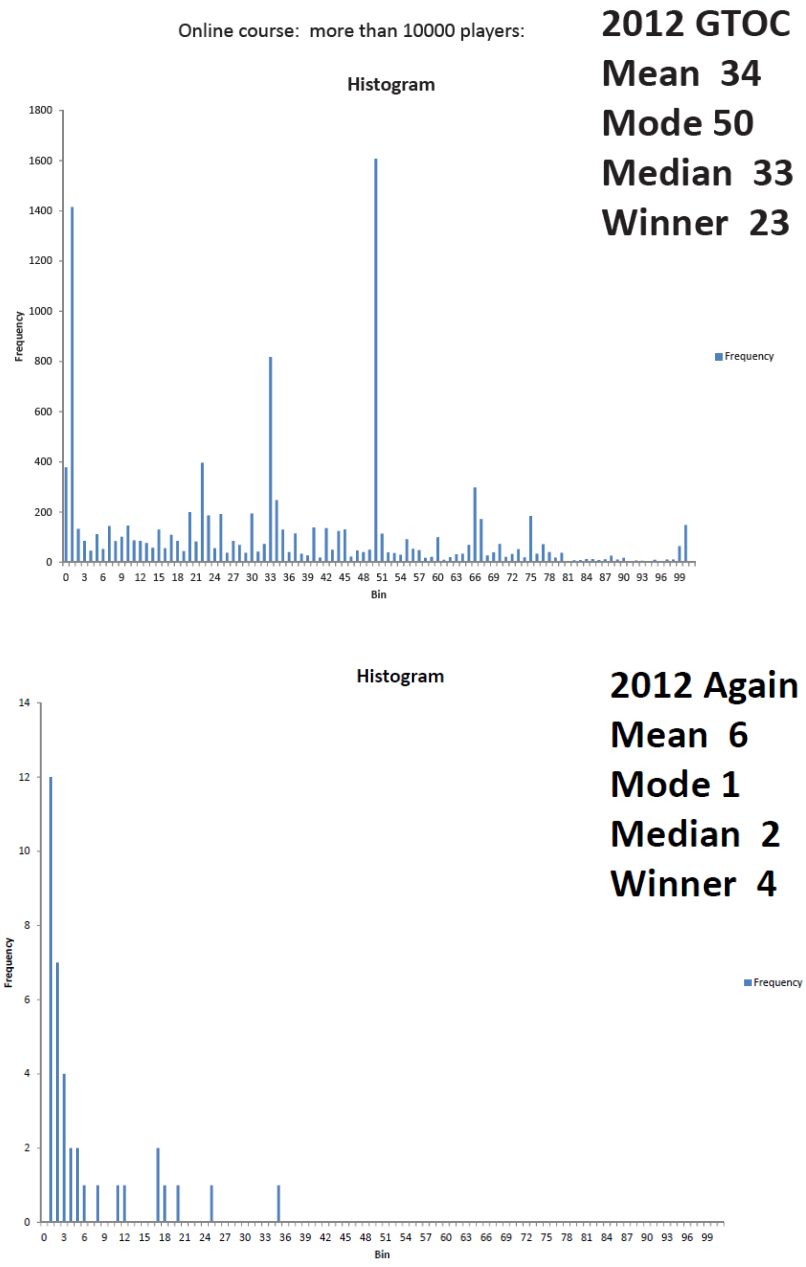


Figure 2.5: Keynes Beauty Contest Game
Histograms obtained from lecture notes at [30]. Histograms for first and second iterations are displayed in order.

As observed in figure 2.5, the Nash equilibrium is not achieved. Certain players have stopped halfway through their reasoning, and some highly naïve players have named impossible values which will never win the game, causing strategic players to suffer a significant loss, considering the winning value. However, even though their learning speeds seem quite different at the second iteration, all players seemed to have learned certain amount of reasoning by the second iteration.

Due to imperfect player evolution taking place at different speeds throughout iterations, certain losses may have to be borne by strategic players. It is a cost that a system must be willing to bear for an eventual equilibrium. But in an automated protocol, this situation may differ due to minimized human interaction. And the noise would be a useful parameter in monitoring influences of malicious agents who want to reap alternative benefits.

The models presented by this paper provides us great insight in simulating attacks on our potential solution. They highlight the importance of interpreting results for what they are, and aiming for practical analysis rather than theoretical convergence. Overall, the paper presents game theory to be a solid solution concept and a rational choice for modelling games, even with noisy agents.

2.6.2 Internet and Game Theory

The suitability of game theory in modelling the Internet, an entity that keeps growing in size, connectivity and architectural and behavioral complexity is discussed in [13]. The author points out that technology has come to a place where it interacts very closely with humans, that modelling it without integrating socio-economic considerations is unlikely to capture its reality.

Game theory has already yielded some fruitful results in computer science, as a tool which studies how certain decisions are bound by rationality, repeated interactions and varying player knowledge. But it has certain limitations which may be unsuitable for a solution oriented field such as Computer science, given that even if a game could be designed, solution concepts are not guaranteed to exist.

For example, a game may not have a Nash Equilibrium; players may not have a rational best response they would not deviate from given that all others are playing their best responses. But given that any mixed strategy game is guaranteed to have at least one, the problem transforms into a modelling decision. More than one Nash equilibria can also exist for a game, where which one will emerge can not be theoretically predicted. The author points that this in itself, is a computational complexity problem which holds interest to computer science community. It remains that both fields have much to gain from interaction to each other, and nothing to lose.

TCP/IP congestion control protocol is a natural Nash Equilibrium, where players have communally decided not to override and congest the network with their own traffic while others obey the rules. This is one scenario where tragedy of the commons has not surfaced, and the game to which this is a Nash Equilibrium remains a mystery. The illustration in [31] where Price of Malice causes players to cooperate beyond the worst Nash Equilibria and towards social optimum despite self interest also resembles this behavior. Thus, the author points out that rationality is not the exclusive common denominator of Internet-based naturally occurring games, but also the volatility of environment within which players must continue to exist, making changes to strategy too costly to afford.

As a complementary measure to Price of Anarchy, the author questions whether a Price of Internet Architecture can be derived, given that routing is based on delay information and bandwidth assignments of network operators, as opposed to a social optimum which uses shortest routes. We find this question intriguing to our research, because it could be reinterpreted as what costs are players willing to afford for the sake of a Nash Equilibrium in a volatile environment (an unstable network, or in our case, an insecure one).

Mechanism design, a practice where a game with rational players is designed such that their equilibrium achieves the designer's goals, ("inverse game theory" as the author reiterates), is presented as a crucial research area which computer science engineers must utilize. While computational savings in latency, storage, processing power can be proposed through any solution, it takes public adoption for the solution to succeed. The emergence of facets such as user interface engineering, and even the criticism of Microsoft's Next Generation Secure Computing Base can be seen as examples of this. The significance of this observation is that no design can escape player concerns, and as the author cleverly puts it "All design problems are now mechanism design problems".

Much of the authors observations provide justification for game theory being applicable in modelling distributed ledger solutions. The requirement of socio-economical insights, hidden player utility in environment stability/instability required for internet based game convergence, and the universality of mechanism design has been observed in most literature that followed, not to mention certain naturally occurring equilibria. However, the paper more importantly proves game theory a viable candidate for modelling internet based research. A direct observation of this is how a concept such as Blockchain, where tragedy of commons and Price of Anarchy could be rampant, remains stable.

2.7 Game Theory In Information Security

In considering game theoretic applications in security domain, the literature heavily favors network security applications. Game theory proves to be a unique tool for this task as well, given that usual analysis is based on vulnerability knowledge and attack surface, wherein the iterative action between two competing players can only be modelled through game theoretic concepts.

[29] provides a summary of game theoretical concepts, followed by a game characteristic based classification of existing solutions in network security problems. [28] includes privacy specific network security solutions, and summarizes them to a game type based classification. We discuss both papers below, and obtain a superset of game characteristics and game-types commonly used in information security literature.

[29] argues that traditional defense strategies of prevention and detection fall far from a sustainable solution in ensuring network security, given that sophisticated attackers would always circumvent them, initiating a never ending recursive process of risk assessments and vulnerability exploitations. This is further complicated by the growth of networks in scale, connectivity, accessibility and as a byproduct, vulnerabilities. The

risk assessments, which are subjective to begin with, never get sufficiently mitigated due to by budgetary constraints. Despite the regular patch releases, zero-day attacks still happen. Thus, the output of a security endeavor ends up being the minimum viable protection. Question remains whether this short-term solution collection is the best long-term strategy we have in protecting internet infrastructure, a fundamental part of today's life.

Secure Software Development and Trusted Computing base is presented as a potential solution for this problem, but its practicality is questioned. The increasingly competitive software industry, fueled by research in project management, keeps looking for faster delivery of usable products. This cannot be criticized, because success depends on who caters to the market first. While eventual reviews and updates, or perimeter filtering would strengthen security, the primary weaknesses in code might remain for years in favor of uptime.

Game theory, however, thrives at making subjective observations quantifiable. Through modelling player reactions, it covers a wide area of adversary actions and provides a best strategy profile for the defender. While such mixed strategy equilibria obtained may not necessarily be favorable to defenders, or sustainable at a first glance, they provide insight that allow for simulation based analysis, where optimizations or loopholes which may go unobserved in human analysis would be made visible. Game theory therefore lends the decision process of a network administrator a certain sophistication which otherwise might be of hit-and-miss nature.

In the classification that follows, [29] restricts its survey to non-cooperative games (since attackers and defenders do not cooperate), and summarizes games depending on their static/dynamic, prefect/imperfect information and complete/incomplete information natures. Dynamic games refer to games that converge over multiple iterations, where in static games players react simultaneously and one time. Given that perfect information suggests all players are aware of all past actions of all other players, all static games are imperfect and can only be further classified by their information completeness. Information completeness considers the property of all players having knowledge of all other players' strategies and payoffs, while the actions already taken place are not known.[7] However, as [28] notes, while Attacker vs defender two player games may limit its focus to non-cooperative games, distributed system solutions can take a cooperation based approach.

Per our purpose of studying security domain applications of game theory, we summarize the literature survey of [29] in table 2.2 by security use cases that were studied and the respective methodologies used. (A slight deviation from the presentation at [28])

The authors point out that the game theoretic solutions proposed in static, complete or perfect contexts might be unrealistic, given that in real life situations, such knowledge by defender, nor the attacker is feasible. We partially agree with this observation.

[7]To further clarify the distinction between perfect information and complete information, we define an action as an exact decision a player takes during an iteration, where strategy is all actions a player could take in each iteration given the possible actions of other players at respective iterations. A strategy is therefore a complete plan of action.

Table 2.2: Game Theory Applications in Network Security

Application	Decisions	Strategy/ Payoff Information	Action Information	Game Details	Result
Quantitative Best Strategy for players in Information Warfare	Static	Complete	Imperfect		Multiple Nash equilibria
Present and future network IS risk assessment	Dynamic	Complete	Perfect	Markov game	Selection of best system repair scheme
Mobile electronic commerce chain Security	Dynamic			Evolutionary Game theory	Investment strategies
Management strategy suggestion for information security investments				Pure and Mixed strategy plays	Nash equilibrium
Reverse engineering protection of Intellectual property	Static	Complete	Imperfect		Quantitative risk assessment for investment efficient strategies
Optimizing of worm propagation speed by Attacker vs Defender	Dynamic	Incomplete	Perfect	Zero-sum min-max problem	Optimal placement of vulnerable high value hosts
Attacker vs Administrator in network with correlated node vulnerabilities	Dynamic	Complete	Perfect	Zero-sum stochastic game	Numeric computation of Optimal player strategy
Hacker vs Defender	Dynamic	Incomplete	Imperfect	Zero-sum game	Nash and Bayesian equilibria based predictions
Attacker vs Defender	Dynamic	Complete/ Incomplete	Imperfect	Classical and Stochastic Non-zero games	Nash equilibrium in different error perceptions
Administrator vs Attacker Two player games in DoS/ defacement/ breach scenarios	Dynamic	Complete	Perfect	Stochastic game	Multiple Nash equilibria
Modeling Incentives of DDoS attacker vs Network Admin	Static	Incomplete	Imperfect		Nash equilibria from ns-2 simulation experiments
Attacker vs Network Administrator vs Sensor System (3 players)	Dynamic	Incomplete	Imperfect	Repeated Game with finite/ infinite steps	Simulation based Nash Equilibrium
Attacker vs IDS in a system	Dynamic	Complete/ Incomplete	Imperfect	Stochastic Markov Game	Numerical analysis for different completeness scenarios
Mobile ad-hoc network intrusion detection (A two player game)	Static	Incomplete	Imperfect	Bayesian Game	Investigation on Bayesian Nash Equilibria
Intrusion detection as a resource allocation problem	Dynamic	Incomplete	Perfect	Continuous game followed by discretized model	Dynamic Algorithm for automatic or administrator response

For example, as observed in table 2.2 modelling attacker vs defender behavior for correlated node vulnerabilities using zero-sum games is highly unlikely to provide intuitive outcomes. Given that one of the most discussed difficulties in defensive security is patching all vulnerabilities while attacker only has to find one, Attacker utility and cost functions will most certainly not be the negative of defenders'. They might even not use the same units in their respective payoffs, making the assumption of complete, perfect information void. However, zero-sum games do become useful in min-max worm propagation example, and also in analyzing DoS impact with respect to bandwidth protection/wastage, as observed in [28]. Our disagreement lies in considering perfect information assumption infeasible. The authors later point that synchrony assumptions are another limitation for obtaining realistic results. While this statement is true, asynchronous actions between attackers and defenders allow for perfect information games.

Incorporating domain knowledge and heuristics as state transition probabilities for stochastic games is pointed out as another limitation in research. The illustration of [26]'s iterated introspection model conforms with this, where the belief intuition is shown to be independent of action probabilities for long-term dynamic games. Certain game expectations of perfect knowledge from used sensors, or outcomes, or perfect correlation between attack steps, all prove to be somewhat unrealistic assumptions. [26] explains this behavior in terms of noise in human preferences during the time of learning until eventual convergence, but we believe this could be generalized to represent machine noise as well. Scalability constraints in existing solutions are also highlighted.

While the authors conclude that game theory offer promising insights in shaping cyber security in the present world, we only agree to this under the condition that research goals must stem from the need for practical insight, as opposed to requirement of theoretical proof of convergence.

[28] supports [29]'s observation of game theory being optimum in analyzing many possible decision scenarios of attackers, and highlights that not only decisions, but the players themselves could range from cooperative, selfish to malicious, allowing a much more sophisticated plane of methodical analysis. Another important aspect is that game theory allows physical objects such as intrusion detection systems to represent players, whose fixed strategy profiles can allow for much more insight regarding human players' dynamic strategy profiles in practical scenarios. Existing literature also purpose that technical design must incorporate incentives to obtain dependability. Dependability, analogous to strategyproofness (or dominant strategy incentive compatibility) must be assured by a protocol to ensure that players never obtain worse by honest participation irrespective of other player responses. Overall, game theory offers a methodical analysis that can be used in mechanism design of better, more socially optimum solutions in resource allocation, risk assessment and incentive influenced cooperation. In competition with methodical analysis is heuristic based decision making (which [29] already established as unsustainable), which does not support concepts such as formalized, multi-stage decision making required for business continuity purposes. While methodologies such as machine learning have made significant advances in data correlation and therein producing accurate predictions, they have yet to produce utility

maximizing suggestions available to defenders and their subsequent consequences.

[28] categorizes their chosen literature by several security domains, an abstracted presentation of which is provided in table 2.3. We avoid a use-case specific representation due to high-level criteria themselves accounting for the use cases of game theory, and direct the interested reader to [28, page. 5] itself, where the authors have provided a compact analysis.

Table 2.3: Game Theoretic Applications in Different Domains of Information Security

Domain	Types of Applications	Types of Games	Outcomes
Physical and Mac Layers	Jamming, Eavesdropping in wireless networks or other communication channels	Zero-sum games, Hierarchical Stackelberg games, Cooperative and coalitional games	
Self-organizing networks	Vehicular networks, revocation in ephemeral networks	zero-sum game, fuzzy games, fictitious play, finite dynamic games	
Intrusion detection systems		Stochastic zero-sum games	Algorithmic design and performance evaluation
Anonymity and Privacy	Mobile device location privacy, Anonymity, trust vs privacy tradeoff	Incomplete information games, Repeated games with simultaneous moves	
Economics of Network Security	Correlated Security investments	Linear influence network based game with payoff functions	Equilibrium analysis
Cryptography	Cryptographic mediators, Multiparty computational protocols	cheap-talk games, repeated games	Incentive-based designs in security assurance

Physical and Mac Layers The authors attribute the frequent use of zero-sum games to their ability of capturing conflicting interests of players, but they also support our observation regarding requirement for contextual application of zero-sum games, noting that attacker utility functions will be different than those of legitimate users. They suggest dynamic games for situations where nodes share a communication medium, as it allows the modelling of sequential interactions. The authors observe that if both parties are either active or passive together, then the Nash equilibrium serves as a solution concept while if one party is active and the other is passive or vise-versa, it follows a Stackelberg equilibrium. As such, Stackelberg games can be designed for Jammer vs defender games or Defender vs Eavesdropper games.

Self-organizing networks Problems in Ephemeral networks, in which wireless devices form ad-hoc network infrastructures and interact with each other in short sessions, have evolved from routing issues to false broadcasts and Sybil attacks. Lack of a Certificate authority and communications being peer-to-peer based have given rise to "free riders", who maximize their utility at little to no cost to themselves but at a perhaps small, but usually higher cost to the social optimum. This has resulted in incentive based defense strategies conformed by majority of peers, as observed throughout the protocol designs for Blockchain protocols in addition to examples in table 2.3. The avoidance of unstable equilibriums is another useful result in game theoretic modelling. Similar to criticism

of Transaction fee based Blockchain's stability, certain equilibria stable in the short run could have disastrous outcomes if not analyzed prior to deployment. [32] provides such an analysis and uses it to determine critical design parameters of their network, as illustrated in its review in section 2.8.2.

Intrusion detection systems The paper notes that game theoretic applications in IDS have significantly evolved from simplistic models where attacker strategy profile was attacking or not, and defender's was defending or not. Current work includes avenues such as Security assessment of interconnected systems through attacker behavior probabilities, dynamically evaluated real-time risk assessment where differing trustworthiness of sensors is captured using Hidden Markov Models, Weakest link targeting attacker vs defender's optimum investment points and access control and security warning systems. Dynamic configuration of policies in IDS systems in response to differing sequences of attacks and Alert based automated intrusion response are some of the other highly interesting and thought provoking game theory based IDS applications.

IDS and attacker are usually modelled as non-cooperative entities, and due to probabilistic transitions in an environment under attack, Stochastic games are mentioned to be the appropriate tool for modelling IDS interactions. They go on to recommend on-line learning based approaches, given that they allow IDS to evolve alongside the data accumulated.

Anonymity and Privacy Due to knowledge limitations by entities in privacy games, most games in this domain are placed in incomplete information category. Two interesting applications mentioned are modelling of customer privacy concerns leading organizations' investment in customer data security as an Stackelberg game, and The Onion routing (Tor) Path selection algorithm developed as a repeated non-cooperative game, which minimizes success probability of entry-exit linking attacks in the network. A contradiction from [29] here is the authors do not chastise complete information games, provided that the assumption is limited to some parameters. They stress that refinement is required in specific applications, but they do not consider it a practically limiting abstraction.

Economics of Network Security Games in this domain are motivated by the equilibrium analysis of security investment, insurance and asset protection. Given the interconnected nature of organizations, one organization's vulnerabilities have a cascading effect on all other organizations that interact with them. Thus, security investments of one company have a direct utility to other companies. This presents a situation where social optimum is beneficial to everyone, and game theory provides the exact toolkit required for modelling such situations. In modelling security of societies by individual investments of players, average protection level vs minimum protection level are considered as two possible approaches, with the latter representing a weakest-link based security model (Refer section 2.8.1).

Cryptography Cryptography games discussed mostly center around the cheap-talk game, where a mediator is introduced to increase necessary cooperation between the players, reducing price of anarchy. Multiparty computations are also discussed, where a repeated game based, unsuccessful protocol with uncooperative nodes is analyzed, highlighting the need for a better design.

In conclusion, authors question the sufficiency of pure game theoretic modelling of security, highlighting the need for algorithm and implementation analysis over system model and equilibrium analysis. This relates to our previous observation regarding research goals being practical interpretation oriented as opposed to simply seeking convergence. They propose that such theoretical analysis by itself cannot ensure security, but the information it provides for mechanism design could more efficiently solve security problems. In fact, they highlight some applications in Self-organizing networks domain and anonymity and privacy domain which either used equilibrium analysis for the mechanism design of protocols, or used it for extraction of significant parameters to be included in future mechanism designs. Similar to [29], they address the need for accurate assessment of player resources in the form of game information. They further assert our previous observation regarding quantification difficulties in network security. Inadvertently emphasizing the importance of security investment related game theoretical solutions, they succinctly summarize this situation with following excerpt.

> ... the attitudes towards security seem to go back and forth between *"we are doomed if we don't invest heavily in security"* and *"no need to worry too much, it all seems fine"* depending on the economic situation.

Quantifiable security analysis and dependability assurance through technical design, reemerge in game theory just as they did in Blockchain protocol discussions (Refer section 2.3). From the domains analyzed in [28], our research spreads into the areas of self-organizing networks and Economics of Network Security. Our model heavily relates to incentive based defense implementation since security investment persuasion is at the heart of our problem. The example of business security investments having an interconnected utility to all connected businesses sounds quite familiar to a peer node's investment in assuring its availability. Modelling of conflicting interests and incorporating active and passive nature of players provide further insight in a model, as observed in our solution in Chapter 3 itself.

We find the choice of social optimum maximization a well establish game theoretical concept which could accurately model interrelated incentive parameters of a secure distributed ledger. A coalitional game of selfish nodes vs malicious nodes will not suffice for our context given that attacker is not strictly rational and is not looking to maximize any observable utility. As such, we model the first scenario and observe the convergence properties in assuring the security objectives of our research problem.

We present our model in Chapter 3 as a dynamic, complete imperfect information game. While previous player decisions are available to each other in the form of shared history, a periodically inactive attacker masks a player's true decision in protection investment,

making our model an imperfect information game. Similarly, while we acknowledge the criticism of complete information games by [29] when considering attacker vs network games, for cooperative games with where only peer behavior is modelled, we find a complete information concept quite suitable, similar to observation by [28].

2.8 Game Theory In Distributed Systems Security

We begin this section by observing system reliability games in existing literature. We proceed to discuss [32] and [31], two papers that detail using game theory in quantifying utility required for better peer to peer cooperation and influencing costs of self-protection. We then examine the mixed strategy equilibria obtained by an infinitely repeated game, an approach we use in modelling our solution in Chapter 3, and conclude the section by reviewing behavior of "rational foresighted" players, whose characteristics are crucial for public goods games which eventually converge into the predicted equilibrium. We draw significant inspiration from this section, given that each topic provides unique modelling options complementary to each other. It further establishes our research problem as an amalgamation of many perspectives, each of which has been modelled differently in existing literature.

2.8.1 System Reliability in Game Theory

A comprehensive theoretical analysis of the impact of cost-benefit ratios of participants in a public goods game in relation to amount of effort put forth by a given agent and which agent therefore would be free-riders is presented by [33]. Both private and social outcomes (equilibrium vs social optimum) are considered for three types of two player games; total-effort where system reliability is dependent on combined efforts and weakest link and best shot, where system reliability is dependent on the agent exerting the least and most efforts, respectively.

The author presents that while in social contexts, effort exerted would be dependent on the financial capacity of a participant (i.e. an alliance of countries), in system reliability and security, costs, benefits and probability of failure are the primary concerns while the financial capacity of participant becomes secondary.

For the total effort games, a unique equilibrium is observed where exerting all effort for the public goods game falls on the agent with highest benefit-cost ratio, where all other agents opt to free-ride. An exception to this case would be when the agents are homogenous, in which the game becomes unstable, much like pure strategy equilibria observed in [34] (Refer section 2.8.4). In considering the social optimum however, agent with least cost exerts all the effort. Therefore when the cost of agent with highest benefit-cost ratio is lower than some other agent, the second agent exerts the effort.

For the weakest link game, a pareto dominating equilibrium is observed where effort is exerted by the agent with lowest benefit-cost ratio. Social optimum scenario requires all agents to exert an equal amount of effort, as such the reliability is considered unstable

compared to a total efforts game, making it more expensive to ensure reliability.

For the best shot game, Nash equilibrium could be either when agent with highest benefit-cost ratio exerts all the effort (similar to the total effort case), or by the agent with lowest benefit-cost ratio when the agent with highest benefit-cost ratio opts to exert no effort. The latter is dubbed as "slacker equilibrium", and is an important observation that supports the potential of mixed strategy equilibria (Refer section 2.9.1) in a repeated game. Social optimum analysis was mentioned to yield the same results as the social optimum analysis of the total efforts game.

Varian proceeds to analyze the effects of varying costs when the benefit is a constant for all agents, and observes contradictory effects in system reliability when it concerns the number of participants for total effort and weakest link cases. Total efforts game increases in reliability while the weakest link game decreases in reliability. This echoes the outcome of social optimum being more expensive to achieve for weakest link games, causing higher number of agents to destabilize the system. For larger number of players, total effort games are considered to be more reliable over weakest link games.

In considering fines and liabilities imposed upon agents responsible for system failure, Varian presents that for the case of total effort, agent with the lowest cost of reducing probability of failure would endure a fine equal to the sum of costs of other agents, and could also lead to agents with higher costs being willingly negligent knowing that they will be compensated. In the weakest link case, agents mutually compensate each other for their respective losses, and as such the cost of failure is distributed. While liabilities for weakest link scenario presents an attractive solution for homogenous players, a negligence rule is preferred in the case of heterogenous players in inducing optimal effort. If an agent is proved to be negligent only are they expected to compensate for the cost of other agents.

Further analysis is conducted regarding the effects of number of agents, sequential moves over simultaneous moves, and adversaries. Payoff uncertainty and its effect on "leadership" influenced gameplay, and how inconclusive information regarding other players could result in lesser free-riding are presented as future research avenues.

An extension of this study decoupling security investments into self protection and self insurance is conducted by [35]. They specifically target security games, such that self-insurance represent reduction of magnitude of loss (e.g. through data archival) wherein self-protection represents the classic reduction of probability of loss (e.g. through perimeter security investments). They limit their scope to homogenous costs and simultaneously moving agents.

The authors also introduce "weakest target" games with mitigation and without mitigation in addition to analyzing the games discussed by [33] in a decoupled security investment context. Weakest target game without mitigation allows the attacker to always compromise the weakest link but allow other agents to remain unharmed (attacker with infinite strength), while successful compromise of weakest link in Weakest target game with mitigation depends on a chosen security level (attacker with finite strength). Attacker motivation in the weakest link games is minimizing costs while causing the

most harm to the system.

For the total effort game they observe multiple equilibria for different environments. Cheap protection, high losses, overpriced insurance causes everyone to invest in only protection. When protection costs are higher than expected losses and losses are higher than insurance costs, everyone invests only in insurance. No protection or insurance is observed for negligible losses. Given the revenue redistribution, they observe increasing number of agents to have a destabilizing effect on above Nash equilibria.

For weakest link games, an inefficient equilibrium is observed where all players invest in same minimal protection when cost of insurance is lower than expected losses and the minimum protection is higher than a constant c.

$$c = \frac{\text{Expected losses} - \text{Insurance cost}}{\text{Expected losses} - \text{Protection cost}}$$

If initial security is inadequate and minimum protection is lower than c all parties converge in self-insurance. For negligible losses, no security investment is made which remains the only Nash equilibrium. However, for the unique case of Minimum protection $= c$ both exclusively protecting or exclusively insuring become Nash strategies. Similar to total effort game, equilibria will destabilize with increasing number of agents. And considering the feasibility of one agent deviating in a large group is higher, exclusively self insuring equilibrium is far likely to emerge than the contending exclusively self protecting equilibrium.

For best shot games authors only observe either exclusive self insurance or no security investment. The equilibria are not affected by deviating numbers of agents.

Interestingly, the weakest target game without mitigation does not have any pure strategy equilibria for differing environmental conditions, but a mixed strategy Nash equilibrium is observed. For weakest target game with mitigation two equilibria are observed; A pure strategy Nash equilibrium where all players invest only in protection when protection costs are higher or equal to insurance costs, and a mixed strategy equilibria where self-insuring probability is identical to mixed strategy equilibrium of weakest target game without mitigation.

In relating [33]'s findings to our problem, we observe that we cannot classify our game completely as either a total-effort or a weakest link game. Furthermore, [35] highlights an attack similar to our problem statement, where the attacker intent is delaying transmission of a given piece of information and subsequently observes that in distributed systems where at least one node, or one open route between two nodes is protected, the information will be kept in circulation, thus resembling a best-shot game. Therefore we find it difficult to categorize our problem under above well-established classifications.

[33]'s observation regarding fines and liabilities for the total effort case presents an unfair scenario to heterogenous participants in a self governing system, and further contradicts the observations of protocols such as Tendermint that has stake integrated

into its security. This gives us reason to remove explicit stake from our initial parameter evaluation. We further find the preference of a negligence rule over a strict liability rule in weakest link case to support our decision in following a reputation based reward system, which essentially encapsulates fairness for heterogenous players.

The analysis in both [33], [35] present individual utility optimization followed by social welfare optimization by a coordinated entity, wherein we incorporate optimization of social good into our model design itself through redistribution of unutilized rewards in favor of introducing a "stake" . We adopt the modelling concept of exogenous attack probability and attacker of infinite strength in weakest link game without mitigation from [35], as well as the conditional reliability functions used for utility measurement. At an initial glance, [35]'s "weakest target" games looks applicable to our problem. However in the case without mitigation, compromising a singular agent has no value to our attacker and in the case with mitigation, our attacker continues to have no motivation of minimizing costs to self.

In considering the 3 main types of games presented in game theoretical literature regarding reliability of public goods games, we observe that our problem present an amalgamated scenario which contains characteristics of all 3 games and some of their more specific extensions. We expect simulation results of our proposed solution to provide insight regarding a possible classification, and reuse applicable concepts in modelling our solution. In addition, we expect our model to provide insight on the effect of inconclusive information in controlling free-riding behavior, a sentiment asserted by both [33], [35].

2.8.2 Service differentiation on peer contribution

[32] discusses the inherent centralization of peer to peer networks due to selfish peers who receive benefits equivalent to others while not contributing to decentralization of the service, and proposes differentiating service in order to incentivize honest participation. They present a game theoretical analysis which yields optimum scaling possibilities in such a peer to peer system's performance.

In general, peer to peer file sharing does not demand any reciprocation from participants, even though it is necessary for its continued operation and increased availability. The downloaded files should be "seeded" so that peers in closer proximity would gain better throughput, and allowing the resulting decentralization of the network to ensure increased availability, independent of geographical origin of the content. But most users disconnect after the file has been downloaded and therefore some users, who might not necessarily be in close proximity end up acting as servers in a traditional client server model. The conflicting time-zones intensifies this problem, as a peer from a different time-zone may not be able to obtain a file when the "serving" peer is offline. Thus, to mediate the unreliability the "selfish" peers project to the entire network, the authors propose differentiating the service provided to a peer according to their contribution.

A utility function is a function which integrates the cost and benefits of a player depending on a chosen strategy, to model the return of investment to a participant

of a game. A peer to peer network where the benefit of a peer is proportional to its contribution therefore can be modelled as a non-cooperative game with selfish players. (In a non-cooperative game each player behaves individually in achieving their goal without forming of any coalitions). If a Nash equilibrium can be achieved, then the service can be perceived to be optimal and self-sustaining.

Considering homogeneous peers, "peers who receive equal benefit from every other peer", the authors obtain two pure strategy Nash equilibria for a repeated game setting in each side of a "critical limit of benefit". A benefit which increases with contribution is only perceived when the initial benefit exceeds this limit, causing the equilibria with better social welfare to be realized by the system. Service differentiation is done through modelling the file request acceptance probability by another peer on the level of requesting peer's contribution, which is attached as metadata on the file request.

The utility obtained, can only yield a non-negative maxima if the benefit is over the critical limit. Furthermore, neither the minimum nor the maximum contributions are shown to yield increased utility. This convergence property is optimum for a purposefully decentralized service, which would otherwise be trivial to centralize by simply increasing the contribution.

