Infnote: A Decentralized Information Sharing Platform Based on Blockchain

Anonymous Author(s)

ABSTRACT

Internet censorship has been implemented in many countries to prevent citizens from accessing information and to suppress discussion of specific topics and subjects ranging from politics, abuse of power, monopolies and so on. Circumvention technologies bring hope to this situation. This paper deep dives into censorship techniques used by regulators ranging from DNS and IP blocks to the latest techniques like Deep Packet Inspection, which can even detect flow of traffic. The paper also discusses and analyses various circumvention technologies that have been used in the past in different censorship environments: from no censorship to without Internet at all.

Our proposed solution, is a platform that helps eliminate the problem of sharing content in these censorship regimes. Our solution is named *Infnote* and it is a decentralized information sharing platform, based on Blockchain and peer to peer network, aimed to provide an easy-to-use medium for users to share their thoughts, insights and views freely without worrying about anonymity, data tampering and data loss. Infnote provides a solution that is able to work on any level of Internet censorship, even in area without Internet. Infnote utilizes multi-chain instead of the conventional single-chain approach to storing and accessing information on the chains, bringing implicit reputation system and ensuring the quality of contents in Infnote.

CCS CONCEPTS

• **Computer systems organization** → **Embedded systems**; *Redundancy*; Robotics; • **Networks** → Network reliability;

KEYWORDS

Web, Blockchain, Decentralization, Peer to Peer Network

ACM Reference Format:

Anonymous Author(s). 1997. Infnote: A Decentralized Information Sharing Platform Based on Blockchain. In *Proceedings of ACM Woodstock conference (WOODSTOCK'97),* Jennifer B. Sartor, Theo D'Hondt, and Wolfgang De Meuter (Eds.). ACM, New York, NY, USA, Article 4, 11 pages. https://doi.org/10.475/123_4

1 INTRODUCTION

Freedom of speech is considered a basic human right under Article 19 of the 'Universal Declaration of Human Rights '[8]. The evolution of the digital age has brought with it both opportunities and challenges for freedom of speech.

On the one hand, we can access or deliver information faster and more reliably. On the other hand, regulators can use both technical or non-technical methods to control or suppress what can be published or viewed on the Internet. While internet users could utilize circumvention technologies to bypass the Internet censorship to access or publish information. However, regulators around the world have significantly increased their efforts to control the information flow on social media, according to Freedom report [19].

The current Internet infrastructure model is heavily centralized. For example, there are only 13 logical root name servers for the Domain Name System (DNS). Another example is that Internet Protocol address (IP address) space is directly controlled by ICANN. These are apex players that control the Internet and are involved in delivering messages to the masses. This is being challenged by complex censorship techniques such as DNS/IP blocking and hacking attacks on content hosting websites (e.g.: blogs, social media platforms and more). Hence, there is an urgent need to solve the challenges related to this blockade by employing better circumventing approaches.

There are many techniques that allow users to circumvent Internet censorship. Methods based on proxy or virtual private network are commonly used. However, it relies on the connections to servers hosted in a country with less or no censorship and the connections may fail due to reasons like deep packet inspection or white-listing on gateway firewalls. Methods like cache webpages or Sneakernet allow a user to access information under any level of censorship, but they cannot guarantee validated information. Peer to peer network provide an approach to transfer data behind a firewall and there is no central node failure, making it difficult to block. ZeroNet [22] is a peer-to-peer web hosting project, based on the BitTorrent protocol, which can be used to circumvent Internet censorship, but the ZeroNet site owners have full control over the content on their website. As noted above, there are multiple methods to evade internet censorship, but they all come with their own pros and cons.

Bitcoin caught everyone's attention since it appeared as a whitepaper in 2008 [27]; various cryptocurrencies based on blockchain have emerged since then - some improving bitcoin and its flaws and others more innovative like Ethereum and Hyperledger. Today, the applications of blockchain and their respective Peer to Peer (P2P) network are designed to work beyond the decentralized currency function.

Blockchain, as an append only global ledger, has already been used in decentralized version of the Domain Name System (DNS) [25] and data storage. The append only ledger is ideal for information sharing platform aimed to circumvent Internet censorship, since no one would have authority to delete the content stored in the ledger. IPFS [10] and Blockstack [3] are two existing data storage platforms based on blockchain. However, IPFS currently

WOODSTOCK'97, July 1997, El Paso, Texas USA

ACM ISBN 123-4567-24-567/08/06.
https://doi.org/10.475/123_4

does not support publish-subscribe pattern, and Blockstack utilizes centralized cloud servers to store data [4].

Infnote, is our answer to the world and its citizens who are looking to share information without thinking about data tampering, data loss, and anonymity. It combines all the learnings and formulas that were developed by open-source architects to provide a reliable and effective solution as part of the anti-censorship movement. The name Infnote means providing **inf**inite power through the **note**s that the user publishes. Infnote will be helpful in providing a tool to content creators, social activists, journalists and others who simply want their voices to be heard.

Infnote, based on Blockchain and P2P technologies, is aimed to provide a platform for users to share their thoughts, insights and views through an easy-to-use medium even with varying levels of Internet censorship and even in areas with no internet. Infnote is a decentralized platform that can provide the user full anonymity (if required) and transparency, and last but not the least allow this content to travel and be viewed freely across the network of users. Unlike conventional blockchain, which uses single-chain mechanism to store information, Infnote uses multi-chain, which brings the implicit reputation system among chain owners and ensures the quality of contents in Infnote.

In this paper, we introduce the current situation of Internet censorship around the world and define the different levels of internet censorship mainly from a technical point of view in Section 2. Different circumvention technologies will then be compared in Section 3. Next, various popular consensus mechanisms are analyzed in Section 4. Finally, a detailed design and implementation of Infnote will be presented in Section 5.

2 INTERNET CENSORSHIP

The optics suggested by Thomson Friedman, the author of 'The World is Flat'can be applied to the dynamics that have been bought to this world by the phenomenon called 'internet'. One of the aptly-named 'flatteners'[17] in Friedman's book alluded to inventions like 'Netscape'and 'Web'digitizing the world and space we live in where everything could be viewed, manipulated, and written to through a screen (mobile phone, PC and so on). Another great thinker, Sir Tim Berners-Lee's vision when he invented 'www' was to create an "open platform that allows anyone to share information, access opportunities and collaborate across geographical boundaries" [29]

Unfortunately, the foresight of these thinkers and inventors is being challenged today by the political systems around the world. Information flow is being manipulated to show propaganda information, source of the real information is either blocked or redacted, and in many cases citizens are kept unaware of the happenings outside their borders. On one hand, we have the concept of 'open', 'decentralized', 'democratized 'internet and on the other hand we also have 'The Great Firewall of China', 'Halal Internet'from Iran, 'Kwangmyong'intranet from North Korea.

