International Journal of Aquatic Science
ISSN: 2008-8019
Vol 12, Issue 02, 2021

Implementation Of Peer To Peer Social Media Network Using Block-Chain

Prof. Muneshwara M.S[1], Dr. Anil G.N[2], Mr. A R Sumukha[3], Mr. Arun R Shenoy[4], Mr. Kevin Biju[5], Mr. Charan Kalshetty[6]

[1]Assistant Professor , Department of Computer Science & Engineering, BMS Institute of Technology and Management, Bengaluru- 560064, Karnataka, India,
[2]Professor , Department of Computer Science & Engineering, BMS Institute of Technology and Management, Bengaluru- 560064, Karnataka, India,
[3,4,5,6] Department of Computer Science & Engineering, BMS Institute of Technology and Management, Bengaluru- 560064, Karnataka, India.

Email: muneshwarams@bmsit.in[1],anilgn@bmsit.in[2],sumuar1@gmail.com[3], arunshenoy99@gmail.com[4] , kevibk@gmail.com[5] ,charan1649kalshetty@gmail.com[6]

Abstract: The Rise of Social media has become quite influential in our lives. This rise has led to a huge influx of user data from all these Social media websites and mining of this user data by corporations owning these platforms have become rampant. The issue of preserving privacy & security of data will remain as long as these platforms keep monopolizing the Social media. Social media is also being used to spread false information and rumors. Large-scale spreading of false information could pose severe social and economic damages. This paper aims to eliminate privacy & security concerns of current social media platforms with the help of a peer-to-peer block chain network as well as limit the propagation of false information in the network. By decentralizing the entire social media network with the help of block chain the aim to make the network truly owned and monetized by the user. A network by the user for the user.

Keywords: Block chain, Decentralized, Social media, False information, Cryptography, Data privacy, Data Monetization.

1. INTRODUCTION

Social media has become a key influence in the lives of many people in the society today. It has provided a way for people to keep in touch with other people in the world and also a means to share informative and interesting content to our loved ones. It has also become a platform used by companies to engage with their customers. But one of the problems that has risen over time is the huge monopolization of data as seen bby huge Corporates that make millions off of unsuspecting user data.

The study models a Peer-to-Peer Social Media Network that makes use of block chain in anonymizing user data interactions. A Block chain is defined as a decentralized digital ledger of transactions that is distributed, most commonly public and can be kept track across multiple node computers. A blockchain is essentially a digital ledger that is distributed, publicly available and decentralized, used to keep track of transactions that take place through various computers. Any altercation of these transaction records cannot take place

without changing out all subsequent blocks therefore there is no central point of vulnerability that can be exploited in the network. A Security method that is widely used in blockchain is the public-key cryptography. The social media network is spread across thousands of devices working together instead of the traditional centralized database approach. Information privacy is the act of maintaining the relationship between the collection of data as well as the distribution/spreading of data. This may include conforming to the public expectation of privacy while also dealing with any legal issues related to dissemination of public data. Each post in the media can be tracked using the block chain where the user can track where the data they posted has travelled. Monetization can be achieved through user's posts on the social media. Here the user makes money in cryptocurrency for each view on his/her post. Hence in the end a decentralized social media platform is obtained that allows us to secure data and maintain privacy of the user as well.

The spreading of rumors that take place at a large scale can also lead to serious social and economic damages. As it is famously said false information spreads faster than any virus in the world. Each post made public on the social media will be analyzed by a machine learning algorithm running on each node to determine which parts of it are fake and which are real. Spreaders and posters of false information are penalized with a reduction in their credit score on the network making their posts less visible on the network.

2. LITERATURE REVIEW

Renita M. Murimi et al, investigates the various methods that have been developed by social media sites to allow users to create and share content. Various social network sites exist today and each of these sites offer different ways for users to express themselves anonymously or not and be part of a community of users around the world. This rise has led to the challenge of how content is disseminated and accessed fairly with the user's acknowledgement. Here the paper talks about a framework that the author has proposed to provide a method to generate value for user created content by making use of blockchain. It discusses how the framework can be used in various fields and challenges that may come overtime when implementing such a blockchain powered social media network. [1] Chao Li et al, looks into Steemit which is a reward based social media platform that allows users to share content and rank them based on votes of other users in the network. It provides cryptocurrency as a reward to users that contribute to popular posts in the network. Though on the outside it may seem that this reward system is used to promote users to produce higher quality, network transfer suggests that majority of the supply is taken away by bots which seem to suggest that there is a huge misuse and manipulation of Steemit platform's reward system. One core limitation this paper has is it does not take into account data lost due to web scraping. [4]

