CCS 363 SOCIAL NETWORK SECURITY

UNIT I. FUNDAMENTALS OF SOCIAL NETWORKING.

Introduction to Semantic Web - Limitations of Current Web - Development Semantic Web - Emergence of the social Web - Social Network Analysis - Development of Social Network Analysis - Key Concepts and measures in network Analysis - Historical overview of privacy and security, Major paradigms - for understanding privacy and security

1.1 Web :-

⇒ It is a communication model or platform that enables the retrieval or exchange of information over the internet.

Semantic Web :-

⇒ It is the application of advanced knowledge technologies to the web and distributed system in general.

⇒ It is the extension of (www)

⇒ Not only improves traditional search.

⇒ But also facilitates more seamless, intelligent, & integrated users experience journeys.

⇒ Semantic web add another Layer to the web.

Information such as text, Image, Audios.

1.2 Limitations of the Current Web :-

Finding Relavent Information } difficult.
Extracting Relavent Information.

What's corong with the Web :-

1. CCS 363 Syllabus Notes : .

Keyword → show the results for the Keyword CCS 363 Social n/w Security using google.

10 Links.

3 links apt
3 Links
3 Links.

Search → document.
→ Photo
→ Music.

2. Search Me to photo of Paris.

1 Page
2 Page → Information.
3 Page → Invalid contents.

Photos pairing is difficult task.

Problems :-

1. Associating Photos with keywords is a much more difficult task than simply looking for keywords in the text of documents.

2. Automatic Image Recognition is unsolved Research Problem.

3. Search Engine difficult to understand the

the meaning of the image from its contents.

3. Search the music.

⇒ Tell me about music player with a capacity of atleast 4 GB

⇒ Search Engine will not know about music player capacity.

⇒ Because It does not read the properties of music player.

There are the pbms in current web.

How to improve current web.

⇒ Increasing automatic linking among data

⇒ Increasing recall and precision in search.

⇒ Increasing automation in data integration

⇒ Increasing automation in the service life cycle.

1.3 Development of Semantic Web :-

Research, Development and Standardization :-

⇒ The vision of extending the current human-focused web with machine processable descriptions of web content has been first formulated in 1996 by Tim-Berners-Lee, the original inventor of the web.

⇒ Semantic web has been actively promoted since by the world wide web Consortium (also led by Berners-Lee), the organization that is chiefly responsible for setting technical standards on the web.

⇒ Semantic web has quickly attracted from funding agencies, because of AI research Agenda in a short period of time.

⇒ In particular, the field of,

- → Knowledge representation
- → Reasoning

} took center stage in Semantic web.

⇒ And also Natural Languages Processing.

⇒ Information Retrieval fields also supports the development of Semantic web.

In Research on the Semantic Web Community

⇒ Includes the researchers who have submitted the publications in the International Semantic web Conferences, (ISWC02, ISWC03, ISWC04)

or Semantic Web working Symposium of 2001 (SWWS01)

Semantic Web community consists of 60s Researchers.

In that (79%.) from academia.

(21%.) from Industry,

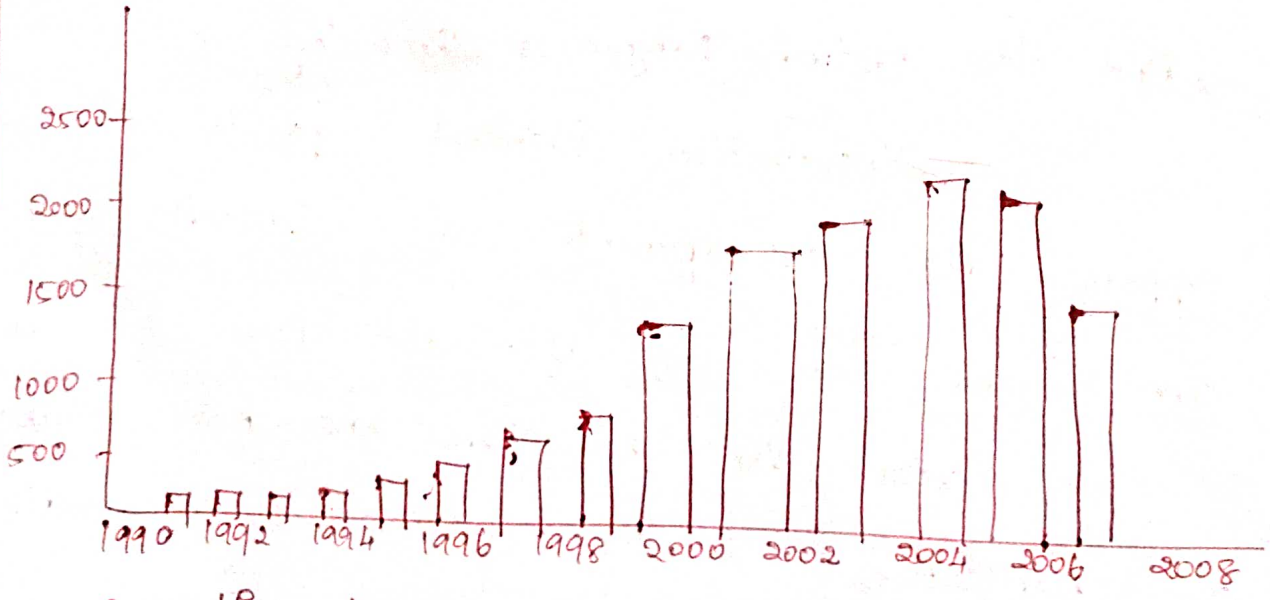
⇒ Community covers Mostly from United State, Europe.