And this brings us to the topic of **censorship**, which this paper focuses heavily on, given that our proposed solution, Infnote is an attempt to solve issues related to censorship in countries like China, Iran, UAE, North Korea and so on.

2.1 Censorship Methods

Several authors have contributed to the growing literature surrounding internet censorship and the methods used by political systems in creating censors that block access to specific websites and content. Below, we highlight a few methods that countries use to block content.

DNS Manipulation or Tampering. In oppressive countries, if the government wants to censor websites, they can employ a technique called DNS manipulation/poisoning/tampering where in once a client requests for an IP address, the DNS server may send back a false IP address intentionally. This means that the client is actually visiting an incorrect website showing completely different or similar content to the user.

Domain and IP Address Blocking. One of the methods to block a user from visiting a website is to block both the domain and IP of websites. This completely deprives the user from visiting the website because in the TCP/IP world, a domain and its IP's is the primary way of reaching it.

Throttling. When an ISP starts to control the traffic and speed, it is known as bandwidth throttling. In some countries, this technique is being used during political events so that the word does not spread out to foreign sources [6] [7]. From a technical perspective, throttling is achieved by slowing down TCP either by dropping packets [16] or by controlling a bandwidth provided to a specific protocol.

Deep Packet Inspection or DPI. Deep packet inspection is another form of packet filtering that is being used heavily in certain countries for purposes of monitoring, blocking and sometimes throttling data flow through the Internet gateway systems. DPI filtering is used by Internet service providers to scan the payload of the Internet packets along with a normal scan of the headers to determine where to move the packet, how to classify and control it, and whether to drop it or not. If the ISP wishes to throttle the connection or drop it altogether, this is possible in real-time with the equipment that is available today.

Content and Keyword Filtering. Politically repressive countries pro-actively block foreign news websites, pornography, propaganda websites and content that do not match their political principles and philosophies. One easy way of censoring websites is based on their content, domain name, and specific keywords. Any website that matches specific criteria / filters, are automatically censored for violation of the government policies.

Distributed Denial of Service or DDoS. This type of censorship method has been used in the past to take down several websites that stand against the regime [26]. From a technical perspective, multiple computers on the network are controlled either deliberately or unwittingly and a coordinated series of traffic is sent to a target server or cluster of servers in the cloud. The traffic could be in the form of either ICMP, UDP packets, SYN flooding or a combination of this type of traffic that will exhaust and probably turn off the computers resources of the target.

2.2 Levels of Internet Censorship

There exists varying degrees of censorship and controls in countries around the world and this section divides Internet censorship into five broad levels from primarily a technical point of view:

Little or No Censorship. (Level 1) Little or no censorship is enforced in these countries. There is no need to use any circumvention technology since majority of the content is open to browse and access. Note: There are several countries who have placed censorship rules against illegal content like child pornography, in a genuine attempt to protect its citizens.

Selective Censorship. (Level 2) A small number of websites are blocked. Simple censorship methods, like IP address blocking or Domain name system (DNS) filtering and redirection are likely to be used. Most democratic countries fall under this category, where websites dealing with illegal or illicit activity may be blocked and freedom of speech is well protected by law systems. Citizens can easily use any circumvention technology to bypass the censorship.

Substantial Censorship. (Level 3) A large portion of content is being blocked. Several censorship methods are implemented simultaneously. A blacklist of IP addresses and domains is likely to be enforced by the firewall, filtering Internet traffic that goes through the border Internet gateway systems. Anti-censorship circumvention tools may also be targets of censorship, making it is extremely difficult for citizens to bypass the censorship.

Pervasive Censorship. (Level 4) In this category, a whitelist is enforced by the firewall, implying that only approved Internet traffic will be allowed to pass the firewall. This makes it theoretically impossible to use any proxy or Virtual Private Network (VPN) to bypass the censorship because the proxy server would not be in the whitelist.

No Internet. (Level 5) In extreme situations, the Internet service may be completely cut off. Any circumvention technology that relies on the Internet will not work. It is very hard for citizens to access or distribute digital information. Currently, most citizens in North Korea cannot access the Internet. During the Arab Spring, the Egyptian government shut down the Internet in Egypt temporarily.

3 CIRCUMVENTION TECHNOLOGIES

Given the censorship scenario and the evolution of such censorship systems on the web, there has been an uptake in the anti-censorship movement as well. Some circumvention technologies are simple, while others require more advanced knowledge of systems to implement and make it work.

Circumvention technologies have been divided into two broad categories for this paper, namely: **Accessing Information** and **Distributing Information**.

3.1 Accessing Information

Evading the firewalls and the censorship network infrastructure (implemented by Governments and Internet Service Providers to censor information) to *access* the information or the website is the ultimate goal with the technologies and techniques listed below:

Cached Pages. Search engines like Google or the Archive.org save or cache pages through its set of crawlers. Users can simply search for the webpage they are looking for and access the cached versions of these. This is an easy and quick way to access blocked and censored content. However, these websites and services can be blocked by censors as well, making this circumvention method effective only in Level 1 (little) and Level 2 (selective) censorship.

Proxy. A proxy server is a server that sits in between a client (requesting information, content, images etc.) and a server (that contains the information). A proxy server needs to be configured on the user's browser or application. A proxy can provide encryption and other forms of security to the user. The regulator can easily ban these proxy servers and hence this method is only effective in select countries with little or no censorship and selective censorship.

Virtual Private Network (VPN). Initially, VPN was being used to access the internal networks (e.g. office intranet) from the public Internet. Recently, we have seen rapid growths in deployment of VPN's [36]. VPN works by creating a virtual end-end connection through virtual tunnel protocols. By using a VPN, a user residing in a censorship regimes can access blocked content by setting up a secure connection to another country with no or only little Internet censorship.

Peer to Peer (P2P) Network. One widely used P2P network is BitTorent, which is a robust protocol for file-sharing that allows users to download content from multiple sources in a swarm that contains seeders, peers and leechers. From a technical perspective, BitTorrent breaks down a single file into several pieces or chunks of data. Peers then pull bits of files from seeders and/or other peers and once they have a hundred percent of the bits they have the entire file. [14]

Anonymous networks like Onion Router (Tor) [32] and Invisible Internet Project (I2P) [1] offer peer to peer communication through its censorship resistant and anonymous networks by relaying traffic through multiple nodes. They are often open-source and circuit based systems that encrypt the user's traffic end to end so that neither the sender nor the receiver need to reveal their respective IP addresses. It supports multiple applications like instant messaging, web browsing, file-sharing and so on. This type of anonymous network cannot function in a whitelist type of censorship.