Sangeeta Kumari et al, looks into the major worries that do relate with respect to social media that include the security and the privacy of users. The important challenge is to protect user's personal data in accordance with the law and policies that have been instated with respect to data rights. There should not be any disclosure of any private data with respect to a user in the social media network. The core challenge with respect to privacy is usually seen in one direction because the privacy of a person is exploited not only on the outside but can also be caused due to negligence of the users themselves. These users are not aware of how data provided by them is made use of. Such methods and awareness are to be made for protecting user's personal data. There is also growing trend where the social networking website themselves sell the user data to advertisement companies to make up the money that they

invested. Many users don't realize this process and their data ends up with the vast range of advertising companies. These company can advertise selectively and can also build a monopoly for their products. By providing a reward system people have a sense of satisfaction as they are paid for the data they provide. They can also view which company acquires data and how they use it. This provides more transparency compared to current system. [3] Shalini Talwar et al, provides an insight on how sharing takes place as well as the motives for people to share. It also discusses these behaviors with association with a framework and third person effect hypothesis [6] Waseem Akram et al, discusses how social media has evolved and how it has affected us in different ways in our lives in positive and negative ways. It lists out all social media networks that are present today and give brief descriptions about it. It also discusses the effects it has in different ways in our lives. A limitation would be it being a rather less technical paper than usual and majority situations discussed in the paper are highly hypothetical with lack of study-based evidence. [5]

J. Zhang et al, discusses and studies fake news spread on various parameters and highly detailed as well. It looks into algorithms and methodologies that can be used for detecting fake news. It also introduces a fake news credibility inference model and compares its results with other existing models. [7] Monther Aldwairi et al, discusses fake news as well as brings up the concept of clickbait and how it has become prevalent in the current internet age. It proposes a solution to help dissipate this problem but the solution is too simple and hugely constrained to that dataset provided in the paper. [8] T S, Steni et al, presents a survey of fake news detection on the basis of community opinion to various user as well as the post itself with the help of machine learning algorithms. [10]

Yize Chen et al, looks to discuss how blockchain could be used to tame and control the spread of false information or rumours. It also discusses the possible new age of social media networks where blockchain can become a key technology in producing networks with trust and also provide a secure way of user data dissemination. [2] Freni, P. et al, talks about the issues with social media today with relation to user data security and talks about how blockchain can be used to resolve such issues. It also discusses a token-based system to provide incentives to provide a reward of sorts to users to gain an incentive for the data they share. [12]

Fran Casino looks to provide a review of how blockchain can be applied in developing solutions to problems that persist in various domains. It investigates how blockchain technology can be used to drastically improve today's practices and how it is used today. It also provides a pathway of how blockchain can be a good fit in each of the applications. [9] Rongen Zhang et al, surveys and discusses few cryptocurrency incentive schemes that exist in blockchain based solutions and have also compared these schemes. It also mainly discusses these schemes around Steemit that is a major blockchain based social network. [13] Blockchain has been a technology being talked about a lot these days. Alexander Pfeiffer et al, provides a brief look into how Blockchain technologies could be used to influence as well as create Social Networks and gives an overview of currently present platforms which facilitate Blockchain technologies and inculcate social media with it. [11] Raffaele Ciriello et al, looks into how blockchain has allowed us to enable and control regular social media practices. It shows how Steem offers various options of social media enabling as well as social media constraining option and how it can have contradictory effect on regular social networking practices. [14] Xie, Peng investigates how incentives in BSNs (Blockchain Social Network) and how the hierarchy of the users for sharing in the network effect the quality of information that is shared on the social network. It also discusses how higher visibility of a post makes it spill over to other cryptocurrency markets easier. [15]

