**REPORT**

**DECENTRALIZED ONLINE SOCIAL NETWORK, PEER-TO-PEER NETWORK, MOBILE PEER-TO-PEER NETWORK**

**PHAM VAN TIEM ITIU09069**

1. **Decentralized online social network:**

**Online social network** is an online platform that allow user building a public profile, declaring the connection between users, sharing information and content, and, interacting and collaborating with both friends and strangers by some social applications.

Almost existing social networking services are centralized and it has some **problems** include both technical and social issues. On technical side, the centralized management of a social network with rapid growth of number of user has led to scalability issues such as the frequent down-time of the Twitter and the slowness and unresponsiveness of Facebook. On social side, the unlimited sharing capability of information did not preserve the privacy of user.

**A decentralized online social network** is an online social network that implemented on a distributed information management platform such as a network of trusted servers or a peer-to-peer system. The advantages of DOSN are to help reducing the cost of the provider (Skype), controlling user privacy better, and enhancing the innovative development. In DOSN, there is no central server, just has a set of peers that take a share of tasks needed to run the system. Therefore, it has some disadvantages like: in terms of privacy – no central data collection, in operation- no central entity that decides or changes the terms of services. Distributed systems help user can use their own storage, delay tolerant social networks, and local treatment of local content.

However, there are still some **challenges for DOSN** such as storage, updates, topology, search and addressing, security, robustness, limited peers, and locality. We are researching and developing all challenges to consider that are general benefits of decentralizing online social network.

* **The case of decentralized social network compare with centralized social networks**

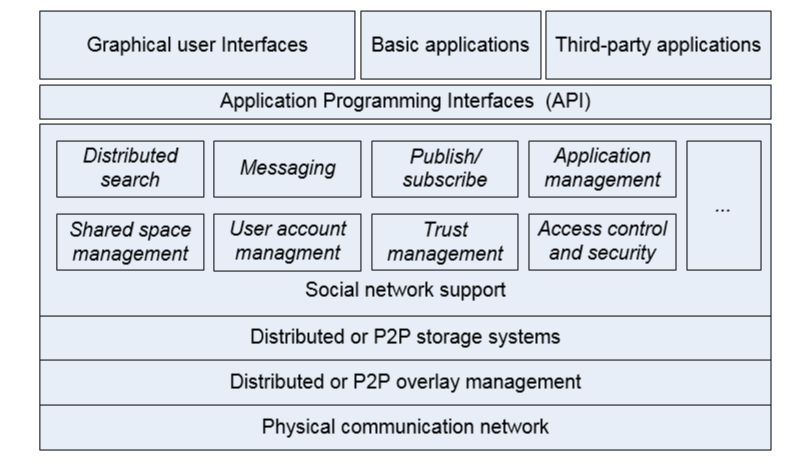
In keeping user data, a centralized depository is more susceptible to virus or malware spreading than mostly local social networks that can be partitioned. DOSN takes inherent advantages as no centralization, no owned by a single entity to decrease privacy concerns with encryption and appropriate key management. To perform well, DOSN mostly use the mechanism of operation of P2P as scalability. This answer the question why uses a decentralized infrastructure for supporting social network while the old good client-server architecture still works fine. In fact, DOSN can give the traditional arguments that P2P scales well when the number of user increase significantly.

Large part of users who enjoy the benefits and fun of social network nowadays also concern to privacy. They want to restrict access to their personal data not only from follow users who happen to be strangers, but also from any provider or indeed the general public. On the other hand, they also want their sharing information (document, photo, video,…) to be secured, mean that just friends who are accepted to see shared files. This disaffected population is expected to be the early adopters of decentralized social networks with encryption. Besides these problems, a decentralized approach also enables content creators to execute greater control over their content, as well as avoid censorship either by the website owner.

With a peer-to-peer approach, the decentralized systems will control over data back to the users and not have one entity access to all personal data of the participants in the social network. In addition, it combines with appropriate encryption users to determine whom they allow access to their data. P2P infrastructure supports the direct exchange of information between devices (between users that meet or between adjacent nodes), and takes advantage of real social networks. In addition to addressing the privacy aspects, we discuss with the controlling data aspects provided by a peer-to-peer infrastructure. User control has consequences beyond privacy and freedom from advertisement. It controls over who can access their content and what they are allowed to do with that content.

Generally, decentralized social networks can improve any constraints that current social network meet, and satisfy some user’s need. The infrastructure of P2P help user-provided content and participatory media creation suit themselves better than client-server model. Peers can carry information for each other in a delay-tolerant fashion and use local access points for local information.

For the **architecture of decentralization of online social network** services, we discuss about six layers and its distribution.