"Cournot Duopoly" model, describes the behavior of competing entities which produce different quantities of an homogenous product. A supply beyond the demand would cause reduction of price by one party, and the other party would retaliate by reducing their price. In a repeated setting, if parties reduce their price beyond their cost of production, the system would collapse. Therefore an equilibrium is obtained through convergence of utility functions that represent each party's costs and benefits.

The two player game presented in the paper behaves quite similar to aforementioned model, where quantity is analogous to contribution and price is analogous to benefit. One player models his contribution in reaction to the other player's contribution, and this reaction function, repeated for iterations, results in equivalent contribution functions for homogenous peers. The critical benefit constant is obtained at this level, wherein the contribution equation has solutions only when the benefit exceeds the constant.

The model continues by generalizing the two player results to an N player game, and the average outcomes of a heterogenous system is compared against the N player game's theoretical predictions. A realistic simulation setup is presented, where the peer benefit and contribution adjustments are modelled against only a subset of peers who represent services of interest. The repeated application of reaction function by different peers who start with random contributions, is seen to converge to the socially optimum Nash equilibrium. Small initial contributions result in an average contribution of 0, indicating system collapse for the unstable Nash equilibrium.

The number of iterations taken for the simulated system to converge (reach socially optimal Nash Equilibrium) is highly reduced when the average benefit increases. System is further shown to be stable for fluctuations of number of users, while more users result in lower time for convergence, and a higher eventual average contribution. When the heterogenous system has lesser benefit yielding peers (lesser peers who share files

of interest), the system takes longer to converge, which further supports the above observation.

Peers who refuse to adjust their contributions and make a constant contribution are seen to disrupt the convergence properties to their advantage, but such a system would have all peers eventually contributing equally and therefore the purposed goal of eliminating benefits for uncooperative peers would still be realized, even if better social welfare would not be achieved.

An interesting use case pointed out by the paper is how the probability function on user contribution can have different outcomes to participating peers depending on the service while the underlying game theoretical analysis remains applicable. While the presented model differentiates on the request acceptance, in the case of Napster, peer's search capabilities (i.e. number of results returned) were differentiated. This was an alternative modelling concept we considered for our design, where the maximum number of peers that a participant can connect to is limited on the level of protection afforded. A differentiation scheme that facilitates random distribution of votes could also be utilized. For example, peers propagate a vote until the minimum number of hops had passed before it is added as a vote to the decision pool, making a large number of unprotected, amateur peers unable to quickly exceed the Byzantine threshold. This gives incentive for the entire network to ensure their availability since the node that accepts the vote may not necessarily be their closest peer. While our proposed model does not explicitly use service differentiation (for reasons also discussed in section 2.8.4), we believe that it is naturally integrated in a reputation based model.

The implementation of self-protection from denial of service is an interesting research area itself, but at the very least this could mean extraction of an availability measure through neighbor auditing of allocated ports that (somewhat crudely) represents the level of protection. (i.e. If all dedicated ports respond to each of few randomly chosen neighbors, the machine is not under attack). The paper proposes a similar scheme in monitoring the contribution level with uptime and diskspace. Even a simpler measure such as previous consecutive votes could suffice in evaluating self-protection for a problem such as ours. Such availability requirements from peers have been mentioned as reasonable in multiple protocols we have previously discussed.

[32] provides valuable insight into application of game theory in ensuring fair contribution by nodes of a distributed system, providing insight regarding modelling self-protection and propagation contribution properties required for our solution. An intuition worthy of note is that equivalent to the critical limit of benefit, similar limits were observed in the game theoretical analysis we present in Chapter 3, making recognition of these limits prior to protocol design crucial for stability and continued security of distributed systems.

2.8.3 Affording costs of self-protection

While [32]'s solution models the behavior of peers whose contribution is made mandatory, a different insight regarding how willing the players would be to afford the required cost

of self protection is provided by [31].

As observed in prior reviews and many game theoretical analysis, rational players seek to maximize their own benefit, even if better alternatives exist when they cooperate with other players. A well known example of this situation is prisoner's dilemma.

		Prisoner B	
		Confess	Lie
Prisoner A	Confess	-2,-2	0,-3
	Lie	-3,0	-1,-1

Figure 2.6: Payoffs for the Prisoner's Dilemma

Looking at figure 2.6, let us consider the payoffs of players A and B, displayed in (A, B) format. For each action of player B, the best option for A is to confess to authorities. If B confesses, A gets a better payoff from confessing than by lying (2 over 3 years of prison), and situation is similarly better for A if B resorts to lying (0 over 1). B's strategy is symmetric, and therefore the Nash equilibrium is when both selfish agents confess to authorities.

The above situation results in 4 years of collective prison time, while both prisoners lying would result in only 2 years of prison time. This represents a well known concept in Game theory called Price of Anarchy[8] (PoA), which is the loss of social welfare due to selfish behavior of utility maximizing agents [36]. It is a well known measure of equilibrium efficiency.

$$\text{Price of Anarchy}_{min(cost)} = \frac{\text{Social Welfare at Worst Nash Equilibrium}}{\text{Maximum Social Welfare}}$$

The "worst case" is when the players show the least coordination among themselves.

Thus, PoA for Prisoner's dilemma with given payoffs would be $\frac{4}{2} = 2$.

[31] extends this analysis by allowing some players in a game to be malicious, and studies the impact of these agents on collective welfare obtained by selfish agents. An important contribution of the paper is the presentation of Price of Malice (PoM), which is the degradation of Social Welfare at Worst Nash Equilibrium due to malicious behavior of byzantine agents.

$$\text{Price of Malice for } b \text{ malicious agents } = \frac{\text{Social Welfare with } b \text{ malicious agents}}{\text{Social Welfare at Worst Nash Equilibrium}}$$

[8]Depending on the game design, the dividend and divisor changes places. For example, cost minimization games such as prisoner's dilemma has maximum social welfare as a divisor, while utility maximization games have maximum social welfare as the dividend.

Measuring social welfare degradation due to malicious agents is highly useful in obtaining realistic resilience parameters. Specially in a networked distributed environment, which is bound to contain malicious agents with arbitrary cost functions, which may not even be in the same units as that of selfish players. Such cost functions inherently complicate the respective utility functions, making PoM reduction more difficult compared to reduction of PoA which can be enforced through rules of participation (As informally observed in [32]). System vulnerability increases with price of malice. Therefore, the paper aims to quantify the number of Byzantine players that could be tolerated by an stable system.

The analysis is based on a virus propagation model borrowed from existing literature. n players are assumed to be located in an undirected grid, each of whom has a choice to risk infection or protect itself through affording the cost of anti-virus software. Infection starts from a random unprotected node which infects the entire connected component, until it is surrounded by protected nodes. Infection results in a loss for a player, which is conditional on the action of risking infection.

The deterioration of system performance by Byzantine agents is represented as maximization of the social cost (sum of all selfish player costs) caused by their presence. Their attack strategy is to infect unprotected nodes upon contact by pretending to be protected nodes. This results in a game of imperfect knowledge, and selfish players either overestimate or underestimate their costs depending on how much they know.

In a model where selfish nodes assume all other nodes to be selfish, they underestimate their eventual cost, and therefore the social cost maximizes with the number of malicious players. The intuition is confirmed through bounds obtained for PoM in oblivious selfish player model.

In a model where number of malicious nodes are known, selfish peers evolve to be more cooperative. A critical contribution of the paper is the outcome that this scenario can yield a PoM below 1, improving on the worst case Nash equilibria. Nodes become more willing to pay for inoculation, and thus better social welfare is obtained, despite the presence of malicious players.

The aforementioned phenomenon is dubbed as the influence of "Fear Factor", a measure of unity against a common enemy, denoted by the inverse of the PoM equation. The measure however is upper bound, due to the fact that PoM in the risk-averse, non oblivious player model is proven to be always equal to or greater than a given constant. It is further dependent on the level of knowledge by selfish players, and is negatively influenced by additional information such as the exact locations of malicious players since it would cause selfish agents to make more precise decision regarding inoculation.

Following the cost maximization attack analysis, stability of system is also measured against the number of tolerable byzantine agents. The system is shown to be tolerant to 1 malicious agent in certain network graphs while in all cases it is unstable for 2 malicious agents.

The possibility of cost functions being of different units for malicious players and selfish players is what primarily threatens the stability of a non-financial blockchain solution.

While this information is discarded in BGP (and to no effect, given that players are loyal), the virus inoculation game outcomes present a situation where a selfish players' likelihood of affording certain costs is increased by imperfect knowledge of Byzantine player presence. As correctly pointed out by the authors, it is quite interesting that an even higher social welfare than a worst case equilibrium can be obtained in non-oblivious model. It should also be noted that the oblivious player scenarios welfare degradation resembles the "Tragedy of commons" scenario discussed in prior sections.

The observation that players reluctant to take risks reducing the PoM contrasts with the [32] described behavior, where agents who contribute the same amount tends to affect the equilibrium in their favor. But it should be noted that utility function is the same in the second scenario, and therefore we consider [31]'s analysis of malicious players more applicable to our problem. However, we also feel reluctant to incorporate the undirected grid node placement to our analysis for obvious reasons. The network topology and the target nodes will be changing with iterations and therefore such constrains may unnecessarily complicate the resulting model.

As discussed earlier, self-protection in our problem would be more complicated than protection through simple installation of software. But collective findings of the aforementioned papers show that such persuasion is possible with carefully tuned parameters of "benefit" and "fear factor", or even a penalty (As observed in the case of Tendermint). Another observation is that the cooperation willingness enforced by "fear factor" results in relatively loyal agents over selfish ones. In the analysis we present in section 3.3.4, we also observe the effects of "fear factor", where a minimum attack probability or a confirmed instability of the network encourages all peers to invest in protection under certain cost-benefit conditions. Thus, if the threat of malicious players and the requirement for self-protection could be properly stressed and implemented, resilience bounds of protocol simulations may closely follow the results of byzantine general's problem.

As an afterthought, another interesting theory coming from duopoly models, the Stackelberg theory (encountered in section 2.7), should be highlighted for its applicability in modeling a coordinated attacker of a distributed system. A subset of selfish players react to the strategy commitment of a leader. The leader aspires to reduce the social cost degradation by non-cooperation, and therefore centralizes the system. While in usual setting the leader is benevolent, this situation also resembles the behavior of agents under denial of service attack by malicious players, where malicious players act as leaders while the influenced selfish players show blind cooperation. We consider this a possible avenue for future research where attacker behavior is also modelled using game theory, limiting their behavior only to an exogenous attack probability within the current scope.

2.8.4 Reputation based service differentiation

[34] obtains a mixed strategy symmetric Nash Equilibria for an infinitely repeated game of homogeneous service providing network of rational peer to peer nodes. Peer ability

to obtain service, (i.e. utility of the game) is proportional to their reputation derived from past actions and their present action. Somewhat similar to [32] (Refer section 2.8.2), they obtain a limit for utility beyond which participation is superfluous. The authors' observe that the pure strategy equilibria of the game is to not cooperate, and justify that mixed strategy equilibria is likely to prevail due to altruism of at least some of the nodes.

The infinitely repeated game is modeled by obtaining equilibria for finite periods of game play. During a given period, a peer requests service from one or more peers and gets served depending on their reputation and in turn services one request of any other suitably reputed peer. An interesting direction the authors take is that different weights are assigned to a player's current decision and the reputation.

Given that players are from a "single homogenous population", a symmetric mixed strategy Nash equilibrium is observed, where players serve with some probability p and does not serve with $1 - p$ probability. Since both actions have support, they obtain the a value for p at which the equilibrium will hold (per property 3.2.1).

Since reputations are calculated per periods of time, p varies in each interval with reputation. But it is seen to converge to a constant over time as players react to each others contributions. Free riding problem is addressed by the action of serving having a positive probability which can be adjusted by adjusting the weight given to previous actions in reputation calculations.

An important contribution by the paper is that when the cost of service provision is negligible, the rational decision is to serve less than 50% of the received requests. The authors argue that when free-riding has a known probability, serving more than 50% of requests is not rational. They further observe that this causes the system to be inefficient and yet fair.

Another contribution by the paper is the characteristics of the weight given to past actions vs current action in reputation calculations. When the weighing is in favor of current action, the low reputed peers are capable of obtaining service and service probabilities all eventually drop to 0, whereas when it's in favor of reputation, peers' service provision probability becomes closer to the 50% limit.

Both aforementioned contributions provide important insight to our research. Firstly a service provision as much as the consensus tolerance is important for the our proposed network to be stable, and therefore cost of availability plays an important role. Given that the equilibrium probability takes the form of $p = \frac{U_0}{-C+2U_0}$ where C is the cost of service provision, and $U_0, C > 0$ and no weight is given to current action, we note that a non-negligible cost satisfying the inequality $C < U_0$ can increase the value of p close to 100% limit. It should be noted that the values are representative and the negative value does not necessarily has to come from cost itself.

Secondly, our completeness requirement stops us from giving any weight to current action while the probabilistic occurrence of attacks in our consensus network prevents us from decoupling current action from reputation (Refer equation 3.4). To elaborate, current action has a collective impact on whether the utility is feasible. It further has

an individual impact on how often the availability was compromised in case players did not actively invest in protection, considering that utility must be provided for them in rounds in which the attacker was inactive. Therefore in the interest of facilitating completeness, we condition benefit on both current action and reputation. As such, rational players have a continued investment in current action in form of future payoffs.

The authors' choice of mixed strategy equilibria over unstable pure strategy equilibria of none of the nodes serving, (i.e. network becoming obsolete) is justified by the assumption that there would always be some altruism present on the network. While this could sustain a general service provision, consensus would not be secure enough to survive through mere altruism and some agents free riding. As such, we do not depend on any means of altruism in modelling our solution. Instead, the mixed strategy would be naturally occurring through the volatility of the internet communication (regular obstructions vs malicious interference).

The paper provides criticism regarding the intentional contribution decisions taken by peers in differential service proposed by [32]. They question the integrity assurances of any meta-data / neighbor audit based contribution monitoring and the rationality of peers wanting of perform such actions at the cost of their resources, and conclude that intentional contribution decisions would therefore be an unrealistic assumption. We agree with this observation, and therefore refrain from using such assumptions in our model. Furthermore, given the context of distributed ledgers, the reputation information are easily calculable with perfect recall and as such remain most suited to proportioning utility.

2.8.5 Future Utility Optimization

Socio-rational secret sharing proposed by [37] provides a social reputation based solution for the problem of multiparty computations only being possible when the actual secret construction round is unknown to rational participants. They observe that when players are conditionally invited to future rounds of game based on their reputations, the social reinforcement that occurs leads to cooperative behavior.

Multiparty computations in secret sharing entails distributing a secret generated with a mathematical property such that a number of parties exceeding or equal to a given constant is required for the successful recovery of the secret. While the game play including honest and malicious agents differ, in the case of rational agents who decide to reveal their shares based on the utility they can obtain, they are shown to deviate during the secret construction round. Having as few as possible parties know the secret will yield the most utility to participants. Therefore it is rational to wait for other players to reveal their share of computations so that the final player can reconstruct the secret on their own. Given the reactionary nature of players, everyone waits until other players move first.

The authors present Rational Secret Sharing and Social Secret Sharing from literature, and proceed to define socio-rational secret sharing. Rational secret sharing is shown to have multiple rounds where arbitrarily only one round is used for secret reconstruction.

The equilibria is achieved by assigning a probability value that satisfies the inequality where the sum of expected utilities of only a given player learning the secret and not learning the secret is lower to equal to the utility earned by one player when all 3 players learn the secret. In Social Secret Sharing, a trust value based invitation concept is introduced where based on it players are categorized as good, bad or new and their behavior is rewarded or penalized accordingly. A type of possible defection is presented where players cooperate consequently while the participation (sharing of rewards) is cheap and defect when it is costly. An additional parameter of "transaction cost" is proposed to adjust the trust value in such cases, enforcing punishment through reduced payoffs of future outcomes.

[37] extends above secret sharing games to "socio-rational secret sharing' , and proposes providing an amount of shares to a player which is representative of their reputation. Therefore if a given party deviates from sharing their computations, they are penalized in future rounds. The game is designed as a repeated game where reputations are periodically updated, similar to the analysis presented by [34]. They dub their players to be "rational foresighted players", and include their concerns regarding future payoffs in their game design in the form of both long-term vs actual-gain utilities. In the interest of fairness (and inadvertently open participation), the selection of shares allocated to players also consider the age of the player (e.g. $x\%$ of shares are given to new players since they do not have pre-existing reputations).

The authors define a "social Nash equilibrium" where in each round, "player cannot gain any benefit by deviating" when future games are also considered (given the rational foresighted players). They use a strict utility assumption to obtain this equilibrium. To elaborate, the players prefer to maintain a high reputation irrespective of whether they learn the secret or not in a given round in favor of long-term utilities (dubbed "greediness"). And then only they consider the "selfish" utility assumptions of rational secret sharing, which is preference to learn the secret and the preference to have as fewer other members as possible learn the secret.

The authors state that their game is a non-cooperative one, and that the cooperation that occurs is enforced through the design of the game rather than player motivations. We relate to this notion in game design, where the players' cooperation is incentivized through a "public trust network" in the interest of avoiding a degenerative pure strategy equilibrium. They also state that the players are given a certain transparency with regard to the reputation assignment, where individually they can optimize their future payoffs through continued collaboration, and other players are fully aware of the nature of a given player thus not having to make decisions on presumed "beliefs" of others.

It should be noted that the mining pool reward distribution (Refer section 2.3.1) and the selfish-mining attacks (Refer section 2.3.3) respectively resemble the concepts of multiparty computations and equilibrium instability in rational secret sharing. Another interesting aspect of the proposed design is the self-sustaining nature where punishment and reward has been integrated into design itself, as opposed to explicit punishment. This can be related to self-governance aspects of peer to peer networks. Uncertainty of future payoffs encouraging the current action irrespective of whether the protection is

necessary for the current round, is another notion we find interesting. In our problem, the attack probability of a given round is unknown, and as such self-protection costs in current round must be blindly afforded by players. Therefore the nature of strict preference to maintain a higher reputation (or rather a constant reputation in our case) irrespective of whether a direct gain was possible in the current round is a quality that should be encouraged in the players of our game design, as noted by our inability to assign any weight to current action in section 2.8.4. It further helps in avoiding the "cheap" consecutive participation and "expensive" defection attacks noted by the paper, which could recur in our problem as well.

In summary, we find the nature of rational foresighted players complementary to the weighted reputation based service differentiation presented by [34], and consider the future payoff or long-term utility optimization of players a valued addition in a possible mixed strategy equilibrium analysis.

2.9 Learning Of Equilibria

When considering the mixed strategy equilibrium obtained by [34], the probability of active participation in a peer to peer service provisioning game, irrespective of converging to a constant over time, varies at each interval with fluctuating reputations of agents. This dynamic behavior resembles learning at individual agent level, specially considering the rationality and network volatility constraints of non-cooperative players. It would specially be applicable for a network with open participation, where the network has to adjust to new players and the players who leave the network with pre-existing reputation contributions.

2.9.1 Reputation Optimization

In considering modelling our problem as a mixed strategy equilibrium similar to [34], we find several unknown variables concerning the environment that players cannot foresee in making their security investment decisions[9] which in turn makes individual utility optimization quite difficult. Attack probability of the current time period (μ^t), and consensus probability of the blocks proposed during the current time period (f^t) are two such variables. Since players must make a decision before finding these values, we propose future reputation optimization in section 3.3.5.1 as the primary learning method of the peers. The intuition behind the decision was that rational peers would want to maximize utility and minimize cost of participation and thus choose the actions which would cause consensus probability of the blocks proposed during the current time period to be 1 ($f^t = 1$) while allowing redistribution of rewards so that cost is also minimized. Following the learning strategy, peers will choose the action that provides the maximum future reputation that is less than a predictable constant (Refer inequality 3.8). It should be noted this reputation value which dictates the action does not require any unknown variable. Peers can calculate their current reputation by previous round's ledger history, compare it with the upper limit constant and choose the next action which

satisfy both the inequality and maximization requirements. We impose that players will be risk-averse and break ties with regard to inequality in favor of security investment.

However, it should be noted that this does not strictly resemble randomization in a mixed strategy beyond the first few initial rounds, and therefore the efficiency of truly random decision making and reputation optimization will be different. When interpreting mixed strategy equilibria as "a stochastic steady state", where the frequencies of finite actions taken are used exclusively to formulate a response [38, p. 37-44], another noteworthy property is that players do not observe/acknowledge any correlation between different stages or actions of different players. While this asserts certain convergence conditions noted by [26] (Refer section 2.6.1), the presence of a shared history begs the question whether players can truly disassociate from such correlation. As such, this learning strategy rather resembles a correlated equilibrium, much like the mechanism explained in following section 2.9.2, which could be argued to serve our completeness assurance constraints better than truly random decisions.

2.9.2 Regret Matching

2.9.2.1 Correlated Equilibrium

A correlated equilibrium is a generalization of a Nash Equilibrium. Essentially, it increases the social welfare of a game with pure strategy Nash equilibria through the use of signals by a coordinating/randomizing entity [38, p. 45-48]. For example, in the traffic game, where the actions available for 2 players are driving and waiting, one player has to wait while other can drive through for the optimum social welfare, but players have no way to coordinate their actions. A correlated equilibrium provides a private signal to each of the players, such as a traffic light, and since following this action provides the best payoff, both players agree to the action represented by the signal they receive. The respective signals are unilaterally followed by players since it guarantees maximum welfare. It is not enforced, but simply followed due to rationality. Given the aforementioned concerns with shared history, it could be argued that a loosened form of equilibria such as a correlated equilibrium would be more appropriate in representing our problem.

2.9.2.2 Learning via Regret Matching

Regret matching, a simple, adaptive procedure is proposed in [39] for learning correlated equilibria, where players probabilistically choose their action depending on what the total utility for the whole history of the game would have been, had they chosen a different action. To briefly define the procedures, let us define a function Γ which represents the regret of not having played an action $\grave{a} \in A_i$ at time t as opposed to the

[9]This section refers to 3.3.1 for notation used in our model formation and 3.3.5.1 for equations mentioned from our solution formation.

actual chosen action a_i^t.

$$\Gamma^t(\grave{a}) = max\Big(U(\grave{a}) - U(a_i^t), 0\Big)$$

Choose the action $\grave{a}$ with the probability $\rho(\grave{a})$ in the time $t+1$.

$$\rho^{t+1}(\grave{a}) = \frac{\Gamma^t(\grave{a})}{\sum\limits_{\hat{a} \in A_i} \Gamma^t(\hat{a})}$$

Note that this methodology heavily complements our reputation optimizing strategy, given that in each step, the players choose "better actions" as opposed to the best action, which is what yielded the mixed strategy equilibria in the first place. [39] presents that the private signal in regret matching could be the common history (as noted above), which is another reason it is applicable to the distributed ledger scenario.

We will explore this learning procedure to compare the results we obtain in reputation optimization learning strategy, in order to establish which methodology is most efficient in completeness assurance and therefore the maximum possible social welfare (Refer section 3.3.2). However, it should be noted that this is simply a learning procedure and therefore is not capable of obtaining design bounds for the game as the mixed strategy representation has, which will remain our main contribution.

2.9.3 Bounded Rationality

Inductive reasoning proposed by [40], argues that humans learn in the form of belief-models which continuously evolve such that they collectively form an environment in which only the most rewarding beliefs survive, which in turn present a state of convergence. The beliefs are player independent and therefore the problem models heterogeneous players. The author presents the "El Farol Bar Problem" (EFBP), which is where a number of people decide to go to bar only if they can collectively keep the crowd below 60% of total capacity. Overcrowding yields worse payoffs than staying home, but given the bar is not crowded, players obtain higher payoff. The learning strategy proposed for this game is starting off with arbitrary "predictors" for the sum of people in the bar at a given day (such as "same as last Tuesday" or "5 less than last week this day") which are distributed randomly among the players (more than one predictor for each player), and then expecting the most accurate beliefs to survive. The findings indicate that on average 40% of surviving predictors predict the crowd will be over 60% (and they stay home) and 60% of the predictors predict it will be less than 60%. Furthermore, the predictors switch their alignment in repeated play, such that the 40/60 split remains the same but the predictors of each group are not the same as in the previous round.

This learning rule presents an incredibly intuitive convergence property which is that an environment with noise will reach an equilibrium pattern by itself [26]. It further accommodates heterogeneous players, similar to our cost minimizing vs utility

maximizing participants who have differing costs of availability. Equilibrium achieved in "El Farol Bar Problem" (EFBP) is the most structure preserving game theoretical representation we have found for our problem consisting of dynamic agents with heterogeneous beliefs. In both scenarios, agents compete for limited resources, and decision making is done without any prior communication [41]. However, given that the convergence state is accurate on average, it would inevitably result in loss of completeness in the context presented in our problem. It could also be argued that the availability of common history (the private signal in regret matching noted in section 2.9.2) complicates the spontaneity of reasoning conducted by EFBP agents, since the only knowledge available to them should be their own previous decisions and subsequent results wherein our agents will have some context on decisions of other agents. This is where we revert back to the reputation optimization as a more suited solution.

Still, it will be worthwhile to observe the predictor survival aspect of this game for the action sequences available to players. For example, a predictor could be "Be $active$ if you were $passive$ in past two rounds". Even the reputation optimization learning strategy can be seen as a unilaterally elected optimum predictor for homogeneous players. Therefore this analysis provides additional insight to the concerns discussed. It would further be helpful to observe the severity of completeness compromise and obtain a realistic price of malice over a theoretically predicted one.

2.10 Evaluation Of Game Theoretical Models

Our problem represents a distributed system affected by both technical aspects such as network instability, and the social aspects such as agent beliefs and preferences. Network instability in itself could be seen as a result of spatial properties and financial capacities of agents, which can be ultimately abstracted to delays at each agent endpoint. Given that the reviewed literature either remains ambiguous to specific simulation technologies used (potentially custom developed simulations), or uses discrete event based simulators such as "ns3" in evaluating distributed networks, we found that agent-based simulation (ABS) would be best-suited for our solution model. In this section we proceed to justify our choice, and establish *NetLogo* [42] to be an optimum tool for this purpose. Our decision was further influenced by the learning methodologies discussed in section 2.9, and facilitating future research avenues such as implementing open participation for a subset of discussed learning methods.

2.10.1 Multi-Agent Based Simulation

A comparison of simulation models used for distributed system simulations is provided by [43] with an emphasis on advantages of Multi agent based simulation (MABS). MABS is compared with Object Oriented Simulation (OOS), Discrete Event Simulation (OOS), and Dynamic Macro Simulation (DMS), and in extension, its applicability in domains beyond purely social scenarios, specifically distributed systems involving human-machine and human-human interactions, is highlighted.

The depths of properties of a given agent that could be represented via an OOS is considered to be shallow, wherein MABS could cover a broader range. Considering that the choice of simulation model depends on importance given to specific properties, and that such properties are context specific, we prefer MABS due to following two properties that are important to our proposed solution. First would be agent specific modelling concepts, which concerns "mentalistic concepts such as beliefs". The second is adaptivity, which concerns static states vs autonomous learning behavior. In OOS, agents do not have a concept of beliefs. OOS agents further have static reactive behavior over pro-active behavior supported by autonomous learning of a MABS agent.

Preference of MABS over DES remains in scalability, in allowing each agent to be a separate piece of software with pro-active behavior as opposed to a collection of objects. It further facilitates changing the number of participants during the simulation, which is optimum for modelling open participation in distributed systems. While DES and MABS both support structure preserving modelling (i.e. how close the simulation is to reality), the aforementioned property allows MABS to swap a real human agent with a simulated agent, facilitating a more dynamic evaluation scope. The decision further depends on whether the simulation is event driven or time driven. While event driven DES are less time consuming, time driven DES provide a more accurate representation of systems that involve human interactions. Similarly, event driven MABS require a central coordinator which contradicts the distributed nature of MABS simulations, and requires time-consuming decision synchronizations. Considering the continuous feedback loop in our system between network instability and agent security investments, we prefer a time driven simulation for our system, and therefore prefer MABS over DES, but we acknowledge the time consumption of synchronization to be a necessary compromise for structure preservation.

DMS obtains statistical properties of a population and redistributes them among simulated entities to represent reality in the simulated environment. The transitions, or decision making of agents over time are determined probabilistically. Considering that such transitions are applied individually without regards to interactions between agents themselves, DMS is not suited for representing our problem. Furthermore, it could be argued that probabilistic modelling does not represent true autonomy in terms of considering cognitive decisions, which is supported by conclusions drawn in EFBP experiment [40]. [43] poses that it is indeed difficult to observe "emergent behavior" through macro simulations, as opposed to MABS where heterogeneity of individuals can be modelled better.

A common advantage of MABS over other simulation techniques is the high level specification allowed for belief/preference modelling, which can be easily interpreted over mathematical or object centric or programming language specific models. It also significantly reduces the development time given the shallow learning curve.

[43] proceeds to demonstrate the advantages of modelling a "socio-technical" system, where a system which interacts with physical devices and humans is designed in order to minimize resource usage while maximizing satisfaction of users who utilize said resources. It is specially mentioned that while the goals seem contradictory, "In a

de-regulated market, the distribution utilities will compete with added value for the customer in addition to delivery of energy". We find this statement to resemble our agent behavior where both cost of security investment minimization and reward maximization are of significance, while the distributed nature of the system can be called a form of "de-regulation", similar to the original intentions of Blockchain and cryptocurrencies.

Therefore, given the modelling requirements of pro-activeness, the representation of mentalistic aspects, adaptivity, scalability, support for open participation, time driven nature, modelling influence of inter-agent interactions and the subsequent emergent behavior, we choose to evaluate our model using Multi Agent Based Modelling. While our simulation will not support open participation at this stage given that the original reputation based learning proposed require a fixed number of agents per session, a secondary focus of our research is choosing the optimum learning methodology for correlated equilibria, and as such, we value the self-organization aspects provided by ABS modelling as a facilitator for future research.

2.10.2 Simulating networks of proactive agents

Given the choice of MABS, our next concern would be the suitability of a given technology in representing agents in our specific model. The shortcomings of simulators which are based on network design in modelling peer to peer systems are discussed in [44], presenting ABS to be a more flexible alternative. They proceed to demonstrate the applications of *NetLogo* tool in network modelling, a tool more commonly used in social simulations.

The authors consider that given the origin of most simulation technologies available was based in representing network structure, they are not adequately capable of modelling complex network properties such as "self-organization, self-healing and self-adaption" that have emerged in pervasive computing and peer to peer networks. They argue that simulators such as "Opnet, OMNET++ and ns2" and "Tiny OS Simulator" build networks from ground up, and therefore could cause the designer to be stuck manipulating low level network parameters (at the level of devices or protocol layers) such that modelling the human/machine interaction is only done as an afterthought, therefore compromising structure preservation. The argument presented in [43] recurs here, which is that complex systems that has human agents, be it a computer network or otherwise, require a higher level of abstraction in its modelling, which is rarely supported in network oriented simulators.

The authors proceed to critically evaluate *NetLogo*, a tool originating from complex systems research, which allows a higher level of abstraction in both network overlay representation and human interactions. Feasibility of modeling heterogeneity, addressability provided at agent type or property level (e.g. all agents who are "peers", all agents with connection strength above a given constant), expressive power facilitating self-organization and observation of micro level behaviors at each agent and macro level emergent behaviors, ease of implementation, extensibility using languages such as Java or Python (which inherently supports swapping a real agent in place of a simulated

entity as discussed in [43]), and the shallow learning curve are stated as advantages of *NetLogo*. Following a brief introduction to *NetLogo* terminology, the authors proceed to present multiple experiments conducted in *NetLogo* for self-organization, peer to peer clustering and modelling overlay networks.

Further to the convincing argument presented for technical capabilities of *NetLogo* in modelling networks as well as agents in [44], our choice of technology was further influenced by [41]. The authors translate the El Farol Bar Problem (EFBP) presented by [40] in natural language to a *NetLogo* model (currently available in the *NetLogo* Library), and proceed to obtain results similar to the original simulation. It highlights the agents' ability to move through space and time, and how novel behaviors can be observed without the constraints of mathematical modelling, where the behaviors are usually influenced by the designer instead.

A noteworthy contribution of the paper is the discussion on how performance of EFBP predictors are measured in choosing the winning predictor by a given agent. Options for this evaluation are listed as 'absolute precision', where the actual prediction of attendees is irrelevant but simply the end result being satisfactory (throughout history) is measured, 'relative precision' which measures predictor efficiency via relative difference between actual number of attendees and the prediction (without consideration to history), and the 'original precision' equation which was used in [40]. When sufficient history is not available, the predictors are chosen randomly.