3. EXISTING SYSTEM

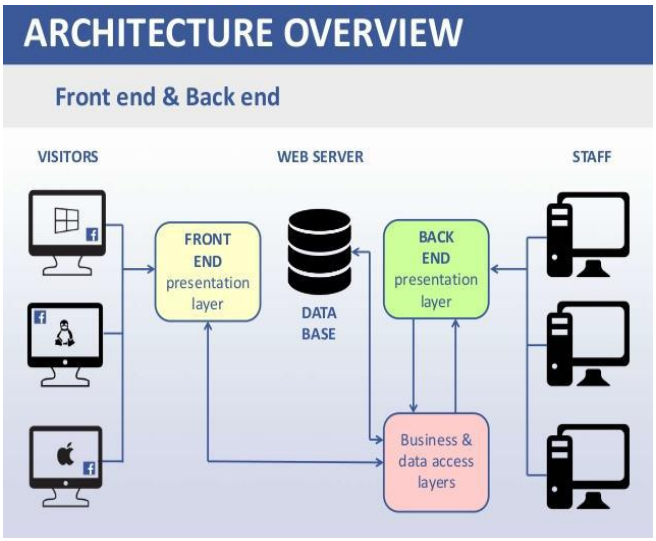

Fig. 1. Facebook Architecture Overview

Various social network sites exist today and each of these sites offer different ways for users to express themselves anonymously or not and be part of a community of users around the world.

As seen above there is a traditional three layer centralized architecture being used by many of the current social media networks. This is a typical client/server architecture where the server exposes certain endpoints through which the user can access the resources provided by the server. Here the server is in control of the resources received by the client. Since the servers are owned by corporations whose main goal is to make profit, there is a high chance of user data being used for targeted advertising, selling bulk data for profits etc.

Limitations

- The media network is run by servers owned by large corporations and not by the user themselves.
- There are hundreds of cyber-attacks each day on the centralized servers which can leak huge amounts of user data and hence cause harm.
- User data is misused by large corporations where the data is used to predict user behaviour and modify it for making profits. The biggest scandal that had come to the public's attention was the Facebook–Cambridge Analytica data scandal. This was a scandal that emerged where personal data of many Facebook users were found to be collected without any permission by the Cambridge Analytica and was mainly used for political promotion and advertising. This is known to be one of the largest leaks of Facebook history.
- A lot of web scraping happens where user data is stolen and used illegally to train machine learning models.
- The user who drives the social media does not make any profit whereas the companies hosting the media make billions of dollars in revenue selling and using data.
- There is no sufficient mechanism to find and reduce false information such as false news, rumours etc. in the network.

4. SYSTEM DESIGN

Block chain design

A Block-chain is a distributed and decentralized digital ledger which is most often made public. It is used to keep track transactions that happen across many computers. Since the Block-chain is expanded in such a manner any record that is considered part of the network cannot be altered, without making those alterations to all subsequent blocks.

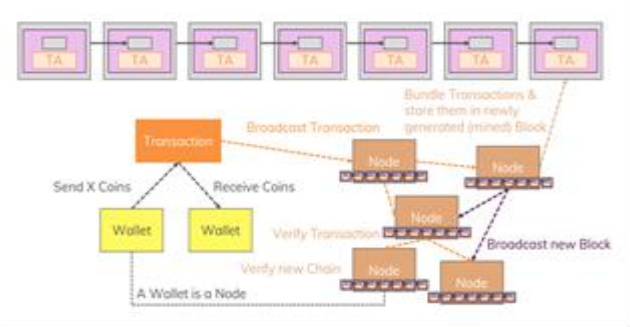

Fig. 2. Block chain Architecture

Block-chain security methods make use of public key cryptic techniques. A public key is a string of random letters and numbers and acts as an address for the nodes on the block-chain. Value tokens sent across the network are recorded as belonging to that address. A private key is acts as a counterpart to a corresponding public key and is used by the owner to access his/her confidential data. RSA is extensively used in the block chain to create a wallet for each user and is described below.

Encryption/Decryption used by the block-chain

Fig. 3. Working of RSA

RSA also known as Rivest Shamir Adleman algorithm is a public key encryption/decryption algorithm and is widely used in networking to transfer data from one place to another securely. The public key also known as the encryption key is available for all users to view and is different from the private key which is also known as the decryption key is kept as a secret. Messages that are encrypted using the public key can only be decrypted with the help of a private key. It is almost impossible to find a private key given a public key and this is at the heart of the algorithm's security. RSA has one disadvantage that it is slow and is hence used to transfer symmetric keys which can then be used for transfer later. Here, RSA can be used to generate public-private key pairs that are uniquely used to identify users in the network. Each public-private key pair is also known as a wallet and hence a user can be

identified in terms of his wallet. The block-chain stores transactions that have occurred between two wallets. Whenever a user views an anonymized and monetized post in his feed there will be a transaction from the latter to the person who posted the content. This is the incentive system used in our model. This is described in [5.3].