Some activity from Japan and Australia.

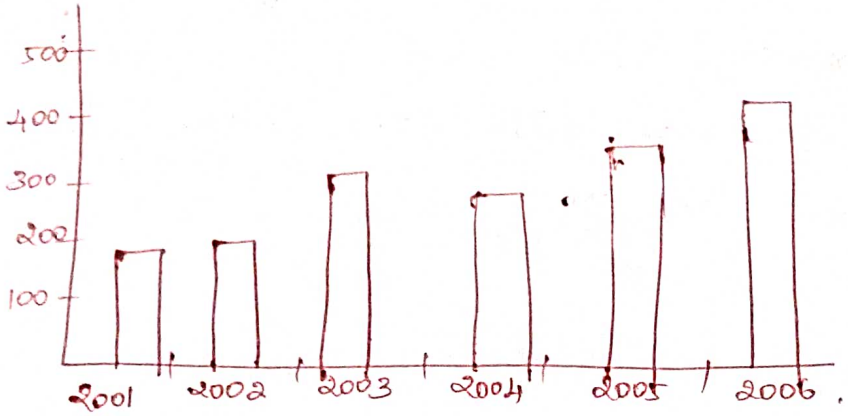
Number of publications written by the members of the community with the keyword "Semantic Web" rising from the beginning.

Core technology of the Semantic Web:-

Logic-based Languages, for knowledge representation and reasoning has been developed in the research field of Artificial Intelligence.



Semantic Web related publications per year (1990-2006)



Participation at the yearly International Semantic Web Events (2001 - 2006)

Technology Adoption :-

⇒ Semantic Web is an extension of Current Web. ie) application of metadata for describing Web content.

⇒ In this vision, the content already on the on the web (Text, but also multimedia) would be enriched in a colloborative effort by the users of the web.

→ Difficult for the average person contribution

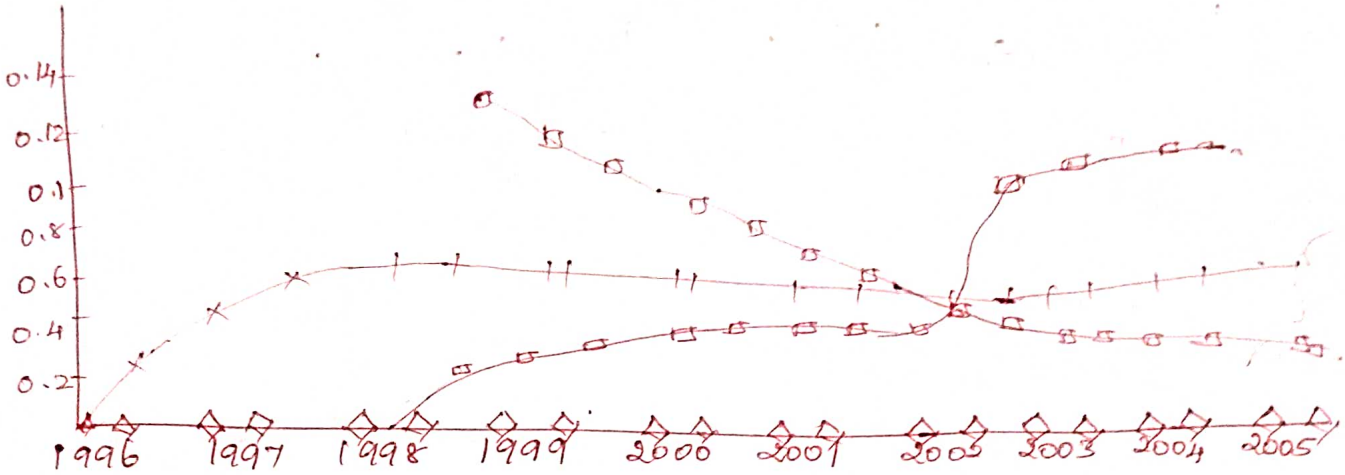
Semantic Web suffers from fax - effect.

Semantic Web.

⇒ At the beginning the price of technological investment is high.

Growth of the Semantic Web by Tracing its Popularity on the Web.

Number of documents, with the terms basketball, Computer Science and XML.



Number of Web pages with the term basketball, Computer Science and XML over time and as a fraction of the number of pages with the term Web.