The authors also highlight that while structure preservation in time-driven simulations with proactive agents that have belief systems is better represented in computational models, the validation of such models to be tricky compared to mathematical models. Comparison of results is presented as the most common technique used in such evaluations. This feeds into our decision in evaluating multiple learning strategies against each other, given that to the best of our knowledge, an existing solution for direct result comparison is not available at the time of writing. It is also highlighted that a "more complete strategy space would reduce the noise of output results", which we believe is achieved by the reputation optimization strategy proposed in our solution.

In addition to *NetLogo*'s capabilities in time driven MABS simulations, it is also freeware, has a rich support community, is fully documented, has a library with an extensive set of simulation models and tutorials, and supports language extensions [41], [42], [44] making it an attractive simulation platform for our research.

2.11 Summary

We focus on literature from multiple domains that relate to our research problem on incentivizing peers to afford self-protection costs against denial of service based selective interference attacks by malicious agents in BFT based consensus protocols. We choose these attacks due to their capability of subtly manipulating the order of transactions added to the immutable ledger, disintegrating the decentralization property promised by Blockchains and compromising the security property of completeness. Since financially

motivated blockchain based attacks must gain more financial output than input to be rational, we focus on distributed ledger applications outside the financial domain where a resourceful attacker has reason to compromise consensus to his/her advantage.

Given this scope, we begin our literature review with establishing the existence and ongoing research efforts in Blockchain applications outside the financial domain. We establish the requirements of these applications to be high throughput, low latency, consensus finality, immutability, chronological insertion and decentralization.

With these requirements in mind, we explore the consensus algorithms in Blockchain and their strengths and weaknesses. We choose Byzantine Fault Tolerance based consensus protocols for further optimization given that they satisfy aforementioned requirements. We accept its limitation in scalability in favor of the high throughput it provides over proof of work based protocols.

We proceed to introduce game theory as a solution concept for our research problem. We discuss its viability in solving problems in general, and then in various security domains. We note that our research can identify with mechanism design as well. But at the current stage, we focus on extracting critical parameters that we must integrate into our solution model in order to obtain realistic resiliency limits when designing a stable and secure consensus protocol for distributed ledgers.

We discuss various game theoretical solutions proposed in literature which have varying levels of applicability to our research, and again extract conclusions which support and justify the solution model presented in Chapter 3. We discuss our intent to use multiple learning strategies in evaluating our game theoretical model, and introduce the chosen learning strategies. We conclude our literature review by introducing the simulation model and the tools we use for evaluation of the proposed model.

CHAPTER 3
METHODOLOGY

3.1 Introduction

As noted in section 1.3, our problem considers the reliability of rational players in assuring completeness in a distributed consensus. We establish two consequential vulnerabilities that must be mediated to ensure this; "selective interference" and "committing to inaction". In "selective interference", powerful adversaries randomly deny service to honest nodes such that not enough votes are obtained for the block to be added to the ledger, causing the honest nodes to "commit to inaction". The adversary proceeds to remove this interference when the block proposed is not of their interest, essentially obtaining centralized authority on ledger history, violating security assurance of completeness. As such, our goal is to obtain realistic boundaries for completeness assurance of a distributed ledger history in the presence of a powerful attacker.

We approach a solution to selective interference through implementing availability at peer level in the form of security investment, thus mitigating the risk of completeness compromise through players committing to inaction. Considering the peers participating in the game have motivations of maximizing their payoffs and minimizing the costs, we use game theory as a solution concept in order to incentivize players to afford the costs of security investment.

In the following chapter we proceed to model our problem in the form of an infinitely repeated game, and observe the possibility of stable pure strategy equilibria, and in the absence of which, proceed to establish a mixed strategy equilibrium. We influence self-governance within the distributed network through a reputation scheme facilitated by the shared history, and implement self-sustainability and eventual convergence to a sustainable equilibrium through optimizing the social welfare of the game. Considering that social welfare is indirectly in the interest of all rational peers in the form of majority vote requirement which make payoffs possible in the first place, we consider this an optimum approach to modelling our problem.

3.1.1 Contributions

We repurpose the use reputation maximization in literature as the source of strategy randomization in mixed strategy equilibria, and simulate the convergence by using it as a learning strategy to observe how closely a simulated scenario would follow to the behavior predicted by our model. We present this as our main research contribution.

Additionally, we consider other learning strategies such as regret matching and bounded rationality that resemble subsets of requirements in our problem statement. We simulate the results of all 3 learning mechanisms to mutually evaluate their convergence properties. Additionally, we observe which learning strategy is capable of providing the most optimum social welfare, considering that it is the most likely to naturally occur in practice.

We expect our contributions to provide practical insight regarding the feasibility of distributed ledger completeness assurance in time and security sensitive non-financial distributed ledgers, and further help protocol designers make informed decisions regarding payoffs and environmental conditions that must be facilitated to obtain a predesignated threshold of consensus.

3.1.2 Design And Analysis

Our solution design is structured as follows. We briefly introduce the standard game theoretic notation and definitions as a precursor to our model, and present social welfare optimization in a network vs attacker game. Secondly we present our proposed model, "Game of peers". In this section, we denote solution specific notation, reputation modifier functions that allow social welfare optimization and utilities, and proceed to establish pure and mixed strategy equilibria for our model. We obtain the conditions under which a mixed strategy equilibrium would be feasible, and establish a learning methodology (reputation optimization), an upper limit for benefit per unit of cost and a stability property of the network where benefit per unit of cost should be chosen specifically according to the tolerance threshold expected of consensus. In the third and final section, we consider how our proposed model serves equilibrium efficiency measurements in literature.

Chapter 4 illustrates our simulation design and proceeds to discuss how simulation results fare against the analysis presented in this chapter.

3.2 Standard Notation And Definitions

Let us define a game $G = \langle N, A, u \rangle$ where $u_i : A \mapsto \mathbb{R}$

N Finite set of Peers/Players in the game. $n = |N|, n > 2$

P_i Player i. $N = \{P_1, \ldots, P_n\}$

A_i Action profile for P_i. Recall that an action is an exact decision a player takes during an iteration of a game.

A All action profiles. $A = A_1 \times \cdots \times A_n$

a_i An action taken by P_i in a given round. $a_i \in A_i$.

a_{-i} Actions taken by $P_j \in N, j \neq i$.

$$a_{-i} = (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$$

a Actions taken by all players in a given round. $a = (a_i, a_{-i}), a \in A$

u_i The utility function for P_i

u The utility functions for all players N. $u = (u_1, \ldots, u_n)$

a_i^* Best response of P_i to any and all actions of other players including itself.

$$a_i^* \in BR_{a_{-i}} \ \textit{iif} \ \forall\, a_i \in A,\ u_i(a_i^*, a_{-i}) \geq u_i(a_i, a_{-i})$$

3.2.1 Nash Equilibrium

Nash equilibrium would be each agent best responding to all other agents, such that no one would wish to deviate. Thus,

$$a = \langle a_1, \ldots, a_n \rangle \text{ is a Nash Equilibrium } \textit{iif } \forall i, a_i \in BR(a_{-i})$$

3.2.2 Properties of Mixed Strategy Equilibria

A mixed strategy game is a game where players pick actions from their action set with some probability as opposed to deterministically deciding on a best response and playing it repeatedly. It can simply be interpreted as the player introducing randomness to his actions in the absence of pure strategy equilibria.

Given that the **action set for our players is finite**, we reiterate the following characterizations of mixed strategy games from [38, p. 33-34].

Theorem 3.2.1 *Every finite strategic game has a mixed strategy Nash equilibrium.*

Property 3.2.1 *Every action in the support of any player's equilibrium mixed strategy yields that player the same payoff.*

The property 3.2.1 dictates that if an action is taken in equilibrium state with positive probability, then the expected payoff it yields is equivalent to that of any other action taken in equilibrium state with positive probability.

3.2.3 Social Welfare of an Attacker vs Network game

Considering that a stable system would be representative of a state of convergence, we employ mechanism design to present a prototype of an attacker vs network game that allows us to proceed to the design of a game of peers.

A stable system would require for positive social welfare to be obtained by non-malicious participants. As such, we consider the following representative utilities of a game. Note that costs are not financial for attacker, which is represented as 0. While the cost of a thwarted consensus could be of a different dimension other than strictly financial, we follow the reasoning in [31], which states that including such costs would cause the social optima to completely be at attacker's mercy, and that there should be no rational reason for the system designer to mind such costs in game design to begin with.

	Network	Attacker
Payoff of Consensus Achieved	2	-2
Payoff of Consensus Thwarted	-2	2
Cost of defense	1	-
Cost of attack	-	0

Table 3.1: Benefits and costs of network and attacker

		Attacker	
		Attack	Yield
Network	Defend	1,-2	1,-2
	Yield	-2,2	2,-2

Figure 3.1: Payoffs for Network vs Attacker

From iterated removal of dominated strategies, attacker's best response would be to Attack (Utilities $[-2, 2] > [-2, -2]$), making Network's best response Defend. Thus, the unique pure strategy Nash equilibrium for network vs attacker game would be (Defend, Attack), despite the cost incurred by the network. A similar observation was stated by [33] in considering adversaries, where an attacker with a higher cost-benefit ratio than the defender will push the defender to exert maximum required effort.

This indicates that defense must to be integrated to any security sensitive system which might be under attack by a resourceful adversary. The social welfare of such a system will always be less than that of a system that is not targeted. Under this context, we proceed to define our model for peer to peer consensus network.

3.3 Game Of Peers

3.3.1 Specific Notations

Similar to [34], we model an infinitely repeated game G^{∞} which is repeated in sets of n consensus rounds. To avoid confusion, we shall refer a block of n rounds as a period of time. The timing notation will be denoted with a superscript of t.

Let us define our model specific notations.

A_i A_i is the finite action set of P_i. The two actions available to P_i are $passive$ voting irrespective of any utility loss caused by malicious interference or investing in $active$ protection to assure participation in all consensus rounds. We denote this finite action set as $A_i = \{0, 1\}$.

a_i^t $a_i^t \in A_i$ where $a_i^t = 0$ indicates $passive$ availability and $a_i^t = 1$ indicates $active$ availability during the period of time t.

α^t Total security investment from all players

$$\alpha^t = \sum_{j \in n} a_j^t$$

m Number of byzantine players tolerable by a game (consensus algorithm) with n players. $0 \leq m < n$. Given that consensus correctness is a solved problem, we consider m to be a constant throughout our analysis.

μ^t During a period of time t, attacks will be mounted with an exogenous probability of μ^t. This resembles the number of blocks that attackers need to thwart/delay consensus on during the n rounds occurring at a given t. A higher μ^t indicates that the completeness of the data is constantly under attack. $0 < \mu^t \leq 1$

c_i Cost of active availability protection (security investment) for P_i per a round of consensus. While this is a constant for a period of time, the period of time is defined by the number of players. Therefore we determine the cost per a period of time to be nc_i. Another argument that could be put forth is that the higher u^t is, the higher the cost of service would be. But since u^t is variable in each period of time, we presume the cost to be constant on average.

b The constant benefit given to a peer for having voted for a block confirmed into the distributed ledger. For a time period t, the maximum benefit possible to be earned by a peer is nb, and the sum of benefit obtained by all peers in the best case

scenario is n^2b. Benefit will be paid at the end of t depending on the contribution made by the peer towards voting for all n blocks. We consider that the block will not be attacked until it is proposed since whether it's of interest cannot be determined beforehand.

f^t The probability of the all n blocks proposed in current period of time obtaining consensus. It is essentially the probability of number of unprotected nodes being less than m. However, note that if our model is capable of providing sufficient incentive to rational participants, $f^t = 1$ would be a natural by-product.

$$f^t = \begin{cases} 1 & \text{if } \alpha^t \geq n - m \\ 1 - \mu^t & \text{otherwise} \end{cases} \tag{3.1}$$

e_i^t Effective availability, denotes the actual availability of a peer determined by chosen action, the attack probability u^t and whether the completeness was preserved (value of f^t) in the period t.

$$e_i^t = \begin{cases} 1 - \mu^t(1 - a_i^t) & \text{if } f^t = 1 \\ f^t & \text{otherwise} \end{cases} \tag{3.2}$$

r_i^t We denote r_i^t to be P_i's reputation during the current period of time. Reputation is a function of e_i^{t-1}, which is equivalent to the number of consensus rounds P_i participated during $t - 1$ period from the total number of rounds where consensus was obtained. Since reputation is calculated for the previous period, a peer can calculate the value through observing the blocks where his signature is not present. This is in fact the reason for the conditional nature of the effective availability function. We assume $r_i^0 = 1$. The specific function formation is discussed in the following section 3.3.2.

3.3.2 Reputation modifier functions R

Reputation should allow for more rewards to be given to peers who have invested in availability over the peers who haven't. This is simply satisfied through $r_i^t = e_i^{t-1}$. However, this wastes a certain amount of social welfare that could be put to better use. Therefore, we impose the additional requirement of distributing the fixed total payoff of the best case scenario. This essentially imposes a decentralization aspect to the game required in our context. For example, a linear function for reputation would always have players who has low c_i values affording the costs of protection, as observed in [33]. It further integrates the total effort requirement required for completeness assurance along with the selfish/rational nature of peers. As a practical consideration, this condition also allows existing solutions to simply introduce a reputation function if any specific properties observed by our work should be introduced into the network without concerns for additional costs.

The reputation modifier function (denoted R in following text) must be defined such

that it takes the effective availability into account, and that the total reputation of peers is equal to nf^t so that the rewards are distributed proportionate to chosen actions and rounds of consensus obtained.

3.3.2.1 Scaled Reputation

Let us consider the following format.

$$r_i^t = e_i^{t-1} R$$

In considering the total reputations by all peers, we get the below equation.

$$\sum_{i \in n} r_i^t = R \sum_{i \in n} e_i^{t-1}$$

For $\mu^{t-1} > 0$,

$$\sum_{i \in n} e_i^{t-1} = \alpha^{t-1} + (n - \alpha^{t-1})(1 - \mu^{t-1})$$

Note that when $\mu^{t-1} = 0$, reputation modifier function must do nothing ($R = 1$) since no benefit redistribution can take place when protection decision of individual peers cannot be differentiated. Similarly, $\sum_{i \in n} e_i^{t-1} = n$. Given that, we consider the below equation for benefit redistribution through reputation.

$$\begin{aligned}
\Big(\sum_{i \in n} r_i^t\Big)_{(\mu>0)} &= \Big(\sum_{i \in n} r_i^t\Big)_{(\mu=0)} \\
R_{(f^{t-1}=1, \mu^{t-1}>0)} \sum_{i \in n} e_i^{t-1} &= R_{(f^{t-1}=1, \mu^{t-1}=0)} \sum_{i \in n} e_i^{t-1} \\
R_{(f^{t-1}=1)} \Big(\alpha^{t-1} + (n - \alpha^{t-1})(1 - \mu^{t-1})\Big) &= n \\
R_{(f^{t-1}=1)} &= \frac{n}{\Big(\alpha^{t-1} + (n - \alpha^{t-1})(1 - \mu^{t-1})\Big)} \\
&= \frac{n}{\Big(n - \mu^{t-1} n + \mu^{t-1} \alpha^{t-1}\Big)} \\
&= \frac{1}{1 - \mu^{t-1} + \mu^{t-1} \frac{\alpha^{t-1}}{n}}
\end{aligned}$$

If α^{t-1} is lower, then $R_{(f^{t-1}=1)}$ becomes higher, providing more rewards. However $\alpha^{t-1} > n - m$ must also be true for this reputation function to come to effect in the first place.

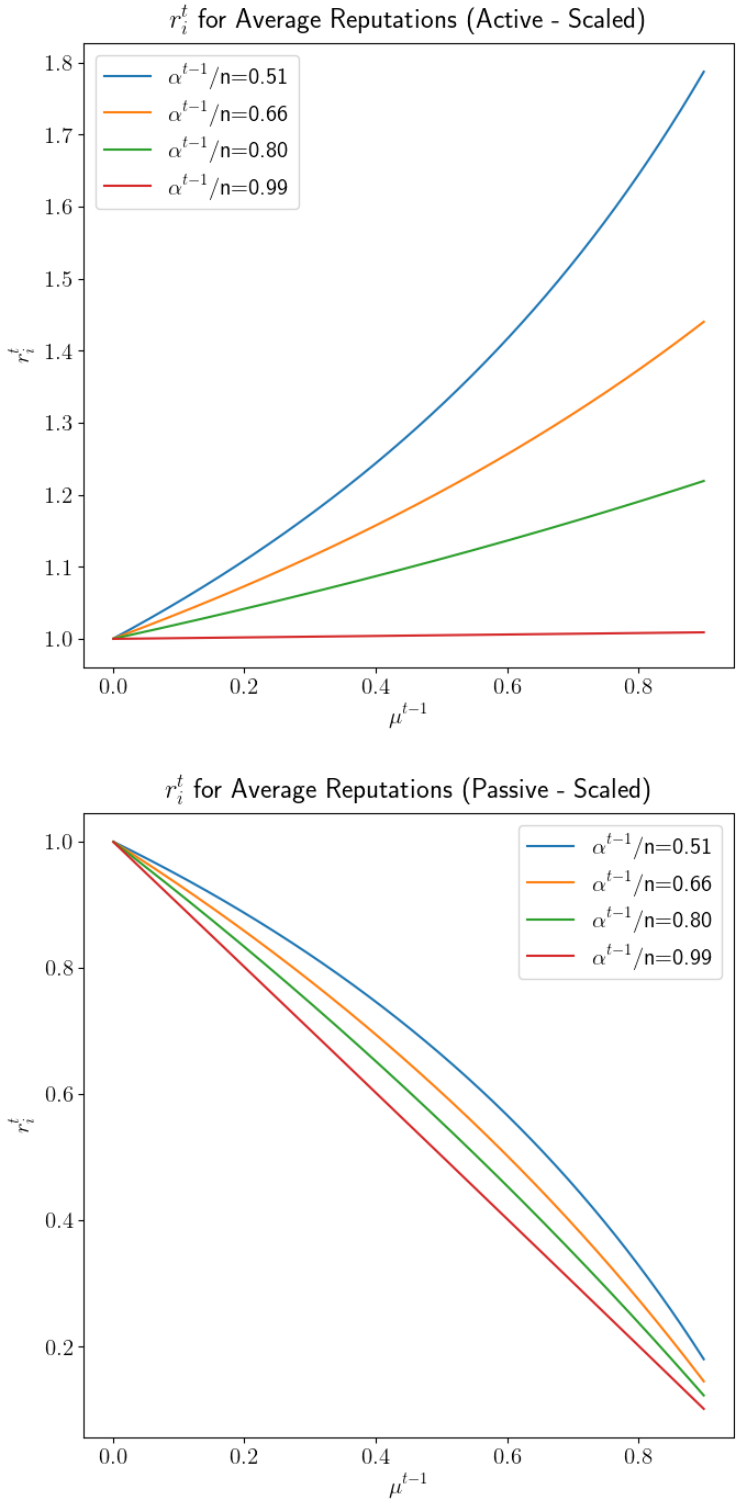


Figure 3.2: Scaled reputations for average availability values against differing attack probabilities

Figure 3.2 represents how r_i^{t-1} behaves for different average availability values for an

$active$ peer and a $passive$ peer. Reputation decreases to 1 as more players invest in $active$ availability but increases as the attack probability goes higher. $\alpha^{t-1} > n - m$ is always assumed in above figure, with $\frac{\alpha^{t-1}}{n}$ providing an indirect reference regarding tolerance thresholds. Note that the range always stays above 0 that even $passive$ peers obtain rewards for participation.

For $f^{t-1} = 1 - \mu^{t-1}$, $R = 1$ must be satisfied. Note that this causes $r_i^t = e_i^{t-1} = f^{t-1} = 1 - \mu^{t-1}$ when the completeness has been compromised during $t - 1$. This represents the worst case scenario since the total of reputations add up to $n(1 - \mu^{t-1})$ as opposed to n, the maximum possible utilization. Function $R_{(f^{t-1}=1)}$ can be slightly modified to result correct values in either case.

$$R_{(f^{t-1}=1-\mu^{t-1})} = \frac{n}{\left(\alpha^{t-1} + \frac{1}{(1-\mu^{t-1})}(n - \alpha^{t-1})(1 - \mu^{t-1})\right)} = 1$$

As such, we can obtain a common function R for both $f^{t-1} = 1 - \mu^{t-1}$ and $f^{t-1} = 1$ scenarios as below, which results in the r_i^t that follows.

$$\begin{aligned} R &= \frac{n}{\left(\alpha^{t-1} + \frac{1}{f^{t-1}}(n - \alpha^{t-1})(1 - \mu^{t-1})\right)} \\ r_i^t &= \frac{n}{\left(\alpha^{t-1} + \frac{1}{f^{t-1}}(n - \alpha^{t-1})(1 - \mu^{t-1})\right)} e_i^{t-1} \end{aligned} \tag{3.3}$$

Expanding above reputation depending on consensus in the previous round,

$$r_i^t = \begin{cases} \frac{1-\mu^{t-1}+\mu^{t-1}a_i^{t-1}}{1-\mu^{t-1}+\mu^{t-1}\frac{\alpha^{t-1}}{n}} & \text{if } f_i^{t-1} = 1 \\ 1 - \mu^{t-1} & \text{otherwise} \end{cases}$$

3.3.3 Utilities

Given above assumptions and notation, we denote utility of a peer as below.

$$u_i = f^t e_i^t r_i^t nb - a_i^t nc_i$$

u_i is the utility for all blocks proposed in time t, therefore the total possible gain for the n blocks proposed becomes nb. However, the gain is conditional on the effective availability e_i^t of the individual peer and the consensus probability f^t of the network. In addition, to differentiate our rewards by previous round participation contribution, we also condition the utility gain by reputation r_i^t. Similarly, if player invests in protection, a cost of nc_i must be reduced from the utility.

Note that if a peer invests in security in two consecutive rounds ($a_i^{t-1} = 1, a_i^t = 1$) where consensus is also achieved for all blocks by the network ($f^{t-1} = 1, f^t = 1$) , this causes $e_i^{t-1} = 1, e_i^t = 1$. In this scenario, utility becomes simply $\mu_i = n(r_i^t b - c_i)$. Per our reputation modifier function R, r_i^t would provide additional rewards if some less than m agents did not invest in protection.

Let us denote $B_i = \frac{b}{c_i}$, the benefit per unit of cost and $U_i = \frac{U_i}{c_i}$, utility per unit of cost.

$$\begin{aligned} u_i &= f^t r_i^t e_i^t nb - a_i^t n c_i \\ &= f^t r_i^t nb(1 - \mu^t + \mu^t a_i^t) - a_i^t n c_i \\ U_i &= f^t r_i^t nB_i(1 - \mu^t + \mu^t a_i^t) - a_i^t n \end{aligned}$$

$$U_i = f^t r_i^t nB_i(1 - \mu^t + \mu^t a_i^t) - a_i^t n \tag{3.4}$$

3.3.4 Pure Strategy Equilibria

For being $passive$ to be the best response in pure strategy case, following inequality must be satisfied.

$$\begin{aligned} U_i^{active} &< U_i^{passive} \\ f^t r_{i_{active}}^t nB_i(1 - \mu^t + \mu^t a_{i_{active}}^t) - a_{i_{active}}^t n &< f^t r_{i_{passive}}^t nB_i(1 - \mu^t + \mu^t a_{i_{passive}}^t) - a_{i_{passive}}^t n \\ r_{i_{active}}^t - \frac{1}{f^t B_i} &< r_{i_{passive}}^t (1 - \mu^t) \\ f^t \Big(r_{i_{active}}^t - r_{i_{passive}}^t (1 - \mu^t) \Big) &< \frac{1}{B_i} \\ B_i &< \frac{1}{f^t \Big(r_{i_{active}}^t - r_{i_{passive}}^t (1 - \mu^t) \Big)} \end{aligned}$$

The higher the f^t is, smaller B_is become capable of satisfying above inequality and subsequently $passive$ action becomes the universal best response, but by itself causes $f^t = 1 - \mu^t$, making the network obsolete and thus invalidating its convergence.

Given that the game represents the pure strategy $f^{t-1} = 1 - \mu^{t-1}, f^t = 1 - \mu^t$ is always true for above inequality to be applicable.

$$B_i < \frac{1}{(1 - \mu^t) \Big(r_{i_{active}}^t - r_{i_{passive}}^t (1 - \mu^t) \Big)} \tag{3.5}$$

Similarly, for being $active$ to be the best response, $f^{t-1} = 1, f^t = 1, \frac{\alpha^{t-1}}{n} = 1$ values can be applied for the following inequality.

$$B_i \geq \frac{1}{\left(r^t_{i_{active}} - r^t_{i_{passive}}(1-\mu^t)\right)} \tag{3.6}$$

If either of the inequalities are true for their respective f^{t-1}, f^t values and for all μ^t, μ^{t-1} values (i.e. We can observe a tangible $B_i > 1$ for all cases), then we obtain pure strategy equilibria. Note that inequality 3.6 uses a weak dominance operator, making $n - m$ votes sufficient for consensus.

3.3.4.1 Equilibrium without a reputation function

If reputation function was not in effect ($\forall\, t; r^t_i = 1$), we get $B_i < \frac{1}{\mu^t(1-\mu^t)}$, $B_i \geq \frac{1}{\mu^t}$ for inequalities 3.5 and 3.6 respectively.

$$\begin{aligned}\frac{d\left(\frac{1}{\mu^t(1-\mu^t)}\right)}{d\mu^t} &= -(\mu^t(1-\mu^t))^{-2}(1-2\mu^t) \\ &= -\frac{1-2\mu^t}{(\mu^t(1-\mu^t))^2}\end{aligned}$$

Inequality function for 3.5 has a local minimum at $\mu^t = 0.5$ where $B_i = 4$. Thus, the inequality can always be satisfied unless a $B_i < 4$ is set by the network designer. Contrarily, for smaller μ^t values, which are likely to be common through continuous play, function 3.6 goes up to infinity resulting in intangible benefits (unless the cost is negligible). Assuming the availability costs are similar such that $B_i \leq 4$ is always satisfied, we observe a pure strategy Nash Equilibrium for $passive$ action. However, as noted by our problem statement, this does not serve as a stable or desirable equilibrium [32]–[35].

3.3.4.2 Results for varying consensus conditions

For scaled reputation in equation 3.3 we observe below for $passive$ action being the best response through inequality 3.5. Recall that $r^t_{i_{active}} = r^t_{i_{passive}} = f^{t-1} = 1 - \mu^{t-1}$ for $passive$ Nash equilibria.

$$B_i < \frac{1}{(1-\mu^t)\Big(1-\mu^{t-1}-(1-\mu^t)(1-\mu^{t-1})\Big)}$$
$$< \frac{1}{\mu^t(1-\mu^t)(1-\mu^{t-1})}$$

Figure 3.3 represents the B_i value limits for differing current and previous attack probabilities. Note that similar to the analysis without reputation consideration, the inequality can be satisfied by defining a sufficiently low B_i. The lowest value above equation can take is when $\mu^t = 0.5, \mu^{t-1} \approx 0$. Defining a $B_i < 4$ would again be sufficient to ensure being $passive$ to be the best response, and it would similarly fail to serve as a desirable equilibrium in completeness assurance.

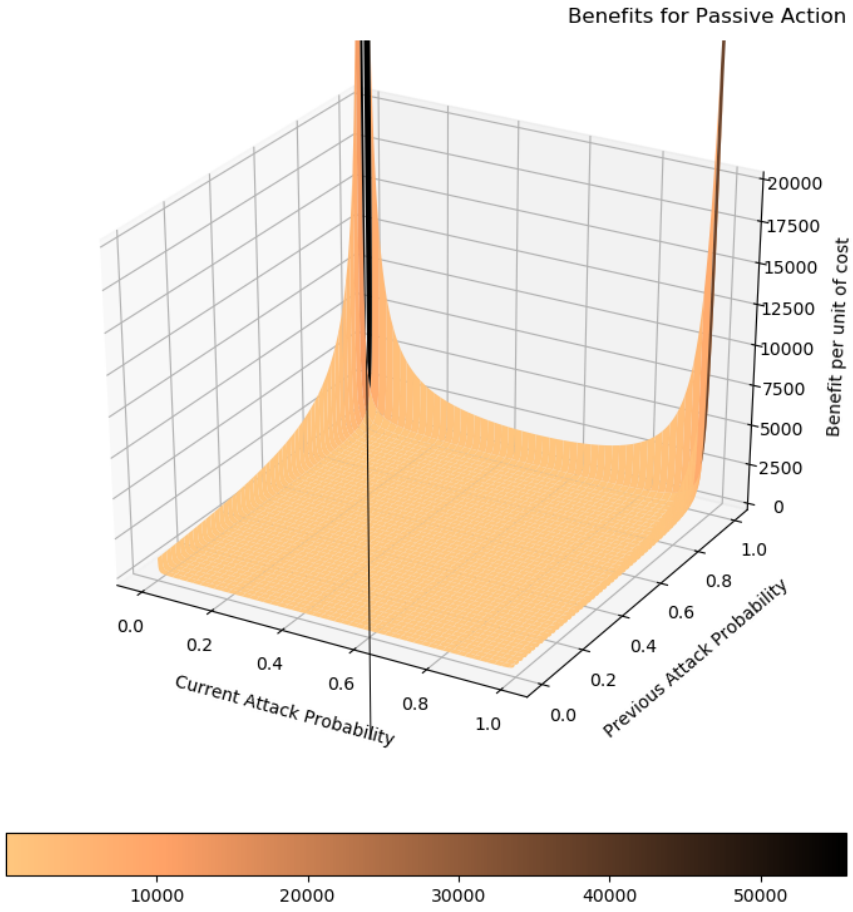


Figure 3.3: Upper limits of benefit for $passive$ action being the best response
Valid values for B_i should be below the surface plot

Similarly, let us proceed to explore possibility of $active$ action pure strategy equilibria

through inequality 3.6 and reputation function 3.3.

$$
\begin{aligned}
B_i &> \frac{1}{\left(\frac{1}{1-\mu^{t-1}+\mu^{t-1}\frac{\alpha^{t-1}}{n}} - (1-\mu^t)\frac{1-\mu^{t-1}}{1-\mu^{t-1}+\mu^{t-1}\frac{\alpha^{t-1}}{n}}\right)} \\
&> \frac{1-\mu^{t-1}+\mu^{t-1}\frac{\alpha^{t-1}}{n}}{1-(1-\mu^t)(1-\mu^{t-1})} \\
&> \frac{1}{1-(1-\mu^t)(1-\mu^{t-1})}
\end{aligned}
$$

As seen in figure 3.4, a tangible upper limit for B_i cannot be observed for lower attack probability values such that $active$ action becomes the best response in all cases. Therefore a pure strategy equilibria where $active$ protection is the best response is also not feasible.

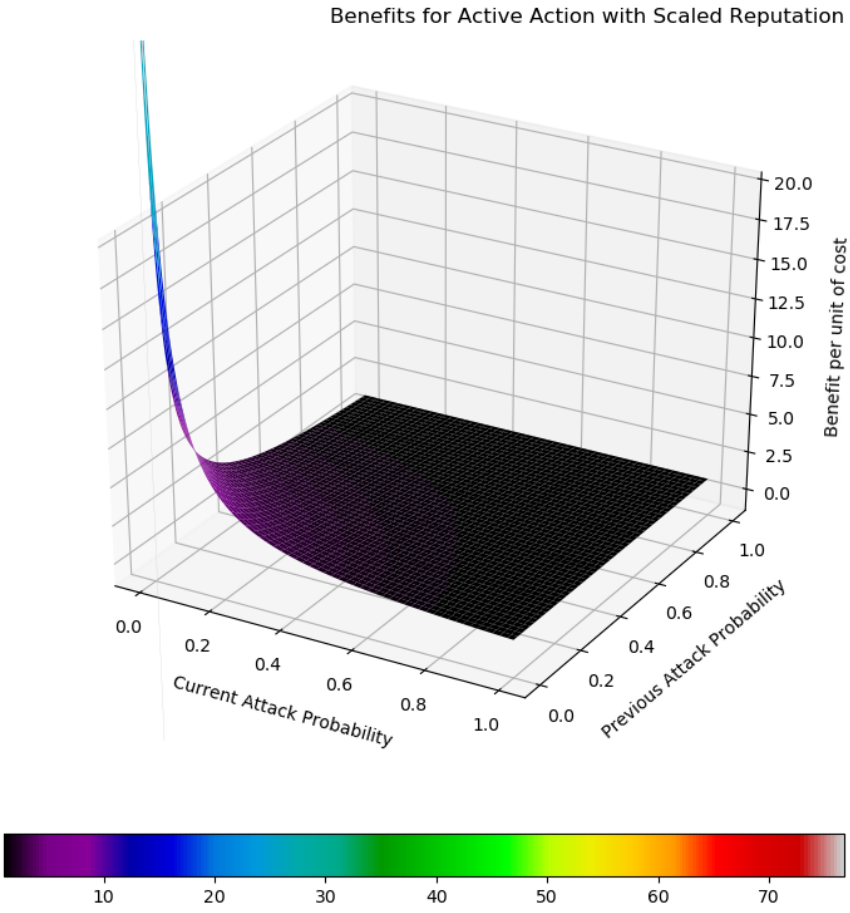


Figure 3.4: Lower limits of benefit for $active$ action being the best response
Valid values for B_i should be above the surface plot.

