A Novel Framework for Evaluating P2P Identity Solutions

Evan Craska[1], Brendan McGlynn[1], Daniel Monteagudo[1], Jacob Peterson[1], and Alex Rosse[1]

[1]Computing Security, Rochester Institute of Technology

December 9, 2020

Abstract

In an increasingly centralized world with rapid proliferation of data, many researchers and believers in individual privacy have created systems for online authentication that do not rely on massive datastores. These systems rely on peer-to-peer (P2P) networks, in which the peers on the systems are responsible for tasks normally carried out by centralized servers. By decentralizing these tasks, not only are provider resource requirements (such as storage and computation) reduced by spreading the responsibility among peers in the network, but also many security advantages are gained, such as eliminating single points of failure and preventing Man in the Middle (MitM) attacks between users and a centralized authority.

Currently, there are no uniform evaluation methods to examine the overall quality of decentralized authentication models inside P2P networks. This paper will focus on defining five categories (*Fairness, Performance, Robustness, Scalability, and Security*) where many existing authentication models inside P2P networks struggle. Within each of these categories, certain criteria will be defined to assist in measuring the effectiveness and capabilities of a decentralized authentication scheme. Finally, the newly created Evaluation Framework will be applied to two well-recognized P2P networks (Microsoft's Identity Overlay Network and Gnutella), and these results will be discussed.

Keywords

Authentication, Peer to Peer, P2P, Blockchain, Network Protocols, Evaluation Framework.

1 Introduction

There have been many attempts to create models and systems for decentralized peer to peer authentication [1][2]. Current frameworks, such as the laws of identity, encompass rules for evaluating identity management frameworks [3]. One notable issue, however, is that there are not many viable methods to determine the overall effectiveness of various decentralized peer to peer mechanisms. Having a method to do this would be highly beneficial for many system administrators and developers wishing to implement these types of mechanisms in an environment.

For our main topic, we will be taking a deep dive into decentralized authentication mechanisms. To analyze these systems, we will also venture into the subtopics of cryptography and systems theory to determine which requirements should be met in strong decentralized trust management systems.

To compare various decentralized authentication mechanisms, we will be designing a practical evaluation framework. We will first compartmentalize existing P2P decentralized mechanisms into individual criteria, and the presence of each criterion can be used to create a score ranking for each tested protocol.

To identify a set of criteria for a strong decentralized authentication mechanism, we can look at existing research for attributes present in highly-respected protocols. One example paper, from Bell Laboratories, highlights a zero-knowledge proof. These proofs are the foundation for strong decentralized authentication protocols, and this specific proof is claimed as proficient, as it has a constant number of rounds, communication linear in length of the statement and the witness, as well as negligible knowledge errors [2]. These three attributes are certainly beneficial to have, and this type of information would be considerable for our framework. With a criteria list generated from a large combination of sources, we will have a large pool of criteria to assess whether a given protocol is up to standard.

The report is split into 7 sections. **Background & Significance** will describe the layout of current P2P networks. **Related Work** summarizes previous work related to P2P networks and compares similarities amongst these projects. **Research Design & Methods** will go into detail about our newly developed "Evaluation Tool", as well as describe each of the criteria defined within it. **Findings** will include an in-depth analysis of two well-known P2P systems using our Evaluation Tool. **Conclusions** will summarize our findings and discuss future work. **Acknowledgements** will give credit to those who contributed to the success of this project. Finally, **Appendix** will contain the GitHub link to where the Evaluation Tool template is stored.

2 Background & Significance

With the increasing use of decentralized P2P networks, there is a need for a uniform framework to evaluate the numerous factors within these systems. Our framework would greatly assist individuals deciding between multiple P2P networks, as various protocols are better suited to address various tasks (security, scalability, etc.). This can also be beneficial for those developing new systems, as they can use this framework as a guideline to create the most accomplished system as possible. To understand what our new framework should assess, a wide variety of current P2P systems must first be be researched.

A P2P network is one where two or more peer nodes are connected without a centralized server, allowing for the share of resources such as CPU cycles, storage, bandwidth, and file storage. These systems gained popularity from file-sharing systems, and many transitioned into usage in mainstream networking. Nodes in P2P networks are responsible for self-organizing, communication, and adapting to their ever-changing environments, and these tasks are usually handled by a centralized server. There have been many changes in the structure of these networks since their first documented use, and this can affect variables within the network [4].

There are many ways P2P networks can be structured to be able to communicate with each other (see Figure 1) [5]. In most cases, these systems use overlay networks defined as a collection of logical links that communicate at the application layer [5]. The overlay networks can be further broken down into single or multi overlay networks.

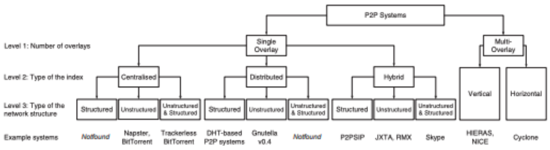


Figure 1: P2P Systems Classifications [5]: Diagram Image Link

2.1 Single Overlay Networks

A single overlay network exists as an isolated network where node communication is contained [5].

2.1.1 Indexing Type

Single overlay networks can be further characterized by the type of indexing used to locate nodes: centralized, distributed, and hybrid. In centralized indexing, the index of the P2P network is stored on one or more centralized servers that are often referred to as “trackers” [5]. Trackers originated in the popular P2P system BitTorrent and is a server used to look up peers holding a given object [6]. In distributed indexing, there is no centralized server and the index is distributed among the nodes. Finally, there is hybrid indexing which combines centralized and distributed indexing. In hybrid indexing, there are “super-nodes”: a small subset of nodes in the network that are responsible for maintaining the indexing of the traditional nodes connected to them [5]. The difference between the trackers (described above) and super-nodes is that trackers are independent devices, whose main purpose is to perform the indexing function. Super-nodes, on the other hand, are user devices that perform additional functions [5].