Incentives/ Gains for the user

The rate awarded to the user on receiving hits on his/her post will be decided by an algorithm that takes the following factors into account such as exposure of the post, amount of false information in it and credit score of the user (determined by the network based on his past activity). It can be as simple as a linear combination of a variety of factors.

Incentive earned = w0 + (w1 * no_of_views) + (w2 * credit_score)

w0 is the base incentive earned on creating a post on the social media and can range from 0 to 100. w1 and w2 are weighted values that determine the contribution of each of the factors into the incentive earned. For example, business accounts may have a higher value for w1 than w2.

credit_score is a real valued number and is described in more detail in [5.4].

Credit Scores

Credit scores are a reflection of the user behaviour on the network. Users who share malicious content or engage in practices that violate the policies of the social media network will be penalized leading to a reduction in their credit scores.

Reduction in credit score = (no of reports for malicious content * a)

a is a variable and can vary in order to reduce or increase the reduction of credit score.

For example say a user X as 100 credits and engages in the malicious practices which are the reported by 3 users. Then the reduction in credit score would be 3a. Let us say that the post was from an account that has a history of penalties, in this case the value of a would be high (say 10) and hence the user's credit score would reduce to 70.

A user can make use of the false information detection mechanism which can aid him in deciding whether he must report a post or not. Detection of false information is done as described in [5.5].

4. ARCHITECTURAL DESIGN

The proposed architecture can be viewed as a stack of 3 layers each equally important for the proper functioning of the entire system

The block-chain layer: Where the transactions of the user data are stored in a systematic matter where data manipulation can be easily detected.

The peer nodes layer: Where each peer node has a copy of the block-chain, these nodes run the machine learning algorithms and communicate with each other to share block data.

The user layer: Also known as the wallet in block chain jargon, this layer provides both the user interface which communicates with the peer nodes. The user can either use the services of another peer node or be a peer node himself

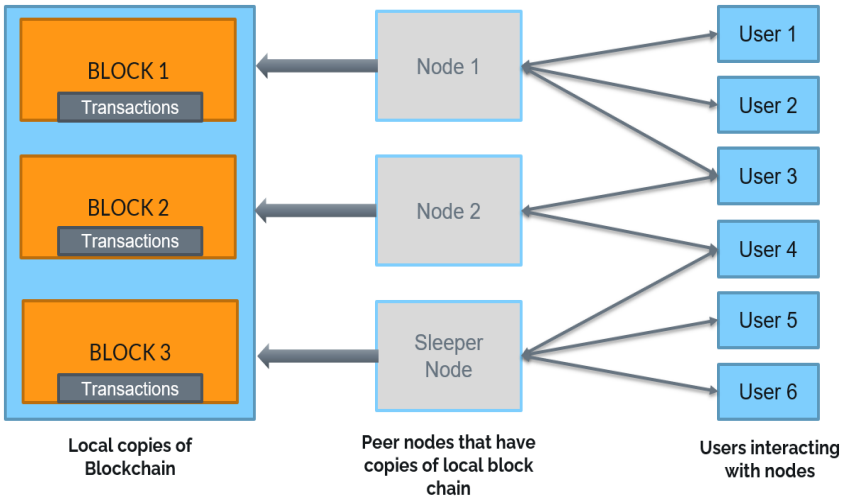

Fig. 4. Proposed System Architecture

The user can create a post and choose to anonymize and track the data (basically sell it for a profit in crypto). The block chain in itself has its own currency which is awarded to the user whenever there are views on his post. If a user chooses to anonymize the post, the node will then remove his name and add his/her public key to the post. Whenever the post is viewed by someone there will be a transaction on the block chain from the viewer's public key to the public key of the original poster which cannot be traced back to the latter (the original posters name will not be known only his public key). The original poster will be the only one who can track the transactions on the block chain using his/her private key and hence know exactly who has his/her data. Hence the user can post his information on the media and track where it travels while making money out of it. If the user feels his data is being used illegally, he can immediately report the transaction and the corresponding holder's account will be put under inspection on the peer to peer network which can reduce his credit score.