We also look at,

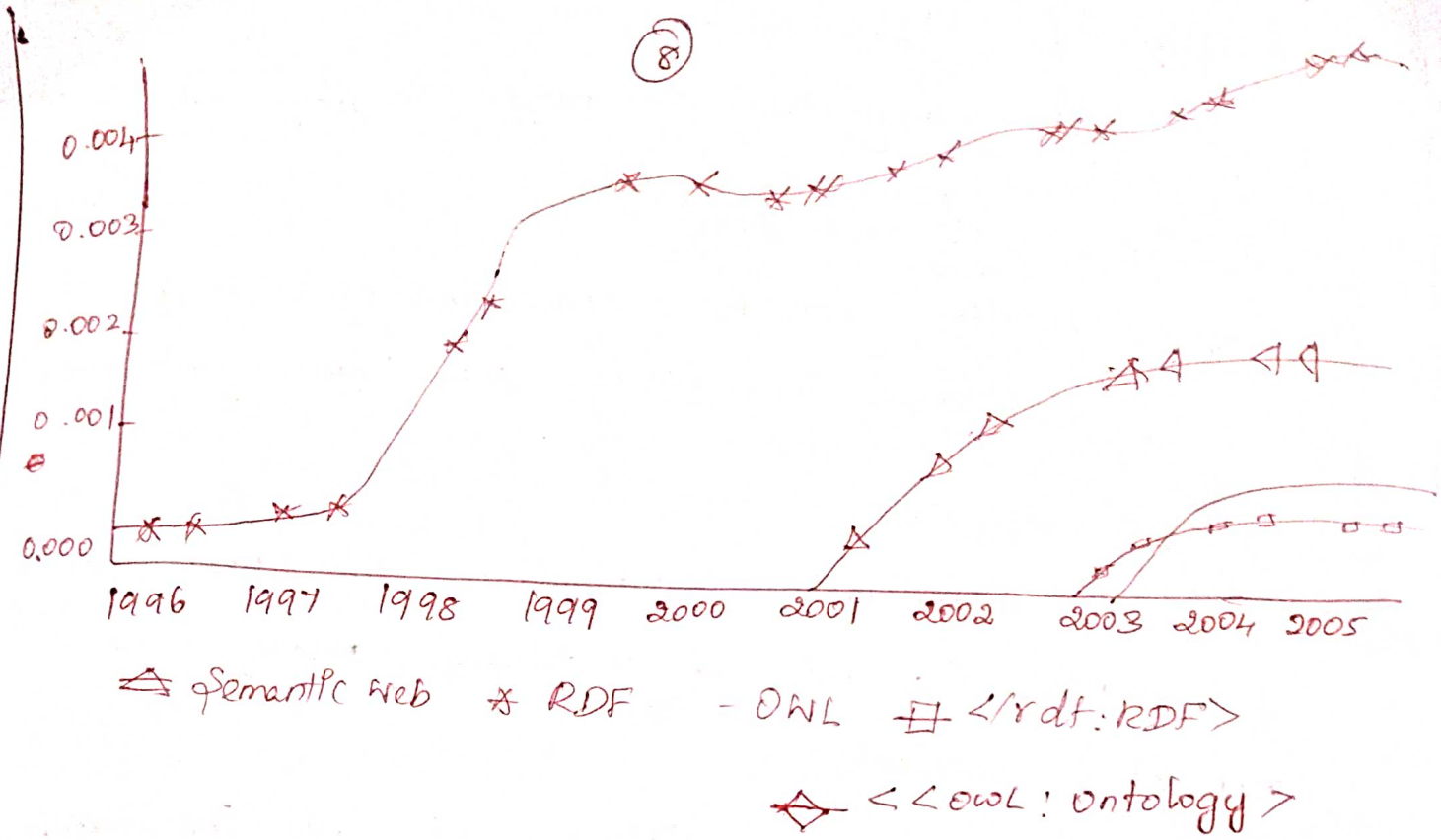
⇒ Share of Semantic Web, related terms and formats, in particular the terms 1) RDF, 2) OWL and the 3) number of ontologies (Semantic Web Documents) in RDF or OWL,

⇒ Most of the curve have flattened out after January, 2004.

RDF → has settled at a relatively high level.

OWL → has very low trajectory

Because, OWL documents only found in June, 2006 onwards.



The curves are very similar to the hype cycle of Garner Research.

1) First Phase :- " Technology Trigger "

Product Launch —

2) Next Phase :- " Frenzy of Publicity "

Generates ⇒ over enthusiasm and unrealistic expectations.

some successful application of technology

~~But~~ More failures.

3) " Trough of disillusionment "

~~Because~~ If they fail to meet expectations and quickly become unfashionable.

4) "Slope of enlightenment ?:- (9)

Press may have stopped covering the technology,

5) "Plateau of Productivity":

=> Benefits can be demonstrated & accepted.

=> Evolves in second and Third generations.

14 Emergence of social Web:-

=> In 1990, Web - combination of phone book, yellow pages (a mix of individual postings and corporate catalogs)

=> The passive attitude toward the web was broken by a series of changes in usage pattern and technology that are now referred to as web 2.0 a buzzword coined by Tim O'Reilly

History of Web 2.0.

=> Set of innovations in the architecture and Usage patterns of the Web ~~lead~~ Led to an entirely different role of the online world

=> A recent survey based on interviews with 2200 adults shows that the internet significantly improves Americans!

=> Capacity to maintain their social n/w. despite early fears about the effects of diminishing real life contact. Indoor games

Blogs and Wikis :-

Blogs (10) are a type of regularly updated website that provide insight into a certain topic

⇒ First wave of socialization on the web, due to blogs, wikis & other forms of web based communication.

⇒ Blogs and wikis → Most popular in 2003.

↳ allowed the individuals and groups to claim their personal space on the web, and fill it with content.

Wikipedia, the online encyclopedia is outstanding, wikis large and small are used by groups of various sizes as an effective (provide) knowledge and management tool for keeping records.

⇒ The significance of Instant Messaging (IC&) is also not just instant communication, but the ability to see who is online. (that has)

Social Networks.

⇒ The first online social n/w also referred to as Social Networking Services.

⇒ These sites allow user to post a profile with basic information.

To invite others to register & to link to the profiles of their friends.