2.1.2 Network Structure

Each of these can be further broken down based on their type of network structure, which determines the type of routing algorithm that needs to be used: Structured, Unstructured, and Combination. Structured P2P systems will maintain a global data structure, which are typically based on algorithms called Distributed Hash Tables (DHTs) [5]. A DHT provides functions (similar to hash tables) that deal with the distributed data in the network. DHT provides a global, abstract key space where all the resources have unique identifiers (IDs). The data is treated as a tuple (K,V): K denotes the key that is mapped from the data by a hash function and V denotes the original data. Therefore, all data and nodes in these systems can efficiently be mapped into the DHT space [7]. In unstructured P2P systems, the nodes and their resources are arranged in an unorganized way. These systems rely on a technique called flooding, which is a function that will propagate through the network by mapping neighbors until the Time-To-Live (TTL) parameter hits zero [5]. The final structuring option is a combination between structured and unstructured and is usually used by hybrid systems. These systems will use structured communication between all super-nodes and use unstructured between super-nodes and their ordinary nodes [5].

2.2 Multi-Overlay Networks

Multi-Overlay networks are comprised of interconnected single overlay networks that form a new entity [5].

2.2.1 Topology Structure

These networks can be categorized based on their topology structure into vertical or horizontal. These systems may use different types of indexing and network structures in each of their independent single overlay networks. In vertical P2P systems, each overlay network forms a layered structure, where each layer is independent. These networks are usually DHT-based, and the communication between each layer is implemented by gateway nodes [5]. In horizontal P2P systems, each overlay network (each of them called a “leaf”) is connected to a single DHT-based P2P network [5].

With ever changing environments, it is important to understand the numerous current categorizations of P2P networks. Each of these different forms has positive and negative attributes which will be considered later in this paper to determine what characterizes an ideal P2P network.

3 Related Work

While researching various criteria for our framework, we decided to research various attempts to create decentralized identity management schemes. Many of these schemes also implement scoring mechanisms, some of which may be considerable for our evaluation framework.

3.1 Seven Laws of Identity

Our first piece of related work involves IdM schemes and the Seven Laws of Identity. In the IEEE Security Privacy Journal "A First Look at Identity Management Schemes on the Blockchain", the researchers compare the functionality and architecture of multiple DLT-based IdMs (Distributed Ledger Technology Identity Management schemes) [8]. This paper considers three specific DLT-based IdMs: uPort, ShoCard, and Sovrin [3].

For our purposes, the most notable section of this paper is the lack of definitive criterion existing to evaluate IdM schemes. Instead of describing a criterion, this paper resorts to the "laws of identity", a well-known framework that addresses "the successes and failures of digital identity systems" [3].

The IEEE Security Privacy Journal describes the seven laws of identity [3]:

1. "User control and consent – Information that identifies the user should only be revealed with that user’s consent."
2. "Minimal disclosure for a constrained use – Identity information should only be collected on a “need-to-know” basis and kept on a “need-to-retain” basis."
3. "Justifiable parties – Identity information should only be shared with parties that have a legitimate right to access identity information in a transaction."
4. "Directed identity – Support should be provided for sharing identity information publicly or in a more discreet way."
5. "Design for a pluralism of operators and technology – A solution must enable the inter-working of different identity schemes and credentials."
6. "Human integration – The user experience must be consistent with user needs and expectations so that users are able to understand the implications of their interactions with the system."
7. "Consistent experience across contexts – Users must be able to expect a consistent experience across different security contexts and technology platforms."

The existence of these laws brings the question: why create our evaluation framework in the first place? While these laws provide strong guidelines for frameworks, they do not provide a means to directly score the effectiveness of decentralized identity management schemes. If a developer was to decide between two identity management schemes, the task would be made easier if each framework had its own effectiveness score.

3.2 Scoring Mechanisms

In the development our own scoring mechanism for The Framework, we researched other forms of scoring systems within decentralized protocols. While the goal of our scoring system is to assess the quality of a protocol itself, other researchers have implemented them to build reputation systems.

In the paper "An Adaptive and Robust Reputation Mechanism for P2P Network", a scoring mechanism is utilized for evaluating nodes' experiences with other target nodes [9]. This reputation system was designed using a combination of trust values, a decay timer for trust decisions, and the total number of transactions in a confidence assessment [9].

In "A Reputation-based Trust Management System for P2P Networks", an entirely different scoring mechanism is built. Here, the researchers propose their own trust management system and utilize simulation experiments to test its overall effectiveness [10]. For this system, a trust score is given to peers based on interactions in the network [10]. The successful experiments from this paper indicate that this management system is proficient at limiting the spread of malicious content [10].

Despite the differences between these scoring systems and ours, we see find this previous research highly valuable. For our evaluation framework, the presence of a reputation system would make for a useful addition to our criteria.

4 Research Design & Methods

Based on our extensive research, we have defined five points of concern the average P2P system will face: fairness, performance, robustness, scalability, and security. The following sections will analyze each of these categories and determine the criteria that will create an ideal P2P system.

4.1 Fairness

Ensuring fairness is an integral part of delegating control across many nodes. Without a centralized authority, all nodes must work in conjunction to ensure proper communication. Many mechanisms exist to prevent malicious actors from taking advantage of a P2P network. The following criteria can be used to ensure fairness in decentralized protocols.

- **Peer Behavior Monitoring**

  The general monitoring of peer behavior is an essential way to restrict malicious activity in a network. With no central authority to determine which actors may be malicious, it is the responsibility of all nodes to monitor actions taken by other nodes.

  The Iowa State University paper "A Reputation-based Trust Management in Peer-to-Peer Network Systems" [10] describes a practical implementation of behavior monitoring and reputation scoring.

  In this paper, three behaviors are monitored by each peer: resource searching, resource uploading, and resource downloading [10]. These behaviors are monitored to allow for each peer to log peer ID numbers, total number of actions, and number of bad actions taken by other nodes [10]. The collection of this type of information is the first major step in managing potentially malicious behaviors on a decentralized network.

- **Reputation Scoring**

  Once behaviors are recorded by nodes in a P2P network, a reputation system would need to be implemented to measure the overall trustworthiness between nodes.