The currency gained by the user from the post can be used by the user to view further posts or buy data from other posts on the social media network which are also recorded as transactions on the block chain.

The rate awarded to the user will be decided by an algorithm that takes the following factors into account such as Exposure of the post, Amount of false information in it, credit score of the user (determined by the network based on his past posts).

For the detection of false information in the evaluation of the rate an analysis of some of the major reasons for the spread of false information which are: false news is more inclined towards a user's beliefs, peer approval and political parties pushing their agenda. Posts made public on the social media will be analysed by a natural language processing algorithm running on each node to determine parts of it that are real and parts that are fake. The viewer of a post can click fact checker button and parts fake will be highlighted in red while parts that seem real will be highlighted in green. The original poster will then be penalized on his credit score and any subsequent viewers who want to share it will be given a warning to not share such a post. If they continue to do so they too will also receive a penalty on their credit score.

Users with lower credit scores will have their post visibility limited and eventually phased out of the network. All these processes described of course require compute and storage in order for the entire system to function in harmony. The network will never run out

of space or compute as nodes communicate with each other and find more nodes to share load with.

A machine learning regression algorithm predicts the requirement in storage that may arise each day and hence the system is always prepared to face expansion on its own. There will be nodes called sleeper nodes on the system which will be activated and deactivated based on the needs of the network.

5. COMPONENT DESIGN

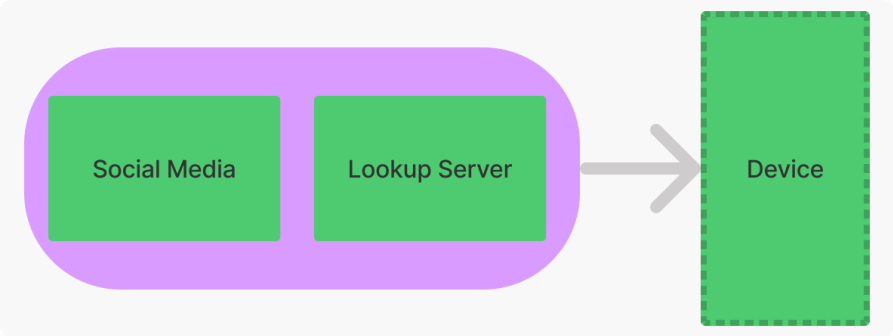

Fig 5: Design of each node of the Network

Blockchain nodes are the core component of the social media network. Each node that is connected in the Block-chain consists of 2 core components as shown in Fig 5.

1. Social Media

The social media server is responsible for the following:

- Deliver the user interface.
- Handle requests from the user interface. These requests can be of various types. i.e., login/signup, data retrieval and uploads.
- The data is stored in a NoSQL database under the collection social-media.
- The distributed database appears as a single central database to the social media server that it can query.
- The social media server has no idea about other peers and functions like it is the only server that is active in the environment

2. Lookup Server

The lookup server lies at the heart of each peer node and is responsible for establishing communication between the nodes, it also performs the following functions:

- **Register itself with the network:** The lookup server registers itself with the master lookup server via a register request. The master lookup returns the set of peers from its database. The lookup server then stores these peers in its database and then registers itself with each of the peers received via a register request.
- **Send a heartbeat when it comes online after a failure:** The node may have gone down for many reasons and hence would have lost connectivity with the network. Any peer connected to the node going offline will mark this node as an inactive node in its database. It is the responsibility of the lookup server to send a heartbeat to all the peers in its database when it comes back online.

- **Retrieve data not present locally:** Whenever the social media receives a request for data not present in its local database, the lookup server takes the responsibility to locate the data. It requests each peer node in its database for the data, if the node has the data it returns an acknowledgement.

  The lookup server then retrieves this data from the address of that node. Once the data is retrieved from a node in the network the lookup server places this data in the social media collection of the database. This data is then retrieved from the collection by the social media server and given to the user interface for display.