3.2 Distributing Information

Platforms, protocols and technologies that are helpful in *distributing* information through the web are listed below:

Web-to-Email. This is a simple service that takes a snapshot of any website and sends it to directly to your email. In geographies with little to selective censorship, this method can be quite effective. One disadvantage being that if the emails and the website to access this service are being blocked on SMTP and IP respectively, this service will cease to exist for the user.

Sneakernets. The name 'Sneakernet'has come about because its easy to carry information in your sneakers and physically transport this digital information from one physical location to another thereby helping distribute information to other users or groups and circumvent surveillance and censorship. This method can work in countries with no Internet or has pervasive censorship deployed given its little reliance on Internet. The major drawback is that the source of this information, and the content itself cannot be fully trusted and this is a slower method of transport.

Peer to Peer (P2P) Network. To distribute or share content in a P2P network like BitTorrent, seeders can start to seed content and invite other users in the network to download this content bit by bit.

Archive Websites and Mirrors. Websites like Archive.org save multiple versions of a webpage and so the user can access the past versions of a specific website. Webpages that have been taken down or gone offline can be accessed via this service as well. It claims to have cached 338 billion web pages as of today.

In **Table 1**, the circumvention methods, technologies and techniques are compared to understand where and how these methods are effective when compared to each other.

3.3 Features of Circumvention Technologies

Difficulty Level - Identify and Block (Rating: Easy, Medium and Hard). This rating refers to how easy or difficult it is for the government, regulators, and Internet service providers to block the respective technique on the Internet. For example: To block foreign websites that provide news, the government has to block the IP Address or the Domain name of the website. This is considered an easy task for the government compared to the feature-set of sophisticated firewalls and backbone systems in their inventory.

Anonymity (Rating: Yes and No). This rating refers to whether anonymity is provided to users by the circumvention tool or method.

Data Tampering Protection (Rating: Yes and No). Data like files, html pages, music, video and so on can be tampered with once they are from their original source. Only in some cases, through algorithms like hashing, one can be assured that the data is from the original source and that it has not been changed or modified in any way shape or form. This feature provides a yes or no answer to the question: is data tampering protection provided using this technology.

Encryption (Rating: Yes and No). To prevent any unauthorized access, users can encrypt data with algorithms. This classification provides a simple yes or no to the question : can the data be encrypted while stored at the source and viewed by decrypting it.

Censorship Category (Rating: Level 1 to 5). Each circumvention technology has its limitations when it comes to deceiving the censors or simply finding another route to access or distribute content. The category level (1 to 5) at which the circumvention technology can be effective at, is mentioned through this feature. Below are categories that have been defined in the previous section. Each level is consecutive in nature and hence includes features of the previous level.

- Level 1: Little or No Censorship
- Level 2: Selective Censorship
- Level 3: Substantial Censorship
- Level 4: Pervasive Censorship
- Level 5: No Internet

3.4 Related Project: Peer to Peer Web Hosting

Infnote is a Peer to peer web hosting project, that uses peer to peer network to distribute and access webpages without the need for any intermediary hosting providers. This feature makes this technique very suitable for circumventing internet censorship . Here are some popular projects based on it.

Interplanetary file system (IPFS). IPFS is a distributed file storage system [10] that tries to combine the power of decentralization and the web by providing features like strong data integrity and availability [34]. Hashes of content would be stored within the blockchain [5]. It is a good platform built for evading censorships because the access to IPFS is difficult to block[2]. Currently, IPFS does not support publishâĂŞsubscribe pattern, which is necessary for real-time information sharing (ipfsâĂŹs pubsub is a experimental implementation for it).

FreeNet. FreeNet, similar to IPFS, uses a distributed data storage mechanism where the storage space is distributed amongst all nodes on the network. It is both anonymized and decentralized thereby making it difficult to take down in censorship driven countries [13]. However, unpopular files might disappear from the network [18], which does not fit the use case of discussion forum in which browsing old posts is needed.

ZeroNet. ZeroNet works on P2P and Decentralized web principles where sites are recognized through a public key unlike the traditional web where sites are recognized by IP addresses [22]. As long as a site is supported by peers, the content is alive and can be accessed by ZeroNet users. However, a site owner has full control over the content of the website, resulting too much power given to the site owner. Currently, the sites cannot directly access from the web browsers unless using the ZeroNet application, and the history of modification would not be stored.

Blockstack. Blockstack is a project that provide decentralized key/value storage, similar to Namecoin [25], built on top of the Bitcoin blockchain. It is a strong solution for deploying a decentralized public key infrastructure (PKI), as demonstrated in [3]. However, one disadvantage is that the values are stored in a centralized cloud storage system [4], which makes it less ideal for Internet censorship circumvention.

4 BLOCKCHAIN

Blockchain or distributed ledger is a technology that is set to change how we currently conduct business in any given sector, be it healthcare, finance, banking, retail, logistics and so on. Blockchain is an innovative and impenetrable stand-alone framework that was popularized by Satoshi Nakamoto's work [27] in 2008. Blockchain, in conventional terms, is a public ledger that records all events, transactions and exchanges that happen between parties or nodes in the network. [30]. Bitcoin popularized the concept of Blockchain, but Blockchain as a baseline platform has far greater implications than Bitcoin itself.

4.1 Advantages of Blockchain

Blockchain technology relies on a given consensus method (described in later section) to add information as blocks in the network. These blocks form the distributed ledger that is shared between

Table 1: Censorship Technologies and Techniques - Comparison Table

Type of Method	Cached Pages	Proxy	VPN	P2P	Sneakernets	Mirror Websites	Web-to-Email
Difficulty (blocking)	Easy	Medium	Medium	Hard	Hard	Easy	Easy
Anonymity	No	No	No	Yes	Yes	No	No
Data Tampering Protection	No	Yes	Yes	Yes	Yes	No	No
Encryption	No	Yes	Yes	Yes	Yes	No	Yes
Censorship (applies to)	Level 2	Level 3	Level 3	Level 4	Level 5	Level 2	Level 2

several computers or nodes within a network. Generally, these blockchains do not have a central authority controlling the information or creation of blocks. The above mechanism brings certain advantages to its users like visibility of data, immutability of data, and the ability to decentralize data over the network and leverage the security aspects of the Blockchain.

Data is Safe, Secure and Resilient. Data on blockchain is stored on a chain of blocks, which is then accessed by the users. Since blockchain is linked using cryptography, it is guaranteed that the information written cannot be tampered with, since it relies on digital signatures and the hashing function. Unless the entire network fails or the cryptographic function is attacked, the information on the blockchain is secure and tamper-proof. This makes the data, which is already encrypted, resilient to outside attacks, power losses to a subset of blocks and so on.