  The Iowa State University paper mentioned above describes a system where each node can score the reputations of other nodes on the network. With behavioral information logged, threshold values can be decided per each node to access the trustworthiness of different nodes. Using a range of 0 to 100, nodes are identified as having complete trust, being generally trustworthy, or being untrustworthy [10]. All nodes rank the trustworthiness of other nodes independently, meaning that a particular endpoint may be trustworthy to some host (and untrustworthy to another).

- **Behaviors can change based on Reputation System**

  Once reputation scores are calculated per each node in a network, the last major progression for a decentralized protocol to take would be behavioral changes.

  The Iowa State University paper describes its own method for changing node behavior. Two trust thresholds are given: X_1 and X_2 [10]. In this model, a threshold exceeding

X_1 would indicate distrust, and a threshold less than X_2 would indicate full trust (with scores in between indicating average trust) [10]. Once a node calculates reputation scores for other existing nodes, it may utilize its own configured trust threshold values to decide whether it will accept or reject any traffic from them.

4.2 Performance

The performance of a P2P network is essential to the proper function of a system. Performance is described as the ability of the network to perform defined tasks with an acceptable delay. It is generally measured using either routing delays or, most commonly, hop count metrics. An end-to-end routing delay is the cumulative delay of the data and nodes on a given communication path. Lookup hop count is the number of logical links the lookup request must go through during the lookup phase [5].

The significance of performance is demonstrated in two types of procedures: those who are performed often and ones who are less frequent and heavy to perform. The types of procedures that occur frequently include resource indexing and the discovery of nodes in the network. If these frequent operations do not run at an acceptable performance metric, their issues can compound to create performance issues for the entire network. If procedures are heavy to perform (such as group establishment or joining groups) do not operate at an acceptable metric, the network will be significantly impacted by delays [5]. The following criteria will help to ensure performance is maintained in P2P networks:

- **Use of Parallel Look-ups**

  P2P networks can benefit greatly if the system uses parallel look-ups when querying nodes, especially in the node finding/discovery processes. During a parallel lookup, a client will simultaneously send multiple lookup requests to different peers and will compile the results [11]. There are multiple advantages associated with using parallel look-ups. First, the time for the first and last data objects are processed faster [5]. Secondly, each node sends out multiple requests, leading to more nodes ultimately being searched. This would increase the probability of finding a closer node for communication, thus decreasing the hop count every time nodes need to communicate [11]. While this does lead to a slightly larger number of requests being produced, research has shown this delay is almost negligible in relation to the whole P2P network's traffic [5]. The specific number of parallel look-ups that best fit a P2P network may vary, but a form of parallel lookup would be beneficial to implement.

- **Use of Analytic Hierarchy Process**

  Analytic Hierarchy Process (AHP) is a method that breaks down aspects of decision-making into categories such as goals, criteria, and plans. It also performs both qualitative and quantitative analysis based on those categories. Studies have shown that having an AHP incorporated into a P2P network will increase performance because it takes into consideration the factors of the real-time environment to determine the most suitable nodes for communication. The example AHP process takes the following into consideration: the performance of bandwidth, CPU, storage, on-life time, and mobility. When compared to a traditional algorithm of choosing communication with peers, the AHP algorithm performs better in elements such as failure rate, query delay, and indexing time [12]. The AHP-based algorithm will use the factors for criteria weighting and evaluation by pairwise comparison. The relative weights of the factors are achieved through calculating the eigenvector of the matrix with the eigenvalue that is closest to the number of factors used [13]. While these specific characteristics do not need to be implemented, some form of an AHP based algorithm should be considered in a P2P network.

- **Nodes are Location Aware**

  Efficiently locating content in a P2P system is a challenging problem. Since there generally are no centralized places holding node location information, it is the responsibility of individual nodes to index the network. Ensuring this process is done in a reasonable amount of time is a great area of research with decentralized systems. In recent years, there have been great successes with implementing location based P2P systems to help the nodes self-organize based on their location, which will ultimately reduce the hop count during communication. One implementation suggests only needing to send a flood request once when a node first joins the network. From this, each node will keep a "shortcut list" in which the peers determine which nodes are closet to them for later communication [14]. Another study showed that when compared to a normal P2P network, the location aware system provides on average a 46% better success ratio [12]. Another research method showed each node will keep a variable within themselves called an "Area Identifier (aID)" to denote the peer's location in the network. From there, all nodes who have the same aID will eventually come together to create their own unique subnet [15]. There are numerous methods to implement location awareness in P2P networks, and there are many additional benefits to be seen with its usage.

While there are numerous other ways to improve performance metrics within P2P systems, the defined criteria above are just three possible forms of improvement. When considering performance improvement, it is important to remember that an increase in performance typically correlates with a decrease in another variable (often security). It is critical to ensure there is good balance of all these variables to develop an efficient P2P network.

4.3 Robustness

In the event of system failures (or network configuration changes) in a P2P network, robustness addresses the ability of a system to maintain proper functionality. Another goal of robustness is being able to contain and restrict the actions of malicious nodes on the network, as these nodes may cause damage to other parts of the network. As such, three robustness challenges addressed in a well-designed P2P network are churning, handling super-peer outages (if super-peers are present in the system), and identifying and blacklisting malicious nodes.

- **Churn**

  Churn is the measure of the number of individual items, in this case nodes, moving out of a collective group. Part of our framework addresses churn in a P2P network over a given period of time. Handling churn properly is important when considering robustness because failing to do so may present a number of issues in authentication systems. If nodes continue to wait for responses from a removed node, computational time would be wasted. A robust system should be able to quickly adapt to and handle a dynamic network, resolving requests seamlessly through the changes without end-users being aware that other nodes have been added to or dropped from the network. Churn is especially relevant to Internet of Things (IoT) device focused networks, as these devices tend to be frequently relocated and re-purposed physically and logically. Many of the proposed decentralized authentication schemes focus on handling authentication for IoT devices, increasing the weight of consideration for churn on the topic of P2P authentication. Additionally, as mobile networks continue to grow in size and ability, a P2P authentication model may become more relevant in that world as technology continues to be developed for the platform.