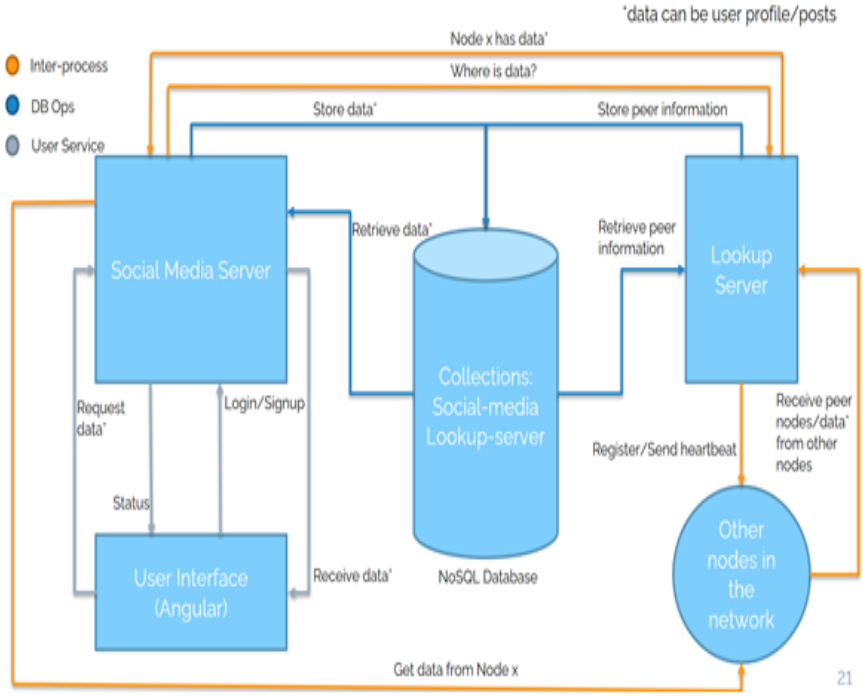

Fig 6: Dataflow Diagram of each peer in the network

6. RESULT AND DISCUSSION

The big picture here is that this is a study on a peer-to-peer social network that is decentralized with the help of block-chain. Since block-chain is a distributed digital ledger and is most often available to the public, any record in the network (in our case transactions of data) cannot be modified or altered without making those modifications to subsequent blocks. This significantly reduces the risk of cyber-attacks which results in a more secure system.

Data Anonymization & Privacy

Data privacy and anonymity is a key outcome. The users are able to completely isolate their identity from any post. By this the other users on the platform would be able to view the post but would not be able to trace the data to any particular user.

Monetization & Security

Users can make profit out of their posts as by choosing the user data to be monetized, so whenever any user views a post or gets access to data of another user, a transaction takes place in which the viewer is pays a certain amount in crypto which is then credited to the owner of that particular data. The parties in this transaction cannot be traced as the

transaction only records the public keys of both the parties. In a way this ensures that the viewer is paying the owner of the data to get access of the data.

Curbing of False Information

Spread of false information is significantly reduced with the help of a credit score system for every user. For any reported post or comment of a user, a check is conducted if the post is an opinion or statement. If it is a statement, a false information check would be done, and on a positive check the user's credit score would go down. If the user avoids such events the score is increased based on his/her activity.

Scalability

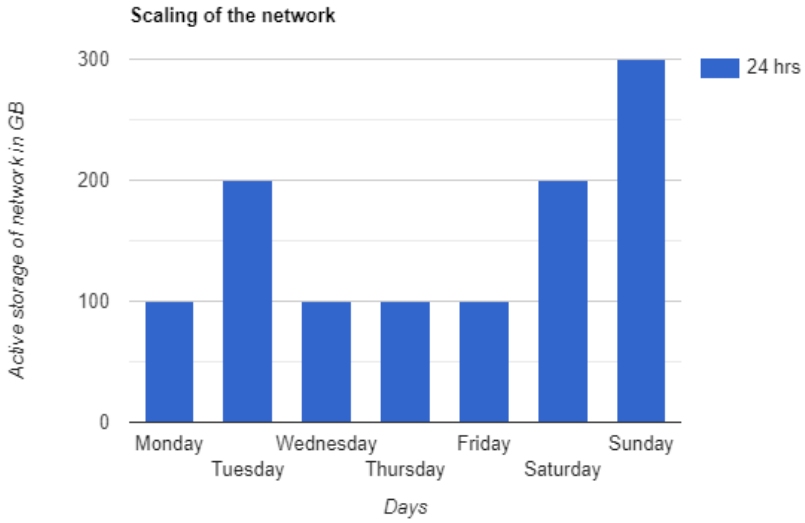

Fig. 7. Graph showing scalability in the storage space

The graph shown in Fig. 11. shows a trend in the storage of the network over a period of one week. Tuesday being a public holiday sees a large increase in the activation of sleeper nodes for more storage as there are a lot of posts coming into the network. This then reduces to 50 GB on the next 3 days. Note: This does not mean that the remaining 150 GB of data was lost, it simply means that the nodes containing this data were never activated nor did they get any requests for this data as the user or any of their viewers were never online. Again on the weekends where there is an increase in the traffic the sleeper nodes are activated in order to store/retrieve the data.