It leads to discover friends in common,

⇒ Friends thought to be lost,

⇒ Potential new friendship based on shared interest.

The latest services are thus using user profiles and networks,

⇒ Photos are shared in Flickr, bookmarks are exchanged in del.icio.us

User Profiles :-

⇒ Explicit User profiles makes it possible for these systems to introduce rating mechanism whereby either the users or their contributions are ranked according to usefulness or trustworthyness.

Implementation :-

⇒ Asynchronous Java script and XML / AJAX.

⇒ Script languages, formats such as JSON, Protocols such as REST.

⇒ Rss Feed.

Web 2.0 + Semantic Web = Web 3.0 ?

Web 2.0 mostly effects how users interact with the web.

Semantic web opens new technological opportunities for web developers in combining data and services from different users.

Web 2.0 is that users are willing to provide content as well as metadata.

⇒ Web pages created automatically from a database can encode metadata in microformats without the user necessarily being aware of it.

⇒ Semantic technology can help in matching users with similar interest as well as matching users with available content.

1.5 Social Network Analysis :-

Social Network Analysis (SNA) is the study of social relations among a set of actors.

It focus on relationship between actors rather than the attributes of individual actors.

SNA -> It is a new method for data collection and Analysis.

Network analysis :- It is a formal description of n/w as graphs.

Social Role or Social group :-

It can be defined on a formal model of networks, allowing to carry out more precise discussions in the literature, and to compare across the studies.

Records of social interaction (publication databases, meeting notes, newspaper articles, documents and databases of different sorts.) are used to build a model of social n/w.

1.6 Development of Social Network Analysis :-

Social Network Analysis.

↳ Convergence of several streams.

of applied research in sociology, social psychology and anthropology.

Social psychologists of the 1940's.

⇒ found a formal description of social group process of group communication.

1950's Anthropologists

have found network Representations from field observations.

Researchers from Harvard looked at the Workgroup Behavior.

(Communication, friendship, helping, Controversy) at a specific part of the factory, the bank wiring room.

The investigators noticed that workers themselves used specific terms to describe who is in "our group" Group structure of the org as it emerged from the individual relationship of the factory coorkers

Development (15)

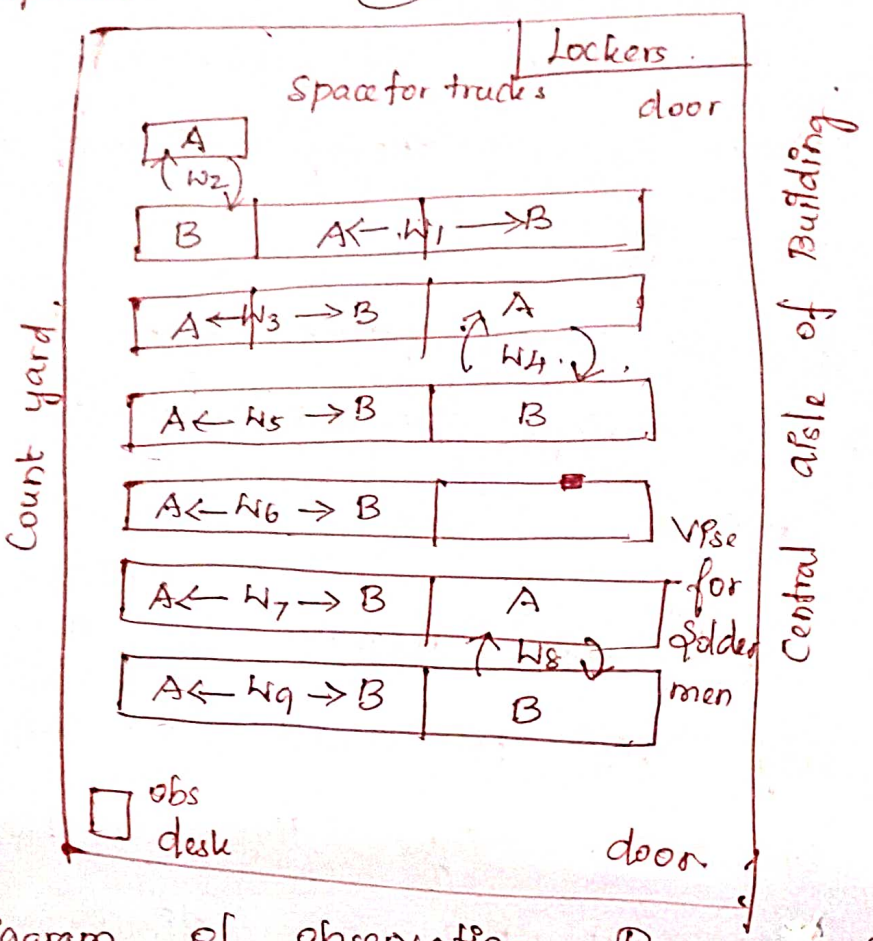
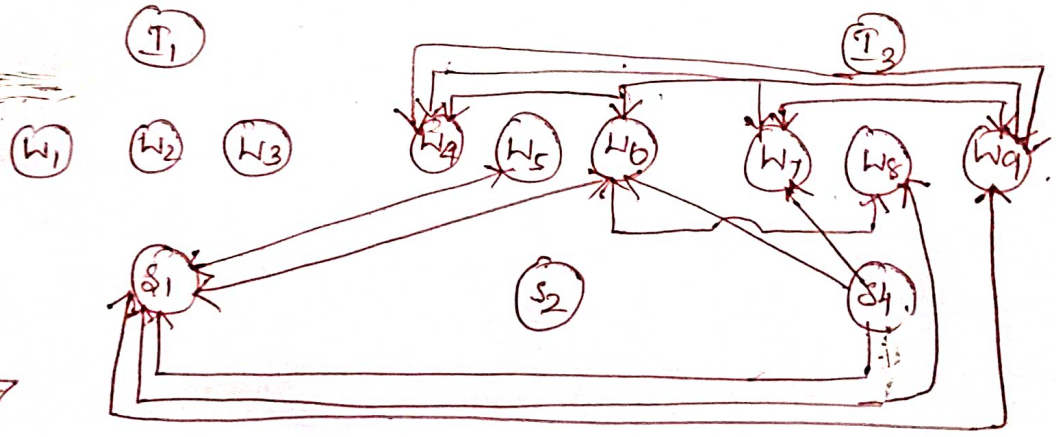