Open And Free-for-all. Not only is the data available for everyone to see on the platform, but anyone with access to the platform can contribute with their work for the users to see and appreciate. On a typical transaction-based Blockchain, it is possible to see each user's (through their respective public address) transaction history, currency exchanged and their holdings.

4.2 Consensus Mechanisms

In a blockchain system, the underlying assumption is that there is no centralized node and nodes generally do not trust each other. A consensus mechanism is a fault-tolerant mechanism to achieve necessary agreement on a single state over the network. In this section, we provide an introduction to some popular consensus mechanisms.

Proof of Work (POW). Proof of work is to solve mathematical puzzles and the answers can easily be verified. Bitcoin uses proof of work [27] to achieve consensus. The node that wishes to insert a block to the chain, is called a miner. The mining process is where a miner needs to scan a value that when hashed, the hash begins with enough number of zero bits. Other nodes can easily verify it by hashing a single value. After a miner produces a satisfying hash value, they have the permission to insert a block (with transactions) into the chain. The bitcoin mining process currently needs huge amount of computational resources as well as electricity to power these computers. The use of application-specific integrated circuit or ASIC can solve the mathematical puzzles much faster than CPU and GPU, in both speed and efficiency, making it is almost impossible for personal computers to join the mining process. In order to resist ASIC, many requirements will not only rely on computational power, but also other computational resources, such as memory and disk space. In POW, nodes having sufficient computational resources are more likely to solve the mathematical puzzles and therefore have a higher chance to insert blocks into the blockchain.

Proof of Stake (POS). Proof of stake states that a node needs to stake an amount of its token so it has chance to insert blocks into the chain. The more tokens a node stakes, the higher the chance of inserting blocks into the chain, because it is believed that more token a user has, the less likely he would attack the network [11].

Instead of competing using computational resources, in proof of stake, nodes compete based on the number of tokens they stake, therefore reducing the energy requirements. Similar to POW, the node with tokens (rather than computational resources) has a higher chance of inserting blocks into the chain.

Delegated Proof of Stake (DPOS). Unlike proof of stake, in which every node has chance to insert blocks to the chain, Delegated Proof of Stake only allows delegated nodes to insert blocks. Delegated nodes are chosen by voting processes. Votes are weighted according to the number of tokes each voter stakes. The first tier of nodes (usually less than 100 nodes) who receive most of the votes will earn the right to insert blocks into the chain.

Voters can acquire the tokens through Initial Coin Offering (ICO) or trade platforms. The price of the tokens is determined by the market. The value of these tokens comes with the power to vote for delegated nodes and deploy smart contracts on the chain. The more tokens a voter has in their possession, the more power they have to elect the delegated nodes and therefore have more indirect control over the chain. To attract votes, potential delegated nodes need to demonstrate their abilities and identities.

Proof of Authority (POA). In Proof of Authority (POA) network, only approved nodes can validate blocks and insert them into the chain. Unlike delegated nodes in DPOS mechanism, approved nodes are not chosen by voting.

Currently, POA is mainly used in private network, where every node knows each other and therefore trust approved nodes to maintain the chain. However, approved nodes have to maintain their computer uncompromised state given the power vested in them. Approved nodes need to gain reputation through their work on the network. However, any negative activity recorded, can destroy the reputation of the approved node as well.

Practical Byzantine Fault Tolerance (PBFT). Numerous protocols have been proposed to solve the problem of Byzantine Fault Tolerance (BFT)[23]. Practical Byzantine Fault Tolerance (PBFT) [12] is one of its solutions, which can handle up to 1/3 of the malicious nodes.

A block will be generated in a round. Each round can be divided into three phases: pre-prepared, prepared and commit. Each node has to receive 2/3 nodes from other nodes in order to enter the next phase [12]. Therefore, PBFT requires every node to be known to the network.

Delegated Byzantine Fault Tolerance (DBFT). Delegated Byzantine Fault Tolerance (DBFT) is another solution for the BFT problem. The whole process is similar to PBFT, except only a small number of delegated nodes are voted to insert blocks into the chain.

4.3 Consensus Mechanism Comparison

Different consensus mechanisms have different advantages and disadvantages. **Table 2** gives a comparison between them and we use the prototype give by [35].

Node identity management (Rating: Permissionless and Permissioned). In POW, POS and DPOS, everyone can download the code and participate in the network generating new blocks by only knowing a single peer in the network. In POA, PBFT and DBFT, only certain identifiable nodes can generate new blocks and each node needs to know the whole nodes list participating in consensus.

Latency (Rating: Low and High). Latency is the amount of time for a transaction to be confirmed and accepted in the network. The blockchain systems based on POW need multi-block confirmations, causing high latency [35]. Current implementations of POS either hybridize with POW or need checkpoints signed under the developer's private key, causing high latency. In DPOS, POA, PBFT and DBFT, the number of nodes participating in the consensus is small, leading to practical network-speed latencies.

Throughput (Rating: Limited and Excellent). Due to possibility of chain forks, POW has limited throughput [35]. Some of the implementations and variations based on POS outperform bitcoin when it comes to throughput but POS still has its limitations. EOS is based on DPOS consensus mechanism, that can support millions of transactions per second [9]. PBFT and DBFT can sustain tens of thousands of transactions [35]. As the throughput of POA is bounded by hardware not consensus, it has excellent throughput.

Energy Saving (Rating: Yes and No). Among all the consensus mechanisms, only POW needs huge amount of energy. Estimated annual electricity consumption for entire Bitcoin network currently is 73.12 TWh, about 30% annual consumption of Australia, as Oct 2018 [33].

Scalability (Rating: Limited and Excellent). Here we discuss the scalability in number of nodes the consensus mechanisms can support. All the consensus mechanisms have very good scalability except PBFT, which tested only small numbers of nodes [35].

4.4 Incentive of Blockchain System

Most public blockchain systems rely on its cryptocurrency to motivate its network and users. The cryptocurrency can be exchanged into fiat money through exchanging platforms. The blockchain system maintains a transaction ledger, where the balance of each account can be calculated. The cryptocurrency can be transfered to another account through a transaction, which will be written into the ledger. By generating new blocks, miners can receive transaction fees and block rewards. It is the cryptocurrency system which keeps most of public blockchain systems working well, since any misbehavior would cause the loss of cryptocurrency and therefore lose fiat money.

5 DESIGN AND IMPLEMENTATION

In this section, we discuss the design of Infnote from consensus mechanism choices, a platform designed from the ground up, protocol usage, architecture design and the overall technology it relies on.