3.3.4.3 Known minimum for Attack probability

An interesting aspect of figure 3.4 is that the higher either of the attack probabilities are, the flatter the surface. Therefore, it's worth exploring whether a pure strategy equilibria could be influenced if the network could induce a minimum attack probability. For example, observe the same inequality behavior for $min(\mu^t) = min(\mu^{t-1}) = 0.1$.

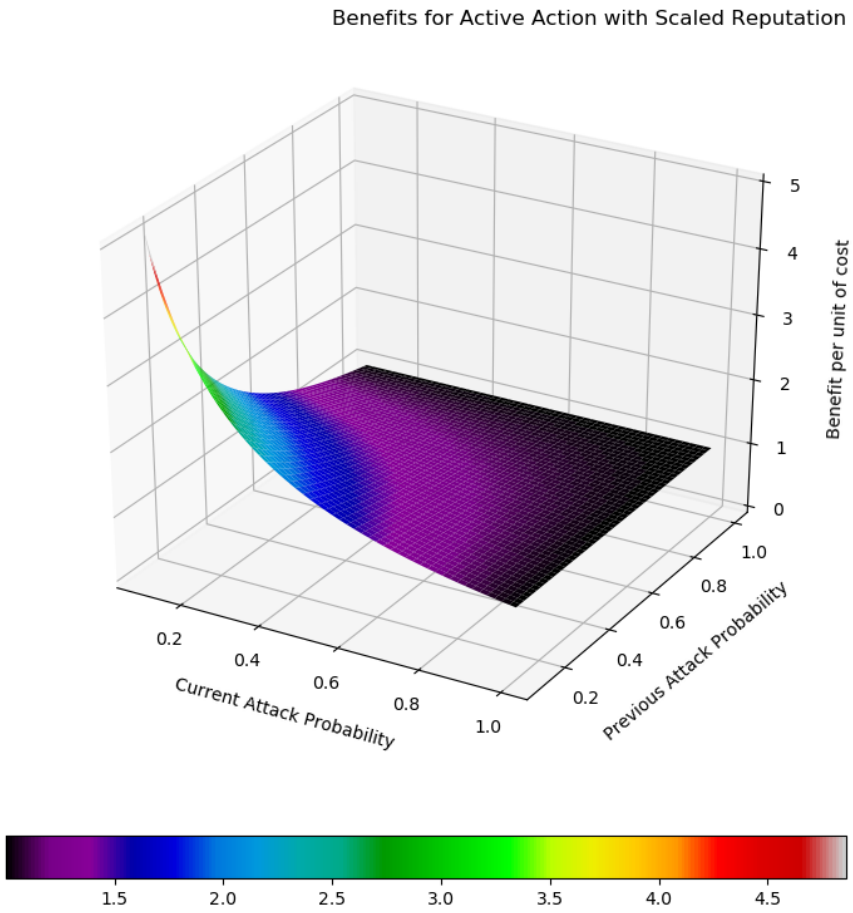


Figure 3.5: Lower limits of benefit for $active$ action being the best response with minimum attack probability

Note that a concise lower limit for B_i can be obtained in this case. Therefore $active$ action could indeed be influenced to become a pure strategy Nash equilibria. This ensures the completeness requirement of the network which is the main objective of the model. It can also be seen as a reasonable assumption in a realistic setting, given the natural volatility of distributed network communication and the timeout in effect. For a higher number of peers, $\frac{n}{10}$ blocks being unable to obtain consensus may not be naturally effective, but the timeout could act as a moderator. For a smaller number of players however, such moderation may not occur naturally.

In any case, even though reputation function was designed for maximum utilization of available social welfare, pure strategies are not suited for that purpose. Either all peers would invest in protection causing unnecessarily higher total cost or no one will invest in protection and peers would be maintaining a completeness compromised ledger.

Another fact that undermines $active$ pure strategy equilibria specifically is that if all players were to invest in protection, and thus the completeness was guaranteed, a rational attacker would not be targeting the network. Without a need for protection, free-riders would appear and completeness would be vulnerable. This speaks to the evolutionary stability of equilibria even if it could be inspired.

Given the shortcomings of pure strategy equilibria and complications in completeness assurance, choosing a dynamic action seems more practical than sticking to the same strategy for the whole of game-play. This leads us to consider mixed strategy equilibria, where the volatility of the environment, diversity of player conditions and utility maximizing nature of peers are better represented through player specific, randomized decision probabilities.

3.3.5 Mixed Strategy Equilibria

Let us assume that the players are willing to invest in $active$ protection with p probability in the infinitely repeated game. Following the theorem 3.2.1 and the subsequent property 3.2.1 we can obtain the following.

$$p\,U_{i_{active}} = (1-p)\,U_{i_{passive}}$$

$$\begin{aligned} p\,n\big(f^t B_i r_i^t - 1\big) &= (1-p)\,n f^t B_i r_i^t \\ p\,\big(f^t B_i r_i^t - 1\big) &= (1-p)\,f^t B_i r_i^t \\ p\,\big(2f^t B_i r_i^t - 1\big) &= f^t B_i r_i^t \\ p &= \frac{f^t B_i r_i^t}{2f^t B_i r_i^t - 1} \end{aligned}$$

$$p = \frac{1}{2 - \frac{1}{f^t B_i r_i^t}} \tag{3.7}$$

The lower $f^t B_i r_i^t$ is, the higher p becomes. However, $f^t B_i r_i^t > 1$ must also be true for such a mixed strategy to exist. Change in p for $f^t B_i r_i^t > 1$ can be observed in figure 3.6. Note that the strategy is capable of always obtaining $p > 0.5$ for any amount of variation, assuring majority consensus (for $n > 100$) with completeness intact. Furthermore, if $f^t B_i r_i^t$ could be controlled to be within a range below the $f^t B_i r_i^t$ for a respective

probability, a tolerance threshold up to that level can be obtained. Lower tolerance thresholds therefore require external conditioning of the network in sustaining the mixed strategy equilibrium, much like induction of minimum attack probability in influencing pure strategy equilibria.

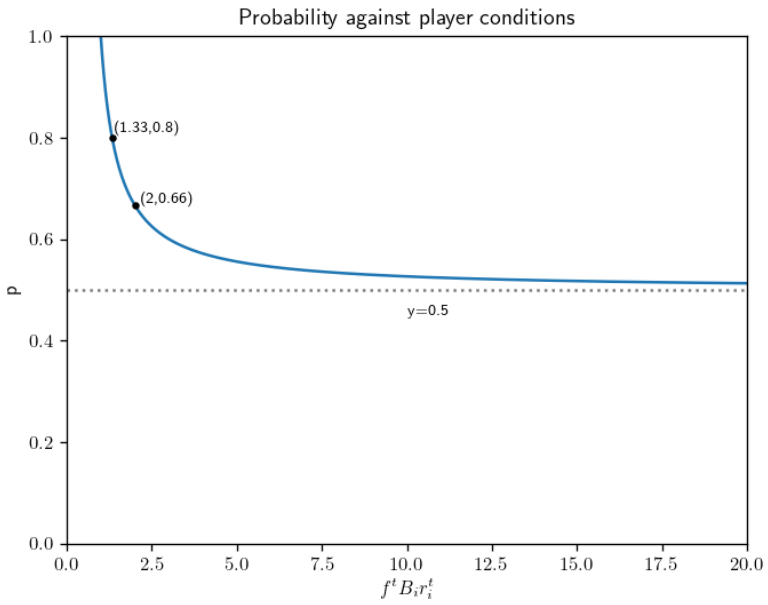


Figure 3.6: Active protection probability for varying environmental conditions
$f^t r_i^t B_i$ denotes collective variation of environment conditions

Figure 3.6 complements the findings of [34] where negligible costs compared to the benefit lead to $p \approx 0.5$ participation with the significant difference that [34] observes less than 0.5 participation due to selfishness of peers while our mixed strategy always provides slightly above 0.5 participation given that precondition for environment is satisfied. Essentially, higher the range available for $f^t B_i r_i^t$, more peers are inclusive of satisfying the above precondition and therefore higher tolerance probabilities become feasible. Note that similar to [34], the convergence at 0.5 can be explained by higher $f^t B_i r_i^t$ values which can also be caused by lower costs. Higher c_i values in fact cause the $active$ protection probability to go higher.

If there is no proper randomization enforced within the game, a manifestation of [33]'s observation regarding free-riding agents in a total-efforts game, where the burden of investment falls on the player (or 50+% of participant players in our case) with highest benefit-cost ratio could likely occur. Incidentally however, we recall the best shot game observation where agents with highest benefit-cost ratios could opt to exert no effort, causing agents with lowest benefit-cost ratios to exert all the effort. Recall that [34] addressed the free-riding problem simply by considering p being a probability causes both cooperation and noncooperation to be valid actions in agents of a homogenous population. We note that given that the distributed ledger is shared history, free-riding agents are capable of noting when the agents with highest benefit-cost ratios will opt out

of exerting the effort, and therefore any role benefit-cost ratios will play will be limited to the initial sequence of actions taken.

3.3.5.1 Reputations

While the above mixed strategy equilibria guarantees majority consensus while ensuring completeness, without knowledge of f^t which is chosen by nature after the actions are committed to, players cannot reason regarding their actions. However, they are aware of the p that would cause $f^t = 1$ which is the utility maximizing scenario for all peers given our reputation function. Therefore we can observe the below inequality.

$$\begin{aligned} 1 - \frac{m}{n} &< \frac{f^t B_i r_i^t}{2 f^t B_i r_i^t - 1} \\ 1 - \frac{m}{n} &< \frac{B_i r_i^t}{2 B_i r_i^t - 1} \\ (n-m)(2B_i r_i^t - 1) &< n B_i r_i^t \\ n B_i r_i^t - n - 2m B_i r_i^t + m &< 0 \end{aligned}$$

$$r_i^t < \frac{1 - \frac{m}{n}}{B_i(1 - 2\frac{m}{n})} \tag{3.8}$$

Since higher r_i^ts cause the most benefit, players will take the action that maximizes r_i^{t+1} up to the value of $\frac{1-\frac{m}{n}}{B_i(1-2\frac{m}{n})}$ in upcoming time periods given the information available from the previous period. This remains the source of randomness in a player's choice to change strategy since r_i^t is calculated for only a single period of time in the past.[1] This nature of utility maximization involving future payoffs resemble Socio-Rational Games where future utility is optimized [37], as noted in section 2.8.5. Key difference would be that socio-rational games use strict preference based utility maximization in obtaining equilibria while our mixed strategy model focuses on randomized learning for optimized social welfare. [37] indeed does not consider or have the need to consider optimizing the expenditure of the entire network.

3.3.5.2 Upper limit for Benefits

In using reputation moderation as a utility maximization strategy, we can observe an upper limit B_i with respect to our scaled reputation function 3.3.

Note that $f^t = 1$ is a continued assumption until otherwise specified.

[1] The implications of reputation based decision making are further discussed in section 2.9.1.

$$
\begin{aligned}
r_i^t &< \frac{1-\frac{m}{n}}{B_i(1-2\frac{m}{n})} \\
\frac{1-\mu^{t-1}+\mu^{t-1}a_i^{t-1}}{1-\mu^{t-1}+\mu^{t-1}\frac{\alpha^{t-1}}{n}} &< \frac{1-\frac{m}{n}}{B_i(1-2\frac{m}{n})} \\
1-\mu^{t-1}+\mu^{t-1}a_i^{t-1} &< \frac{1-\frac{m}{n}}{B_i(1-2\frac{m}{n})}\Big(1-\mu^{t-1}+\mu^{t-1}\frac{\alpha^{t-1}}{n}\Big) \\
&< \frac{1-\frac{m}{n}}{B_i(1-2\frac{m}{n})}\left(1-\mu^{t-1}+\mu^{t-1}\Big(1-\frac{m}{n}\Big)\right) \\
a_i^{t-1} &< \frac{1-\frac{m}{n}}{\mu^{t-1}B_i(1-2\frac{m}{n})}\left(1-\mu^{t-1}\Big(\frac{m}{n}\Big)\right)+1-\frac{1}{\mu^{t-1}}
\end{aligned}
$$

For $passive$ action to be the choice by all players in optimizing r_i^t, (i.e for $a_i^{t-1} < 1$ to always be true), below inequality must be true.

$$
\begin{aligned}
0 &> \frac{(1-\frac{m}{n})(1-\frac{m\mu^{t-1}}{n})-B_i(1-2\frac{m}{n})}{\mu^{t-1}B_i(1-\frac{2m}{n})} \\
0 &> \Big(1-\frac{m}{n}\Big)\Big(1-\frac{m\mu^{t-1}}{n}\Big)-B_i\Big(1-\frac{2m}{n}\Big) \\
B_i &> \frac{\Big(1-\frac{m}{n}\Big)\Big(1-\frac{m\mu^{t-1}}{n}\Big)}{\Big(1-\frac{2m}{n}\Big)}
\end{aligned}
$$

Therefore to inspire $active$ protection, a $B_i \leq \frac{\Big(1-\frac{m}{n}\Big)\Big(1-\frac{m\mu^{t-1}}{n}\Big)}{\Big(1-\frac{2m}{n}\Big)}$ must be chosen.

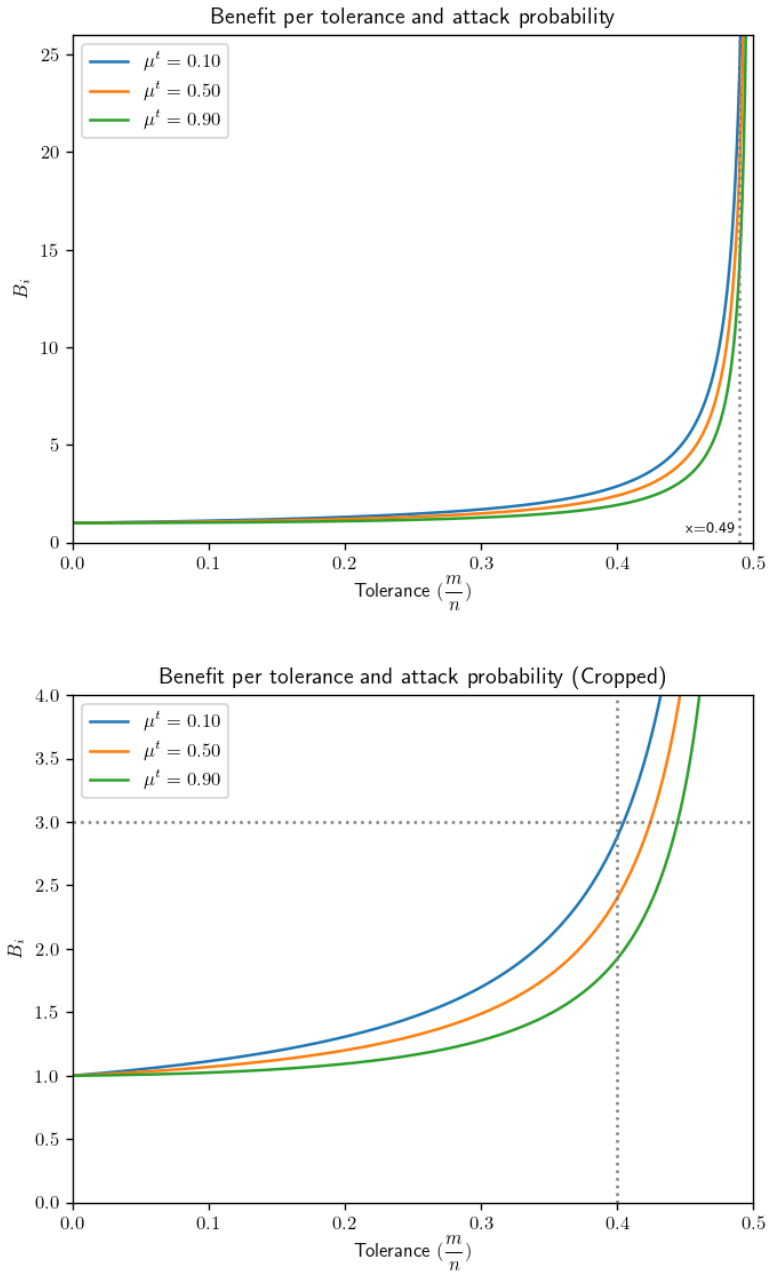


Figure 3.7: Maximum benefits feasible for various tolerance thresholds
Benefit per unit of cost per corresponding attack probabilities must be lower or equal to the plot line.

As shown in figure 3.7, lower the benefit per unit of cost is, more resistant the network is to varying attack probabilities and tolerance conditions. A higher benefit can be provided for higher tolerance conditions while retaining protection for varying attack probabilities, which is necessary in mediating the contradictory nature of benefits provided for active protection. $B \leq 3$ is for example, is capable of incentivizing $active$ protection for a 40% tolerance threshold and maintain equilibrium for any μ^t values.

Higher μ^t values could be sustained with lesser B_i values, similar to the observation for pure strategy equilibria when considering known minimum attack probability.

3.3.5.3 Utilities

Recall from inequality 3.8 that higher protection costs will allow for higher reputations. Applying the inequality value to our utility equation, we observe below.

$$
\begin{aligned}
U_i &< nB_i(1-\mu^t+\mu^t a_i^t)\left(\frac{1-\frac{m}{n}}{B_i(1-2\frac{m}{n})}\right) - a_i^t n \\
&< n(1-\mu^t+\mu^t a_i^t)\left(\frac{1-\frac{m}{n}}{1-2\frac{m}{n}}\right) - a_i^t n \\
U_{i_{active}} &< \frac{n}{\frac{n}{m}-2} \\
U_{i_{passive}} &< n(1-\mu^t)\left(\frac{1-\frac{m}{n}}{1-2\frac{m}{n}}\right)
\end{aligned}
$$

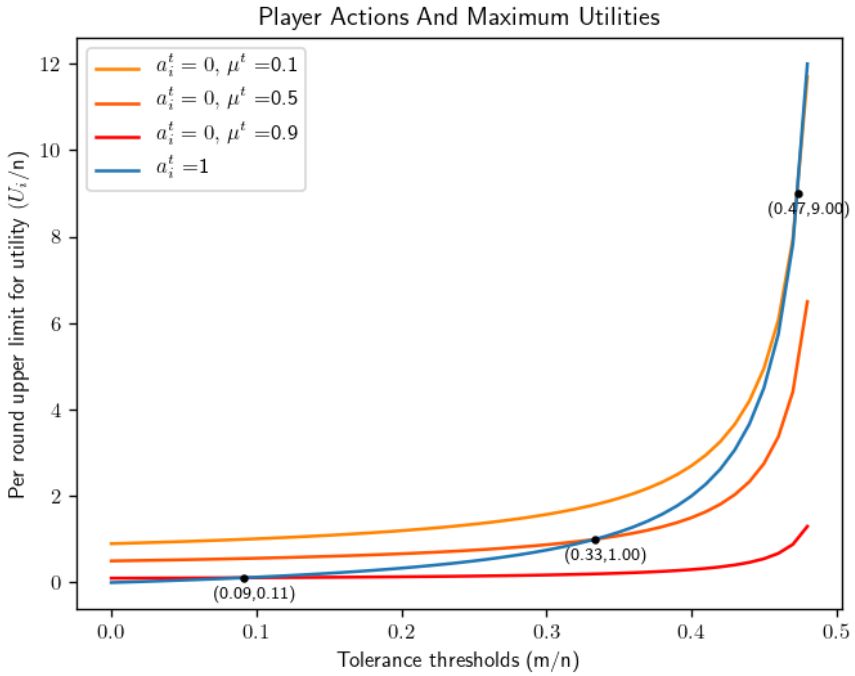


Figure 3.8: Utilities for different actions in Mixed Strategy Equilibria
Values below the lines represent the maximum possible utility while maintaining mixed strategy equilibria

Figure 3.8 shows the upper limits of utility obtainable by actions chosen by peers at different attack probabilities. Note that each attack probability has a tolerance threshold which must exceed for $active$ players to obtain higher utility than $passive$ players. This value remains lowest for the highest attack probability, which provides an interesting contrast to the choice of reputation function and necessity of its benefit redistributing nature.

For lower tolerance thresholds and lower attack probabilities, the model presents a free-riding scenario. However, as noted by [34], by nature of being a mixed strategy it provides fairness for when the same peers choose $passive$ action with $(1-p)$ probability in upcoming time periods. Continued compliance in this regard is further influenced by peer reputation investment not being rewarded until the next round is completed.

3.4 Equilibrium Efficiency Measurements

3.4.1 Social Optimum Welfare

The maximum utility the system can provide occurs when the system is coordinated to always obtain consensus, $\forall\, t;\, f^t = 1$, and yet spend the minimum cost in doing so. In such an arrangement $n - m$ players will be choosing to be $active$ and m players will be choosing $passive$ protection. However, the reputation values of peers will differ.

Consider $\sum r_i^t a_i$, where a system simply assigns players to play fixed strategies for the sake of maximizing social welfare without consideration for fairness.

$$\begin{aligned}\left(\sum r_i^t a_i\right)_{fixed} &= (n-m)\frac{1}{1-\mu^{t-1}+\mu^{t-1}\frac{n-m}{n}} \\ &= \frac{n(n-m)}{n-\mu^{t-1}m} \\ &= \frac{n(n-\mu^{t-1}m-m(1-\mu^{t-1}))}{n-\mu^{t-1}m} \\ &= n-\frac{nm}{n-\mu^{t-1}m}\left(1-\mu^{t-1}\right)\end{aligned}$$

Calculating the total utility of the system where it is coordinated to bear the minimum cost,

$$\begin{aligned}U_i &= f^t r_i^t n B_i(1-\mu^t+\mu^t a_i^t) - a_i^t n \\ \sum_{i\in n} U_i &= nB_i(1-\mu^t)\sum r_i^t + nB_i\mu^t \sum r_i^t a_i - n(n-m) \\ &= n^2 B_i(1-\mu^t) - n(n-m) + n^2 B_i \mu^t\left(1-\frac{m(1-\mu^{t-1})}{n-\mu^{t-1}m}\right) \\ &= n^2 B_i - n(n-m) - n^2 B_i\mu^t\left(\frac{m(1-\mu^{t-1})}{n-\mu^{t-1}m}\right)\end{aligned}$$

Note that the reduction of $n^2 B_i\mu^t\left(\frac{m(1-\mu^{t-1})}{n-\mu^{t-1}m}\right)$ is conditional on μ^t. If $\mu^t = 0$, then the remaining cost $n(n-m)$ becomes unnecessary (to be borne by peers) and a total benefit of $n^2 b$ will be earned collectively by all peers. However, social welfare of our system depends on endurance to availability attacks. Therefore in considering the maximum possible social welfare, we choose the cost inclusive version.

3.4.2 Price of Anarchy

Sum of payoffs in the worst case equilibria would be all players choosing $passive$ action and therefore $\forall\, t;\, f^t = 1-\mu^t$.

$$\begin{aligned}\sum_{i\in n} U_i &= n\Big(f^t r^t_{i_{passive}} n B_i(1-\mu^t)\Big) \\ &= n^2 B_i(1-\mu^t)^2(1-\mu^{t-1})\end{aligned}$$

Recall that for a utility maximization game, the price of anarchy is as below.

$$\text{Price of Anarchy} = \frac{\text{Maximum Social Welfare}}{\text{Social Welfare at Worst Nash Equilibrium}}$$

$$\begin{aligned} POA &= \frac{n^2 B_i - n(n-m) - n^2 B_i \mu^t \left(\frac{m(1-\mu^{t-1})}{n-\mu^{t-1}m} \right)}{n^2 B_i (1-\mu^t)^2 (1-\mu^{t-1})} \\ &= \frac{B_i - 1 + \frac{m}{n} - B_i \mu^t \frac{m}{n} \left(\frac{1-\mu^{t-1}}{1-\mu^{t-1}\frac{m}{n}} \right)}{B_i (1-\mu^t)^2 (1-\mu^{t-1})} \end{aligned}$$

Both optimum social welfare and social welfare at worst case Nash Equilibrium increase as attack probabilities approaches 0. In such a scenario affording the cost of protection is unnecessary, therefore social welfare takes a lower value than welfare at worst case equilibria causing a $POA < 1$. However, as $\mu^t \approx 1$, optimum social welfare is reduced only linearly, while welfare at worst case equilibria is reduced exponentially, causing $POA \to \infty$. Since it is likely that lower attack probabilities are more common, we can conclude that on average, a system that linearly reduce social welfare with attack probability would be considerably efficient.

3.4.3 Price of Malice and Fear Factor

Due to the mixed strategy equilibria, it is not feasible to calculate social welfare for our system since choices depend on player histories and the randomized selection at a given moment (even though probability remains consistent). For example, there could be players whose reputation is $r^t_{i_{passive}}$, and yet choose to be $passive$ for the current round also. There could also be occurrences where the number of $active$ players are larger than $n - 2m$. We consider these unlikely since our rationality condition of maximizing r_i^{t+1} within $max(r_i^{t+1}) < \frac{1-\frac{m}{n}}{B_i(1-2\frac{m}{n})}$ would not cause these outcomes. Assuming these conditions are not true, let us consider the maximum social welfare feasible through our model.

Considering $\sum r_i^t a_i$ for our system, where $n - m$ includes players who changed their strategy from $passive$ to $active$ (i.e. m), and the $active$ players who did not change their strategy.

$$\begin{aligned}\Big(\sum r_i^t a_i\Big)_{reputation\ opt} &= \frac{m(1-\mu^{t-1})}{1-\mu^{t-1}+\mu^{t-1}\frac{n-m}{n}} + \frac{n-2m}{1-\mu^{t-1}+\mu^{t-1}\frac{n-m}{n}} \\ &= \frac{n\big(n-m-\mu^{t-1}m\big)}{n-\mu^{t-1}m} \\ &= n - \frac{nm}{n-\mu^{t-1}m}\end{aligned}$$

Calculating total welfare,

$$\begin{aligned}\sum_{i\in n} U_i &= nB_i(1-\mu^t)\sum r_i^t + nB_i\mu^t \sum r_i^t a_i - n(n-m) \\ &= n^2B_i(1-\mu^t) - n(n-m) + n^2B_i\mu^t\left(1 - \frac{m}{n-\mu^{t-1}m}\right) \\ &= n^2B_i - n(n-m) - n^2B_i\mu^t\left(\frac{m}{n-\mu^{t-1}m}\right)\end{aligned}$$

The loss of social welfare from the reputation optimized system as opposed to fixed strategy system would be as follows.

$$\big(\sum U_i\big)_{fixed} - \big(\sum U_i\big)_{reputation\ opt} = \mu^t\mu^{t-1}n^2B_i\left(\frac{m}{n-\mu^{t-1}m}\right)$$

Note that this difference is the cost of fairness, since fixed actions would allow for permanent free-riders. Cost of fairness would be proportional to attack probabilities and grow exponentially with number of peers.

Recall the price of malice equation from section 2.8.3 for a cost minimization game.

$$\text{Price of Malice for } b \text{ malicious agents } = \frac{\text{Social Welfare with } b \text{ malicious agents}}{\text{Social Welfare at Worst Nash Equilibrium}}$$

We rewrite this equation as following to suit our utility maximization model.

$$\begin{aligned}\text{Price of Malice for } \mu^{t-1}, \mu^t \text{attack probabilities} &= \frac{\text{Social Welfare at Worst Nash Equilibrium}}{\text{Social Welfare at } \mu^{t-1}, \mu^t \text{ attack probabilities}} \\ &= \frac{n^2 B_i (1-\mu^t)^2 (1-\mu^{t-1})}{n^2 B_i - n(n-m) - n^2 B_i \mu^t \left(\frac{m}{n-\mu^{t-1} m}\right)} \\ &= \frac{B_i (1-\mu^t)^2 (1-\mu^{t-1})}{B_i - 1 + \frac{m}{n} - B_i \mu^t \left(\frac{\frac{m}{n}}{1-\mu^{t-1}\frac{m}{n}}\right)}\end{aligned}$$

Figure 3.9 depicts the values for Price of Malice and Fear Factor for different attack probabilities in a reference setting within our convergence conditions. Price of Malice retains a value less than 1 for most attack probabilities. However, it becomes higher for lower attack probabilities, which might be more frequent. Fear factor rises exponentially for higher attack probabilities, due to the rapid reduction of welfare at worst case equilibria. In any case, we can assume that an optimized social welfare could be obtained at least 50% of the time.

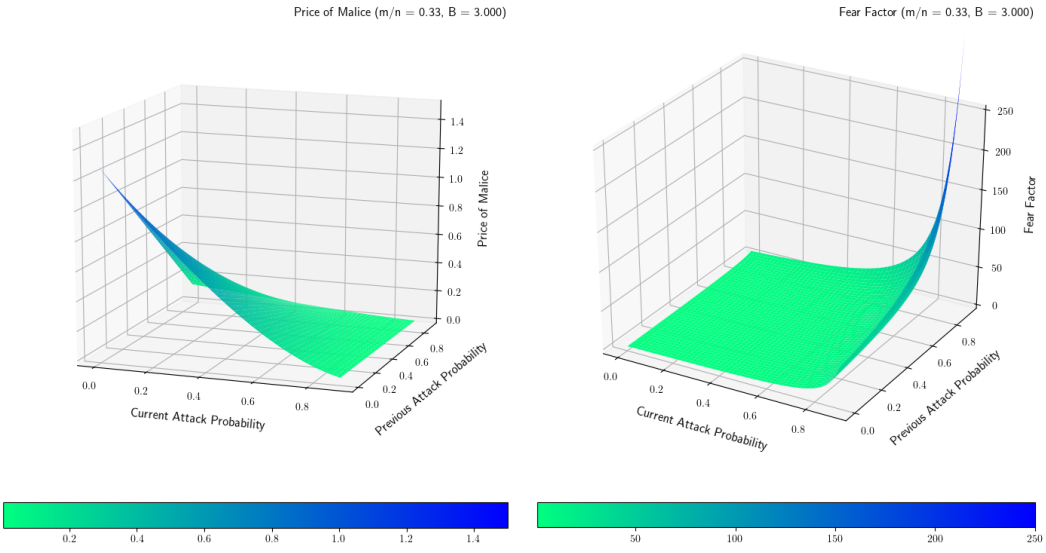


Figure 3.9: Price of Malice and Fear Factor

Price of malice and fear factor against differing attack probabilities in a reference setting $(\frac{m}{n} = 0.33, B = 3)$ μ^t, μ^{t-1} up to 0.9.

Figure 3.10 depicts how Fear Factor changes with differing conditions using the same reference range as a baseline. When increasing benefit per unit of cost, fear factor seems to yield slightly higher minimum values, indicating higher benefits to be capable of better tolerating lower attack probabilities. However, with higher tolerance ranges, range of fear factor seems to shrink, indicating lesser stability, which is consistent with our findings.

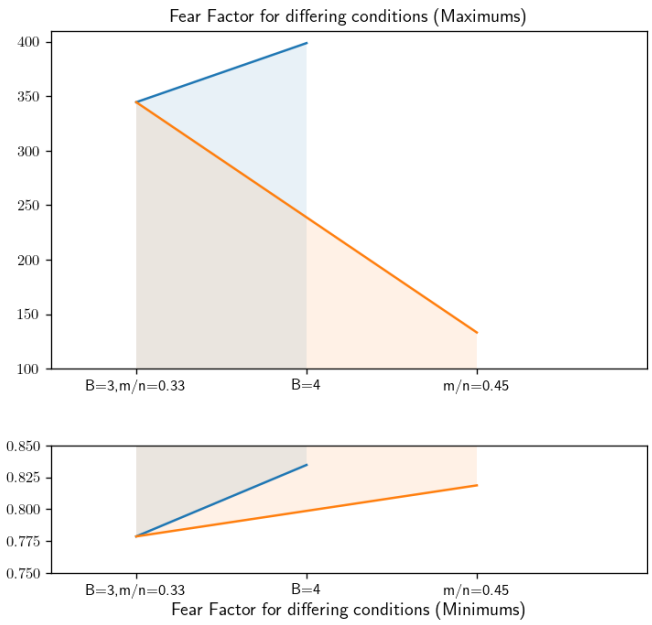


Figure 3.10: Fear Factor

Fear Factor in differing conditions. Reference condition ($\frac{m}{n} = 0.33, B = 3$) is used as the baseline. Note that y axis is scaled separately in each graph.