Diagram of observation Room showing wireman's Position.



Participation In Controversies about windows.

Illustration from an early social network study at the Hawthorne works of Western Electric in Chicago. The upper part shows the location of the workers in the wiring room, while the lower part is a network. Image of fights about the windows between workers (W) solderers (S), and Inspectors (I)

Development of Social Network Analysis:-

⇒ In the southern US researchers looked at the network of overlapping "clique" defined by race and age.

⇒ They also went further than the Hawthorne study in generating hypotheses about the possible connections between cliques.

Eg) The lower-class members of a clique are usually only able to connect to higher class members of another clique through the higher-class members of their own clique.

Social Network :-

⇒ It can be defined as a graph in which vertices are people or groups of people and edges represents some kind of social relationship b/w or social interaction between vertices.

⇒ The structure of a social network is commonly framed as a graph. Network (N, g)

Net (N, g)

Social Network Analysis :-

⇒ It is the process of investigating social structure through the use of networks and graph theory. It characterizes networked structures in terms of nodes (individual actors, people, or things within the network) and the ties, edges, or links (relationships or interactions) that connect them.

Community :-

⇒ A community in the social network is a set of nodes or users which can be potentially grouped and examined together such that the interactions or ties between the communities are quite dense.

⇒ Discovering sub graphs of vertices which

→share common properties is known as community Detection.

1.7 Key Concepts and Measures in Network Analysis.

Social Network analysis has developed a set of concepts and methods specific to the analysis of social networks. In the following, we introduce the most the basic notions of network analysis and the methods we intend to use later in this book.

Global structure of Networks:-

A (social) Network can be represented as a graph $G=(V, E)$ where V denotes the finite set of vertices.

$E \rightarrow$ denotes the set of Edges.

such that $E \subseteq V \times V$

Recall that each graph can be associated with its characteristic matrix $M := (m_{i,j})_{n*n}$

where $n = |V|$, $m_{i,j} = \begin{cases} 1 & (v_i, v_j) \in E \\ 0 & \text{otherwise} \end{cases}$

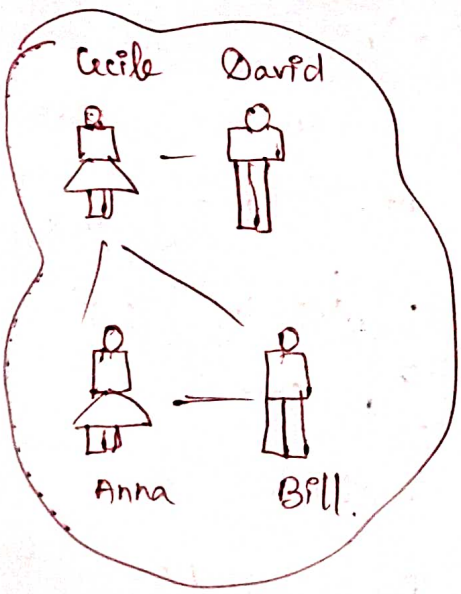
Some network analysis methods are easier to understand when we conceptualize graphs as matrices. Note that the matrix is symmetrical in case the edges are undirected.

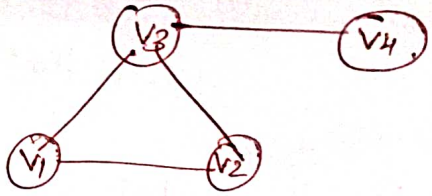
valued graph when we are also given a real valued weight function $w(e)$ defined on the set of edges ie) $w(e) := E \times R$.

In case of a valued graph, the matrix is naturally defined as

$$m_{i,j} = \begin{cases} w(e) \\ 0 \end{cases} \Bigg| \begin{matrix} (v_i, v_j) \in E \\ \text{otherwise} \end{matrix}$$

Typically, we assume that the network is connected, there is single (weak) component in the graph. Otherwise we choose only one of the components for analysis.





	V_1	V_2	V_3	V_4
V_1	0	1	1	0
V_2	1	0	1	0
V_3	1	1	0	1
V_4	0	0	1	0

Graph Based Representation of real world Networks.

At this point the question of what else can we say about the structure of society.

American psychologist Stanley Milgram experiment about the structure of the social network.

Milgram estimated is the size of the average shortest path of the network, which is also called characteristic path length.

A simple path in a graph is a sequence of vertices $v_{i0}, v_{i1}, \ldots v_{in}$ such that $\forall j = 0 \ldots n-1$ $(v_{ij}, v_{i(j+1)}) \in E$ and $\forall j$, ⇒ Every vertex connected to the next vertex and no vertex is repeated on the path.

