International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 10 Issue 4 ǁ 2022 ǁ PP. 53-57

A review on Decentralized Online Social Network

Uvais Mon V V N [1], Sumukh R [2], Vignesh V [3], Zabiulla Sheriff [4], Yashpal Gupta S [5]
Department of Information Science and Engineering
Vidyavardhaka College of Engineering, Mysuru, India

Abstract: *Today, Online Social Networking is the main cause of targeted marketing. All major OSN providers are free to use. To generate revenue, they mine information about users' interests and sell it to potential advertisers. Due to this, users are becoming reluctant to use existing giant OSN networks. A Decentralized Online Social Network(DOSN) can resolve this issue. A DOSN ensure data privacy. Users will own the entire data they generated. They can decide where to store it and with whom to share it. Here, we are reviewing some of the existing approaches to decentralize the way how internet users socialize digitally. We review different approaches to identify the most optimal implementation technique for a DOSN. Later in the paper, we discuss our future work to improve the existing DOSN concepts.*

Keywords: *Online Social Network, peer-2-peer networking, Ethereum, InterPlanetary File System.*

Date of Submission: 06-04-2022 Date of acceptance: 22-04-2022

I. Introduction

Online Social Network(OSN) is a continuously growing network of users interconnected with each other. Today, there are over 4.5 billion social media users worldwide. They keep on generating data everyday. They socialize with their connections by sharing posts containing texts, images, audio, or other files, and liking, commenting, or sharing these posts. All these data are stored on a centralized data store. These data are owned by a central authority. The data generated by the users no longer belongs to the user. They will have to trust the moral ethics of the central authority with the privacy of their data.

Most of the 3rd party OSNs analyze the data generated by the users to understand more about the user. This information about the user is then used to deliver targeted advertisements. Users no longer feel their data is private at a centralized data server. These 3rd party OSNs sell user data to advertising companies. Though they sell only metadata of a user profile to the advertisers, which does not cause any direct harm to the users, it is still a violation of user privacy. These metadata can be used to understand the interests of a user which can be exploited by the advertising company.

To overcome the issue of privacy at a centralized OSN, a better concept of OSN that runs as a decentralized network called Decentralized Online Social Network (DOSN) is introduced. Here, every user is a participating node of a decentralized network. Each user will have complete control over the data they generate. Each node establishes a peer to peer connection with its peer nodes for data distribution. Different techniques of data distributions and encryptions are reviewed later in the paper.

Many peer to peer networks exist today, such as torrents. These networks are used mainly for file sharing. Same kind of protocol with a few tweaks for the privacy policy could be used to establish connection and share user data in a DOSN. The main concern in a DOSN is access control management. In a centralized OSN, authorization of a user to access the data of another user is checked by the centralized server. In a DOSN, Access Control Management is achieved through distribution and encryption policies.

In a DOSN, users can decide who can view their data. Distribution policy may be either set at the network level or at the node level. If it is set under the network level, every node will have to follow the same policy. If it is set under node level, each node can decide a policy for the distribution of its data. The distribution policy only determines the availability of the data. Privacy of the data is guaranteed regardless of the policy chosen.

In this paper, we review a few of the existing DOSNs, their proof of concept discussed, implementation details, features and shortcomings, if any.

II. Literature Review

A. Peer to peer

Peer to peer(P2P) is the interconnection of peers connected to a network. Peers can communicate with each other without any direct involvement of a 3rd party server. This enables users to interact with their peers more securely.