- **Super-Peer Outages**

  Many proposed decentralized systems implement super-peers which hold additional responsibilities. Super-peers are nodes which manage higher tasks in the network than regular operating nodes, including managing authentication requests, requesting trust information from other nodes, and handling puzzle distribution. Considering the higher responsibility of these nodes in the system, it marks them as potential targets for malicious actors, as bringing them down may prevent operations for the authentic nodes in the network. It follows that a robust system is able to handle outages (both network and hardware) for these super-peers in order to continue system operation in the event that one or more super-peers are unable to perform their duties. If a super-peer were to be brought down alongside a decentralized system, many of the benefits of a P2P design are lost. In this situation, a single point of failure (or marginally more) would be present. Decentralized responsibility is a large reason why research is being done into P2P authentication schemes, as many current authentication schemes for IoT devices are centralized and create a single point of failure [16]. If they are present in the model, it is important to maintain focus on covering super-peers, as otherwise, the goal of the research in reducing single points of failure and decentralization may be unfulfilled.

- **Malicious Nodes**

  Malicious nodes are able to create many problems inside a P2P network if not properly managed and contained. Examples of these problems include targeting important nodes, such as super-peers or highly trusted nodes, creating difficulty in operation or reducing the average trust among nodes, potentially raising themselves higher in the hierarchy of trust, and gaining a high trust value and intentionally providing false responses or condemning other highly-trusted but authentic nodes. With these possibilities available to maliciously acting nodes, it is evident that a robust system should be able to correctly identify malicious nodes and prevent them from succeeding in their goal. A system's accuracy in identifying malicious nodes is also important to be considered, as marking an authentic node as malicious would ultimately hurt the system by lowering the amount of authentic nodes in the network.

4.4 Scalability

“Scalability is the ability to accommodate an increasing number of nodes in a P2P system without severely degrading the performance of the system" [17]. The measurement of hop count in relation to the network size is how scalability is evaluated. The following criteria correlates to the possible scalability of P2P networks:

- **No single authorities are used**

  To remove the capacity limits of a centralized server, a decentralized distributed delegation mechanism can be implemented. This would vastly impact the scalability of a P2P network. Allowing multiple authorities the ability to grant peer group membership allows for the avoiding of a single point of failure, in addition to the reduction of overhead (and the response time of an authority). In "An adaptive and robust reputation mechanism for P2P network"[9], a proposed voting scheme is implemented to grant permissions and accept new members by votes passed by existing peer groups [9]. The designated groups allow each individual peer a vote in the ability to accept a new peer[9]. If the desire to not allow every peer the voting ability, then multiple authorities or super-nodes must be present in a group[9]. A combination of the two authorization mechanisms would establish a secure cooperative process to enhance scalability.

- **Decisions made based on attributes of authenticated peers**

  The largest concern with scalability is the increase of malicious users. As a network grows, so does the risk of malicious users joining. To combat malicious peers, proper reputation management must occur. Reputation management has been widely developed by allowing peers to interact with other peers based on a peer's attributes or global reputation. The reputation of a peer can be created from trust surveys or scores from previous interactions within the network, and interaction can be limited to only reputable peers. Malicious or poorly trusted peers will be given restricted access.

- **Graceful Scaling Test**

  A hybrid network with a centralized server could be used not to authorize, but to cache resources and provide a mapping (such as the routing to certain groups of the network). In a proposed architecture from the Swiss Federal Institute of Technology, developed algorithms for storage proved that a search can be performed in O(log n) where n = number of agents, O(log n). = how storage space scales at each agent.[18]

4.5 Security

The security of a P2P network is highly correlated to its robustness and fairness, as disruptions to either of these categories can constitute a security threat to the network in the form of denial of service. Specific to security, however, are certain features that ensure the safety of nodes operating on the network and the protection of private data in transit on the network. For this reason, the three criteria addressed include the following:

- **Use of Cryptographically Secure Primitives for Access Control**

  Nodes attempting to access resources on a system often require authentication to ensure access to restricted resources remains private. This criterion ensures that this authentication is conducted in a cryptographically secure manner, such as through the use of Zero-Knowledge Proofs or Merkle Puzzles [1].

- **Minimal Disclosure of System Information**

  Current research surrounding the evaluation of P2P identity management systems frequently cites a Microsoft paper describing the "Seven Laws of Identity" [3]. One of these laws describes "Minimal Disclosure for a Constrained Use" [19]. To prevent theft of information from nodes on the system, it is important that the system is designed to store and reveal as little information as possible, only that which is necessary to authenticate users.

- **Encryption of Data in Transit**

Due to the decentralized nature of P2P systems, it is very common for nodes on the network to multicast or broadcast information to the rest of the network. Furthermore, specifically in authentication systems, it is especially common to defer authorization tasks to another network node or group of nodes [20]. Sensitive data that is broadcast or multicast could potentially be observed by malicious nodes on the network, and therefore it is important that such information is encrypted in transit [21].

While the robustness and fairness aspects of our model serve to protect integrity and availability of data on the P2P network, these security considerations are critical to maintain the confidentiality of private information used in a decentralized authentication mechanism.

4.6 Designing the Evaluation Tool

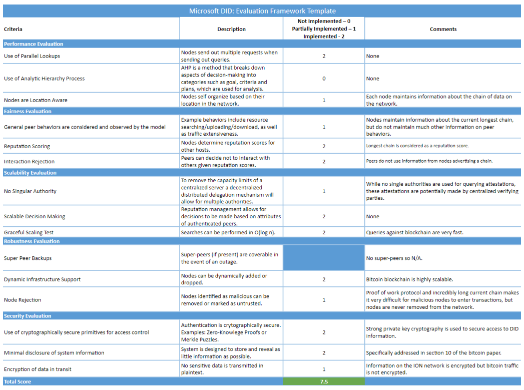
The criteria defined in each of the five sections above is used to create a tool which can be used to audit P2P systems and see how they rank when compared to other systems. The Evaluation Tool is designed as an Excel spreadsheet and allows an auditor to rank each criterion on a zero to two basis. A user should put a zero in if the the P2P system does not implement what is being asked. A one should be used if the system has the criterion partially implemented, also if this is the case, a user should add a comment explaining what they have currently implemented. Finally, a two will be used if the criterion is fully being met. Based on each of these scores, the tool will produce a final score out of ten. The tool will rank the P2P as being compliant if the score is greater than seven, and if below it will be deemed non-compliant. If the score is non-compliant, we suggest the user implement some of the lacking criterion to raise the score.