Infnote is an information sharing platform that offers flexibility to its users. It can be used as a portal to share information, a blog to write short and long form articles, or even a discussion forum to discuss specific topics.

Infnote's architecture was designed with the above use cases in mind. On a high-level, Infnote uses blockchain and peer to peer (P2P) technology to achieve features like decentralization and immutability.

5.1 Blockchain

Infnote utilizes blockchain technology to store information. When a user wishes to publish a post on the platform, the post will be signed with the user's private key. Later, the post will be bundled with other posts and additional information (like timestamp etc.) together into a block. The *chain owner*, who has the authority to insert blocks into the blockchain, will sign the block with his private key. **Figure 1** demonstrates the process of posts inserted into the blockchain. The block will be broadcasted to the P2P network, and every node in the network will verify it.

Cryptography. We utilize a Digital Signature Scheme (DSS) implemented using *ECDSA* with *secp256r1* curve [20] and a cryptographic hash function *SHA-256* [28].

Multi-chain. Infnote supports a multi-chain structure, which means there are several independent parallel chains. Each chain is controlled by its chain owner and everyone can become a chain owner simply by creating a new chain. However, whether it will be maintained by enough number of nodes depends on the reputation of the chain owner and the quality of the information in the chain. The community of Infnote would maintain a default list of chains recommending the users to follow. With this mechanism, the chain owners are given incentive to follow the code of conduct. Any chain owner who violates general rules set by the community would be removed from the default list. However, a user can simply override the behaviors set by the default list, if the user disagrees with the community's decision. In this model, every participant only has limited power and the ultimate decision is made by the users themselves. **Figure 2** demonstrates the multi-chain structure of the Infnote.

Consensus mechanism. Infnote, as an information sharing platform, must ensure its quality of information, such as not allowing machines or bots to automatically send advertisements onto the platform. As Infnote does not include any currency system, sending advertisements to the platform is almost free. Unlike the automated verification process utilized during transactions in cryptocurrency

Table 2: Consensus Mechanism Comparison

Command	POW	POS	DPOS	POA	PBFT	DBFT
Node identity management	Permissionless	Permissionless	Permissionless	Permissioned	Permissioned	Permissioned
Latency	High	High	Low	Low	Low	Low
Throughput	Limited	Limited	Excellent	Excellent	Excellent	Excellent
Energy consumption	No	Yes	Yes	Yes	Yes	Yes
Scalability	Excellent	Excellent	Excellent	Excellent	Limited	Excellent

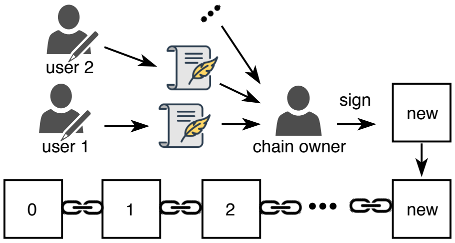


Figure 1: The process of publishing a post and inserting into the blockchain.

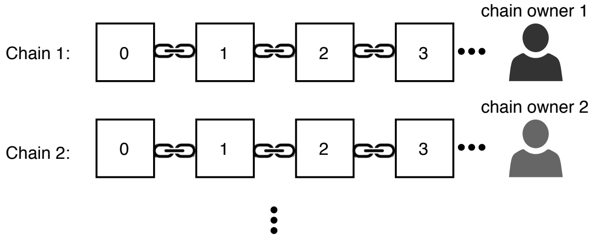


Figure 2: Multi-chain structure

systems, there are no standards defined to determine if a post should or should not be published to the Infnote platform. The two common consensus mechanisms POW and POS are not compatible with Infnote, since there is no guarantee who (which node) will generate the next block and insert it into the blockchain, therefore no guarantee what kind of posts will be published to the platform. PBFT does not allow many nodes to participate in the network, thus does not suit Infnote's requirements either. In DPOS and DBFT, only delegated nodes can insert blocks into the blockchain. However, when determining whether a post should be published to the platform or not, delegated nodes may have conflict making it harder to reach a consensus. This would cause a significant delay in writing information into the blockchain.

Infnote uses POA as its consensus mechanism. POA only allows authorized nodes to insert blocks into the blockchain. For a chain, only its chain owner can insert blocks into it and therefore control the information on the platform. Just like a miner in bitcoin, a chain owner's role is to generate new blocks signed with his private key and broadcast it to the nodes that are connected with him. Due to the append-only property of blockchain, the chain owner's power is limited. Once the chain owner decides to insert a block into the blockchain, it will be broadcasted to the peer to peer network, and thus become impossible to remove it from the blockchain. The chain owner can still soft delete a post by inserting another block to mark the deletion of that post, however the history will be permanently recorded in the blockchain and there is no way to remove it.

Fork. It is possible that a chain owner signs two blocks causing the blockchain to diverge into two paths, like a fork in bitcoin [24]. However, this is strictly prohibited in Infnote. If any node detects that two blocks of the same height are signed with the correct signature of the chain owner, it will stop trusting the chain owner and stop broadcasting its blocks. Without the support of peer to peer network, the chain owner cannot send the information out. In case of a scenario where the chain owner's private key is stolen, this provides a termination method for the chain owner to permanently close its blockchain.

Implicit reputation system. Unlike traditional information sharing platforms, like Facebook or Twitter, the identity of the owners is open to all. A chain owner can choose to hide its identity by using anonymous communication technology like Tor [31]. Each chain owner gains its reputation on the network by the work conducted so far. Even if a chain owner decides to hide its identity, the users are able to observe the chain owner's behavior in the blockchain and decide whether or not to use the services. Generally, it is expected that the higher the reputation, the more peers will join the network.

Incentive. The cryptocurrency based platforms are not an ideal solution against censorship. On the one hand, governments can easily spend money to buy the cryptocurrency and use them to destroy the blockchain system, by sending too many transactions (similar to DDoS) into the network, or by investing huge amounts of computational resource to become miners and therefore control the generation of new blocks. On the other hand, to gain entry into the platform, users must obtain or hold the cryptocurrency or token before using the platform, adding an additional barrier for entry for normal users.

With Infnote, our underlying assumption is that users should not be dependent on money or any other monetary factor. Instead, the unifying factor should be the sharing and bringing topics into discussion that will help build societies in a constructive way. Similar to the incentives driving the Linux Foundation and many other open-source communities around the world where the contributions are recognized and the work is shared amongst people as part of the community. We do not need a driver like cryptocurrency to run this free-for-all information sharing platform where freedom of speech is paramount and everyone's view is equally important.