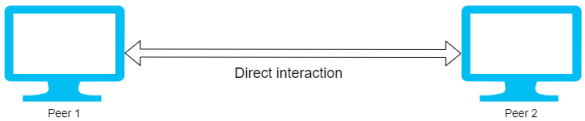

Fig. 1: Peer to peer communication

In [1] Leila Bahri, et al. compared centralized OSNs and DOSNs. They talk about online privacy and offline privacy. Online privacy refers to providing access control management. Offline privacy refers to protection against metadata analysis, mining user information and targeted marketing. They claimed centralized OSNs inherently provide online privacy, while DOSNs provide offline privacy. They further proposed 2 ways to achieve decentralization. One way is to have different individual federated servers. Users can then join any one such federated server, and migrate over different servers without losing any user data. The other approach is to use peer to peer communication to make it truly decentralized. They propose distributing data over nodes that are already authorized to access the data. They conclude the paper by highlighting that though DOSNs can solve some of the privacy issues caused by centralized OSNs, it opens up new issues and technical challenges, like instant messaging and real time data sharing.

In [2] Nashid Shahriar, et al. have talked about data availability and replication in a P2P based OSN. They proposed a mathematical approach to calculate the beta availability of a group of nodes. Beta availability is the probability that beta number of nodes will be up and running within a group of nodes at any given time. The main goal of this approach is to ensure maximum availability with minimum replication overhead. They proposed a structured approach for Diurnal Availability by Temporal Assemblage (S-DATA) to achieve this. S-DATA uses a Distributed Hash Table(DHT) protocol called plexus protocol to construct globally optimized availability groups. Using the S-DATA protocol and beta availability, they were able to formulate a mathematical model that informs about the node and content availability in a given P2P group. Using this model, the network is able to group nodes, identify most available and dependable nodes in the network, calculate how many replicas of the data has to be created, and on which all nodes these replicas have to be stored. Their experimental simulation showed a promising solution to the content availability problem in a P2P network.

In [3] Giuliano Mega, et.al have given the disadvantages of centralized social media systems such as data security and proposed decentralized social media using the p2p .The users can run p2p on the local server to view contents posted by their friends and they can also post their contents.The p2p uses distributed hash tables.The alternative is also proposed in the paper that is friend 2 friend the disadvantage of this is data transfer can take place only if the owners know each other. The authors also give a brief about the three main principles of the protocol i.e use of message histories,anticentality selection heuristic and fragmentation awareness.They also proposes how can the current proposed system can be improved by including more realistic workload ,evaluation of the anti-entropy mechanism and an integrated analysis encompassing overlay maintenance.

In [4] Anandhakumar Palanisamy, et al. compared centralized and decentralized social media networks. They talk about privacy and trust related issues of centralized social media platforms.They present the ARTICONF approach to a car-sharing use case application, as a new collaborative peer-to-peer model providing an alternative to private car ownership. ARTICONF addresses issues such as privacy,trust and time criticality to fulfill the privacy, robustness, and autonomy related promises that proprietary social media platforms have failed to deliver so far. The technologies that are used are: Peer to peer network, hyperledger fabric, TIC, Tac. The main goals are to create an open and agile social media ecosystem, detect interest groups and communities, to autoscale time-critical social media applications and enhance monetary inclusion in collaborative models through cognitive and interactive visualization.

B. IPFS

InterPlanetary File System is a network of storage nodes. It leverages the concept of P2P communication to enable the storage of large data files on a decentralized network.

In [5] Quanqing Xu, et al. talks about how the decentralized system can improve the online social media by proposing a system using Ethereum and IPFS.Ethereum is an open source blockchain which provides runtime environment for smart contracts and this is known as EVM. IPFS is used to store immutable data and remove duplicate data.The advantage of using this is that it provides security and also make the system available

during server downs and cannot be censored by anyone.The user need a separate individual contract for registration of his/her accounts and this user address will be sent to account manager for the record purpose.The proposed system has three main components a backend private blockchain,IPFS for storage and a frontend UI for user interaction.The disadvantage of proposed system is that the current smart contract does not provide return type for complex data structures.