7. CONCLUSION

With the help of a block chain, a network that is self-scaling and distributed was implemented. The system also provides a model to limit the propagation of false information. The users on the network are rewarded for sharing their data through the built-in currency. Data privacy was achieved by replacing user identities with public keys for the transactions on the block chain. The outcome of building a distributed social network that scales itself based on need and usage, a network owned by the user and not by any corporation and data privacy and security by isolating user data from user identity was also achieved. A mechanism that allows users to monetize their data on the network was explored and implemented. Spreading of false information on the network was reduced drastically by assigning credit scores to each user on the network.

8. REFERENCES

[1] R. M. Murimi, “A Blockchain Enhanced Framework for Social Networking”, ledger, vol. 4, 2019.
[2] Yize Chen, Quanlai Li, Hao Wang, “Towards Trusted Social Networks with Block chain Technology”, Symposium on Foundations and Applications of Blockchain Proceedings, University of Southern California, 2018.
[3] Sangeeta Kumari, Shailendra Singh, “A Critical Analysis of Privacy and Security on Social Media”, Fifth International Conference on Communication Systems and Network Technologies, 2015.
[4] Chao Li, Balaji Palanisamy, “Incentivized Blockchain-based Social Media Platforms: A Case Study of Steemit”, 145-154. 10.1145/3292522.3326041, 2019.
[5] Akram, Waseem. (2018). A Study on Positive and Negative Effects of Social Media on Society. International Journal of Computer Sciences and Engineering. 5. 10.26438/ijcse/v5i10.351354.
[6] Talwar, Shalini & Dhir, Amandeep & Singh, Dilraj & Virk, Gurnam & Salo, Jari. (2020). Sharing of fake news on social media: Application of the honeycomb framework and the third-person effect hypothesis. Journal of Retailing and Consumer Services. 57. 102197. 10.1016/j.jretconser.2020.102197.
[7] J. Zhang, B. Dong and P. S. Yu, "FakeDetector: Effective Fake News Detection with Deep Diffusive Neural Network," 2020 IEEE 36th International Conference on Data Engineering (ICDE), 2020, pp. 1826-1829, doi: 10.1109/ICDE48307.2020.00180.
[8] Aldwairi, Monther & Alwahedi, Ali. (2018). Detecting Fake News in Social Media Networks. Procedia Computer Science. 141. 215-222. 10.1016/j.procs.2018.10.171.
[9] Fran Casino, Thomas K. Dasaklis, Constantinos Patsakis, “A systematic literature review of blockchain-based applications: Current status, classification and open issues”, Telematics and Informatics, Volume 36, 2019.
[10] T S, Steni & P S, SREEJA. (2020). Fake News Detection on Social Media-A Review. Test Engineering and Management. 83. 12997-13003.
[11] Pfeiffer, Alexander & Kriglstein, Simone & Wernbacher, Thomas & Bezzina, Stephen. (2020). Blockchain Technologies and Social Media: A Snapshot. 10.34190/ESM.20.073.
[12] Freni, P. & Ferro, E. & Ceci, G.. (2020). Fixing Social Media with the Blockchain. 175-180. 10.1145/3411170.3411246.
[13] Zhang, Rongen & Park, Junyoung & Ciriello, Raffaele. (2019). The Differential Effects of Cryptocurrency Incentives in Blockchain Social Networks.
[14] Ciriello, Raffaele & Beck, Roman & Thatcher, Jason. (2018). The Paradoxical Effects of Blockchain Technology on Social Networking Practices.
[15] Xie, Peng and Xie, Peng (2020) "The Effect of Incentive Hierarchy System of Social Media In the Delivery of Quality Information," Journal of International Technology and Information Management: Vol. 28 : Iss. 4 , Article 1.