While the influence of B_i's freefall could not be quantified, it can be presumed to challenge the stability of the equilibrium (Recall transaction withholding related discussion in section 2.5.1.2 [5]). Therefore using maximum benefits per respective tolerance levels could be advised.

3.5 Evaluation Strategy

We simulate a distributed network of peers to evaluate our proposed model. The peers will have benefits and costs assigned to them, and they will choose an $active$ or $passive$ action depending on the output from the learning strategy they have chosen. An attacker will influence their connectivity depending on their action at some of the blocks at each round of the consensus (blocks chosen randomly by attack probability), and the effects of varying attack probabilities on the equilibrium at predefined levels of tolerance thresholds will also be observed.

3.5.1 Evaluating effects of Noise

In the interest of ensuring completeness in the presence of noise, we consider multiple sources of noise against which our model must be resilient. The most prominent generator of noise in our model would be the varying interest of the attacker in blocks that are proposed, represented by the exogenous attack probability which will vary at each round. Secondarily, we introduce noise through technical heterogeneity of peers,

represented by differing delays at node endpoints and the varying number of connections a given peer could maintain with other peers. Note that all results obtained will be subjected to both technical heterogeneity of peers and varying interest of an attacker. Additionally, we assign differing costs to each peer to evaluate how our system would fare for socially heterogeneous agents.

Considering the environmental conditions we have observed for our mixed strategy equilibrium to exist (Refer figure 3.6), we recognize two possible approaches in evaluating effects of noise. We start off with homogenous players with equivalent costs, and consider their actions in a network with varying attack probabilities for certain tolerance thresholds. Secondly, we assign costs as a normal distribution and observe any possible deterioration of completeness requirement for the same tolerance thresholds previously evaluated. This allows us to measure the resilience of our proposed solution, and observe how applicable [26]'s observations will be in a practical environment.

3.5.2 Evaluating Efficiency

As previously noted in section 2.10.2, to the best of our knowledge, an existing solution for direct result comparison for our proposed model does not exist. Therein, we consider multiple learning strategies observed in section 2.9 in evaluating our reputation optimization based mixed strategy equilibrium.

The simulated results of the reputation optimization methodology will therefore be compared with results of learning methods Regret Matching (Refer section 2.9.2) and Bounded Rationality (Refer section 2.9.3) at differing tolerance thresholds. This will provide an additional efficiency measurement for our proposed model other than the measures discussed in section 3.4, and determine which learning strategy provides least deterioration while serving the completeness requirement.

3.6 Summary

Our solution model was presented in this chapter. We translated our problem into a game theoretical representation and derived the utility equations required to analyze possible equilibria.

We observed the shortcomings of available pure strategy equilibria in a repeated setting. $passive$ action is unsuitable due to leaving completeness constraint vulnerable. $active$ action will not fully utilize available social welfare and inspire free-riding in the long run, again compromising completeness. If at least 10% of blocks were to be attacked in each round, a moderate benefit per unit of cost was seen to be capable of influencing an $active$ pure strategy Nash Equilibrium.

We proceeded to analyze mixed strategy equilibria, which yields environmental conditions which must be met for it to be stable. Lower tolerance thresholds call for more controlled environments in sustaining the equilibrium but despite the variance, majority consensus is always feasible. We observed that network resistance for varying attack

probabilities and tolerance conditions can be influenced by low B_i values. A larger range of μ^t values could be sustained through lower B_i values. Free-riding was seen to occur if lower tolerance thresholds are chosen for networks which have low attack probabilities. It was recognized that per a given tolerance level, choosing an exact maximum benefit value would best ensure stability. We further discussed the efficiency measures available for evaluation of our model.

A correlated balance of the varying parameters was noted to be required for the mixed strategy equilibrium to be stable, leading to the belief that the player actions would be based on learning rather than randomization. With that insight, we briefly discussed our evaluation strategy of observing how both noise (volatility) and differing learning mechanisms will impact the security properties expected of our model.

In summary, we presented reputation optimization as a preliminary learning mechanism in obtaining completeness preserving mixed strategy equilibria, which was designed for both individual utility maximization and social welfare optimization.

CHAPTER 4
IMPLEMENTATION AND EVALUATION

4.1 Simulation Design

Simulation design was conducted in three phases. The first two phases included simulating peer-based and server-based consensus with a default security investment protection decision in order to establish NetLogo as a valid candidate in emulating the dynamic aspects of distributed ledger based consensus protocols. NetLogo provided a rich interface for the abstract network representation of the protocol. The peers in the network communicated with a python application to perform the distributed ledger functions such as transaction propagation, block generation, block verification and voting. While the peer-based design evaluated the practical design of a distributed consensus protocol with its inherent network volatility, the server-based design evaluated the node scalability with regards to constant back and forth communication with python application without the burden of network volatility.

However, our problem statement (Refer section 1.3.2), and the subsequent game theoretic solution model requires a higher level of abstraction for an insightful evaluation. The time required for evaluating an infinitely repeated game with a significant number of peers is considerable, and subsequently a simulation that monopolizes available resources to perform expensive and sophisticated functions performed in a distributed consensus protocol is not ideal for the evaluation of multiple learning strategies.

The practical limitations of peer-based and sever-based simulations therefore led us to the third and final phase of simulation, where the distributed consensus functions performed through python application was simplified and migrated to NetLogo code itself, reducing inter-process communication costs through abstracting the voting process to simply recording the list of active peers during a "block proposal", and using these records as "ledger history" for subsequent reputation and reward calculations. This phase narrowed its focus to be learning-based, since it serves our problem statement of studying "relative relationships between incentives, costs and peer security". It allows the network volatility impact of the peer-based design to be observed via block timeout, number of peer connections and delay at each peer in receiving messages, providing insight regarding the increased volatility in the dynamic peer-based simulation, while further facilitating higher scalability than the server-based design while retaining the peer oriented network design.

NetLogo's default duration measure of 'ticks' was used during the simulations to demark the passage of time. During first two phases, message scheduling was done via

the NetLogo *Time extension*[1] . During the learning-based simulation, NetLogo *Rnd extension*[2] was used for weighted random draws.

4.2 Peer-Based Simulation Design

4.2.1 Design

The peer-based simulation uses direct messages from server for transaction propagation[3] and a NetLogo facilitated gossip protocol for block propagation among peers. Peers form a number of connections with neighboring peers that range from a necessary minimum and a maximum. A connection delay value between an arbitrary minimum and a maximum is assigned to all peers, and at receipt of a message they schedule their response at the time of $(\text{now} + \text{connection delay})$. Both connection delay based scheduling and variation of allowed peer connections emulate the dynamic and volatile nature of peer to peer networks. Additionally, connection delay, along with the timeout that is in place successfully abstracts emulating how the stability of connection affects/regulates peer participation and reputation of the consensus protocol.

Block proposal is conducted by an active peer in the network, and once the block is voted for (irrespective of being added to the chain or not), the leadership is handed over to another active peer for the next block proposal. The proposing peer sends the block to all its neighbors, and they then vote for the block and broadcast it to their respective neighbors until the required number of votes are obtained, representing a gossip protocol.

[1] https://github.com/NetLogo/Time-Extension

[2] https://github.com/NetLogo/Rnd-Extension

[3] while peer based transaction propagation was implemented, it was substituted with server based propagation due to heavy resource costs.

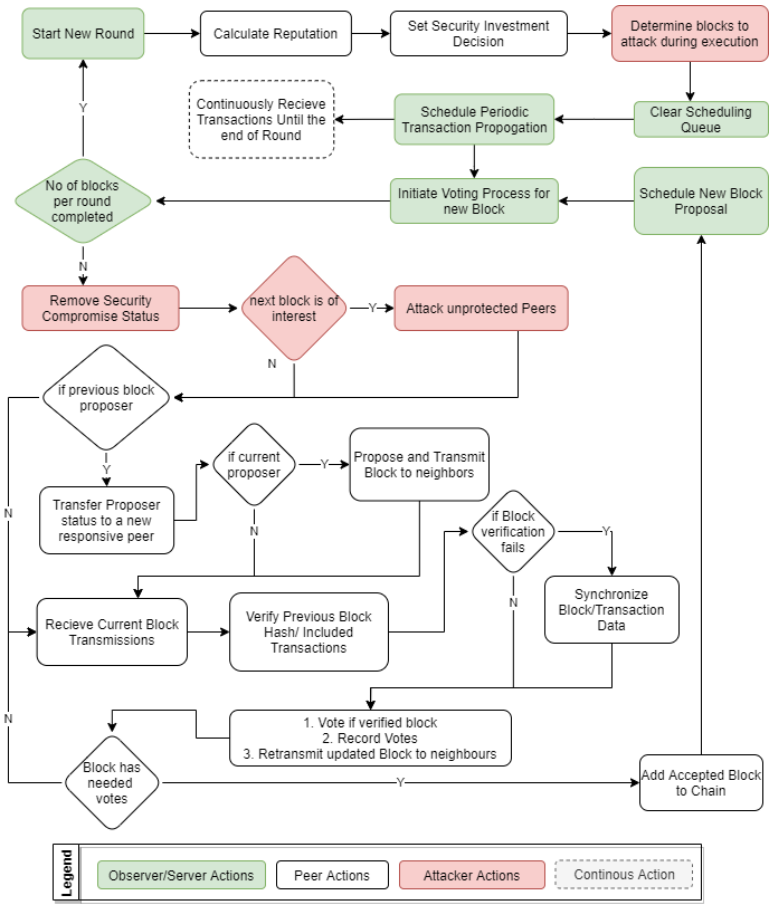


Figure 4.1: Peer-based simulation design

4.2.2 Implementation

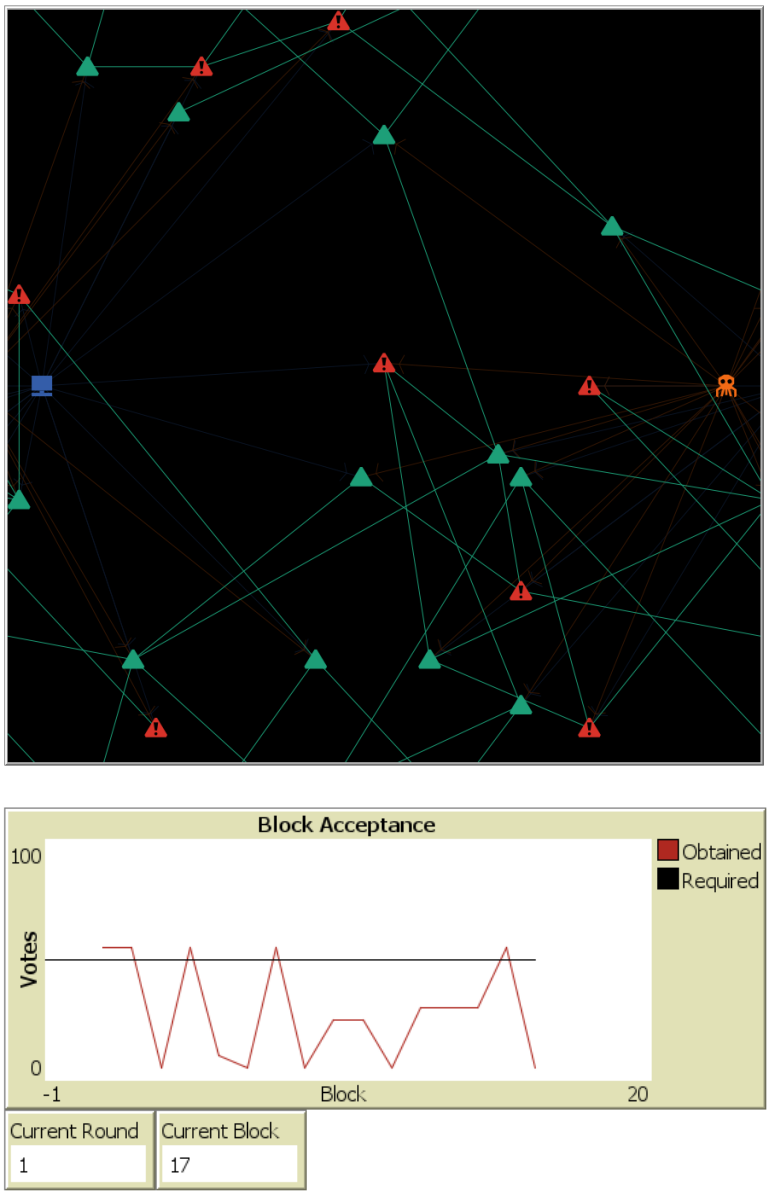


Figure 4.2: Peer-based simulation implementation
Block acceptance results after 1000 'ticks' where 20 peers use gossip protocol for consensus propagation with 45 'tick' timeout, 1-5 alternating connection delays, 2-4 alternating allowed peer connections, and higher than 0.5 probability of a block being attacked. > 50 votes are required for consensus and protection is invested in by default for > 50 peers.

4.2.3 Limitations

One of the limitations in the peer-based design was that event queue cannot be cleared for individual users or specific events. Since the gossip propagation requires continuous scheduling, without periodical clearing of the queue, events are scheduled exponentially

causing the simulation to lag. As such the transaction stream scheduled also had to be cleared and reinitiated after voting for each block is conducted. However, the impact of this limitation was quite minor considering that only mock transactions were used. Note that in reality, peers clearing a specific event from the schedule would be quite trivial.

Another limitation of the peer-based design was the inability of integrity verification of the voter list (beyond a simple check on whether the voter list is a subset of registered users). For reputation calculations, votes received after block acceptance would also have to be recorded, resulting in dynamic voter lists at each peer. A hash of such a list therefore could not be included in the block itself without another round of gossiping, which apart from being expensive, would not represent the reputation variance that ultimately causes network volatility and completeness stability, which is a property we needed to retain.

Design of the server-based simulation (Refer section 4.3) was influenced by the scalability limitations of the peer-based design, where a successful run was limited to about 20 peers and 1000 ticks. Using a higher number of peers for longer durations caused the simulation to be unstable. This limited us from evaluating our game theoretical model, considering that learning had to be done over a large number of rounds (representing an infinitely repeated game).

A more prominent limitation was the forking that occurred at peer endpoints. Since the previous block hash must be same as what current block indicates it was, various peers who do not have the up-to-date chain (despite the update requests sent to other peers, who due to scheduling conflicts or previous forks may also not have up-to-date chains) are unable to successfully vote for the block despite their availability. To overcome this limitation, peers were made to request the recent chain from the peer who gossiped the block, and additionally for the block to be communicated to all available peers once the required number of votes are reached. It should be noted that this would not represent a truly distributed consensus, but implementing sophisticated fork handling scenarios were too expensive considering the pre-existing scalability limitations at this phase of simulation.

4.3 Server-Based Simulation Design

4.3.1 Design

Server-based simulation design was considerably more straightforward than peer-based simulation, and was by design made to use less resources during the execution. The transaction propagation was kept similar to the peer-based design while direct communication with server was used for block propagation. It represents the behavior of a two-phased protocol of voting and acceptance, where the proposer is static and invulnerable.

The server generates a block and broadcasts it to all peers. The peers verify that transactions were identical to the ones received to them alongside the previous block

hashes, and accordingly communicate their vote to the server. After collecting the votes until the timeout, the server generates the voter-list and the voter-list hash and broadcasts the finalized block to all peers who then add it to their respective chains.

While this design is incapable of simulating the decentralization of the network, it provided insight regarding mutual limitations of both peer-based and server-based designs, determining the requirement for a more scalable and model evaluation oriented simulation design, while simultaneously confirming how network volatility caused by timeout and connection delay features could provide varying consensus outcomes and thus facilitate the dynamic environment required for successfully evaluating learning strategies.

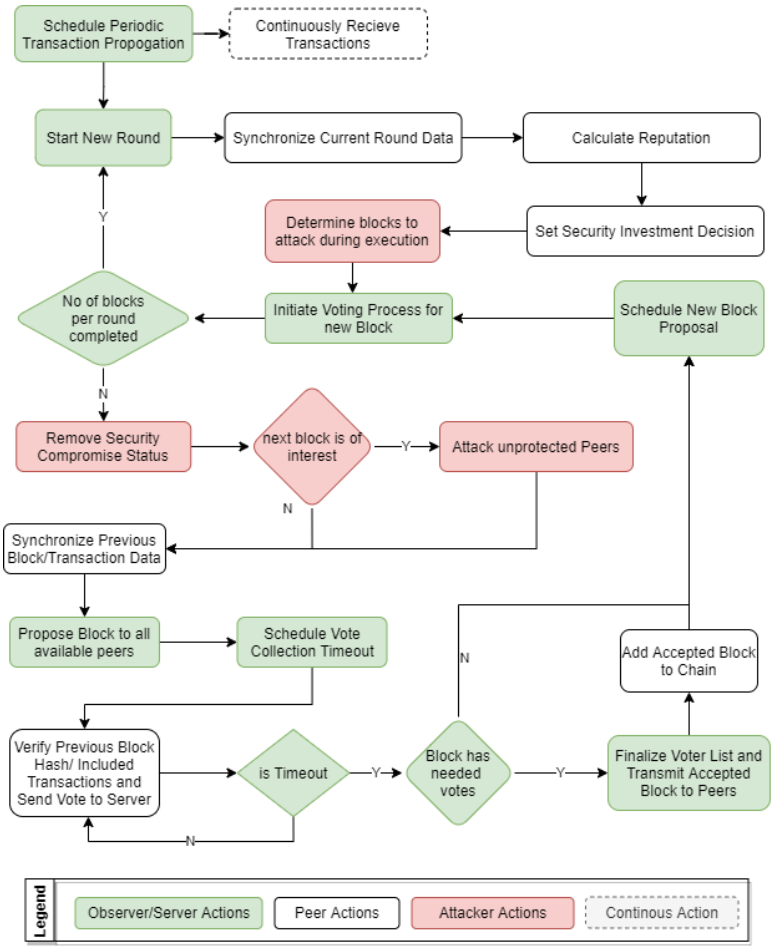


Figure 4.3: Server-based simulation design

4.3.2 Implementation

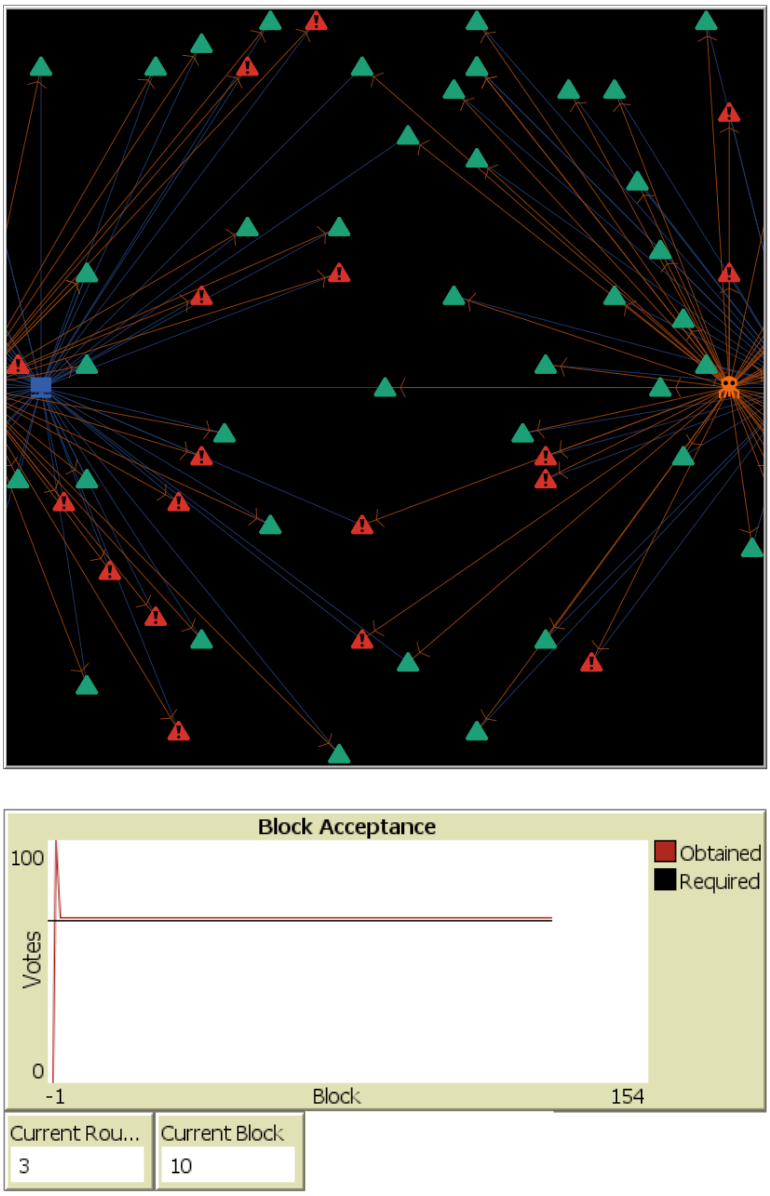


Figure 4.4: Server-based simulation implementation
Block acceptance results after 30000 'ticks' where 60 peers use direct communication with a server for consensus propagation with 30 'tick' timeout and higher than 0.8 probability of a block being attacked. $> 66.66\%$ votes are required for consensus and protection is invested in by default for $> 66.66\%$ peers.

4.3.3 Limitations

Certain limitations such as reduced network volatility, and lack of true decentralization were present by design to increase scalability of the application. However, considering that the distributed ledger is meant to be constant between all peers, reputation calcula-

tions were not affected by proximity to neighbors and the timeout value, which can be seen as a disadvantage in evaluating how network volatility aspects would impact peer learning strategies.

In addition, scalability limitations of the peer-based design were persistent in server-based design as well, albeit to a lesser extent. A successful run was limited to about 60 peers and 30000 ticks, which yielded over 3 rounds of learning which was still insufficient to observe the patterns of learning between the 60 peers.

Forking was still seen to be present considering the scheduling conflicts and lacking of graceful termination in cases of dynamic edge cases. This was however significantly lesser than that of peer-based design, as made evident from the increased scalability from the peer-based simulation (Refer figures 4.2 and 4.4).

4.4 Learning-Based Simulation Design

Considering scalability limitations of peer based and server based simulations (Refer sections 4.2 and 4.3), it could be observed that heavy inter-process communication with the python application hinders uninterrupted execution of the NetLogo simulation which must support over 100 peers through a considerable number of round executions until the learning strategy could stabilize the collective security investment of the network. Therefore the learning-based simulation specifically focused on a higher level of abstraction that could support extensive scalability and uninterrupted execution.

4.4.1 Design

Transaction propagation functionality was detached for the learning-based design. A mock block was proposed by a randomly selected peer who was not under attack, reducing the previous proposer handover process to a single step. Block and the voter list was propagated to all available peers through gossip protocol, and their re-transmission was conducted after the respective delay at each peer has passed.

This implementation also simplified certain calculations such as deriving the total number of available peers since the network variables were readily available for peers to use during their respective calculations. As such, certain liberties were taken where a realistic implementation would have to iterate through the block history of each peer in order to determine which peers have invested in security in a given round (i.e. peers whose signature was present in all voted in blocks in the given round).

This implementation was successfully able to retain the advertised scalability aspects of NetLogo. Considering that NetLogo is primarily a social simulation tool, despite supporting extended functionality through plugins, its requirement for keeping the model representation relatively simple in order for uninterrupted executions must be noted.

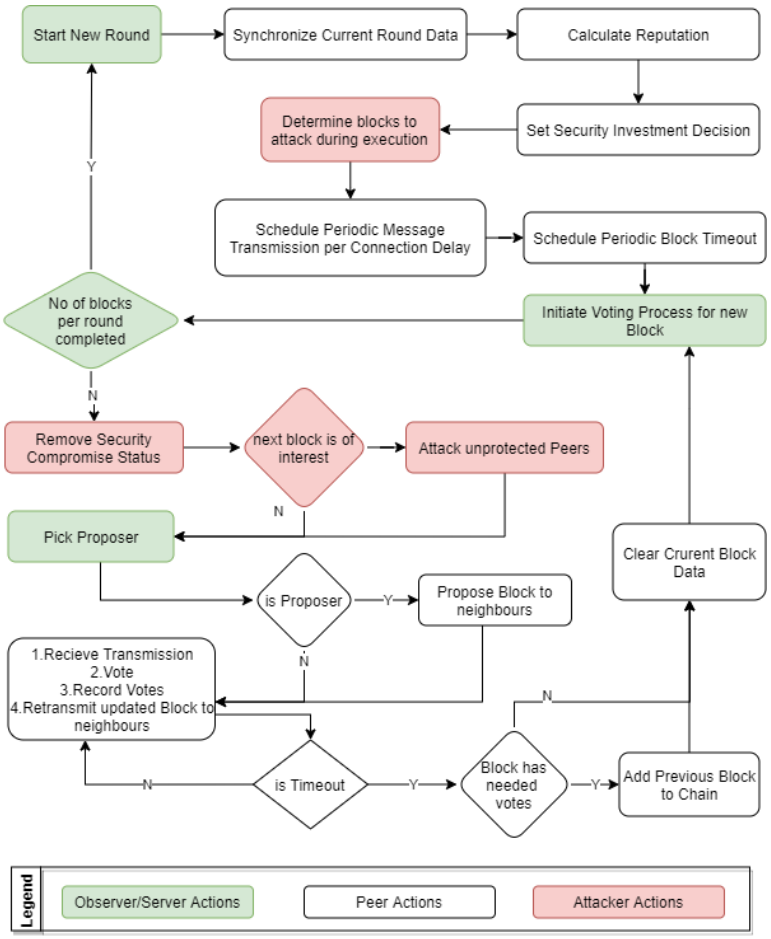


Figure 4.5: Learning-based simulation design

4.4.2 Implementation

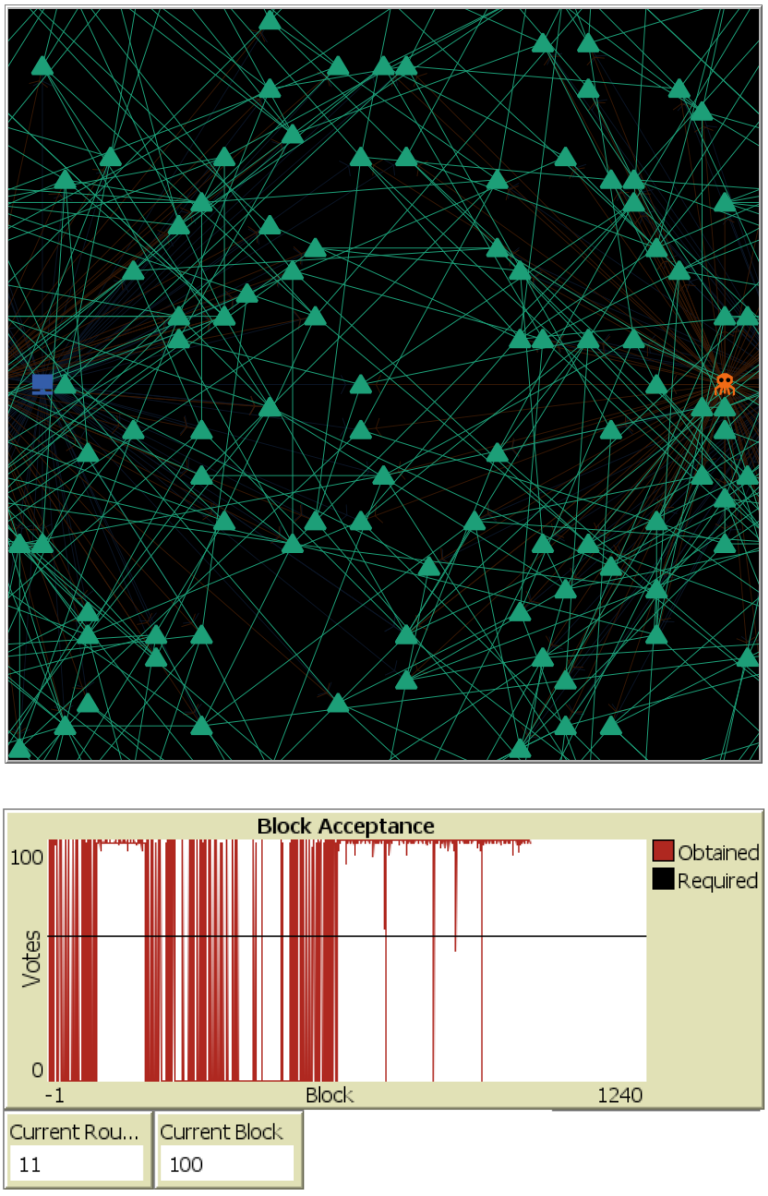


Figure 4.6: Learning-based simulation implementation

Block acceptance results after 10 'rounds' where 100 peers use gossip protocol for consensus propagation with 15 'tick' timeout, 2-9 alternating connection delays, 2-6 alternating allowed peer connections, and higher than 0.5 probability of a block being attacked. $> 60\%$ votes are required for consensus.

4.4.3 Limitations

Learning based simulation did not use scheduling functionality of the *Time extension*, but used native NetLogo functions to induce connection delays and block timeouts. Some amount of time allocated to the first block timeout was consumed during round initiation calculations. Therefore in the Block Acceptance graphs, first block of each round is not accepted by the network due to the remaining amount of time not being

sufficient for gossip propagation. We advise the reader to interpret the results with this in mind, and to consult the Total Protection graph in doubt.

In case all peers were under attack, the simulation was tweaked for one random peer to come online and propose a block irrespective of their protection decision. Considering that this block would never be accepted, the impact of this was negligible.

The main limitation of the learning-based simulation was the level of simplification required for continued execution of the simulation. Through scrapping the transaction and block propagation as well as the block verification aspects of the system, the network volatility induced to the system, and therein the level of noise present which ultimately affects the learning outcomes and investment probabilities were reduced. However, studying the effects of such noise could also be induced through lower timeout values, higher ranges of connection delays and smaller ranges of connections allowed between peers.

4.5 Design Constraints And System Level Limitations

Our simulation was designed to accommodate at least 100 peers, guided by evaluations present in existing literature. Once the learning strategy started, random attack probabilities were considered in each round in order to emulate the varying interest of an attacker depending on the content of the blocks proposed.

Given the nature of the game theoretical model proposed, the effects of open participation are not incorporated in our analysis. While this is a limitation of the reputation-based learning based on our round-based infinitely repeated mixed strategy equilibria, considering that our implementation used NetLogo, it should be possible to extend our model to evaluate the effect of peers dropping and joining the network during the simulation execution. We hope to evaluate the impact of varying number of peers in future work.

Another noted design constraint was the incentive for message propagation. In a practical implementation (and in Peer and server based designs), this is baked into the solution by the requirement of peers having up to date transactions for block hash verification and subsequent voting purposes. In addition, since it is a known fact that there will be some peers who will not invest in security, reaching everyone available in the interest of obtaining the required number of votes before timeout is also required. Another aspect that influences this would be the necessity of accurate calculations regarding reputation, benefits and system-wide availability during learning processes. Learning based simulation design however, does not include transactions in it's implementation, and thus is disadvantaged in representing the amount of network volatility that would be present in a real network.

In bounded rationality learning methodology implementation, parameters of memory size and the number of strategies for peers were kept constant. A comparison of differing parameters for this methodology alongside the existing parameters of our model was considered out of scope for our evaluation. Note that the implementation of this learning

strategy was extracted from the NetLogo model library as is [42], [45], [46].