5.2 Nodes

We fully expect multiple type of devices to join the network, i.e laptops, desktops, servers, smart-phones and so on. The front end interface can be through a web browser (e.g. safari, firefox, chrome etc.), a program or a smart-phone app. However, different devices have different capacities. It is necessary to analyze the features of each kind of device and design different strategies for them. In Infnote, there are two kinds of nodes, a full node and light node.

Full Node. For personal computers or servers, they can be full nodes. Same as bitcoin, full nodes are for devices that have sufficient bandwidth and computational resources to support all the functions of Infnote. Functions include: store all the data in blockchain, provide logic to view and publish content and act as a server by listening for connections and providing services to clients. People and organization can run full nodes by using their spare resources.

Light Node. Many devices, such as smart-phones or web browser cannot be full nodes, due to limited resources and processing power. Hence, they must rely on the full nodes to provide comprehensive services. At the same time, light node can use its limited resources to contribute to the system.

Smart-phones usually do not have enough storage space, therefore it is unreasonable for a smart-phone to store the whole data in the blockchain. It is also unlikely that a smart-phone will run the Infnote software for a long time. Most of the smart-phone platforms allow software to make the Internet connection, making it possible for a smart-phone to join the P2P network. For a smart-phone device, it can cache some recent blocks and broadcast them to peer to peer network. By caching recent blocks, the phone user is able to view the data stored in recent blocks.

Web browsers are restricted environments. A program written in *JavaScript* can be run in web browsers, but with more restrictions. We had two main issues that need resolving: storage and communication protocol. Before *HTML5*, application data must be stored in cookies, which would be sent to the server on every request. *Web storage* is a more secure method to store data locally, supporting larger data to be stored, and the data will never be sent to the server. Today, most web browsers would support web storage. Same as smart-phones, it is impossible for web browsers to store entire blockchain data, but by utilizing web storage, the program running on web browsers can cache some recent blocks and the user can view the information in those blocks. The communication protocol is strictly restricted in web browsers. As *UDP* and *TCP* protocols are not directly allowed in most of web browsers, Infnote uses *Websocket* as its communication protocol. Websocket is currently supported in most major browsers.

Multi-layer Structure.

Infnote introduces high level functions on top of the blockchain, that are supported by the multi-layer structure. A full node has three layers, while a light node may only have two layers. **Figure 3** shows a typical three layer structure. An arrow in the figure represents the direction of a data flow.

(1) Blockchain Layer: All nodes would have this layer. Blockchain layer serves two purposes: It stores all the data of the Infnote in sequence and provides the advantages of using the blockchain as described in earlier sections. A full node is expected to store all the blocks while a light node is only expected to cache a few recent blocks, due to limited storage space.

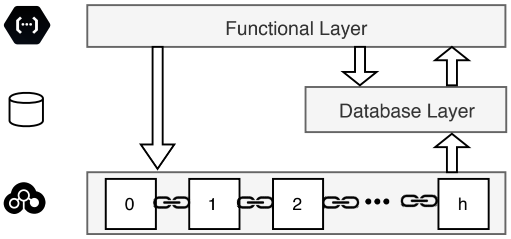


Figure 3: Three layer structure

(2) Database Layer: All full nodes would have this layer. It is necessary to reorganize the data into database, since relying only on sequence data, a full node cannot provide services efficiently. It is possible to let some of the light nodes, like smart-phones, run a database to improve the efficiency.
(3) Functional Layer: Both types of nodes would have this layer. This layer provides high level operations of the Infnote, like publishing or viewing an article in Infnote. The function layer verifies the operation based on the data in the database, but updates are applied to both the database and the blockchain accordingly. For a light node, since it is only expected to cache some recent blocks, the functions it can provide is limited.

5.3 Network

A network allows for nodes to interact with each other. In this section, we discuss the specific characteristics of our network.

Decentralized Network. The peer to peer network plays a vital role when developing the entire system and laying down the architecture. Same as Bitcoin, Infnote's architecture does not rely on a centralized server. For a censorship resistant platform, this is a necessary condition, since any single server would easily be blocked by censors.

Broadcasting Blocks. Similar to bitcoin, whenever chain owners generate a block or nodes receive a new block, they will immediately send out the new block(s) to the peers that they have direct connections with so that every node can obtain the new block in a short time. It follows the same principle as the publish-subscribe pattern, where publishers (chain owners) send blocks and subscribers receive these blocks. This feature allows nodes to automatically obtain new posts in Infnote in a short time.

Peer Discovery. Peer discovery is extremely crucial for a peer to peer network to circumvent Internet censorship. How to find the initial peer when a new node wants to participate in the network is a difficult task. For a pure decentralized peer to peer network, it seems that the only way is to search on the Internet and send a handshake message to millions of addresses hoping to find one peer who has already joined in the peer to peer network. In reality, it is not a practical method. The solution is to centralize, making the initial peer discovery the weakest link in the entire system. Authorities may simply block the initial seeds and thus prevent new nodes from joining the network. Infnote provides several methods for a node to initially find the peers in the network to relieve this issue.

- Hard-coded nodes: The developing community would hard code several recommended nodes around different geographies by indicating their addresses in the software. This method, however, may increase the workload of those nodes and they are likely to be blocked.
- DNS Seeding: DNS seeding servers would run a web crawler exploring the stable nodes in the peer to peer network and maintains a list of them. Whenever a node request is sent to a DNS server, it would return multiple node addresses. DNS protocol is a light protocol, therefore, it would not result in heavy workload for DNS seeding servers. However, the DNS seeding servers might also be blocked.
- From other nodes: Once a node joined the peer to peer network, the node can send requests to other nodes asking for more nodes' addresses.
- Address database: A node would store the addresses of nodes in its local database. On the next runtime, the node may not need to do the initial peer discovery given that nodes in the address database are still available.
- User specified address: The users can manually specify a node address into the software. The users can enter the address or simply scan a QR code. Although this method seems less efficient, it is the hardest for authorities to prevent initial peer discovery. This method allows users to join the P2P network by relying on real life connections, which seems like the only solution in pervasive censorship countries.

Obfuscation. Censors may use DPI to detect the protocol deeper inside the network packets. By using obfuscation, our goal is to avoid detection of Infnote packets. The Infnote currently uses two approaches:

- Mimicry: In this method, packet payloads are made to look like something that will be allowed by the DPI. A common example would be making the payloads look like HTTP packets, which are rarely blocked, because of its ubiquity [15]. Infnote directly uses WebSocket as its underlying communication protocol. Same as HTTP, WebSocket is a commonly used protocol, therefore should not be blocked by censors.
- Encryption: Similar to HTTPS, the WebSocket protocol supports encrypted connection, indicated by the prefix wss in URI. By using encrypted connection, the censor would not be able to obtain the content of the packets by intercepting network traffic.