⇒ The shortest path between two vertices V_s and V_t

It is a path that begins at the vertex V_s, and ends with the vertex V_t and contains the least possible number of vertices.

Geodesic:

⇒ The shortest path between two vertices is called a geodesic.

Diameter of the graph:

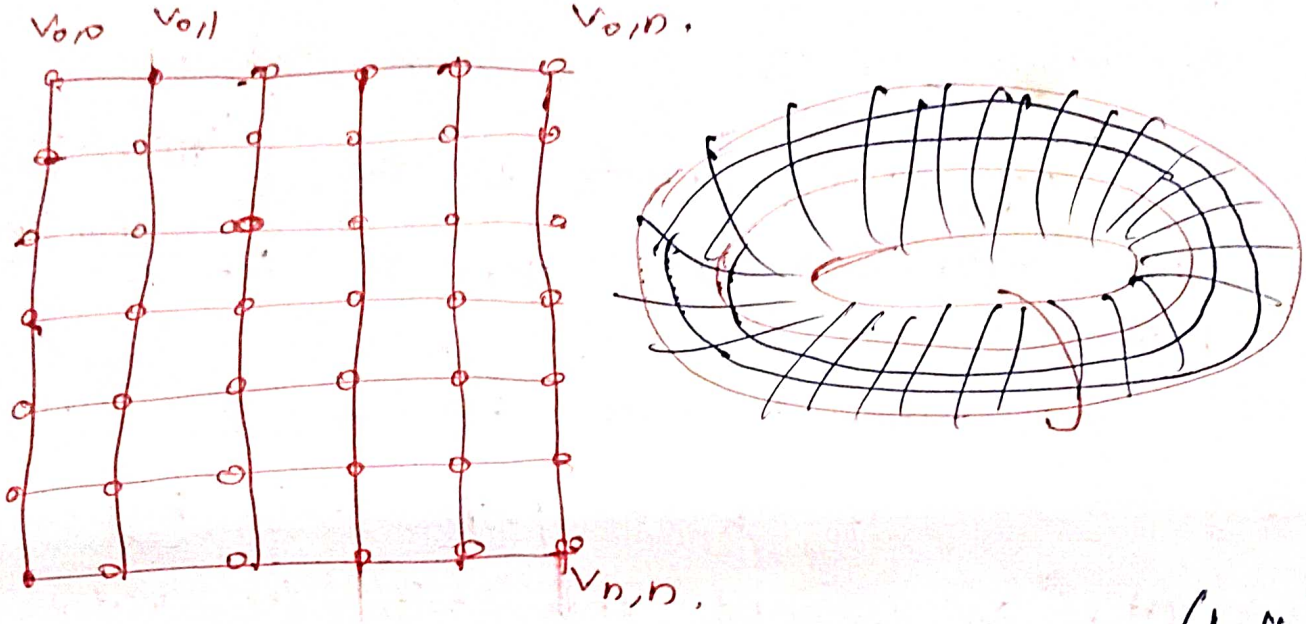
⇒ Longest geodesic in the graph is called the diameter of the graph. That is the maximum number of steps that is required between any two nodes.

Average shortest Path:

Average of the Length of the geodesic's between all the pairs of vertices in the graph.

from Milgrams finding.

We can include contain kind of structures as possible models for social networks.



The 2D Lattice model of networks (Left)

By connecting the nodes on the opposite borders of the Lattice we get a toroidal Lattice (right)

Another simple model could be the tree graph

Clustering:-

It is measure the degree's of nodes to decide which nodes in a graph tends to be clustered together.

Clustering for a single vertex can be measured by the actual number of the edges between the neighboars of a vertex by the possible number of edges between the neighbors.

Clustering Coefficient:

⇒ Clustering coefficient measure is to quantify how close its neighbors are to being a complete graph.

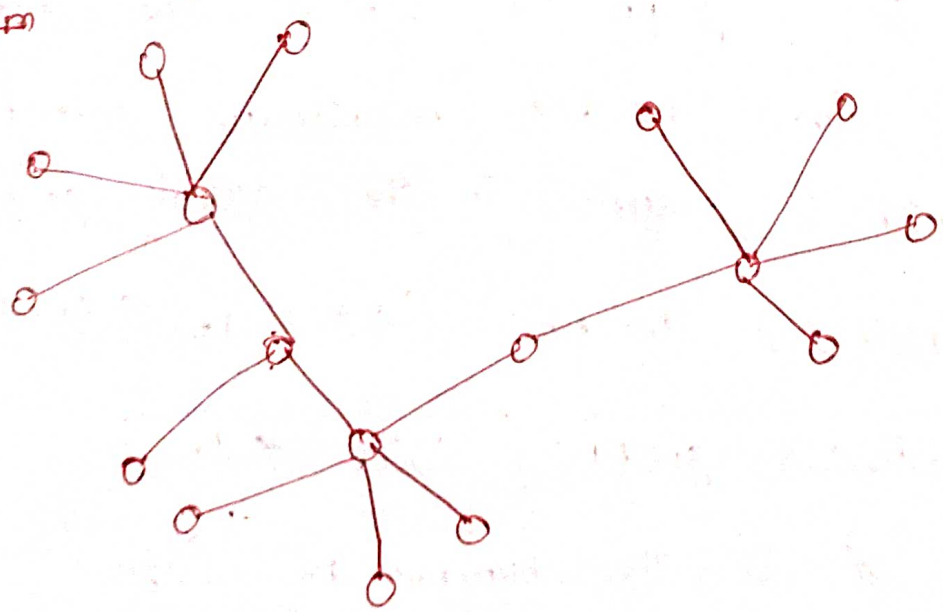
$$\text{clustering Coefficient} = 3 * TC / CT$$

TC = Triad Closure.