4.7 Framework Application

Finally, the Evaluation Tool will be used in practice to examine two well know P2P systems: Microsoft's Identity Overlay Network and Gnutella. The **Findings** will show the in-depth analysis of the two networks to show the strengths and weaknesses of each system.

5 Findings

To demonstrate the application of our P2P evaluation framework, we demonstrate its application to two P2P networks: the new Identity Overlay Network by Microsoft's Decentralized Identity Foundation, as well as the established P2P file sharing system Gnutella.



Microsoft DID: Evaluation Framework Template			
Criteria	Description	Not Implemented – 0 Partially Implemented – 1 Implemented - 2	Comments
Performance Evaluation			
Use of Parallel Lookups	Nodes send out multiple requests when sending out queries.	2	None
Use of Analytic Hierarchy Process	AHP is a method that breaks down aspects of decsion-making into categories such as goal, criteria and plans, which are used for analysis.	0	None
Nodes are Location Aware	Nodes self organize based on their location in the network.	1	Each node maintains information about the chain of data on the network.
Fairness Evaluation			
General peer behaviors are considered and observed by the model	Example behaviors include resource searching/uploading/download, as well as traffic extensiveness.	1	Nodes maintain information about the current longest chain, but do not maintain much other information on peer behaviors.
Reputation Scoring	Nodes determine reputation scores for other hosts.	2	Longest chain is considered as a reputation score.
Interaction Rejection	Peers can decide not to interact with others given reputation scores.	2	Peers do not use information from nodes advertising a chain.
Scalability Evaluation			
No Singular Authority	To remove the capacity limits of a centralized server a decentralized distributed delegation mechanism will allow for multiple authorities.	1	While no single authorities are used for querying attestations, these attestations are potentially made by centralized verifying parties.
Scalable Decision Making	Reputation management allows for decisions to be made based on attributes of authenticated peers.	2	None
Graceful Scaling Test	Searches can be performed in O(log n).	2	Queries against blockchain are very fast.
Robustness Evaluation			
Super Peer Backups	Super-peers (if present) are coverable in the event of an outage.		No super-peers so N/A.
Dynamic Infrastructure Support	Nodes can be dynamically added or dropped.	2	Bitcoin blockchain is highly scalable.
Node Rejection	Nodes identified as malicious can be removed or marked as untrusted.	1	Proof of work protocol and incredibly long current chain makes it very difficult for malicious nodes to enter transactions, but nodes are never removed from the network.
Security Evaluation			
Use of cryptographically secure primitives for access control	Authentication is crytographically secure. Examples: Zero-Knowlege Proofs or Merkle Puzzles.	2	Strong private key cryptography is used to secure access to DID information.
Minimal disclosure of system information	System is designed to store and reveal as little information as possible.	2	Specifically addressed in section 10 of the bitcoin paper.
Encryption of data in transit	No sensitive data is transmitted in plaintext.	1	Information on the ION network is encrypted but bitcoin traffic is not encrypted.
Total Score		7.5	

Figure 2: Evaluation of Microsoft Identity protocol: GitHub Image Link

5.1 Microsoft's Identity Overlay Network

This network uses the sidetree framework, designed by the Decentralized Identity Foundation, as the backbone which anchors attestations regarding identity information [22]. The sidetree framework, as well as the bitcoin blockchain, are being evaluated to some degree when examining this new network.

5.1.1 Performance

- This Identity Overlay Network (ION) is built on the bitcoin network, in which every node on the network is location-agnostic and maintains a full history of transactions in a chain of blocks, all linked by hashes [23]. Because of this structure, parallel lookups are not a problem as each full node can reference the data stored in its own blockchain, so this network received two points.
- The blockchain does not use an analytical hierarchy process, due to the location-agnostic nature of the nodes, so this aspect of the evaluation received zero points.
- The platform received one point for nodes being location-aware, as awareness of location in the network is rendered somewhat unimportant by the blockchain system.

5.1.2 Fairness

- Nodes on the bitcoin blockchain observe the behavior of other nodes in order to determine the longest chain being broadcast, but they do not pay much attention to other behaviors of the other nodes, so this section receives one point.

- In the ION, nodes inherently trust the information on the bitcoin blockchain due to attestations given by (potentially centralized) identity verification parties. The attestations are stored on the bitcoin network, in which the longest chain is always considered to be correct, while other chains are incorrect. This could be considered a form of ranking peers, so the network received 1 point in this category.

- The bitcoin protocol explicitly states that nodes do not accept information from nodes with a less than maximal blockchain, which can be equated to a reputation score, so the network received two points in this regard.

5.1.3 Scalability

- While the part of the ION on the blockchain does not require any single authorities, attestations of identity that are stored on the blockchains can be issued by a central authority, potentially causing centralization issues [24]. For this reason, this section received only one point.

- As the decision on whether or not to accept a block is based solely on the length of the blockchain, the decisions are being authenticated in a way. For this reason, this criterion received two points.

- Due to the large number of nodes in the bitcoin network (and the fact that each has a full copy of the blockchain stored), queries within the P2P identity network defined by Microsoft are very fast, and their use of a sidechain is used to make these lookups even faster than would be supported directly on bitcoin's network. For this reason, the network received two points for this criterion.

5.1.4 Robustness

- As super-peers are not a part of the network structure of ION, this section was left blank and the comment indicated that the criteria was not applicable.

- The ION, and the underlying bitcoin network, are very flexible when it comes to nodes adding and dropping from the network. This holds true in large volumes, mostly due to the underlying blockchain technology. For this reason, this criterion received a full two points. It is worth noting that if a sufficiently large number of nodes were to drop from the network, it might be possible to mount a 51% attack on the underlying blockchain. With the current number of nodes on the network, this would be unlikely enough to preclude consideration [25].