Anonymity. Similar to other public blockchain platforms, everyone can download the blockchain and view the data stored in it. This feature makes it is possible to reveal the true identity of the users. In countries where substantial censorship rules exist, a user's identity may need to remain anonymous. If more nodes join the peer to peer network, the difficulty of finding out the owner of a post would increase.

For the users who need a higher level of anonymity, they can combine the Onion Router (Tor) [32] with Infnote. Once a user uses Tor to make a connection, the data packets will be relayed multiple times over distinct intermediary servers and each server only knows limited information of the packets, making it extremely difficult to trace back the source.

5.4 Modes

Infnote, depending on the need and requirements, can work in different modes. In essence, Infnote provides a solution for different degrees of Internet Censorship.

Direct Connect Mode. This mode is same as the traditional client-server architecture, in which the client is the requester while the server is the service provider. In a client-server model, the server will handle the requests and return the information to the client [21]. In Infnote, a full node could be a server, which can provide comprehensive functions. In an area where the server can be directly accessed by users, direct connect mode is the most efficient method. The server normally has powerful computing capacity and more network bandwidth, therefore, can support more clients and provide more functions. The user can easily connect to the server by using the HTTP protocol or HTTPS protocol which further encrypts the transmission. By using Tor, the user can even establish anonymous communication to servers [32].

Owing to all the data being stored in the Blockchain, the servers are able to provide comprehensive services based on the data in the Blockchain. Any node which has enough capacity can download the Blockchain and become a server to handle requests from the client. This feature enables Infnote's architecture to support multiple servers, making the system much more robust and censorship resistant. **Figure 4** demonstrates an example of multiple server handling requests from multiple clients.

Peer to Peer (P2P) Mode. In areas with substantial censorship level or above, a direct connection may not work. The common approach would be to use a proxy service. The proxy will relay the data sent from the users to the servers. This means users must use additional proxy services. If the whitelist type of Internet censorship (like in the case of pervasive censorship category stated in Section 2.2) is being implemented, the proxy server may not be allowed to access.

Infnote can utilize the P2P network to transfer or receive data. Every node in the P2P network would actively broadcast and receive blocks. Once a node receives a block, the node is able to extract and validate the data in the block. Similar to bitcoin, in which the user can send a transaction to any full node and it will be broadcasted to the whole network and written into the blockchain, the user can send their data into the P2P network and it would be permanently written into the blockchain later. **Figure 5** demonstrates a possible scenario of network structure in P2P mode.

No Internet Mode. In extreme situations, access to Internet may be cut off. However, users could still send blocks to others by connecting to the same internal network, or by physically spreading Blockchain files, like Sneakernets. Once the users receive the blocks, they would still be able to verify and therefore trust the data in the Blockchain.

6 EVALUATION

The evaluation focuses on two aspects of the system: throughput and latency. Throughput is how many posts the system can sustain

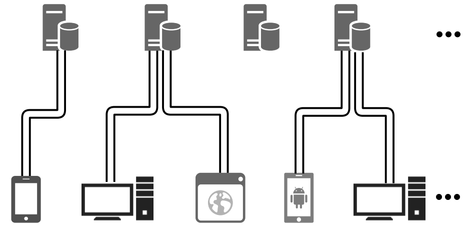


Figure 4: Direct Connect Mode with multiple servers

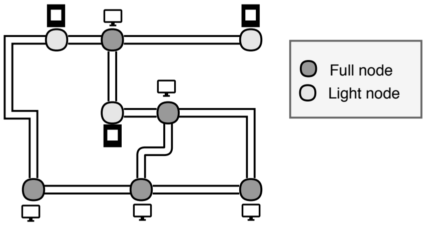
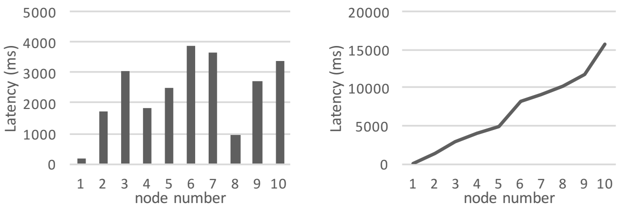


Figure 5: Peer to Peer Mode



(a) latency with different number of nodes (b) latency with different network diameters

Figure 6: Latency of the Infnote

per second. Latency is the amount of time for a block to be confirmed by all nodes on the P2P network. Here we assume that the size of a post in Infnote is 250 bytes, around the same size of a basic bitcoin transaction with 1 input and 2 outputs. We also assume that the size of a block is 1 megabytes.

For the throughput, we only calculate the time of validating block and saving it to the database, with the assumption that the network latency is less than validating and saving time, therefore transferring a block in P2P network would not influence the throughput of the system. Our experiment shows that throughput is 35544 posts per seconds, running the Python version of the Infnote program on a 2015 version of Macbook Pro. The result could be further improved by deploying better hardware or optimizing the code.

For the latency, we hope to understand the system in a global scale. As it is impossible to deploy a P2P network on a large scale due to limited resources, we try to speculate the performance of a system by using only a small number of nodes. We utilize ten nodes around different geographies in the world with eight full nodes and two light nodes [1]. Node No.6 is a light node running on iPhone 8 and node No.10 is another light node running a JavaScript program of Infnote on a Chrome web browser.

The first part is to evaluate the latency performance of the system up to ten nodes. It assumes that the connections between nodes in the P2P network have already been established. The **Figure 6(a)** shows the result of this experiment. On average, the latency is 2376.8ms, for every node on the network to receive a 1 megabyte block, which basic matches the network latency.

The second part is to understand the latency performance of the system on a large scale by using only a small number of nodes. We formed a linear topology network, in which each node is connected one after the other in a sequential chain. The experiment would represent the results of the large network with different network diameters. The first node in the chain is the chain owner who generates new blocks. We simulate the situation where two internal networks are connected only by the No.6 node, a smart-phone. The **Figure 6(b)** shows the result. It took 15665ms to transmit a 1 megabyte block on the network with 10 diameters.

7 CONCLUSIONS

In this paper, we began with an objective and defined Infnote's capabilities to meet these objectives which is to provide a platform to users around the world to share their views and opinions with an underlying assumption that the content shared will remain intact, unchanged and be protected.

We defined a few levels and types of censorship that can be used to put different types of countries into 'buckets' for comparison purposes. From our research, we understood that each country has its own ways of controlling the Internet traffic, political narrative online and sometimes even sabotage those actors that do not match the ruling government's political views and principles.