The general architecture of a distributed online social network

The lower layer is the physical communication network, which can be the Internet or a ad hoc network. The distributed or P2P overlay management provides core functionalities to manage resources in the supporting infrastructure of the system. The distributed or P2P storage is the decentralized data management layer, which implements functionalities of a distributed of P2P information system to query, insert, and update various persistent object to the systems.

The social networking layer implements all basic functionalities and features that are provided by contemporary centralized social networking services.

The top layer of the architecture includes the user interface to the system and various applications built on top of the development platform provided by the DOSN.

1. **Peer-to-peer network:**

***What is P2P?***

Firstly, we should know P2P is a network that can share computer resources and services by direct exchange of information between peers. It takes advantage of resources: storage, cycles, content, human presence that are available at edges of the Internet.

P2P networks are distributed systems without centralized control. A distributed network architecture may be called a P2P network if the participants share a part of their own resources. Clients in P2P networks are also servers and routers. Each peer is autonomous (no administration), means that peers form self-organizing.

P2P networks have some benefits like: efficient storage of resources, easy search of data items, selection of nearby peers, trust and authentication, anonymity-privacy, self-organizing, massive scalability, and fault-tolerant.

***P2P overlay networks***

P2P overlay networks have two classes that this report just focuses: Structured and Unstructured.

1. *Structured p2p overlay networks:*

Features: - the p2p overlay network topology is tightly controlled.

* The content is placed not at random peers but at specified locations that will make subsequent queries more efficient.
* Using Distributed Hash Table (DHT) to locate the data object (value) and unique identifier (key).

Each peer in Structured topology maintains a small routing table consisting of its neighboring peers’ NodeIDs and IP addresses.

Based on DHT, there were the models of structured p2p overlay networks such as: CAN, Chord, Kademlia, Tapestry, Pastry, and Viceroy.

* CAN:
* The CAN is a distributed decentralized P2P infrastructure that provides hash table functionality on Internet-like scale.
* The architectural design is a virtual multi-dimensional ID coordinate space.
* A CAN peer maintain a routing table that holds the IP address and virtual zone of each neighbor. CAN has a routing performance of O(d.) and its routing state is of 2.d bound.
* The virtual coordinate space is used to store {key, value} pairs. The lookup protocol to retrieve an entry corresponding to key K is to map K onto point P and retrieve the corresponding value V from the point P by using uniform hash table.
* The system is partitioned among all the peers (N number of peers) in network d-number of dimensions.
* A new peer that joins the system must have its own portion of the coordinate space allocated. The peer looks up in the DNS a CAN domain name to retrieve a bootstrap peer’s IP address to choose the new random peer. After obtaining its zone, a new peer learns of the IP address of its neighbor and adds to the previous peer itself. A peer leaves CAN network, then an immediate takeover algorithm ensures one of the failed peer’s neighbor takes over the zone and start takeover timer. Individual peers maintain a list of 2.d neighbors.
* Failure of peers will not cause network-wide failure.
* Pros/cons: CAN designs to be scalable, fault-tolerant, and self-organizing, but it is limited about security.
* Chord:
* Using the decentralized P2P infrastructure that provides DHT. This decentralized scheme tends to balance the load on the system.
* The architecture design use uni-directional and circular NodeID space.
* For N peers in the system, each peer maintains routing state information for about only O(logN) other peers (N number of peers in the system).
* The lookup queries involve the matching of key and NodeID. A given identifier could be passed around the circle via successor pointers until they encounter a pair of peers that include the desired identifier, the second peer will be the peer that query maps to.
* When a peer n joins the network, keys previously assigned to n’s successor now need to be reassigned to n. When peer n leaves the Chord system, all assigned keys are reassigned to n’s successor. So the peers join/leave the system with performance.
* Failure of peers will not cause network-wide failure. Replicate data on multiple consecutive peers.
* Pros/cons: Chord system is designed to let peers enter and leave network with minimal interruption. Similar to CAN, the disadvantage of Chord is about security, it is low level.
* Kademlia:
* The Kademlia P2P decentralized overlay network takes the basic approach of assigning each peer a NodeID space.
* The routing algorithm will be used to locate peers near a destination key.
* The architecture design use XOR metric for distance between points in the key space. XOR is symmetric and it allows peers to receive lookup queries from the same distribution of peers contained in their routing table.
* The looking up match key and NodeID based routing.

1. *Unstructured p2p overlay networks*: in this category, the overlay networks organize peers in a random graph in flat and use flooding on the graph to query content stored by overlay peers.

Features: - fully distributed: no central server.