- While nodes are never removed from the network, the proof-of-work system underlying the bitcoin blockchain makes it very unlikely that a malicious node could compromise the network, so this criterion received one point.

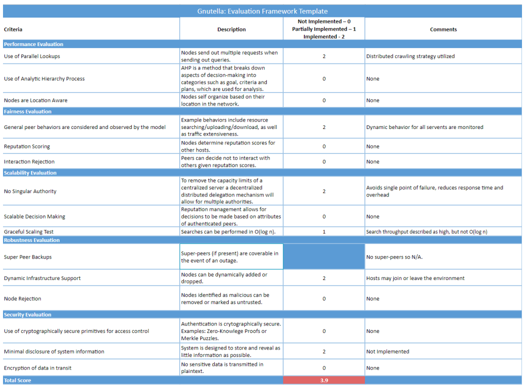
5.1.5 Security

- The ION uses public key cryptography to protect attestation information that is placed on the blockchain, so this criterion received the full two points [22]. While the system is currently not secure against quantum computer attacks, it is currently secure enough to garner full points. There is also work currently being performed to upgrade the bitcoin blockchain's resistance to quantum computing attacks [26].

- As the ION only places attestations on the bitcoin sidechain, it can be said to reveal a minimal quantity of information. This privacy is further protected by the structure of the bitcoin network (as the number of parties involved in storing information on the bitcoin blockchain is minimized by design) [23]. A full two points can be awarded for this criterion.

- While parts of the ION use encryption in transit, the bitcoin network does not encrypt traffic in transit, so this criterion only receives one point.

Overall, the framework gave a score of 7.5/10, or 75%. This is over the passing threshold we defined at 7/10, so this network would be considered acceptable. The elements in which Microsoft's Identity Overlay Network performed worst were largely related to overall network oversight, which is heavily lacking in blockchain-based systems. In addition, the potential for centralization in verification authorities caused their network to lose points in the decentralization criteria. The suggestions for ION that could be gleaned from the use of this network would be that more careful monitoring of nodes in the sidechain could yield better performance and resistance to attack, and that Microsoft should attempt to promote decentralization of



Gnutella: Evaluation Framework Template			
Criteria	Description	Not Implemented – 0 Partially Implemented – 1 Implemented - 2	Comments
Performance Evaluation			
Use of Parallel Lookups	Nodes send out multiple requests when sending out queries.	2	Distributed crawling strategy utilized
Use of Analytic Hierarchy Process	AHP is a method that breaks down aspects of decsion-making into categories such as goal, criteria and plans, which are used for analysis.	0	None
Nodes are Location Aware	Nodes self organize based on their location in the network.	0	None
Fairness Evaluation			
General peer behaviors are considered and observed by the model	Example behaviors include resource searching/uploading/download, as well as traffic extensiveness.	2	Dynamic behavior for all servents are monitored
Reputation Scoring	Nodes determine reputation scores for other hosts.	0	None
Interaction Rejection	Peers can decide not to interact with others given reputation scores.	0	None
Scalability Evaluation			
No Singular Authority	To remove the capacity limits of a centralized server a decentralized distributed delegation mechanism will allow for multiple authorities.	2	Avoids single point of failure, reduces response time and overhead
Scalable Decision Making	Reputation management allows for decisions to be made based on attributes of authenticated peers.	0	None
Graceful Scaling Test	Searches can be performed in O(log n).	1	Search throughput described as high, but not O(log n)
Robustness Evaluation			
Super Peer Backups	Super-peers (if present) are coverable in the event of an outage.		No super-peers so N/A.
Dynamic Infrastructure Support	Nodes can be dynamically added or dropped.	2	Hosts may join or leave the environment
Node Rejection	Nodes identified as malicious can be removed or marked as untrusted.	0	None
Security Evaluation			
Use of cryptographically secure primitives for access control	Authentication is crytographically secure. Examples: Zero-Knowlege Proofs or Merkle Puzzles.	0	None
Minimal disclosure of system information	System is designed to store and reveal as little information as possible.	2	Not Implemented
Encryption of data in transit	No sensitive data is transmitted in plaintext.	0	None
Total Score		3.9	

Figure 3: Evaluation of Gnutella protocol: GitHub Image Link

the verifying authorities (as too much data passing through centralized authorities for attestation could defeat the purpose of the decentralized system altogether).

5.2 Gnutella Protocol

The Gnutella Protocol is a P2P decentralized model made up of endpoints called servents [27]. A set of descriptors are utilized to manage communication between servents, and the protocol is described as "highly fault-tolerant" due to being decentralized [27]. One massive difference between Microsoft's Identity Overlay Network and the Gnutella protocol is that the ION protocol is designed as a decentralized identifier system, and Gnutella is not. Because of this, we are expecting to see Gnutella's performance to not be best suited for authentication purposes, as identity management is crucial to this. Despite the claims of Gnutella being a highly fault-tolerant protocol [27], there are many potential flaws to be uncovered by our evaluation framework.

5.2.1 Performance

- One of the strongest aspects of Gnutella is the protocol's searching mechanisms. A distributed client/server crawling strategy helps reduce the amount of time needed for performing lookups [8]. In this strategy, the server's responsibilities include managing a list of nodes to be contacted, assigning client work, and constructing the final graph [8]. The role of clients is to receive a list of initial points and analyze the network surrounding these points [8]. A full two points can be awarded for this mechanism.

- The usage of an analytic hierarchy process is unfortunately not present for this protocol. While Gnutella may be utilized to divide a topology graph into clusters [8], an analytical hierarchy process is not utilized in any form (and thus, no points are awarded here).

- Location awareness is also unfortunately absent from in Gnutella. While direct neighbors are accounted for by each servent in a Gnutella network, there are no additional awareness mechanisms present. No points are awarded for this category.

5.2.2 Fairness

- As part of the Gnutella procedure, the network topology, generated traffic, and dynamic behaviors are captured to perform macroscopic analysis [8]. The capturing of this information not only assists in fairness, but also scalability. A full two points can be awarded to this criterion.