In [6] Van-Duy Pham, et al. compared centralized and decentralized storage systems. The main focus is on the disadvantages of the centralized systems such as single point failure and privacy concerns. A decentralized storage system is proposed by the author to eliminate the disadvantages of the centralized system, the new proposed system is aimed at being secure and transparent. To present the secure and transparent characteristics of the decentralized system they have used a combination of IPFS, ABE, MA-ABE and Ethereum blockchain . Two use cases are shown in regard to the decentralized system and there is ongoing research to enhance the security features.

In [7] Koushik Bhargav Muthe, et al. highlighted the complications in the current internet architecture, it mainly points out that very few organizations control most of the data on the internet. They also point out the consequences of such a system such as data manipulation, lack of privacy and data misuse. It proposes a new architecture where it is focused on a fully secure and decentralized network. It uses technologies such as IPFS, Peer-to-peer, Ethereum and smart contracts. The proposed architecture also uses zero knowledge proofs and proxy re-encryption mechanism for privacy of the nodes in the network.

In [8] Barbara Guidi, et al. mainly focuses on analyzing the problem of data persistence in decentralized applications by considering decentralized social media as a case study. They used the IPFS protocol to store and share data with other people.They have discussed IPFS technology, its limitations and possible solutions for the same such as Private networks,encryption of data to overcome privacy issues, data replication for data availability issues and "Pinning service" for data permanence.
They also provided preliminary analysis of the IPFS network and this showed that cloud infrastructure services are most commonly used nodes which are used for Pinning service, gateway service and others.

C. Hybrid architecture

Hybrid architecture collaborates the best features of both centralized and decentralized storage media. Decentralized storage media maintains privacy of communication while centralized storage media provides real-time response.

In [9] Thomas Paul, et al. have proposed a system called Lilliput to store data in a P2P network for OSNs. This system is particularly focused on OSNs and its highly dynamic data. In the proposed system, they store static contents like videos and images on cloud after encrypting it. Only metadata is stored on the P2P network. This reduces the network load by a great extent. Since no metadata is stored on the cloud, it will not be possible to mine any information from the media data. This gives better performance compared to a completely decentralized OSN, but suffers from a central point of failure.

In [10] Giuliano Mega, et al. proposed a system which is the integration of cloud with distributed social media,this approach is Serverless and is inspired by p2p computing,it requires a computing device,installation software and internet connection to run the social media.The advantage of this is cost of maintaining the centralized servers is reduced and blocking of any contents cannot be done.Cloud-assisted profile dissemination over social overlays (CLOPS) is a hybrid system proposed here. It sues the highly-available cloud infrastructure to support the social overlay, without sacrificing any properties clops relies two things 1) takes the social overlay structure into account, allowing quick update dissemination while respecting clustering and degree heterogeneity; 2) resorts to information stored in the cloud only when and where required

D. Distributed data structure

Distributed data structure uses conventional data models such as trees, graph to construct a distributed data structure that provides advanced user control over the data. Different parts of a larger data structure could be assigned different access policies.

In [11] Jens Janiuk, et al. proposed a system based on Distributed data structures such as lists, trees and sets. The main characteristics of DDS is that it not only store the payload but also have the pointer for the next element .In order to make the proposed system secure and access control,we need to enable user authenticate each user and send only encrypted,signed and authenticated data.The system tells how we can make User ID to be public key, and password and usernames to be hashed to obtain a private key.To use of access control is to support read and write operation for DDS. The qualitative evaluation of the proposed system is done based on security and access control and quantitative evaluation is based on the cost of creating a DDS.

In [12] Andrea De Salve, et al. have proposed a system to store user data in decentralized P2P OSNs. They suggested storing user data on those nodes that already have permission to access the given data, hence avoiding need for encryption. Each users' data is organized as a tree. Each node of the tree corresponds to a

particular piece of user data. The system allows users to specify who all can access the data based on some attributes. The owner elects a trusted replica. This node is also responsible for verifying authorization of other nodes accessing user content. Primary trusted replicas will have the privilege to elect other trusted replicas. A user's content could soon become unavailable if the user has few trusted nodes.