* Each peer indexes the files it makes available for sharing.
* It can limit per-node state and can be fault tolerant. However, it takes high bandwidth usage when the request of client sent to all neighbor nodes, takes long time to locate item, and unbalanced load.
* Gnutella:
* Gnutella is a decentralized protocol for distributed search on a flat topology of peers.
* Gnutella’s distinction is its peer to peer, decentralized model for document location and retrieve. In this model, every peer is server or client. Query message sent over existing TCP connections, peers forward Query message.
* Peer join network must find another peer in Gnutella network using list of candidate peers. A new peer attempts TCP connections with candidate peers until connection setup with its neighbor.
* Flooding: a peer sends Ping message to a other peer, then the other forwards Ping message to his overlay neighbors. After that a new peer will receives many Pong message, and can then setup additional TCP connections.
* Pros/cons: flexibility in query processing, complete decentralization, simplicity, fault tolerant and self-organization. However, it takes severe scalability problems and it is susceptible to attacks.

1. **Mobile peer-to-peer networks:**

Mobile phone nowadays becomes popular with every people. It is known as a new dimension of social interaction to enable people communicate anywhere and at any time. With this development, mobile phone created a new connection in social. The enhanced connectivity afforded by mobile phones has facilitated the creation and maintenance of social networks. People who use mobile phone also have some demands such as securing their call, text, message, or data; sharing information (photo, video, and document).

To serve these demands, we can replace the old mobile network model (centralized network model on mobile) by mobile peer-to-peer model to overcome the constraints as the capacity of the terminal, the connectivity technologies available on the device, and restriction policies applied by network operators. It allows mobile terminals to share contents they are now able to produce: contextual information, multimedia documents,…

How the principles of peer-to-peer networks apply to mobile networks, and how they can be used to serve mobile communication or contextual applications. We should understand about applying p2p network in mobile. Same with p2p network, mobile p2p systems are system where mobile devices can collaborate together and fixed devices like computers without the intervention of a central server. These systems can connect together spontaneously in an ad-hoc fashion, or the telecom operator’s mobile network.

Most of mobile cellular networks (ad-hoc, MANET, GPRS, 2G, 3G,…) have the architecture similar to pyramidal structure. This has an influence to traffic and signal of system when mobile peer try to connect with other. In mobile networks, the possibility for a given mobile to establish a data connection to an IP network is a costly resource for mobile network operators, which restrict the use of some protocols or transport methods on their networks. To tackle these limitations, first the peer-to- peer systems have to be efficient and to avoid sending too many discovery or network maintenance messages on the network in order to reduce its footprint. Thus, the discovery mechanism and the architecture of the peer-to-peer network have to be adapted to this constraint. Besides, in order to maintain its reach ability in networks where network address translation is often done in a very ephemeral way, mobile peers have to refresh regularly their connection to their neighbors in order that they know how to reach them.

In fixed-mobile peer-to-peer networks, mobile peers could collaborate with a fixed peer for tedious operations such as the discovery of resources within the network or maintaining the connectivity of the mobile peer through the operator’s network address translation system. In case the mobile operator restricts the use of some protocols, the fixed peer can also take in charge the network protocol translation operations. This way, all the operations that are made difficult by the structure of the mobile data network or by management policies are done by the fixed peers. Those peers then provide the mobile peer with the capability to fully interact with the other peers in the network. They should be selected according to an automatic mechanism, based on their stability in the network or on the amount of available resources they share. As we have seen before, this super node concept is available on both centralized and distributed hash table networks. Given the less centralized nature of distributed hash table networks, this kind of system is particularly adapted to the implementation of peer-to-peer systems in a mobile cellular data network environment.

Peer-to-peer systems can be used to set up real-time collaborative applications. For instance, mobile peer-to-peer systems may be used for multiplayer gaming in order to allow players involved in the same game to exchange information together on their positioning and actions. Mobile p2p also has some advantage of p2p network such as reducing the loads on the server, high scalability. In mobile p2p, sharing information becomes easier because mobile phone is considered multiple content production endpoints since they have a camera and microphone. Besides that, the privacy and security is a major concern. P2P system protects and secure at each node when it upload information on the centralized platform.

1. **Comparison of mobile p2p network and p2p network:**

After learning and find out information about mobile p2p networks and p2p networks. We see some similar points and different points between them. For similar, mobile p2p mostly operate same with p2p (directed exchange between nodes, no centralization, easy to scale,…). In this part, we mainly focus on the differences between them.

* In joining network of nodes, P2P networks just contain fixed peers like PC, laptop. Mobile p2p has participation of mobile devices and fixed devices.
* In mobile p2p, it is easy to use in sharing information in anywhere and at any time.
* In mobile p2p, it is easy to connect between devices which can access internet.
* Because p2p network is just connection of computers, at each node, computer has the role of sharing information with faster speed than devices in mobile p2p network.