CT = Connected Triple.

⇒ The clustering coefficient of the tree is zero. which is easy to see if we consider that there are no triangles of edges (triads) in the graph.

~~Random~~



A tree is a connected graph where there are no loops and paths leading from a vertex to itself.

Random Graph :-

(24) proposed by Hungarian mathematicians

It can be generated by taking a set of vertices with no edges connecting them.

Subsequently, the edges are added by picking pair of nodes with equal probability.

Node Degree :-

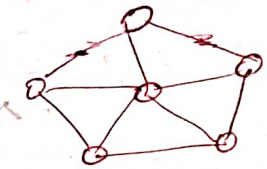
The degree of a node in a graph, is the number of Edges incident to that node.

It can be calculated using the formula.

⇒ Undirected Graph $N*(N-1)/2$

⇒ Directed Graph $N*(N-1)$

where N is the number of nodes presents in the graph.



Node Density :- (tightness)

It is a graph in which the number of edges is close to the maximum number of edges (which is present in the actual graph)

⇒ Undirected Graph : $(2*E/N)*N-1$

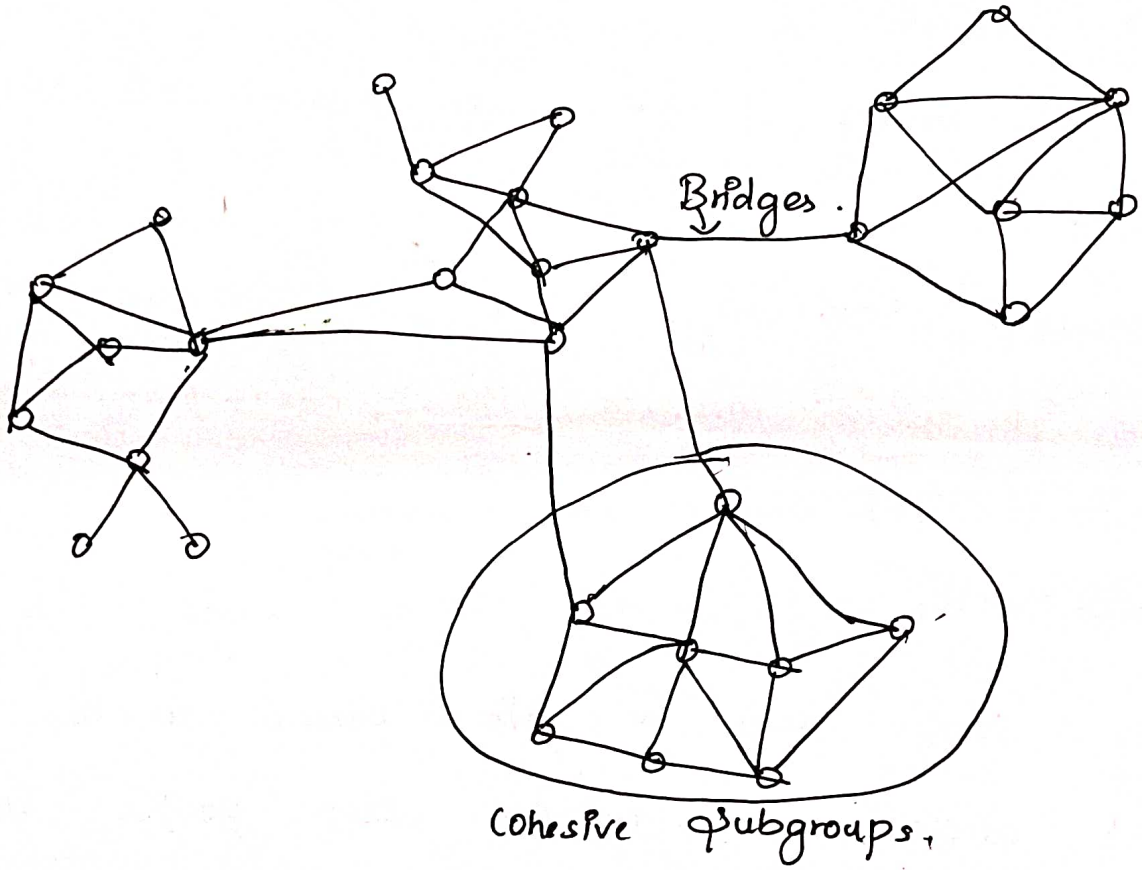
⇒ Directed graph : $(E/N)*N-1$

Where E is the number of Edges. in the network, and N is the ~~N~~ Number of nodes in the network.

Macro - structure of social Network :-

To find the global characteristic of social network

Most real world networks show a structure where densely connected subgroups are linked together by relatively few bridges.



Centrality :-

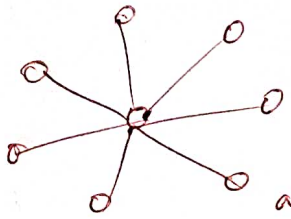
It is a measure indicating the importance of node in the network.