E. Replication protocol

Distributed data could be misused by the nodes that hold a replica of the data. To prevent this, replication protocols are used. Replication protocols decide who can store a replica of the data and what they can do with that replica.

In [13] Rammohan Narendula, et al. proposed a system called "My3" which is a privacy friendly decentralized system and an alternative for Online social networking.They outlined the system architecture and proposed a number of replication system which can be independently chosen by the users according to their recruitment. The proposed My3 system exploits several properties of OSN. They used real data traces of Facebook and Twitter for their experiment and proved the effectiveness of their replication algorithms towards their respective goals when jointly or independently chosen by users. According to the results of the conducted experiment a total online time of 40 minutes of a user is enough for higher availability with 4-5 replicas.

In [14] Anna Kobusinska, et al. have proposed a socialrank protocol which is used for content replication that can be used by services deployed in p2p networks. The proposed protocol extends Easyrank replication protocol by using knowledge that is specific to social networks. In this protocol they have expanded the list of nodes taking part in the replication. Due to these changes it was found out that it works very well in networks where nodes have a low average number of neighbors during the simulation tests. They also compared the social rank protocol with Easyrank protocol and found out that the proposed system delivered the best results.

In [15] Mohammad A Khan, et al. have presented an effective content replication scheme for p2p OSN.

They have defined the topology of p2p OSN by the social network of participants. The proposed replication method prevents the skewness of available replication storage across the network and improves the replication success/fairness without depending upon the global knowledge of social networks. Here they have developed a new centrality metric called Easy rank. This metric is calculated at each neighboring node which finds out the underlying connectivity structure that is responsible for the skewness in storage. In the proposed replication scheme the replicas are stored based upon the Easy rank scores and the currently available storage. The results showed that this replication scheme stores the replicas in a fair and balanced way among the tested methods and it also provided the highest replication success rate.

In [16] Richard Gay, et al. focused on Decentralized Online Social Networks. They point out that Decentralized social networks have strict rules and regulations for sharing data with friends of friends, sharing in such a manner is totally prohibited or resharing is prohibited with certain conditions. The author presents a reinforcement mechanism for resharing DSON's by relation based access control. The author addresses that DOSNs are controlled by multiple providers. A prototype of such DOSN is presented which permits resharing in a controlled manner, it enforces private security policies of the user, it enables the authors to have an enhanced control on how their messages spread during resharing, this enables the users to geta better outreach and connect to new networks and users who have received their message through a trusted or private network.

III. Conclusion

Several techniques of decentralization are used for file sharing use cases. Data pertaining to an OSN is highly dynamic. Hence, traditional decentralized file sharing protocol cannot be used for DOSNs. DOSNs should be able to guarantee content availability with minimum required replication. It is optimal to minimize the number of replications to prevent the overhead of updating all the replicas of a given piece of data. Access control management is also a major concern. Since the entire service is decentralized, it will now be the duty of the individual nodes to ensure authorization before sharing data with any other node. Implementing a decentralized online social network will resolve many offline privacy issues. In a decentralized system, it will be hard to minimize network traffic and achieve real time response, which is very much important for OSN services such as instant messaging. A practical system will have to choose between optimizing network traffic and eliminating a single point of failure.

IV. Future work

In future, we are planning to implement a DOSN that uses a decentralized network to store metadata and a group of federated servers to store media files. The use of a decentralized network to store metadata will ensure data privacy. All the files stored in the data store will be encrypted and will have a unique identifier. No other data pertaining to the media file will be stored in the data store. When a node needs to fetch any media file, it fetches the unique identifier and the decryption key pertaining to the media file from the decentralized

network, fetches the file from the data store and then decrypts it. Using federated servers improves the overall performance of the system. It reduces network traffic. Most of the home network will have lower uplink speed. So it is ideal to use federated servers to serve media files and a decentralized network to store metadata.