Once we defined the levels of censorship, we also compared and contrasted existing circumvention methods and technology that allow to bypass these censorship blockades. Each method was analyzed on the basis of effectiveness against each level of censorship and ranked accordingly.

Therefore, this presented us a unique opportunity to create a platform based on experience and learnings from existing ones that are decentralized, provide data tamper proofing, and allow to control one's own data.

Show summary of performance results of ours vs others here

REFERENCES

[1] 2003. I2P: The Invisible Internet Project. https://geti2p.net/
[2] Faten Adel Alabdulwahhab. 2018. Web 3.0: The Decentralized Web Blockchain networks and Protocol Innovation. In *2018 1st International Conference on Computer Applications & Information Security (ICCAIS)*. IEEE, 1–4.
[3] Muneeb Ali, Jude C Nelson, Ryan Shea, and Michael J Freedman. 2016. Blockstack: A Global Naming and Storage System Secured by Blockchains.. In *USENIX Annual Technical Conference*. 181–194.
[4] Muneeb Ali, Ryan Shea, Jude Nelson, and Michael J Freedman. 2017. Blockstack: A new decentralized internet. *Whitepaper, May* (2017).
[5] Muhammad Salek Ali, Koustabh Dolui, and Fabio Antonelli. 2017. IoT data privacy via blockchains and IPFS. In *Proceedings of the Seventh International Conference on the Internet of Things*. ACM, 14.
[6] Collin Anderson. 2013. Dimming the Internet: Detecting throttling as a mechanism of censorship in Iran. *arXiv preprint arXiv:1306.4361* (2013).
[7] Simurgh Aryan, Homa Aryan, and J Alex Halderman. 2013. Internet Censorship in Iran: A First Look.. In *FOCI*.

[1] The nodes are located in Tokyo (node No.1), Kuala Lumpur (node No.2), Sydney (node No.3), Singapore (node No.4), Mumbai (node No.5), Hong Kong (node No.6), Dubai (node No.7), Hong Kong (node No.8), Silion (node No.9), Hong Kong (node No.10)

[8] UN General Assembly. 1948. Universal declaration of human rights. *UN General Assembly* (1948).

[9] Leo Bach, Branko Mihaljević, and Martin Žagar. 2018. Comparative Analysis of Blockchain Consensus Algorithms. In *41st International Convention for Information and Communication Technology, Electronics and Microelectronics (MIPRO 2018)*.

[10] Juan Benet. 2014. IPFS-content addressed, versioned, P2P file system. *arXiv preprint arXiv:1407.3561* (2014).

[11] Iddo Bentov, Ariel Gabizon, and Alex Mizrahi. 2016. Cryptocurrencies without proof of work. In *International Conference on Financial Cryptography and Data Security*. Springer, 142–157.

[12] Miguel Castro, Barbara Liskov, et al. 1999. Practical Byzantine fault tolerance. In *OSDI*, Vol. 99. 173–186.

[13] Ian Clarke, Oskar Sandberg, Brandon Wiley, and Theodore W Hong. 2001. Freenet: A distributed anonymous information storage and retrieval system. In *Designing privacy enhancing technologies*. Springer, 46–66.

[14] Bram Cohen. 2008. The BitTorrent protocol specification.

[15] Lucas Dixon, Thomas Ristenpart, and Thomas Shrimpton. 2016. Network traffic obfuscation and automated internet censorship. *arXiv preprint arXiv:1605.04044* (2016).

[16] David Fifield. 2017. *Threat modeling and circumvention of Internet censorship*. Ph.D. Dissertation. UC Berkeley.

[17] Thomas L Friedman. 2005. *The world is flat: A brief history of the twenty-first century*. Macmillan.

[18] Ragib Hasan, Zahid Anwar, William Yurcik, Larry Brumbaugh, and Roy Campbell. 2005. A survey of peer-to-peer storage techniques for distributed file systems. In *Information Technology: Coding and Computing, 2005. ITCC 2005. International Conference on*, Vol. 2. IEEE, 205–213.

[19] Freedom House. 2017. Freedom on the Net 2017. (2017).

[20] Don Johnson, Alfred Menezes, and Scott Vanstone. 2001. The elliptic curve digital signature algorithm (ECDSA). *International journal of information security* 1, 1 (2001), 36–63.

[21] Channu Kambalyal. 2010. 3-tier architecture. *Retrieved On* 2 (2010).

[22] Tamas Kocsis. 2015. ZeroNet Project. https://zeronet.io

[23] Leslie Lamport, Robert Shostak, and Marshall Pease. 1982. The Byzantine generals problem. *ACM Transactions on Programming Languages and Systems (TOPLAS)* 4, 3 (1982), 382–401.

[24] Iuon-Chang Lin and Tzu-Chun Liao. 2017. A Survey of Blockchain Security Issues and Challenges. *IJ Network Security* 19, 5 (2017), 653–659.

[25] Andreas Loibl and J Naab. 2014. Namecoin. *namecoin. info* (2014).

[26] Bill Marczak, Nicholas Weaver, Jakub Dalek, Roya Ensafi, David Fifield, Sarah McKune, Arn Rey, John Scott-Railton, Ron Deibert, and Vern Paxson. 2015. An analysis of chinaâĂŹs âĂIJgreat cannonâĂİ. *FOCI. USENIX* (2015), 37.

[27] Satoshi Nakamoto. 2008. Bitcoin: A peer-to-peer electronic cash system. (2008).

[28] FIPS PUB. 2012. Secure hash standard (shs). *FIPS PUB 180* 4 (2012).

[29] O Solon. 2017. Tim Berners-Lee on the future of the web:âĂŸThe system is failingâĂŹ. *The Guardian* (2017).

[30] Alexandru Stanciu. 2017. Blockchain based distributed control system for edge computing. In *Control Systems and Computer Science (CSCS), 2017 21st International Conference on*. IEEE, 667–671.

[31] Paul Syverson, R Dingledine, and N Mathewson. 2004. Tor: The secondgeneration onion router. In *Usenix Security*.

[32] Inc The Tor Project. 2002. Tor Project. https://www.torproject.org

[33] Alex the Vries. 2014. Digiconomist. digiconomist.net

[34] David Vargas, Robert Tran, and Omar Gonzalez. [n. d.]. Censorship-Resistant File Storage. ([n. d.]).

[35] Marko Vukolić. 2015. The quest for scalable blockchain fabric: Proof-of-work vs. BFT replication. In *International Workshop on Open Problems in Network Security*. Springer, 112–125.

[36] Zhensheng Zhang, Ya-Qin Zhang, Xiaowen Chu, and Bo Li. 2004. An overview of virtual private network (VPN): IP VPN and optical VPN. *Photonic network communications* 7, 3 (2004), 213–225.