- The lack of reputation scoring and an overall reputation system begins the overall downfall of Gnutella, as this information is not considered by the protocol. No credit can be given for this specific criteria.

- Because there is no reputation system present in the Gnutella protocol, there is no reputation-based method for peers to assess each other (and thus, no points can be awarded).

5.2.3 Scalability

- Although Gnutella supports the option of a centralized server search paradigm, the model does not rely on a central authority [27]. A full two points can be awarded here.

- Our next criteria is based on the decision making of peers. While local decisions are made within Gnutella, there is no decision making based on the behaviors of authenticated peers. No points can thus be awarded for this criterion.

- The next test for Gnutella is for Graceful scaling. Our criteria strictly defines graceful scaling as O(log n), and Gnutella's searching efficiency does not meet this criterion. Gnutella's file availability, however, is described as linear when more nodes are added (with file throughput is described as growing) [8]. Given this, we decided to give one out of two points for graceful scaling.

5.2.4 Robustness

- Despite Gnutella's specification describing the protocol as fault-tolerant, there are no super-peer backup mechanisms in place [27]. While some nodes may function if a subset of servents go offline, this is not done via super-peers [27]. Because no super-peer system is in place, we will neither award nor penalize Gnutella for this criterion.
- One beneficial aspect of Gnutella is its dynamic infrastructure. Servents can use local information to dynamically drop or add peers, so we can award a full two points for this criterion [8].
- One massive flaw of the Gnutella protocol is the lack of mechanisms to identify and react upon potentially untrustworthy servents. No mechanisms exist to identify and remove malicious peers, so we cannot award any points for this criterion.

5.2.5 Security

- The entirety of our security section is where Gnutella fails considerably. There is no mentioning of access control within the Gnutella Protocol Specification, let-alone cryptographically secure primitives for access control [27]. Given this, we can not award any points for this criterion.
- As for system information shared between servents within the Gnutella Protocol Specification, there does not appear to be any failures regarding excess information disclosure [27]. Example information stored in the Gnutella query payloads include port number, IP address, speed result, and servent identifier [27]. Because of this, we award a full two points for minimal information disclosure.
- One of the biggest disadvantages of the Gnutella protocol is a lack of encryption for data in transit. In the case study "Peer-to-Peer Architecture Case Study: Gnutella Network", the researchers describe that they were able to eavesdrop on their own network traffic [8]. This is highly concerning to any developer who wishes to store sensitive information in a decentralized manner, making Gnutella traffic susceptible to eavesdropping. The lack of encryption for data in transit makes Gnutella a poor choice to use for authentication purposes. Given these flaws, we cannot award any points here.

Our framework's resulting score for the Gnutella protocol is a 3.9/10 or 39%. While this score may be staggeringly low, the Gnutella developers likely prioritized pure functionality and convenience over security (and other categories). Similar low-scoring protocols may very well have been created as lab experiments or proof of concepts, and it is important to recognize that not all protocols are developed with our criteria in mind.

Looking at our five main categories, Gnutella performed the best for scalability, and all other categories were tied for second place. For the categories of performance, fairness, robustness, and security, there are many places where improvements can be made. Whether Gnutella is a great framework to implement in an application or back-end depends entirely on the developer's use case. While slight flaws in performance or scalability may not be seen as an issue to some developers, security concerns are likely most important for many developers. If credentials or other forms of authentication information are to be stored in a decentralized model, Gnutella may be a poor selection as a protocol.

6 Conclusion

Of the many lessons learned form building our evaluation framework, one of the most vital ones is that different protocols are built to suit different needs. Microsoft's Identity Overlay Network far outperformed Gnutella when tested with our evaluation framework, but that is to be expected given ION's focus on decentralized identities (and Gnutella's focus on acting as a file sharing protocol). Before judging any protocol based on the compliance determined by our framework, it is vital to remember the intent of a protocol before quantifying its characteristics.

Another conclusion is that P2P mechanisms are not well suited for all situations. One example for this is described in the paper "On using peer-to-peer communication in cellular wireless data networks" [28]. At first glance, P2P network statistics showed promise for improved data rates, connect-ability, and overall performance when implemented in a cellular network. Research simulations of 100 nodes found that host mobility and a multi hop approach did not benefit a peer's spatial reuse features, and instead yielded worse per flow throughput [28]. With only 100 nodes being tested (and negative results occurring) the thought of a large-scale implementation for cellular networks would not be ideal [28]. While the topic of this issue is not directly associated with comparing two decentralized protocols, it is important to note that traditional centralized mechanisms may be better suited to address certain problems.

For future development on our current implementation, one potential addition would be the consideration of purely negative aspects of decentralized mechanisms.

Many developers of decentralized P2P mechanisms admit damaging flaws in their implementations, and these flaws could be considered for our overall effectiveness score. One IEEE paper highlights issues of a newly developed side chain structure, with future improvements to be brought at the network and asset level [29]. In a paper published by Springer, it is mentioned that relations with a three-move honest-verifier zero-knowledge (HVZK) proof-of-knowledge are flawed in multiple ways. These protocols which achieve linear communication are not perfect, as they are often reliant on a one-way function or the hardness of a discrete log [2]. To consider specific weaknesses in various implementations, we may choose to create a negative set of criteria to detract from the overall quality score.

7 Acknowledgements

We would like to thank Professor Pelletier for the feedback and encouragement towards our research project. Finally, we would like to thank all the developers for the P2P systems we have mentioned in this paper.

8 Appendix

The appendix section was created to share the GitHub link to our Evaluation Tool. GitHub Link

This includes the Evaluation Tool template which can be used to examine P2P networks, along with the example uses of the tool discussed in this paper.

References

[1] Wierzbicki A, Zwierko A, Kotulski Z. Authentication with controlled anonymity in P2P systems. In: Sixth International Conference on Parallel and Distributed Computing Applications and Technologies (PDCAT'05); 2005. p. 871–875.