References

[1]. Leila Bahri, Barbara Carminati, and Elena Ferrari, "Decentralized privacy preserving services for online social networks", Online Social Networks and Media 6 (2018) 18–25, Elsevier.

[2]. Nashid Shahriar, Shihabur Rahman Chowdhury, Reaz Ahmed, Mahfuza Sharmin, Raouf Boutaba, and Bertrand Mathieu, "Availability in P2P based Online Social Networks", 2017, IEEE.

[3]. Giuliano Mega, Alberto Montresor, and Gian Pietro Picco, "Efficient Dissemination in Decentralized Social Networks", 2017, IEEE.

[4]. Anandhakumar Palanisamy, Mirsat Sefidanoski, Spiros Koulouzis, Carlos Rubia, Nishant Saurabh, and Radu Prodan, "Decentralized Social Media Applications as a Service: a Car-Sharing Perspective", 2020, IEEE.

[5]. Quanqing Xu, Zhiwen Song, Rick Siow Mong Goh, and Yongjun Li, "Building an Ethereum and IPFS-based Decentralized Social Network System", 2018, IEEE.

[6]. Van-Duy Pham, Canh-Tuan Tran, Thang Nguyen, Tien-Thao Nguyen, Ba-Lam Do, Thanh-Chung Dao, and Binh Minh Nguyen, "B-Box - A Decentralized Storage System Using IPFS, Attribute-based Encryption, and Blockchain", 2020, IEEE.

[7]. Koushik Bhargav Muthe, Thiru Srinivasa Teja Vemuru, Khushboo Sharma, and Nilofer Sultana Mohammad, "DECENTRANET - AN ETHEREUM, PROXY RE-ENCRYPTION AND IPFS BASED DECENTRALIZED INTERNET", 2020, 11th ICCCNT IEEE.

[8]. Barbara Guidi, Andrea Michienzi, and Laura Ricci, "Data Persistence in Decentralized Social Applications: the IPFS approach", 2021, IEEE 18th Annual Consumer Communications & Networking Conference (CCNC).

[9]. Thomas Paul, Niklas Lochschmidt, Hani Salah, Anwitaman Datta, and Thorsten Strufe, "Lilliput: A Storage Service for Lightweight Peer-to-Peer Online Social Networks", 2017, IEEE.

[10]. Giuliano Mega, Alberto Montresor, and Gian Pietro Picco, "Social Overlays Meet the Cloud: A Hybrid Architecture for Profile Dissemination in Decentralized Social Networks", 2018, IEEE.

[11]. Jens Janiuk, Alexander Macker, and Kalman Graffi, "Secure Distributed Data Structures for Peer-to-Peer-based Social Networks", 2014, IEEE.

[12]. Andrea De Salve, Paolo Mori, Laura Ricci, Raed Al-Aaridhi, and Kalman Graffi, "Privacy-Preserving Data Allocation in Decentralized Online Social Networks", IFIP International Federation for Information Processing 2016, Springer.

[13]. Rammohan Narendula, Thanasis G. Papaioannou, and Karl Aberer, "A Decentralized Online Social Network with Efficient User-Driven Replication", 2012 ASE/IEEE International Conference on Social Computing and 2012 ASE/IEEE International Conference on Privacy, Security, Risk and Trust.

[14]. Anna Kobusinska, Michał Boron, Beata Szturemska, and Yue-Shan Chang, "Data Replication Based on Common Interests in P2P Social Networks", 2018 IEEE 11th International Conference on Service-Oriented Computing and Applications.

[15]. Mohammad A Khan, Hillol Debnath, and Cristian Borcea, "Balanced Content Replication in Peer-to-Peer Online Social Networks", 2016 IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom).

[16]. Richard Gay, Jinwei Hu1, Heiko Mantel, and Sogol Mazaheri, "Relationship-Based Access Control for Resharing in Decentralized Online Social Networks", 2018, Springer.