4.5.1 Evaluating influence of varying parameters

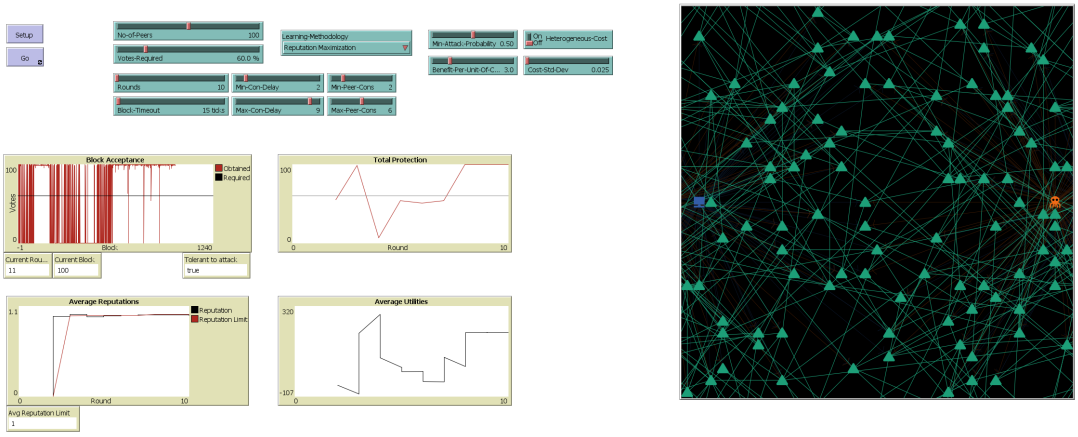


Figure 4.7: NetLogo Interface

As indicated in figure 4.7, evaluation of how each parameter influences the results was facilitated through *BehaviorSpace* experiments. However, given the number of parameters available, an exhaustive analysis was considered strenuous. Therefore the experiments conducted were compared to the performance of a baseline experiment.

In evaluating the heterogenous peer behavior (through varying costs), a normal distribution of a mean of 1 and a standard deviation parameter within the range of $0.025 - 0.125$ was used for cost assignments. Benefit per unit of cost variable by each peer was modified per the cost value during peer initiation. A higher variation for cost was considered unwise since it would violate our mixed strategy existence condition of $f^t B_i r_i^t > 1$ (discussed alongside equation 3.7).

4.6 Simulation Results

We follow the evaluation strategy discussed in section 3.5 in organizing our experiment outcomes. First we focus on the reputation optimization strategy under differing noise inducing conditions, and evaluate how it fares alongside our theoretical predictions. Secondly, we proceed to evaluate our learning strategy itself via comparing the simulation results of reputation optimization with regret matching and bounded rationality learning methodologies. We further correlate the implications of our results with practical implementation constraints at each aforementioned evaluation stage.

An additional set of Simulation Results are included in Appendix B, in order to confirm the repeatability of the simulation results despite the differing random parameters in different runs.

4.6.1 Effects of Noise

4.6.1.1 Homogenous peer behavior at differing environment conditions

In this section, we proceed to evaluate the predictions in figure 3.6, where as long as the votes required for consensus is over 50%, the conditions in the environment should not matter for convergence.

Varying Benefits As indicated by figures 4.8 and 4.9, it can be observed that differing benefits have no impact on the Reputation Optimization strategy convergence.

Reputation Maximization Rounds: **10**

No-of-Peers	Votes-Required	Block-Timeout	Benefit-Per-Unit-Of-Cost	Min-Attack-Probability
100	50.1	25	3	0

Min-Con-Delay	Max-Con-Delay	Min-Peer-Cons	Max-Peer-Cons	Heterogeneous-Cost	Cost-Std-Dev
2	9	2	6	false	-

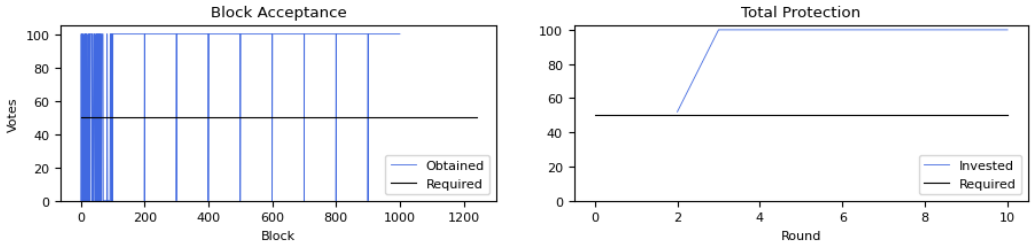


Figure 4.8: Homogenous peer behavior at Benefit per unit of cost 3

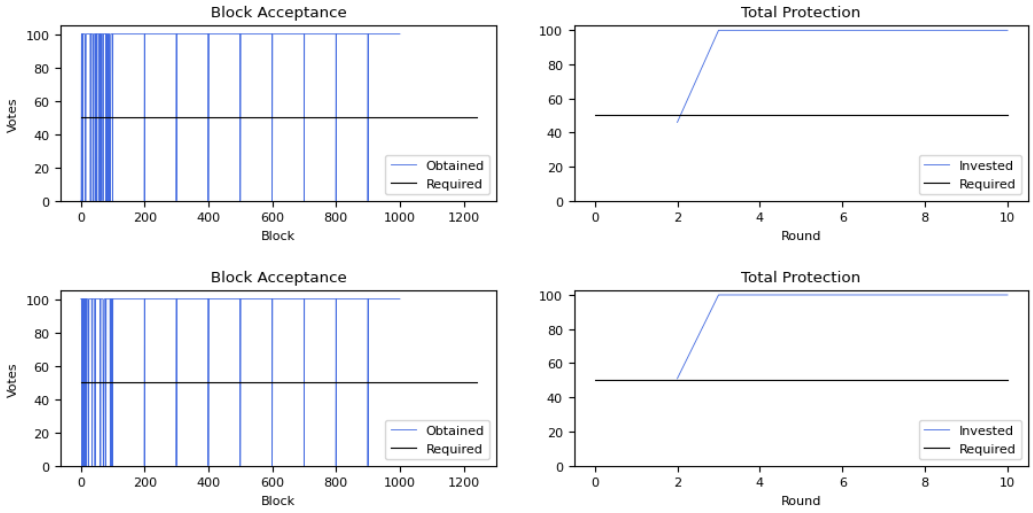


Figure 4.9: Homogenous peer behavior at Benefit per unit of cost 1.5 (top) and 4 (bottom)

Varying Minimum Attack Probabilities As indicated by figure 4.10 differing minimum attack probabilities have no impact on the Reputation Optimization strategy convergence. The time for convergence could differ when concerning higher minimum attack probabilities, but considering that attack probability at each round is chosen randomly, this could simply be an effect of randomization.

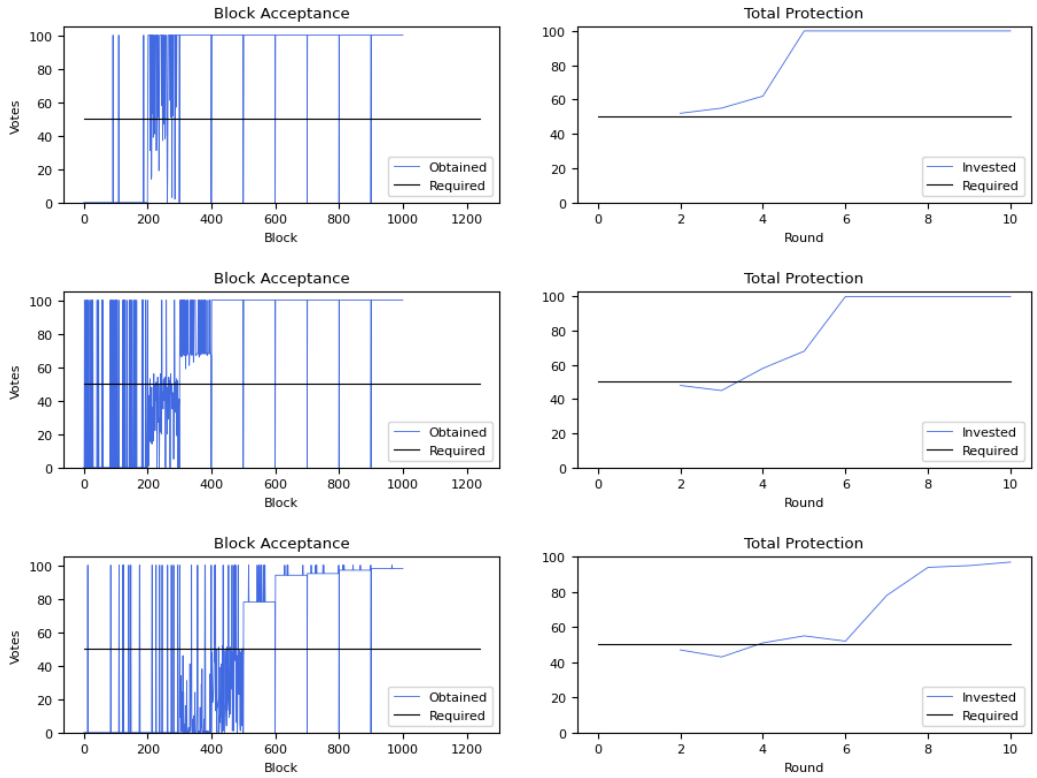


Figure 4.10: Homogenous peer behavior at varying Minimum Attack Probabilities
Results for Minimum Attack Probabilities 0.1, 0.5 and 0.9 are included from top to bottom.
Figure 4.8 provides the reference environment used when Minimum Attack Probabilities is 0

Timeout As indicated by figure 4.11 differing timeouts have no impact on the Reputation Optimization strategy convergence. When the timeout is low, it introduces some noise in player reputations, inducing a variance to security investment decision in initial rounds. However, it defaults to the expected behavior eventually.

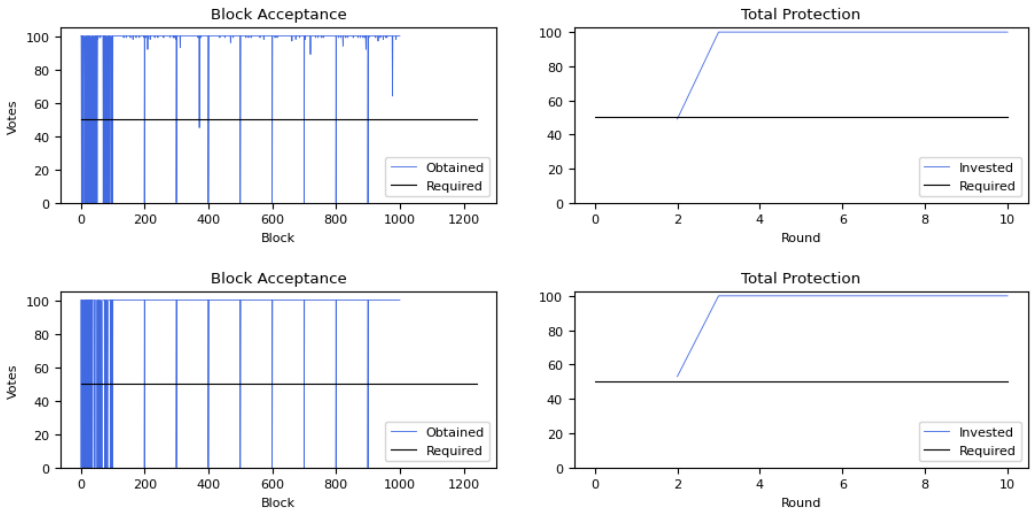


Figure 4.11: Homogenous peer behavior at varying Timeout values
Results for Timeout values 15 ticks (top) and 30 ticks (bottom).
Figure 4.8 provides the reference environment used when Timeout is 25

This simulation also relates to possible observations regarding varying ranges of number of allowed peer connections and connection delay assignments at each peer. Therefore we refrain from producing separate evaluation results for differing ranges for the above parameters.

Varying Number of Peers As indicated by figure 4.12 differing number of peers have no impact on the Reputation Optimization strategy convergence. However, it should be noted that increased number of peers result in additional noise, and therefore time for convergence could be higher compared to lesser number of peers.

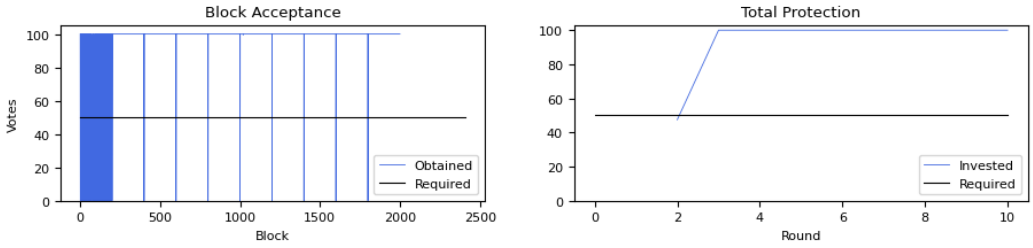


Figure 4.12: Homogenous peer behavior at differing number of peers
Results for 200 peers.
Figure 4.8 provides the reference environment used when no of peers is 100

4.6.1.2 Homogenous peers vs Heterogenous peers

In studying the effects of Heterogenous peer behavior, figure 4.13 indicates that there is no significant difference in converge properties for homogenous and heterogenous peers.

Per figure 4.14, increasing the range of costs allowed also fail to impact the convergence properties in a significant manner.

Note that the Heterogenous players and Homogenous players differ only in their utility derivations. Previously evaluated peers cannot be strictly named as homogenous due to random assignment of number of peer connections and connection delays. Therefore we can conclude that completeness requirement is not compromised even if the costs of participating peers differ. However, this condition might not hold for lower tolerance thresholds which require manipulation of benefit per unit of cost, as discussed in sections 4.6.1.3 and 4.6.1.4.

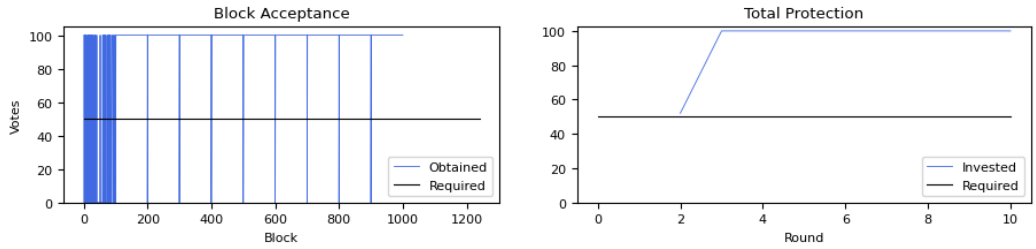


Figure 4.13: Heterogenous peer behavior

Results for Heterogenous cost with Benefit per unit of cost 3. Cost is assigned as a normal distribution with mean $\mu = 1$ and standard deviation $\sigma = 0.025$.
Figure 4.8 provides the reference environment for homogenous peers.

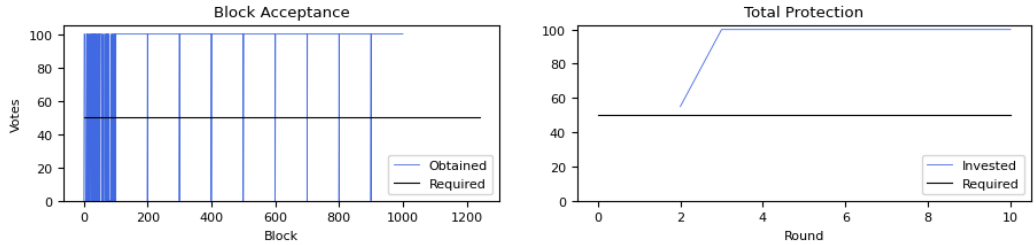


Figure 4.14: Heterogenous peer behavior for larger range of costs

Results for Heterogenous cost with Benefit per unit of cost 3. Cost is assigned as a normal distribution with mean $\mu = 1$ and standard deviation $\sigma = 0.125$.
Figure 4.13 provides the reference environment for when cost is assigned with $\sigma = 0.025$.

4.6.1.3 Homogenous peer behavior at differing tolerance thresholds

Varying Tolerance Thresholds In evaluating how the proposed model fares against differing tolerance thresholds, as depicted in figure 4.15, we find lower tolerance thresholds fail to converge when the benefit per unit of cost is constant. As indicated in figure 3.6, $f^t r_i^t B_i$ value needs to be carefully managed in such cases to incentivize higher aggregate requirement of committing to security investment. In observing 4.16, it can be seen that average player reputations are always higher than the reputation limit (an average value for right hand side of inequality 3.8), thus inspiring $passive$ security investment action. Another noteworthy observation is that the utilities converge

to a constant when the learning strategy achieves convergence, while for the tolerance threshold of 33.4%, utilities keep varying throughout the duration.

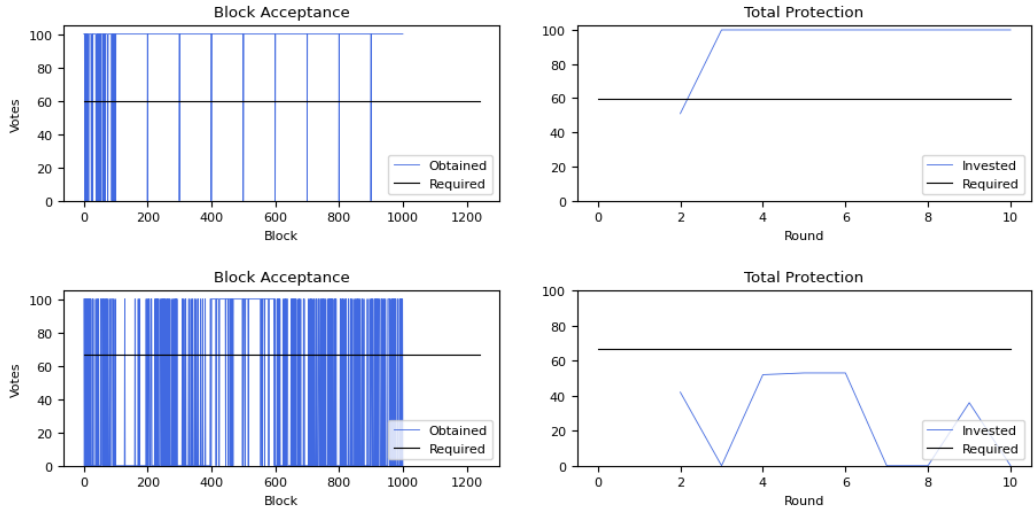


Figure 4.15: Homogenous peer behavior at differing tolerance thresholds

Results for tolerance thresholds 40.1% (top, votes required 59.9%) and 33.4% (bottom, votes required 66.6%) . Figure 4.8 provides the reference environment for when tolerance threshold is 49.9% (votes required $= 50.1\%$)

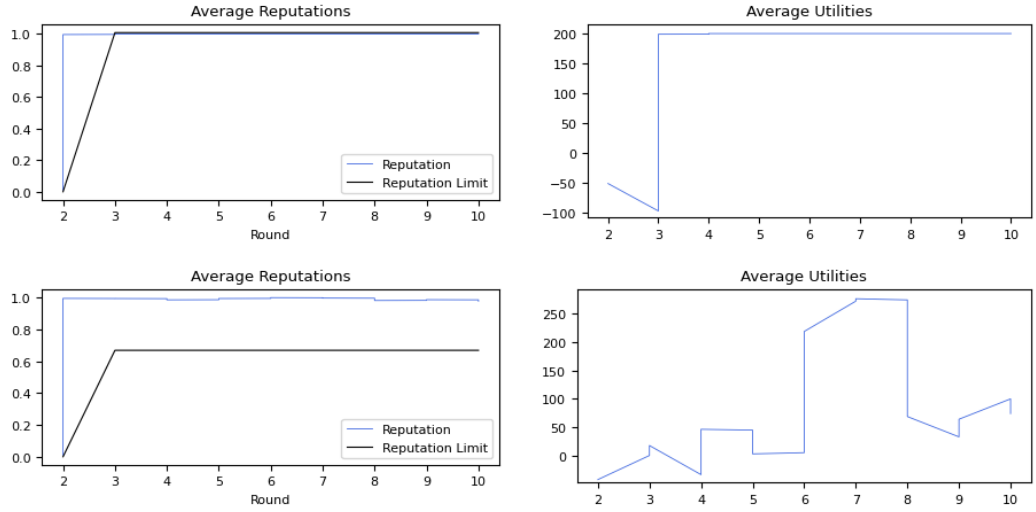


Figure 4.16: Peer reputations at differing tolerance thresholds

Reputations for tolerance thresholds 40.1% (top) and 33.4% (bottom). The reputation limit denotes the average value of RHS of inequality 3.8, providing insight regarding protection decision throughout the rounds.

Varying Tolerance Thresholds and Benefits In managing the $f^t r_i^t B_i$ value for respective tolerance thresholds, figure 4.17 indicate that convergence is feasible for tolerance threshold 33.4% when the benefit per unit of cost is reduced to 1.5. Similarly, figure 4.18. achieves convergence for a tolerance threshold of 20% when the benefit per unit of cost is further reduced to 1.2.

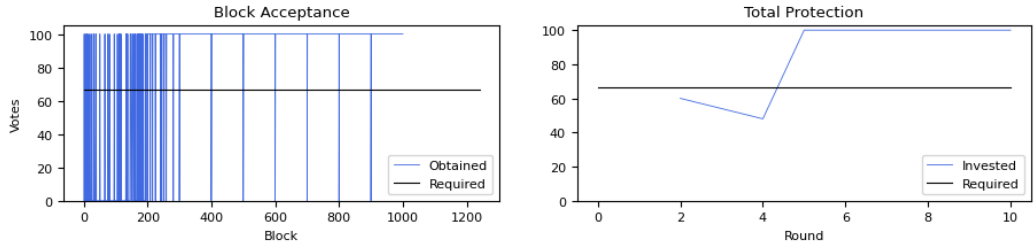


Figure 4.17: Convergence for tolerance thresholds 33.4% when benefit per unit of cost 1.5

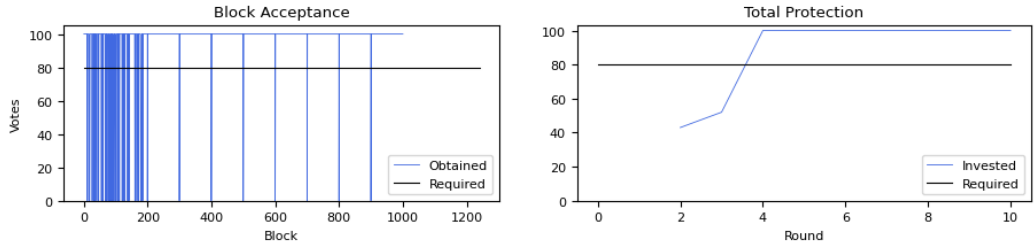


Figure 4.18: Convergence for tolerance thresholds 20% when benefit per unit of cost 1.2

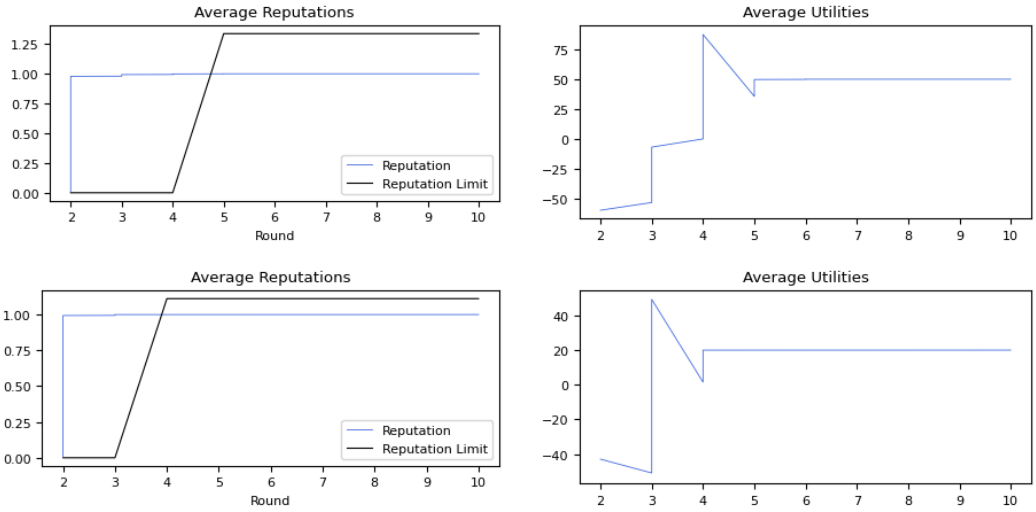


Figure 4.19: Peer reputations at differing tolerance thresholds

Reputations for tolerance thresholds 33.4% (top) and 20% (bottom).

Further to the above observation, 4.19 confirms that the reputation limit must be carefully managed through the Benefit provided by the network designer in order to make higher number of peers partake in $active$ protection. Even though average utility values differ, unlike average utilities for in tolerance threshold of 33.4% in figure 4.16, both tolerance threshold have eventually reached constant average utilities in time, indicating stability of convergence.

Varying Tolerance Thresholds and Attack Probabilities Convergence present in figure 4.18 cannot be observed in figure 4.20 for when the minimum attack probability is increased. Interestingly, this correlates with our predictions in figure 3.7, where higher attack probability values has to be sustained with lesser benefit per cost values.

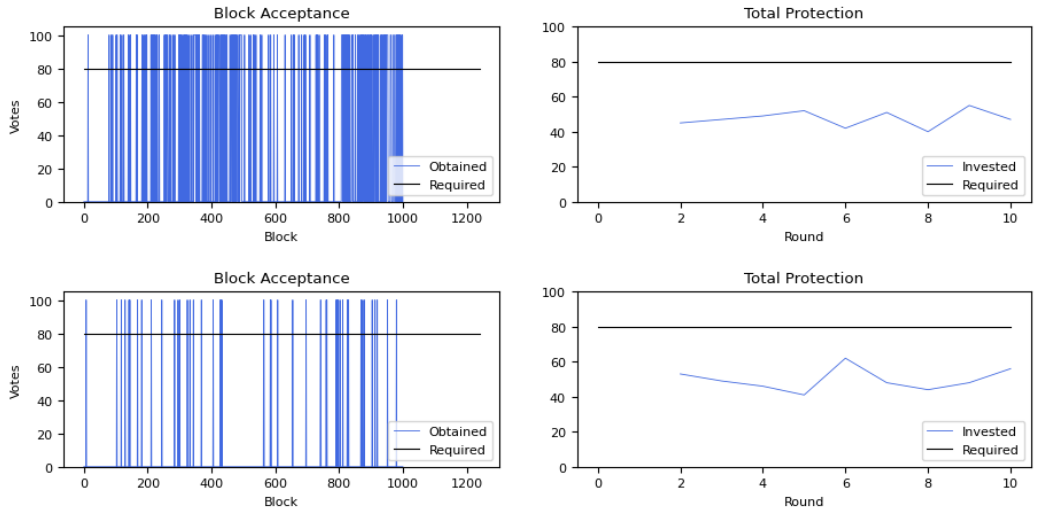


Figure 4.20: Convergence for tolerance thresholds 20% at differing attack probabilities
Results for when the minimum attack probability is 0.5 (top) 0.9 (bottom). Benefit per unit of cost is held constant at 1.2. Figure 4.18 provides the reference environment for when minimum attack probability remains 0.

4.6.1.4 Discussion

The simulations conducted in section 4.6.1.1 follows the predictions made in our theoretical model as indicated in figure 3.6. For most environment variations, the model is capable of obtaining completeness when the requirement is majority consensus. While the time for convergence differ due to various randomization parameters such as attack probabilities randomly chosen at each round, all successful simulations were capable of reaching consensus within 10 rounds of participation.

It is interesting to note that the results converge to all peers committing to $active$ investment. We observe the cause of this to be the reputation limit. As observed in section 4.6.1.3, it is evident that the benefits have to be carefully selected by consulting figures 3.6 and 3.7 in order to balance the environment variation of $f^t r_i^t B_i$ alongside the feasibility of tolerating differing attack probabilities.

A concern of this learning strategy is that the utility optimization does not occur in consulted scenarios, since all peers invest in protection causing an unnecessary additional cost for the network. However, finetuning the benefit per unit of cost value to circumvent the polarizing nature of inequality function (a bulk of peers who invested in current round not investing in the next round) while catering to differing attack probabilities and retaining randomization might be quite difficult. It is further complicated by the fact that in practice, the costs of players will be heterogenous.

Therefore, even though reputation optimization learning strategy is derived from our

Mixed Strategy presentation, quantifying randomness that would result in both $active$ and $passive$ actions with some probability via benefit per unit of cost might not be possible. However, given that players are indeed unaware of the environmental conditions, they will have to commit to a learning strategy that considers both $active$ and $passive$ actions. Considering that average utilities remain consistent throughout multiple rounds once convergence is reached, the stability of the learning strategy could also persuade participants in choosing it over other options. Any concerns regarding differing tolerance thresholds yielding differing utility per block limits where $active$ protection action will obtain higher utilities than $passive$ action (Refer figure 3.8), will also be made redundant.

While simulation results in section 4.6.1.2 indicate that cost of individual peers does not affect majority consensus, it would be interesting to observe the behavior of heterogenous peers in highly specific scenarios such as with higher minimum attack probabilities, contained benefits and lower tolerance thresholds. At present we leave this as an avenue for future research.

4.6.2 Efficiency of Learning Strategies

In evaluating the efficiency of reputation optimization strategy over other well known learning strategies, we increase the number of rounds to 20 in order to provide the simulation with a longer time to converge (if convergence is feasible). While it is possible that convergence is possible in a higher number of rounds, we believe this would at least give sufficient insight regarding possibility of eventual convergence. Figure 4.21 indicates the parameter values used in the baseline experiment.

Reputation Maximization Rounds: **20**

No-of-Peers	Votes-Required	Block-Timeout	Benefit-Per-Unit-Of-Cost	Min-Attack-Probability
100	50.1	25	3	0

Min-Con-Delay	Max-Con-Delay	Min-Peer-Cons	Max-Peer-Cons	Heterogeneous-Cost	Cost-Std-Dev
2	9	2	6	false	-

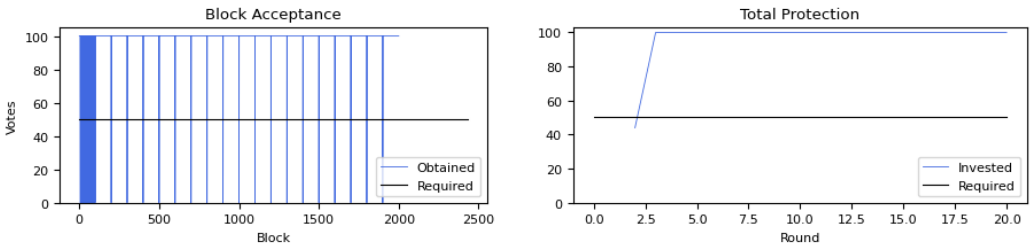


Figure 4.21: Reputation Optimization Learning Strategy execution for 20 rounds

4.6.2.1 Regret Matching

Figure 4.22 presents the block acceptance graph and the total protection graphs for Regret Matching learning strategy. Block Acceptance varies heavily, and the cause for this can be seen in total protection graph, where the total protection always fails to reach the required threshold. Simply augmenting the weight of regrets to include the weight of regrets in history as well, it can be seen that this strategy behaves worse than purely random decisions. Figure 4.23 indicate that weighted probability always favors towards $passive$ protection. However, concerning our utility observation in figure 3.8, in a dynamic environment, it is expected for peers to favor $passive$ action. Additionally, our reputation function was optimized for the mixed strategy and social welfare optimization through benefit redistribution. Regret Matching fails to account for the requirement of social welfare in maintaining the completeness of the ledger, and therefore failure to converge is to be expected. Figure 4.24 also indicates that differing benefits per unit of cost does not have any effect on Regret Matching learning strategy's failure of convergence.