[2] Cramer R, Damgård I, MacKenzie P. Efficient Zero-Knowledge Proofs of Knowledge without Intractability Assumptions. Springer, Berlin, Heidelberg; 2000. Available from: https://link.springer.com/chapter/10.1007/978-3-540-46588-1_24.

[3] Dunphy P, Petitcolas FAP. A First Look at Identity Management Schemes on the Blockchain. IEEE Security Privacy. 2018;16(4):20–29.

[4] Buford JFK, Yu HH, Lua EK. P2P Networking and Applications. Morgan Kaufmann series in networking. Elsevier/Morgan Kaufmann; 2009. Available from: https://books.google.com/books?id=MyLcnQEACAAJ.

[5] Koskela T, Kassinen O, Harjula E, Ylianttila M. P2P Group Management Systems: A Conceptual Analysis. ACM Comput Surv. 2013;45(2).

[6] Jia AL, Chiu DM. Designs and Evaluation of a Tracker in P2P Networks. In: 2008 Eighth International Conference on Peer-to-Peer Computing; 2008. p. 227–230.

[7] Zhang H, Wen Y, Xie H, Yu N. Distributed Hash Table. Springer Science+Business Media. Springer New York; 2013.

[8] Ripeanu M. Peer-to-Peer Architecture Case Study: Gnutella Network. The University of Chicago; 2001. Available from: http://people.cs.uchicago.edu/~matei/PAPERS/gnutella-rc.pdf.

[9] Wang M, Tao F, Zhang Y, Li G. An adaptive and robust reputation mechanism for P2P network. In: 2010 IEEE International Conference on Communications. IEEE; 2010. p. 1–5.

[10] Stakhanova N, Ferrero S, Wong J, Cai Y. A reputation-based trust management in peer-to-peer network systems. Iowa State University; 2004. Available from: http://www.cs.unb.ca/~natalia/201.pdf.

[11] Stutzbach D, Rejaie R. Improving lookup performance over a widely-deployed DHT. In: Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications. IEEE; 2006. p. 1–12.

[12] Gross C, Stingl D, Richerzhagen B, Hemel A, Steinmetz R, Hausheer D. Geodemlia: A robust peer-to-peer overlay supporting location-based search. In: 2012 IEEE 12th International Conference on Peer-to-Peer Computing (P2P); 2012. p. 25–36.

[13] Liu Y, Zhou X, Ren S, Yang L, Ci S. Peer selection in mobile P2P networks based on AHP and GRA. In: 2012 18th IEEE International Conference on Networks (ICON); 2012. p. 179–184.

[14] Sripanidkulchai K, Maggs B, Zhang H. Efficient content location using interest-based locality in peer-to-peer systems. In: IEEE INFOCOM 2003. Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (IEEE Cat. No.03CH37428). vol. 3; 2003. p. 2166–2176 vol.3.

[15] Shen Xin-peng, Zeng Lei-jie, Zhao Xiao-nan. Two-layer P2P system based on location. In: 2010 Second International Conference on Communication Systems, Networks and Applications. vol. 1; 2010. p. 151–154.

[16] Almadhoun R, Kadadha M, Alhemeiri M, Alshehhi M, Salah K. A User Authentication Scheme of IoT Devices using Blockchain-Enabled Fog Nodes. In: 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA); 2018. p. 1–8.

[17] ANDROUTSELLIS-THEOTOKIS S, SPINELLIS D. A survey of peer-to-peer content distribution technologies.. vol. 36 of ACM Computing Surveys. Athens University of Economics and Business; 2004.

[18] Karl Aberer ZD. Managing Trust in a Peer-2-Peer Information System. Swiss Federal Institute of Technology (EPFL); 2001. Available from: http://irl.cs.ucla.edu/~yingdi/pub/papers/Aberer-TrustInP2P-CIKM-2001.pdf.

[19] Cameron K. The laws of identity. Microsoft Corp. 2005;5:8–11.

[20] Andersen MP, Kumar S, AbdelBaky M, Fierro G, Kim HS, Culler DE, et al.. WAVE: A Decentralized Authorization Framework with Transitive Delegation. University of California, Berkeley; 2019. Available from: https://people.eecs.berkeley.edu/~raluca/WAVEFinal.pdf.

[21] Oh B, Lee S, Park H. A Peer Mutual Authentication Method using PKI on Super Peer based Peer-to-Peer Systems. In: 2008 10th International Conference on Advanced Communication Technology. vol. 3; 2008. p. 2221–2225.

[22] Blockchain-Agnostic LT. decentralized-identity/ion: DID Method implementation using the Sidetree protocol on top of Bitcoin. Leveraging The Blockchain-Agnostic; 2020. Available from: https://github.com/decentralized-identity/ion.

[23] Nakamoto S. Bitcoin: A Peer-toPeer Electronic Cash System. bitcoin.org; 2020. Available from: https://bitcoin.org/bitcoin.pdf.

[24] Microsoft. Decentralized Identity. Microsoft; 2018. Available from: http://aka.ms/DIDWhitePaper.

[25] Sayeed S, Marco-Gisbert H. Assessing blockchain consensus and security mechanisms against the 51% attack. Applied Sciences. 2019;9(9):1788.

[26] Semmouni MC, Nitaj A, Belkasmi M. Bitcoin security with post quantum cryptography. In: International Conference on Networked Systems. Springer; 2019. p. 281–288.

[27] (GDF) TGDF. The Gnutella Protocol Specification v0.41 Document Revision 1. The Gnutella Developer Forum (GDF); 2001. Available from: http://rfc-gnutella.sourceforge.net/developer/stable/.

[28] Hsieh HY, Sivakumar R. On using peer-to-peer communication in cellular wireless data networks. On using peer-to-peer communication in cellular wireless data networks - IEEE Journals; Magazine. 2004 Aug;Available from: https://ieeexplore.ieee.org/document/1261817.

[29] Li M, Tang H, Hussein AR, Wang X. A Sidechain-Based Decentralized Authentication Scheme via Optimized Two-Way Peg Protocol for Smart Community. IEEE; 2020. Available from: https://ieeexplore.ieee.org/document/9013015/.