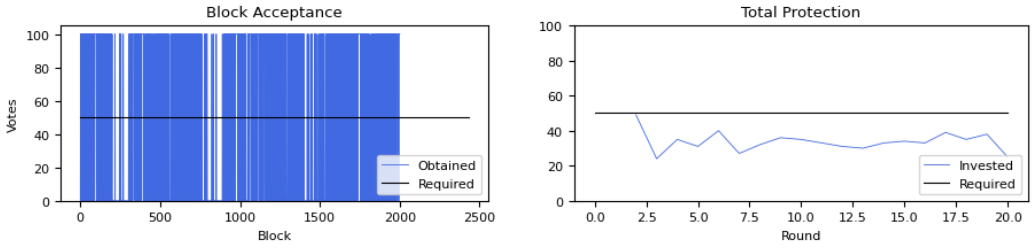


Figure 4.22: Regret Matching Learning Strategy execution for 20 rounds

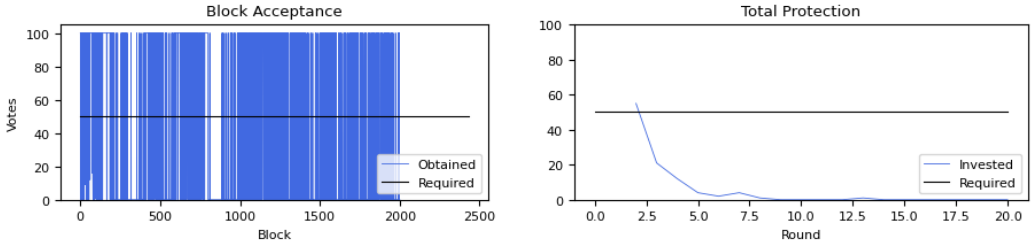


Figure 4.23: Regret Matching Learning Strategy (with History) execution for 20 rounds

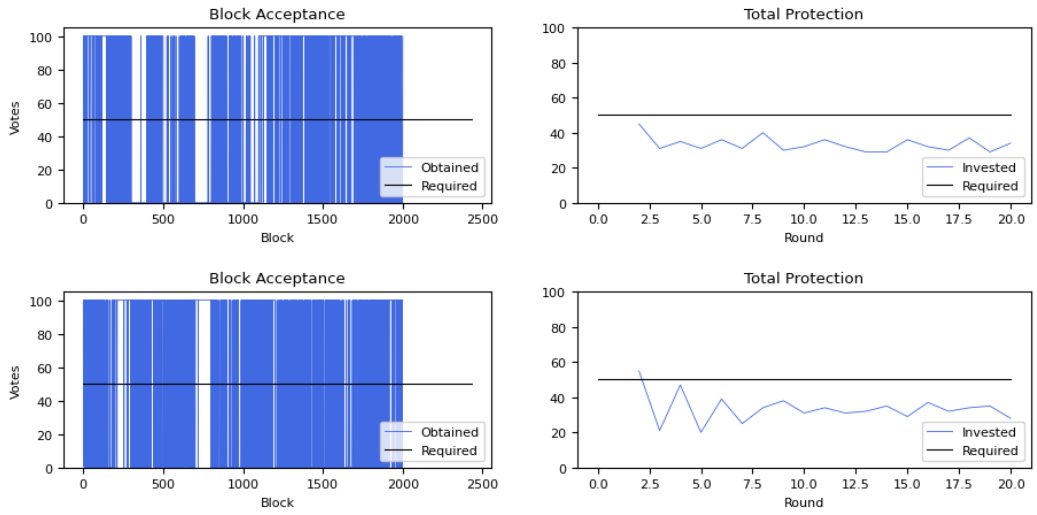


Figure 4.24: Regret Matching Learning Strategy execution for differing benefits
Benefit per unit of cost 1.5 (top) and 4 (bottom).
Figure 4.22 provides a reference for Benefit per unit of cost 3

4.6.2.2 Bounded Rationality

In evaluating bounded rationality, the simulation was designed with a memory of 10 previous actions and a 100 random strategies per peer. Bounded rationality learning strategy can be seen to fare better than Regret Matching learning strategy, wherein it succeeds in obtaining total protection during some of the rounds as observed in figure 4.25. However, as discussed in section 2.9.3, this strategy is only capable of obtaining the required threshold at average, and therefore falls short in providing the completeness assurances required. Figure 4.26 further indicates that differing benefits per unit of cost does not have any perceivable impact on the outcome of Bounded Rationality learning strategy in general.

It should be noted that this strategy does not take account for any of the environmental parameters in decision making but assigns weights to purely random strategies depending on the final output (successful predictions obtained in previous rounds). Even the utilities (and therein the reputation function) are not considered during decision making. Therefore this erratic behavior is to be expected. It would be interesting to observe the convergence properties of this learning strategy under differing conditions such as a larger memory and a larger number of strategies per peer. However, we believe this to be already present in existing literature, even if it may not be in the specific context presented in our research.

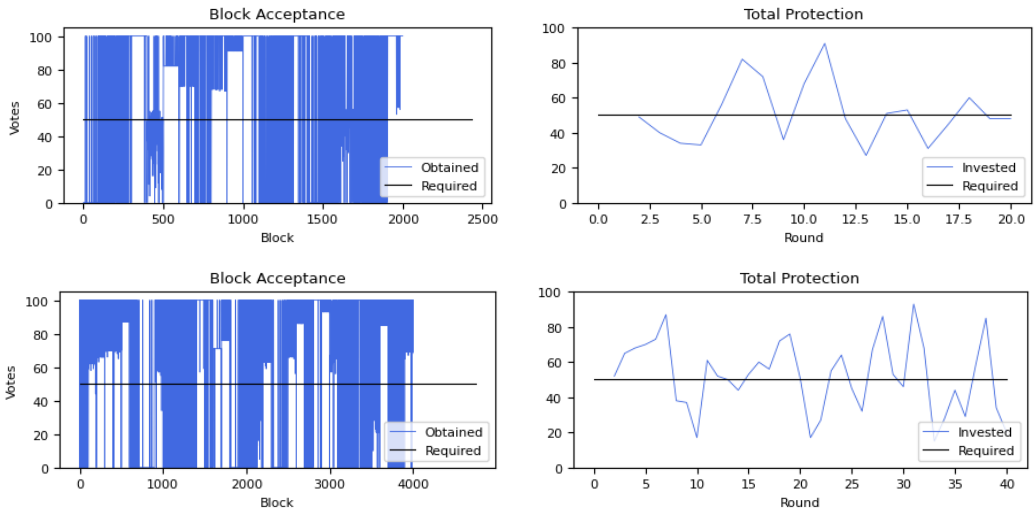


Figure 4.25: Bounded Rationality Learning Strategy execution
Results of simulations executed for 20 (top) and 40 (bottom) rounds.

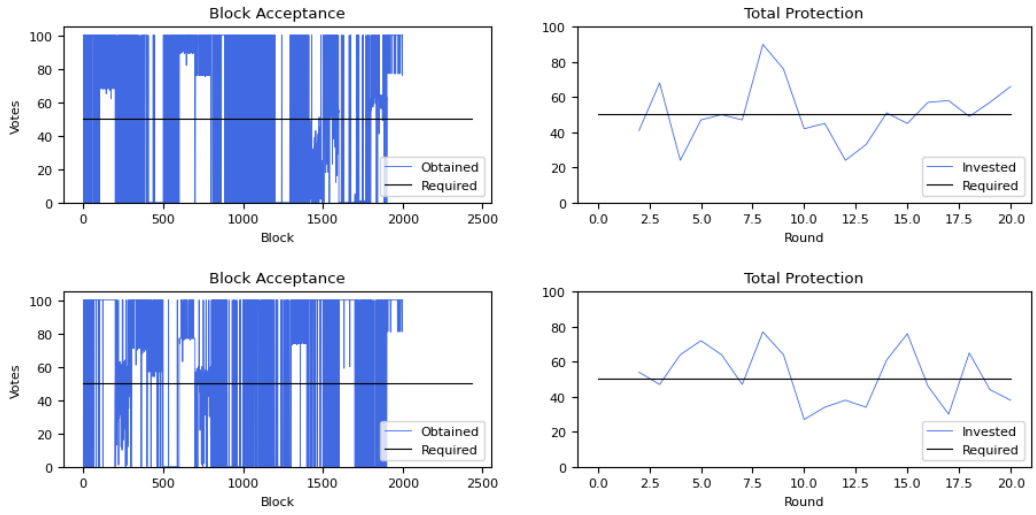


Figure 4.26: Bounded Rationality Learning Strategy execution for differing benefits
Benefit per unit of cost 1.5 (top) and 4 (bottom) for bounded rationality.
Figure 4.25 provides a reference for Benefit per unit of cost 3

4.6.2.3 Payoffs

The stability of the equilibrium depends on the social welfare and the constancy of rewards obtained by peers. Therefore, the payoffs obtained by peers could provide a reasonable insight into which learning strategy will be eventually adopted by peers participating in distributed consensus. This was discussed in section 2.3.2, where alternative mining methodologies were introduced in order to eliminate reward variance induced security concerns.

Figures 4.27, 4.28 and 4.30 indicate that only the Reputation Optimization strategy

provides an stable average utility for the peers throughout multiple rounds of gameplay. Figures 4.29 and 4.31 further attest to this observation, considering that variations of benefit per unit of cost neither affect the convergence nor the desired properties of average utility throughout the duration of learning via Regret Matching and Bounded Rationality methodologies.

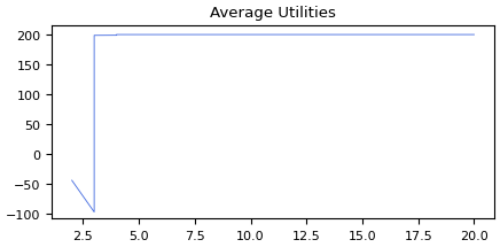


Figure 4.27: Reputation Optimization Learning Strategy Utilities for 20 rounds

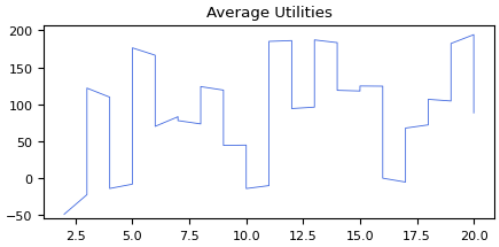


Figure 4.28: Regret Matching Learning Strategy Utilities for 20 rounds

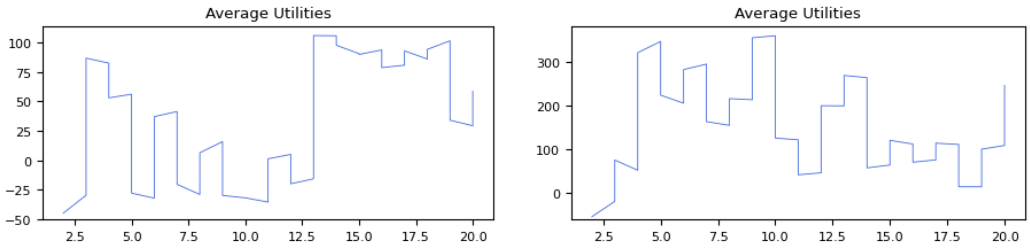


Figure 4.29: Regret Matching Learning Strategy Utilities for differing benefits
Benefit per unit of cost 1.5 (left) and 4 (right) for regret matching without history.
Figure 4.28 provides a reference for Benefit per unit of cost 3

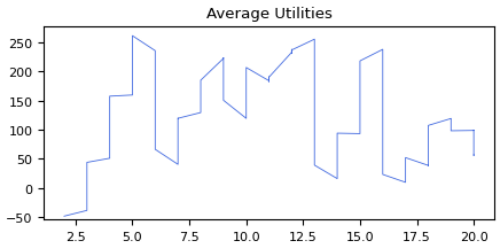


Figure 4.30: Bounded Rationality Learning Strategy Utilities for 20 rounds

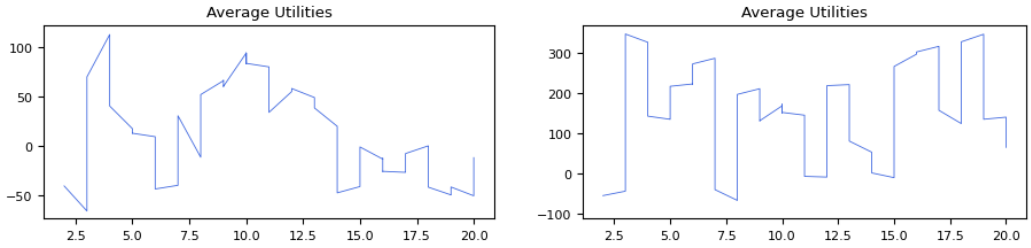


Figure 4.31: Bounded Rationality Learning Strategy Utilities for differing benefits
Benefit per unit of cost 1.5 (top) and 4 (bottom) for bounded rationality.
Figure 4.30 provides a reference for Benefit per unit of cost 3

4.6.2.4 Discussion

Both regret matching and bounded rationality learning strategies fail to converge in a manner that provides completeness assurance for distributed ledgers. However, our reputation optimization strategy is always seen to converge within 10 rounds. It should be noted that the reputation depends strictly on the performance during the immediately prior round, and therefore after a period of adjusting to noise, the learning strategy is quick to converge. We do not proceed to further evaluate differing learning strategies at differing tolerance thresholds, considering that majority consensus could not be assured by the other learning strategy candidates.

A shortcoming of reputation optimization strategy is its tendency to represent the pure strategy equilibria behavior, even though the design was influenced by a mixed strategy game play. However, it should be noted that the randomization is not entirely impossible, given that benefit per unit of cost is adjusted according to the noise present in the environment, even though practical feasibility of assigning such a benefit remains questionable. It would be interesting to observe the effects of benefit per unit of cost when it is assigned as a function of the attack probability of the previous round in future research.

As noted above, the regret matching learning strategy is influenced by our reputation function. Extracting properties of a reputation function, which will cause regret-matching to converge, would be another interesting avenue for future research.

As discussed during the literature review, Bounded rationality falls as close to an existing game theoretical model that addresses our problem of cooperation between peers who are unable to coordinate with each other despite having a common goal. This strategy could be extended to provide a more sophisticated evaluation through facilitating human/belief oriented strategies, as opposed to completely random strategies. Such beliefs could be influenced by the environmental parameters the users are aware of, and therefore could provide a set of successful best-performance beliefs that could inspire convergence. However, considering the range of beliefs possible, such a task would be strenuous in both modelling and evaluation/computation efforts.

4.7 Summary

In this chapter, we present the results for the experiments conducted in order to evaluate our proposed model. We discuss the implementation process of our simulation, and the pros and cons of design choices that were made successively in order to obtain a robust evaluation environment. Our discussion continues on system level design constraints and limitations and the considerations that have to be given in interpreting presented results.

We discuss the results of our simulations per the strategy presented in section 3.5. We observe that the Reputation Optimization learning strategy could withstand noise of environment for majority consensus, and that the benefit per unit of cost has to be carefully adjusted for lower tolerance thresholds. We evaluate the reputation optimization learning strategy alongside regret matching and bounded rationality strategies and conclude that reputation optimization strategy to be the most stable in assuring completeness, given the costs, benefits and tolerance thresholds are carefully considered.

We conclude the chapter with the observation that even though achieving our social welfare optimization goals require extreme finetuning of environment parameters, our proposed model is highly capable of assuring completeness in a distributed consensus maintained by rational participants.

CHAPTER 5
CONCLUSION AND FUTURE WORKS

5.1 Summary

Our research focused on quantifying the effects of relative relationships between incentives, costs and peer security investment requirements a consensus protocol must implement in order to ensure lasting stability and completeness. We presented the current landscape of Blockchain and its alternative use cases, and how non-financial use cases that require higher throughputs would not be sufficiently protected in case the participating peers are selectively interfered by a resourceful and malicious attacker. We highlighted how such systems would be vulnerable to completeness compromise even though agreement would be reached in all consensus rounds due to the vulnerability of committing to inaction.

Through our literature review, we evaluated the consensus algorithms in current Blockchain implementations, their security implementations and subsequent limitations. We extracted a set of requirements that a consensus protocol must support, and the social obligations the reward scheme must facilitate in order to sustain the security of the network. We discussed game theory and its applicability in solving our research problem.

We introduced a game theoretical model which would incorporate the social requirements of network sustainability, in form of an infinitely repeated game. In observing the convergence properties of pure and mixed strategy equilibria, we established mixed strategy equilibria to better serve our security requirement of influencing the aggregate security requirement while catering to the utility maximization / cost minimization nature of rational peers. Subsequently, we introduced a learning methodology, Reputation Optimization, which could be used by distributed peers without any prior coordination in decision making to achieve the required aggregated security investment that ensures completeness in a repeated setting.

We evaluated the resilience of Reputation Optimization learning methodology under differing noise inducing environment variables. We considered how differing peer incentives, costs, tolerance thresholds and variance of attack probabilities present in the network will impact the stabilization of security investment and therefore completeness assurance. We further evaluated how our learning strategy fares against two popular learning methodologies present in game theoretical literature, Regret Matching and Bounded Rationality.

We simulated several experiments using NetLogo to evaluate aforementioned conditions.

Our findings indicated our reputation maximization learning strategy to follow the predictions of our model, and to be resilient to a range of differing environment conditions in obtaining majority consensus. Benefits provided by the network must be finetuned to support lower tolerance thresholds for consensus. We observed that while our learning methodology was derived off of mixed strategy equilibria with the interest of optimizing social welfare of the network, convergence property of Reputation Optimization may not provide the desired amount of decision randomness capable of maximizing social welfare.

In evaluating our learning methodology against Regret Matching and Bounded Rationality, we observed that our methodology provides certain attractive attributes over other learning methodologies. It is capable of fast convergence, providing stable utilities and preserving completeness. Regret matching fails in both convergence to an equilibrium and providing stable utilities, but we observe that this might be a result of our custom reputation function being designed specifically for the proposed mixed strategy model. While Bounded rationality provides similar results as indicated in literature (successful of average), it also fails in providing stable utilities and preserving completeness.

In conclusion, our research provides sufficient insight on the features that must be implemented by a security sensitive distributed ledger, the security of which is solely dependent on participant peers. It establishes the parameters and environmental conditions that must be considered when designing a reward scheme that could incentivize rational peers in collectively committing to the required information security investment. We hope our research will be helpful in designing future consensus protocols that could withstand selective interference of powerful adversaries who target ledger history completeness.

5.2 Future Work

Our solution design focused on self-sustainability of the system considering the nature of non-financial distributed ledgers and therefore did not consider the additional complications in stake-based systems where peers could attack each other. However, we find it interesting to evaluate the influence of direct stake as opposed to the self-governance presented in our system. We hypothesize that such peer-on-peer attacks would eventually lead to global security investment, and therefore would converge to pure-strategy equilibria (Refer section 3.3.4) discussed in our solution model. Interestingly, peer-on-peer attacks would be consistently available in the system, thus the collapse of equilibria due to tragedy of commons might not occur in such a situation.

Our presentation only models the peers that participate in the consensus, while the attacker remains one-dimensional. An interesting future research avenue would be distributed attacker modelling (as in the case of peer-on-peer attacks) and modelling the learning that would be undertaken by an attacker. However, we do not consider this strictly possible through game theoretical models, considering the complexity of timing in decision making by both peers and attackers, as observed via the many equilibrium models we have discussed in our solution. For instance, the tragedy of commons would

occur when the attacker retreats, and then the system would return to being vulnerable, as per our observation in pure-strategy equilibria. We still leave it as an afterthought for an interested reader.

Considering the limitations of our phase based simulation design, and the limitation of reduced noise in abstracting the transaction/block acceptance process, we observe that a singular language based simulation would be more suitable for a more comprehensive evaluation. Representing the network overlay through the python application itself would support the scalability required while not compromising on the blockchain functionality. Similarly, implementing the network functionality in a compiled language as a NetLogo extension would emulate the network volatility and present a more accurate distributed ledger while retaining the scalability. We leave an evaluation of our model at this level to be considered in the future.

It would be interesting to extend our evaluation of reputation optimization learning strategy in the presence of noise considering the benefit per unit of cost optimization observed for retaining completeness at lower tolerance thresholds. Behavior of heterogenous peers in highly specific scenarios such as higher minimum attack probabilities, contained benefits and lower tolerance thresholds could provide insight regarding inducing further noise to the learning strategy output, resulting in better social welfare for the network. Further to this, as noted in section 4.5, the variance of network scale and how it influences our results were also excluded from the current evaluation scope due to the nature of our game theoretical model. We intend to extend our NetLogo simulation to support joining and dropping of participant peers during the execution and observe the effects of this through future research.

We also consider how the output of Regret Matching and Bounded Rationality learning methodologies could be further optimized to suit our security goals. In bounded rationality, an evaluation using human/belief oriented strategies, as opposed to completely random strategies might lead to better converge results. However, the range of beliefs to be considered in such a case would be extremely high, and therefore such an investigation could prove to be strenuous. Investigating properties of a reputation/utility function that will cause regret matching to converge might also prove to be insightful avenue of research.

We conclude with the hope that our NetLogo model of a distributed consensus network would aid future research in evaluating other learning methodologies or game theoretical models which were designed with differing security goals in mind.

REFERENCE LIST

[1] I.-C. Lin and T.-C. Liao, "A survey of blockchain security issues and challenges.," *IJ Network Security*, vol. 19, no. 5, pp. 653–659, 2017.

[2] S. Nakamoto, "A peer-to-peer electronic cash system," *Bitcoin.–URL: https://bitcoin. org/bitcoin. pdf*, vol. 4, 2008.

[3] M. Vukolić, "The quest for scalable blockchain fabric: Proof-of-work vs. bft replication," in *International workshop on open problems in network security*, Springer, 2015, pp. 112–125.

[4] C. Cachin, "Architecture of the hyperledger blockchain fabric," in *Workshop on distributed cryptocurrencies and consensus ledgers*, Chicago, IL, vol. 310, 2016.

[5] I. Bentov, C. Lee, A. Mizrahi, and M. Rosenfeld, "Proof of Activity : Extending Bitcoin's Proof of Work via Proof of Stake [Extended Abstract]," *ACM SIGMETRICS Performance Evaluation Review*, vol. 42, no. 3, pp. 34–37, 2014.

[6] B. Clark, *Blockchain and ip law: A match made in crypto heaven?* 2018. [Online]. Available: `https://www.wipo.int/wipo_magazine/en/2018/01/article_0005.html`.

[7] M. Sharples and J. Domingue, "The blockchain and kudos: A distributed system for educational record, reputation and reward," in *European conference on technology enhanced learning*, Springer, 2016, pp. 490–496.

[8] S. Dean, *$69 million for digital art? The NFT craze explained*, 2021. [Online]. Available: `https://www.latimes.com/business/technology/story/2021-03-11/nft-explainer-crypto-trading-collectible` (visited on 05/24/2021).

[9] M. O'Dair, Z. Beaven, D. Neilson, R. Osborne, and P. Pacifico, "Music on the blockchain," 2016.

[10] N. Kshetri, "Can blockchain strengthen the internet of things?" *IT professional*, vol. 19, no. 4, pp. 68–72, 2017.

[11] A. Dorri, S. S. Kanhere, and R. Jurdak, "Towards an optimized blockchain for iot," in *2017 IEEE/ACM Second International Conference on Internet-of-Things Design and Implementation (IoTDI)*, IEEE, 2017, pp. 173–178.

[12] M. Mettler, "Blockchain technology in healthcare: The revolution starts here," in *2016 IEEE 18th international conference on e-health networking, applications and services (Healthcom)*, IEEE, 2016, pp. 1–3.

[13] C. Papadimitriou, "Algorithms, games, and the internet," in *Proceedings of the thirty-third annual ACM symposium on Theory of computing*, 2001, pp. 749–753.

[14] K. Biswas and V. Muthukkumarasamy, "Securing smart cities using blockchain technology," in *IEEE International Conference on High Performance Computing and Communications*, IEEE Xplore, 2016, pp. 1392–1393.

[15] J. Bonneau, A. Miller, J. Clark, A. Narayanan, J. A. Kroll, and E. W. Felten, "Sok: Research perspectives and challenges for bitcoin and cryptocurrencies," in *2015 IEEE symposium on security and privacy*, IEEE, 2015, pp. 104–121.

[16] R. Pass and E. Shi, "Fruitchains: A fair blockchain," in *Proceedings of the ACM Symposium on Principles of Distributed Computing*, 2017, pp. 315–324.

[17] I. Eyal and E. G. Sirer, "Majority is not enough: Bitcoin mining is vulnerable," in *International conference on financial cryptography and data security*, Springer, 2014, pp. 436–454.

[18] J. R. Douceur, "The sybil attack," in *International workshop on peer-to-peer systems*, Springer, 2002, pp. 251–260.

[19] M. Babaioff, S. Dobzinski, S. Oren, and A. Zohar, "On bitcoin and red balloons," in *Proceedings of the 13th ACM conference on electronic commerce*, 2012, pp. 56–73.

[20] L. Lamport, R. Shostak, and M. Pease, "The byzantine generals problem," *ACM Transactions on Programming Languages and Systems*, vol. 4, no. 3, pp. 382–401, 1982.

[21] M. Castro, B. Liskov, *et al.*, "Practical byzantine fault tolerance," in *OSDI*, vol. 99, 1999, pp. 173–186.

[22] M. J. Fischer, N. A. Lynch, and M. S. Paterson, "Impossibility of distributed consensus with one faulty process," *Journal of the ACM (JACM)*, vol. 32, no. 2, pp. 374–382, 1985.

[23] D. Schwartz, N. Youngs, A. Britto, *et al.*, "The ripple protocol consensus algorithm," *Ripple Labs Inc White Paper*, vol. 5, no. 8, p. 151, 2014.

[24] S. King and S. Nadal, "Ppcoin: Peer-to-peer crypto-currency with proof-of-stake," *self-published paper, August*, vol. 19, p. 1, 2012.

[25] J. Kwon, "Tendermint: Consensus without mining," *Draft v. 0.6, fall*, vol. 1, no. 11, 2014.

[26] J. K. Goeree and C. A. Holt, "Stochastic game theory: For playing games, not just for doing theory," *Proceedings of the National Academy of sciences*, vol. 96, no. 19, pp. 10 564–10 567, 1999.

[27] M. O. Jackson, "A brief introduction to the basics of game theory," *Available at SSRN 1968579*, 2011.

[28] M. Hossein Manshaei, Q. Zhu, T. Alpcan, T. Basar, and J.-P. Hubaux, "Game theory meets network security and privacy," *ACM Computing Surveys*, 2011.

[29] S. Roy, C. Ellis, S. Shiva, D. Dasgupta, V. Shandilya, and Q. Wu, "A survey of game theory as applied to network security," in *2010 43rd Hawaii International Conference on System Sciences*, IEEE, 2010, pp. 1–10.

[30] M. O. Jackson, K. Leyton-Brown, and Y. Shoham, *Game Theory Course: Strategic Reasoning*, 2013. (visited on 12/20/2018).

[31] T. Moscibroda, S. Schmid, and R. Wattenhofer, "When selfish meets evil: Byzantine players in a virus inoculation game," in *Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*, 2006, pp. 35–44.

[32] C. Buragohain, D. Agrawal, and S. Suri, "A game theoretic framework for incentives in p2p systems," in *Proceedings Third International Conference on Peer-to-Peer Computing (P2P2003)*, IEEE, 2003, pp. 48–56.

[33] H. Varian, "System reliability and free riding," in *Economics of information security*, Springer, 2004, pp. 1–15.

[34] R. Gupta and A. K. Somani, "Game theory as a tool to strategize as well as predict nodes' behavior in peer-to-peer networks," in *11th International Conference on Parallel and Distributed Systems (ICPADS'05)*, IEEE, vol. 1, 2005, pp. 244–249.

[35] J. Grossklags, N. Christin, and J. Chuang, "Secure or insure? a game-theoretic analysis of information security games," in *Proceedings of the 17th international conference on World Wide Web*, 2008, pp. 209–218.

[36] E. Koutsoupias and C. Papadimitriou, "Worst-case equilibria," in *Annual Symposium on Theoretical Aspects of Computer Science*, Springer, 1999, pp. 404–413.

[37] M. Nojoumian and D. R. Stinson, "Socio-rational secret sharing as a new direction in rational cryptography," in *International Conference on Decision and Game Theory for Security*, Springer, 2012, pp. 18–37.

[38] M. J. Osborne and A. Rubinstein, *A course in game theory*. MIT press, 1994.

[39] S. Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, vol. 68, no. 5, pp. 1127–1150, 2000.

[40] W. B. Arthur, "Inductive reasoning and bounded rationality," *The American economic review*, vol. 84, no. 2, pp. 406–411, 1994.

[41] M. Garofalo, "Modeling the 'el farol bar problem'in netlogo," *PRELIMINARY DRAFT, Dexia Bank Belgium*, 2006.

[42] U. Wilensky, "NetLogo," *Center for Connected Learning and Computer Based Modeling, Northwestern University, Evanston, IL*, 1999. [Online]. Available: `http://ccl.northwestern.edu/netlogo/`.

[43] P. Davidsson, "Multi agent based simulation: Beyond social simulation," in *International workshop on multi-agent systems and agent-based simulation*, Springer, 2000, pp. 97–107.

[44] M. Niazi and A. Hussain, "Agent-based tools for modeling and simulation of self-organization in peer-to-peer, ad hoc, and other complex networks," *IEEE Communications Magazine*, vol. 47, no. 3, pp. 166–173, 2009.

[45] U. Wilensky and W. Rand, "NetLogo El Farol model," *Center for Connected Learning and Computer-Based Modeling, Northwestern Institute on Complex Systems, Northwestern University, Evanston, IL.*, 2007. [Online]. Available: `http://ccl.northwestern.edu/netlogo/models/ElFarol`.

[46] ——, *Introduction to Agent-Based Modeling: Modeling Natural, Social and Engineered Complex Systems with NetLogo*. Cambridge, MA. MIT Press, 2015.

APPENDIX A GAME THEORETIC DEFINITIONS

In this section we define some game theoretic terminology we have used within this report. Note that this in no means is an exhaustive list, nor a highly descriptive one, but rather a concatenation of definitions spread throughout the literature review, alongside few other concepts we assumed to be prior knowledge.

Game

A game is the collection of strategic interactions between players, whose actions are motivated by maximizing their respective payoffs.

Solution

A game play analysis where players play their best strategies such that ultimate payoffs become predictable.

Utility function

A utility function is a function which combines the costs and benefits of a player, which are usually dependent on a chosen strategy. The function represents the return of investment to a participant of a game.

Action

An exact decision a player takes during an iteration of a game.

Strategy

All actions a player could take in each iteration of a game given the possible actions of other players at respective iterations. A strategy is therefore a complete plan of action

Pure Strategy

A pure strategy is a set of choices a player can choose from. (E.g. I will play action A)

Mixed Strategy

A mixed strategy is the set of choices a player can take which is dependent on some other random event. (E.g. I would play action A, if the coin flip resulted in "Tails")

Nash Equilibrium

Nash Equilibrium is a state of convergence where all players continue to play the action that provides them most utility (best response), given that other players are also playing their best responses. All players choose to not defect from their strategy because they cannot benefit from doing so. A game can have more than one Nash equilibria, or none. Mixed strategy games have been proved to have at least one Nash Equilibrium.

Symmetric Nash Equilibrium

In a Symmetric Nash Equilibrium, the state of convergence occur when all players choose the same action. This often occurs in mixed strategy equilibria. It's an attractive solution for when the players are of a homogenous population [34].

Social Welfare

The total utility obtained by all participants of a game, usually maximized through external coordination by a benevolent agent. This is the summation of the utility function of all agents.

Dynamic Games

Games that converge over multiple iterations, where players can take different actions in each iteration (Provided that it is not a repeated game).

Static games

A game where players act simultaneously and for just one time.

Repeated Game

A subset of dynamic games where a static game is repeated, and therefore the player has to commit to a fixed action during the first iteration itself, which will affect their and other player payoffs throughout the repetition.

Perfect Information Game

A game where all players are aware of all past actions of all other players.

Imperfect Information Game

A game where at least one player is unaware of any past actions of any of the other players. All static games are imperfect games.

Complete Information Game

All players have knowledge of all other players' strategies and payoffs, but the actions already taken place are not known.

Incomplete Information Game

At least one of the players is unaware of any one of the other player's strategies and/or payoffs.

Zero Sum Game

A game with a utility function where one player's payoff (for both costs and gains each) is the negative of the other's, causing the social optimum of the game to be 0.

Saddle-point equilibrium

A Nash equilibrium for two player zero-sum games, where a single objective function is present, which is minimized by one player and maximized by the other.

Coalitional games
: Games where players can form groups and work to maximize the collective utility of the group against all other groups and their best responses.

Non-cooperative game
: No coalitions are formed due to external coordination factors in these games. If coalitions exist they have been formed by players themselves.

Bayesian Game
: An incomplete information game, where players have beliefs with known probability distribution. This allows players to predict what players with different beliefs would do, and therein base their decision on that introspection.

Dominant Strategy
: A strategy that provides a better payoff for the player no matter what action the other player might choose to take.

Weakly dominant Strategy
: A strategy that provides a better or equal payoff for the player no matter what action the other player might choose to take.

Iterated removal of dominated strategies
: If a player has strategies which yield higher utility irrespective of other player's strategy, those strategies are kept while others are removed from the first player's strategy profile. The remaining strategies are considered for other players and this repeated removal leaves with strategies which dominate all other possible (Now removed) strategies. The resulting strategy profile contains payoffs which cannot be further maximized.

Price of Anarchy (PoA)
: The degradation of social welfare due to rational agent behavior of individual utility maximization.

$$\text{Price of Anarchy}_{min(cost)} = \frac{\text{Social Welfare at Worst Nash Equilibrium}}{\text{Maximum Social Welfare}}$$

The "worst case Nash equilibrium" is when the players show the least coordination among themselves, either minimizing the utility or maximizing the cost. Depending on the game design, the dividend and divisor changes places. For example, cost minimization games such as prisoner's dilemma has maximum social welfare as a divisor, while utility maximization games have maximum social welfare as the dividend.

Tragedy of the Commons
: A phenomenon from game theory and economics, when each participant of a system assumes that others would follow the protocol and therefore he or she can maximize their utility without any repercussions to the social welfare, ultimately destroying the equilibrium and the environment.

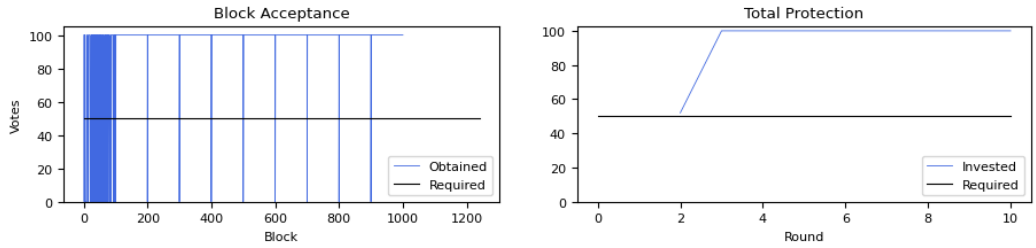
APPENDIX B ADDITIONAL SIMULATION RESULTS

This section contains a second run of the experiments conducted in section 4.6 to confirm the repeatability of the simulation results despite the differing random parameters in each simulation run such as the attack-probabilities in each round, connection delays of peers, number of connections allowed per each peer, and the heterogenous cost assignments in applicable simulations. Note that the ranges of the above parameters are kept the same during the second run.

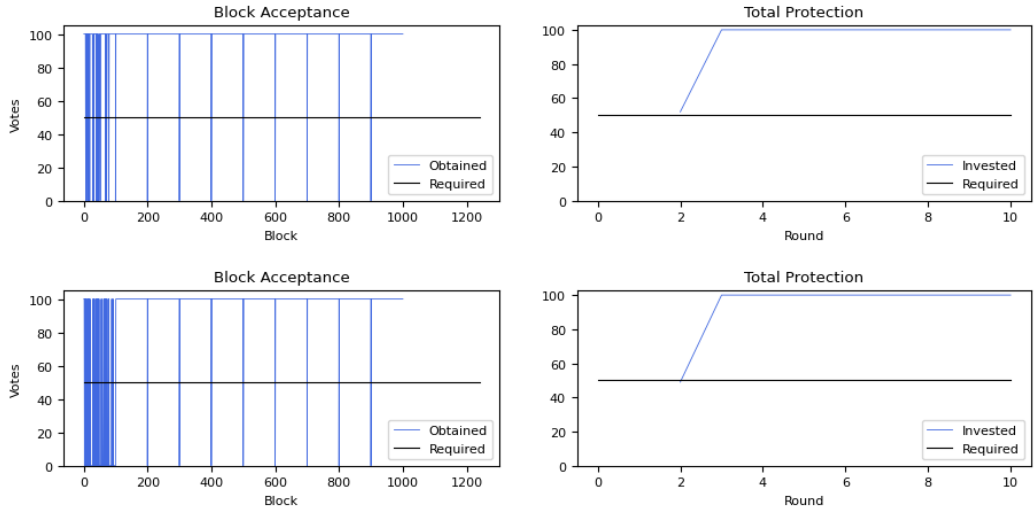
Reputation Maximization Rounds: **10**

No-of-Peers	Votes-Required	Block-Timeout	Benefit-Per-Unit-Of-Cost	Min-Attack-Probability
100	50.1	25	3	0

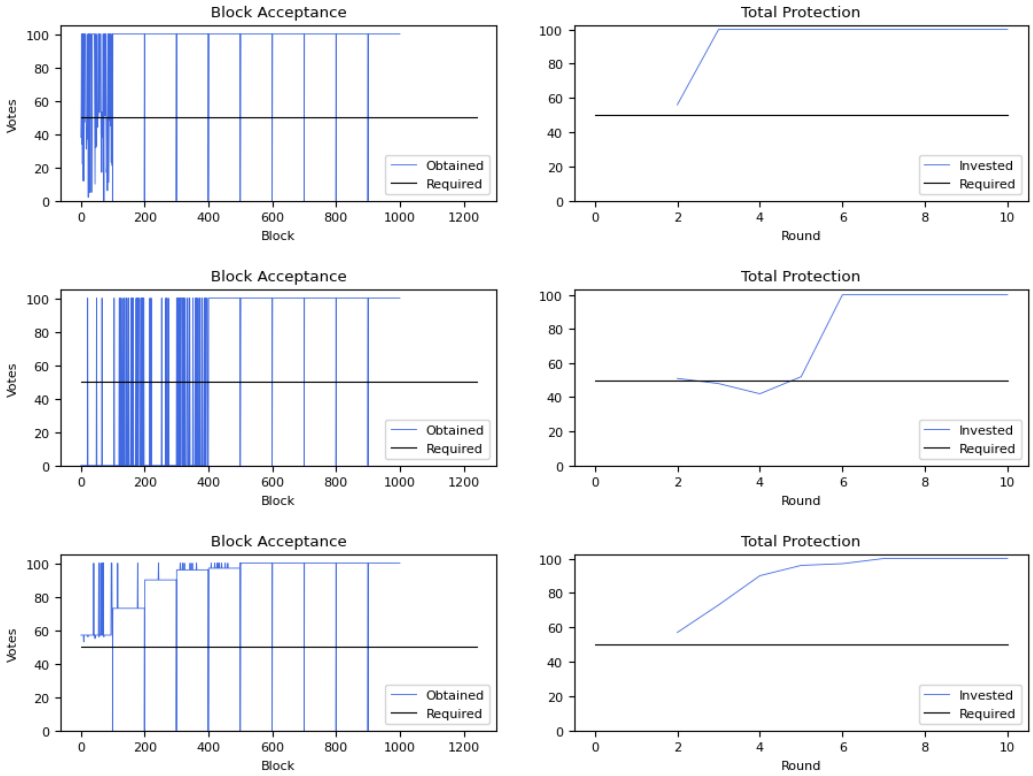
Min-Con-Delay	Max-Con-Delay	Min-Peer-Cons	Max-Peer-Cons	Heterogeneous-Cost	Cost-Std-Dev
2	9	2	6	false	-



Homogenous peer behavior at Benefit per unit of cost 3.
Repeated results for figure 4.8

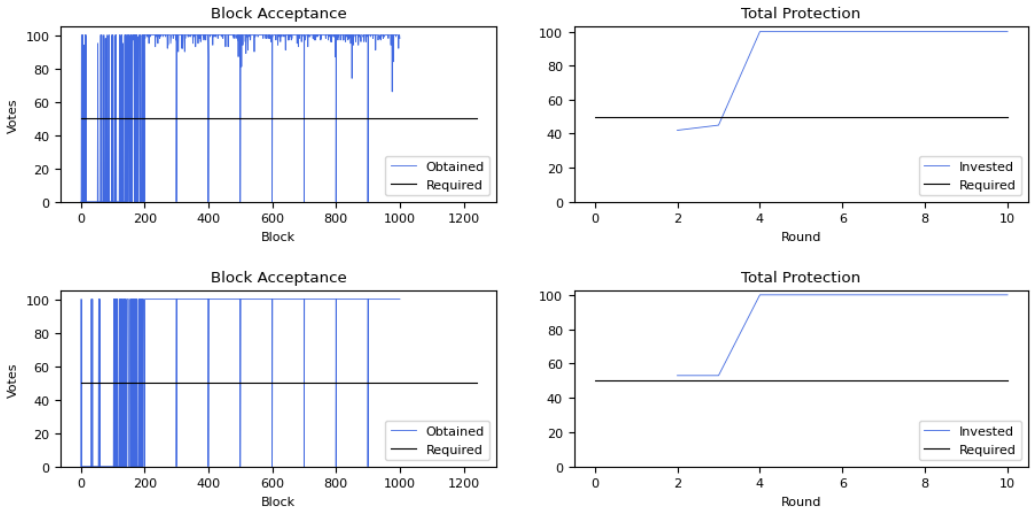


Homogenous peer behavior at Benefit per unit of cost 1.5 (top) and 4 (bottom)
Repeated results for figure 4.9



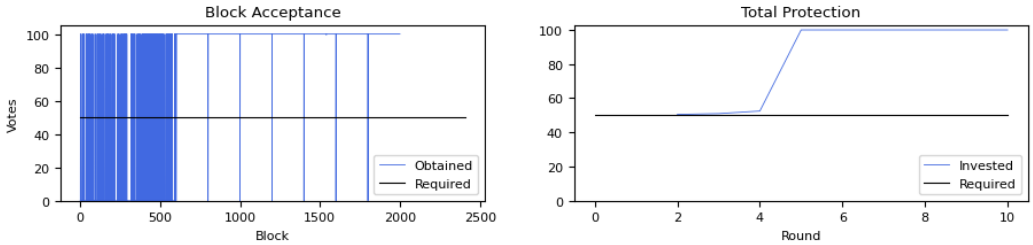
Homogenous peer behavior at varying Minimum Attack Probabilities

Repeated results for figure 4.10



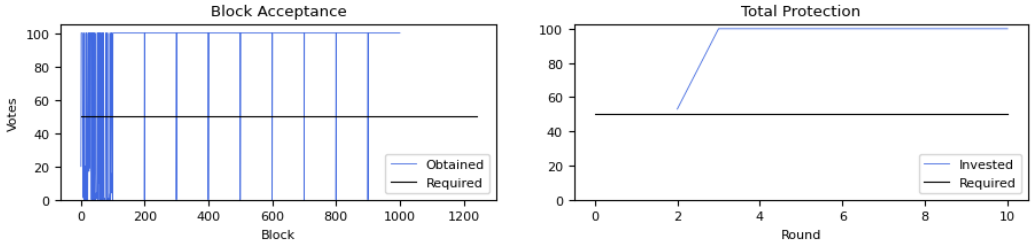
Homogenous peer behavior at varying Timeout values

Repeated results for figure 4.11



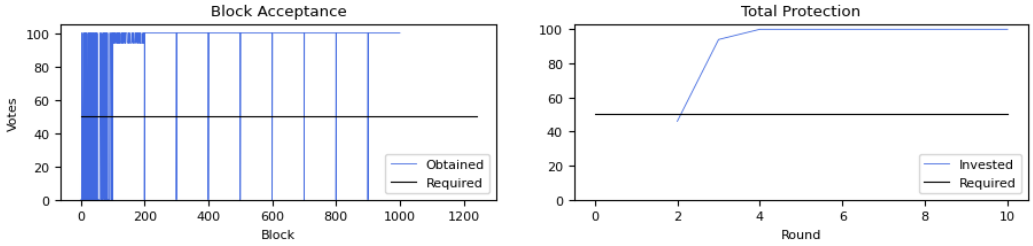
Homogenous peer behavior at differing number of peers

Repeated results for figure 4.12



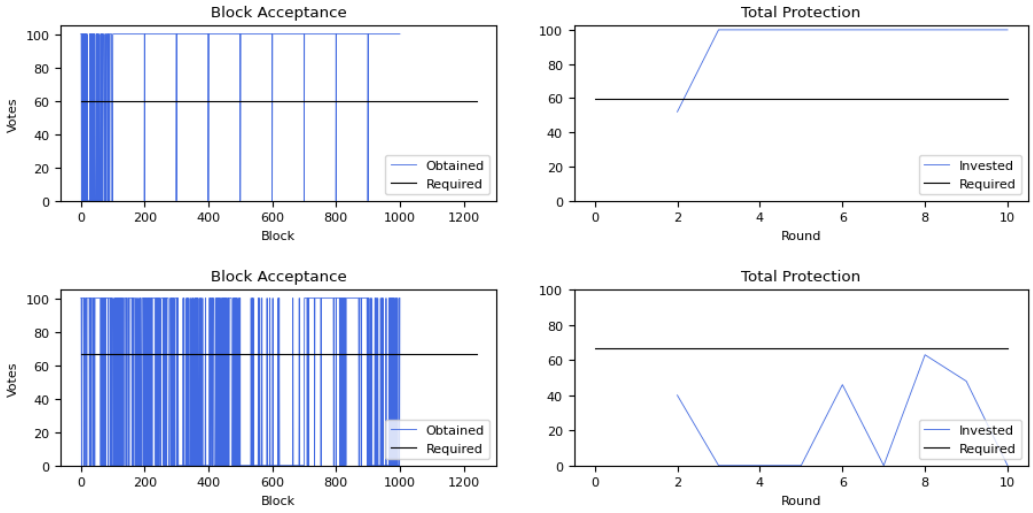
Heterogenous peer behavior

Repeated results for figure 4.13



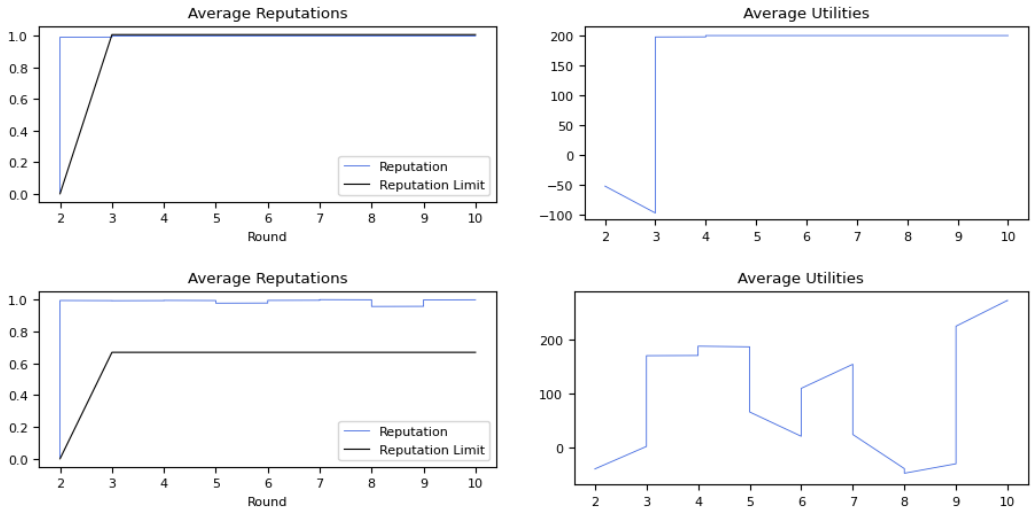
Heterogenous peer behavior for larger range of costs

Repeated results for figure 4.14



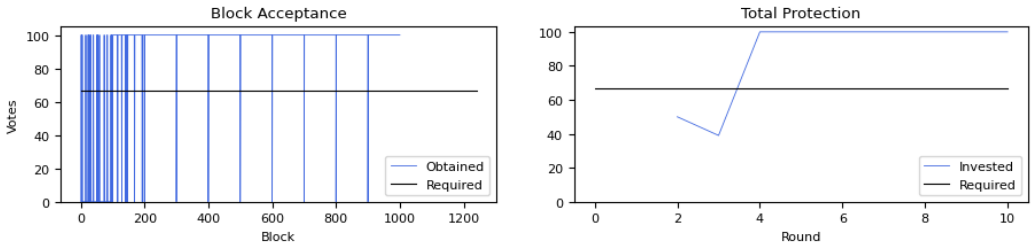
Homogenous peer behavior at differing tolerance thresholds

Repeated results for figure 4.15



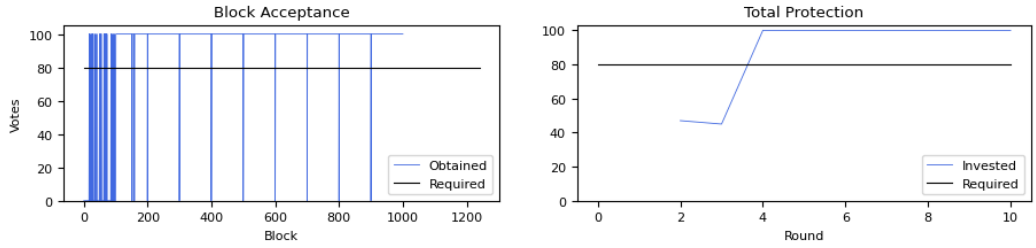
Peer reputations at differing tolerance thresholds

Repeated results for figure 4.16



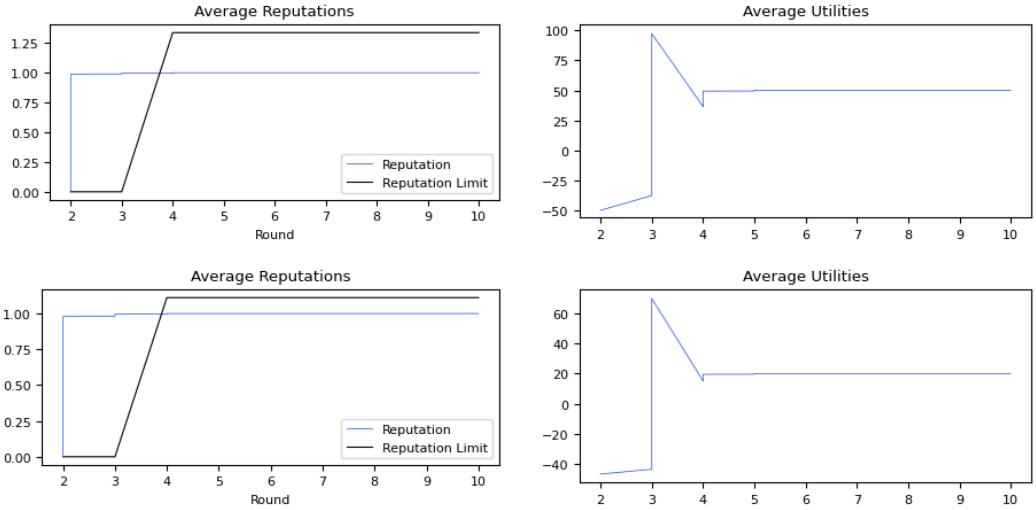
Convergence for tolerance thresholds 33.4% when benefit per unit of cost 1.5

Repeated results for figure 4.17



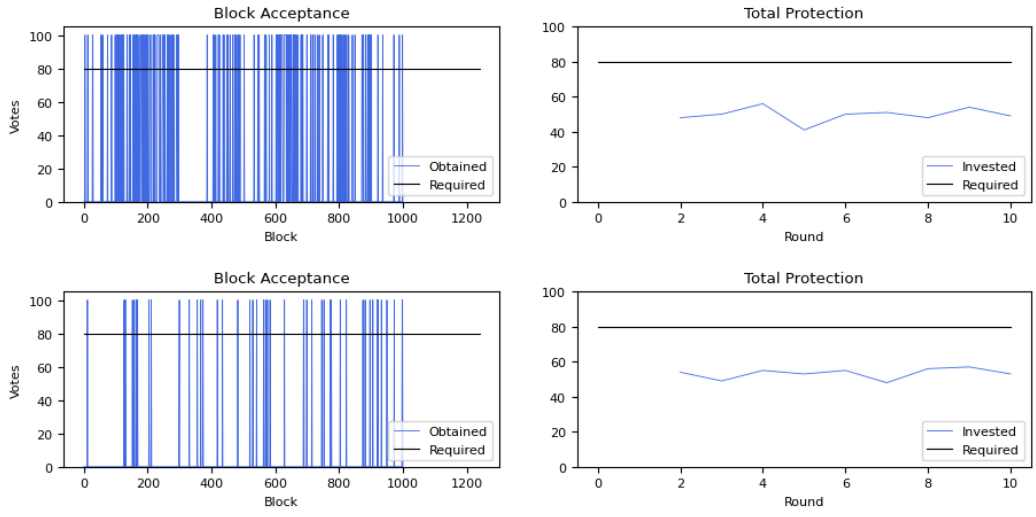
Convergence for tolerance thresholds 20% when benefit per unit of cost 1.2

Repeated results for figure 4.18



Peer reputations at differing tolerance thresholds

Repeated results for figure 4.19



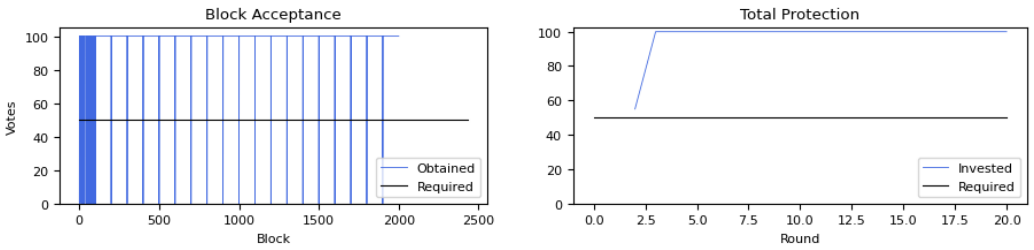
Convergence for tolerance thresholds 20% at differing attack probabilities

Repeated results for figure 4.20

Reputation Maximization Rounds: **20**

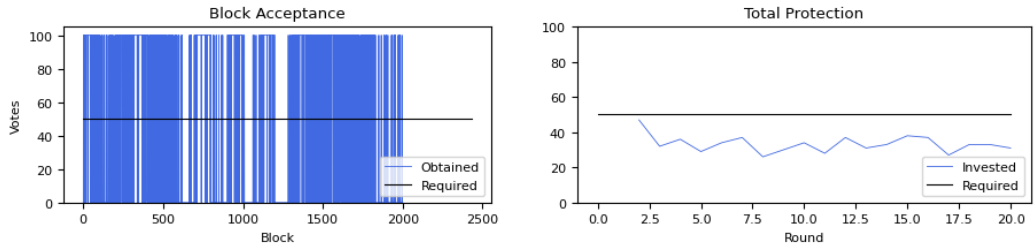
No-of-Peers	Votes-Required	Block-Timeout	Benefit-Per-Unit-Of-Cost	Min-Attack-Probability
100	50.1	25	3	0

Min-Con-Delay	Max-Con-Delay	Min-Peer-Cons	Max-Peer-Cons	Heterogeneous-Cost	Cost-Std-Dev
2	9	2	6	false	-



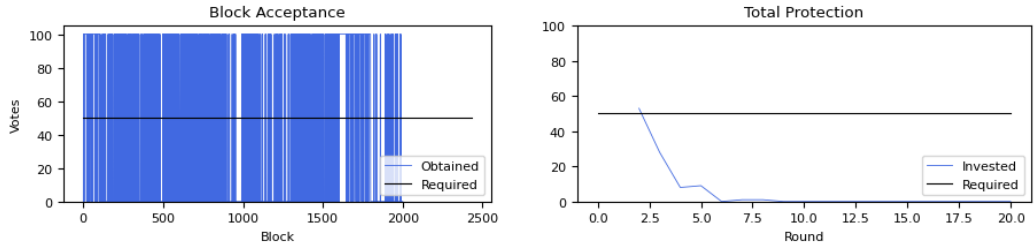
Reputation Optimization Learning Strategy execution for 20 rounds

Repeated results for figure 4.21



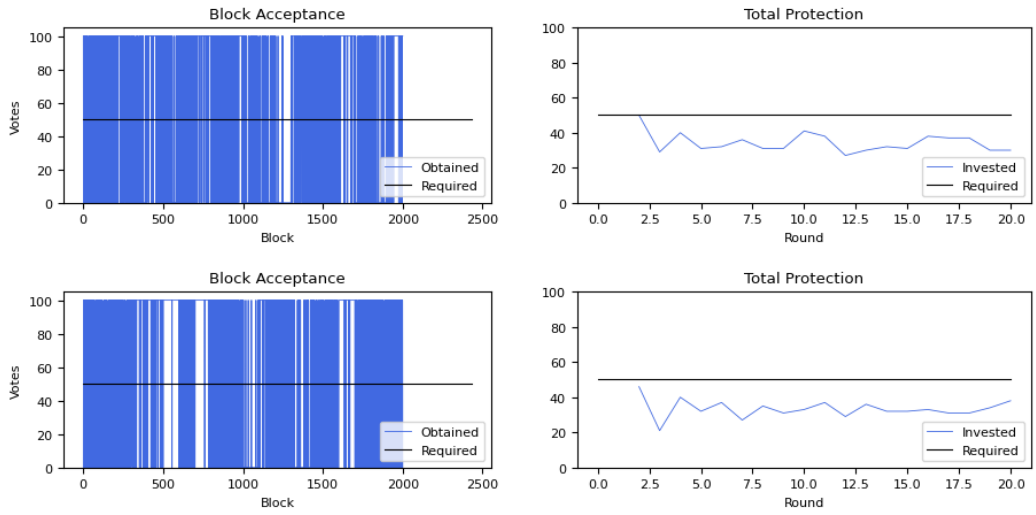
Regret Matching Learning Strategy execution for 20 rounds

Repeated results for figure 4.22



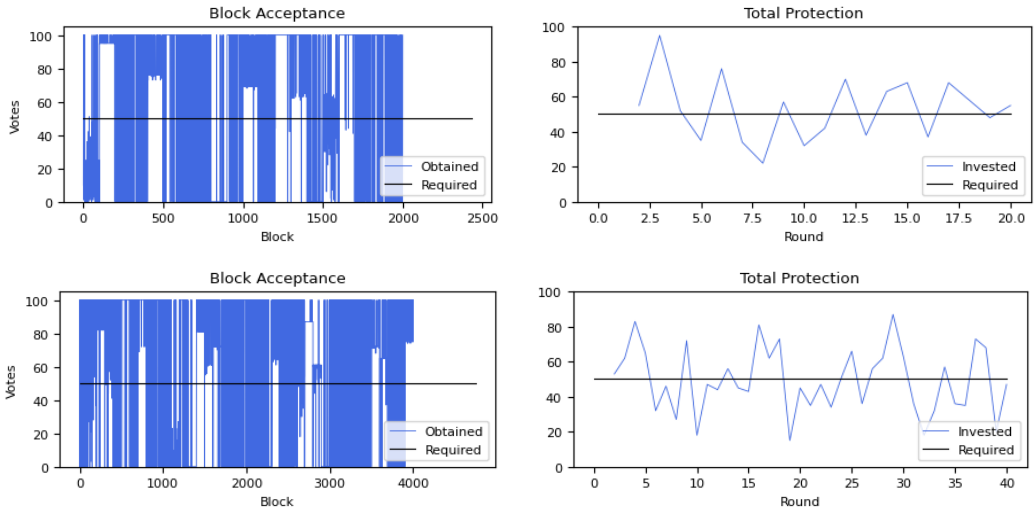
Regret Matching Learning Strategy (with History) execution for 20 rounds

Repeated results for figure 4.23



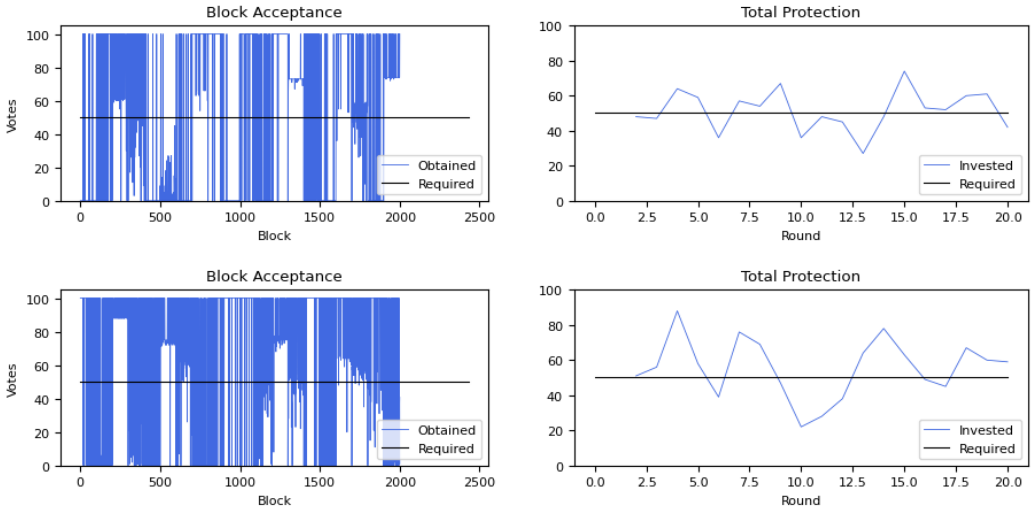
Regret Matching Learning Strategy execution for differing benefits

Repeated results for figure 4.24



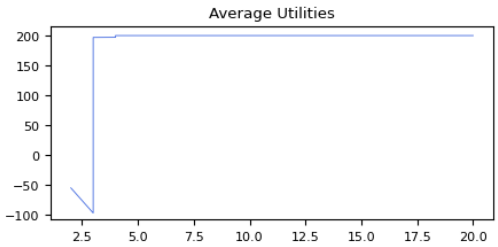
Bounded Rationality Learning Strategy execution

Repeated results for figure 4.25



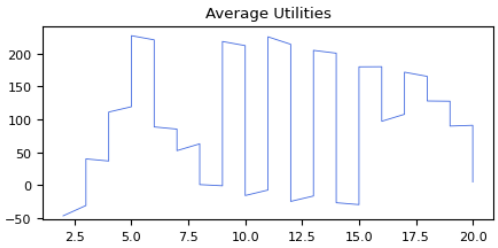
Bounded Rationality Learning Strategy execution for differing benefits

Repeated results for figure 4.26



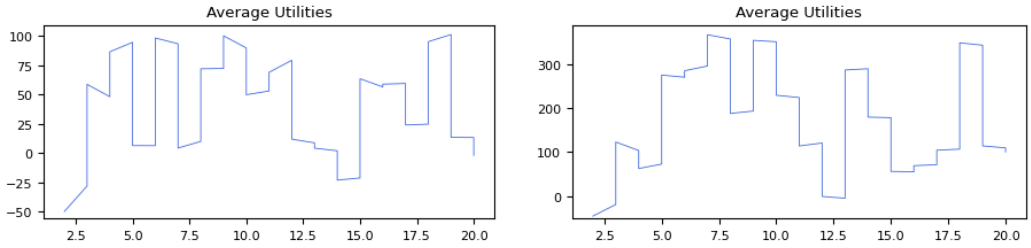
Reputation Optimization Learning Strategy Utilities for 20 rounds

Repeated results for figure 4.27



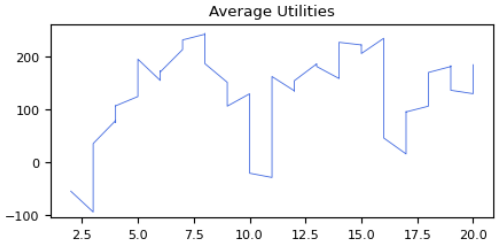
Regret Matching Learning Strategy Utilities for 20 rounds

Repeated results for figure 4.28



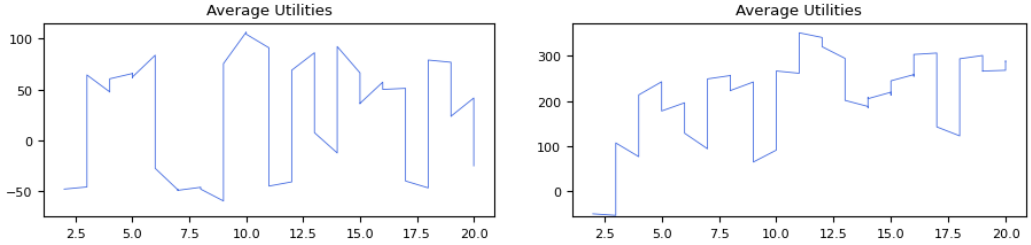
Regret Matching Learning Strategy Utilities for differing benefits

Repeated results for figure 4.29



Bounded Rationality Learning Strategy Utilities for 20 rounds

Repeated results for figure 4.30



Bounded Rationality Learning Strategy Utilities for differing benefits

Repeated results for figure 4.31

APPENDIX C DIGITAL DOCUMENT AND SIMULATION CODE

This section contains a compact disk that includes the below resources.

- Dissertation in PDF format
- Simulation source code for the NetLogo model