It is further divided into Degree Centrality, Betweeness Centrality and Closeness Centrality.

Degree Centrality :-

It is defined as the number of edges incident upon a node and thus it is usually the first way to calulate the nodes that are most potential to determine other nodes which is present in the network.

* How many people can reach this person directly ?



How many neighbors are directly connected to that node.

Betweeness Centrality :-

It is a measure the connectivity of the neighbor's of your node and to give a higher value for nodes which bridges clusters.

* How likely is this person to be the most direct route between two people in this network.

Closeness connectivity :- How fast reach others in the n/w

* How distinct a node is to the other nodes in the network.

* How fast can this person reach everyone in the network.

Core - periphery (C/P) structure :-

A C/P structure is one where nodes can be divided in two distinct subgroups :

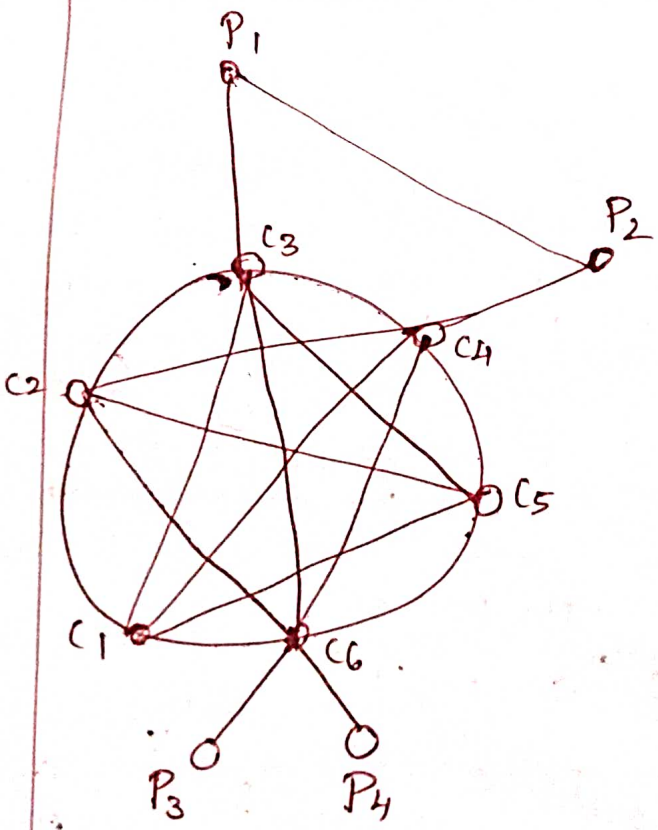
1) Nodes in the core are densely connected with each other

2) Nodes in the Periphery :

While peripheral nodes are not connected with each other, only nodes in the core.

The matrix form of a core periphery structure.

$$\begin{pmatrix} 1 & \cdot \\ \cdot & 0 \end{pmatrix}$$



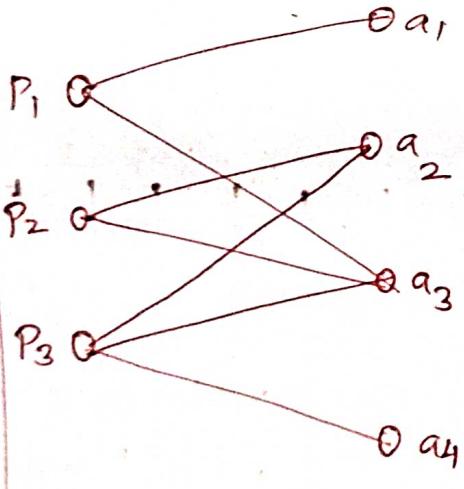
	C1	C2	C3	C4	C5	C6	P1	P2	P3	P4
C1	1	1	1	1	1	1	0	0	0	0
C2	1	1	1	1	1	1	0	0	0	0
C3	1	1	1	1	1	1	1	0	0	0
C4	1	1	1	1	1	1	0	1	0	0
C5	1	1	1	1	1	1	0	0	0	0
C6	1	1	1	1	1	1	0	0	1	1
P1	0	0	1	0	0	0	0	0	0	0
P2	0	0	0	1	0	0	1	0	0	0
P3	0	0	0	0	0	1	0	0	0	0
P4	0	0	0	0	0	1	0	0	0	0

Affiliation Networks:-

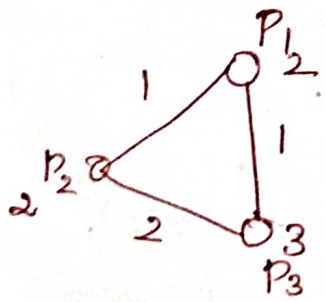
Affiliation networks contain information about the relationship between two set of nodes:-

* a set of subjects.
* a set of affiliations.

An affiliation network can be formally represented as a bipartite graph also known as a two-mode network.



	P1	P2	P3	a1	a2	a3	a4
P1	0	0	0	1	0	1	0
P2	0	0	0	0	1	1	0
P3	0	0	0	0	1	1	1
a1	1	0	0	0	0	0	0
a2	0	1	0	0	0	0	0
a3	1	1	1	0	0	0	0
a4	0	0	1	0	0	0	0



	P1	P2	P3
P1	2	1	1
P2	1	2	2
P3	1	2	3

Personal Networks.

Social Capital. Definition.

⇒ Social capital is a set of shared values or resources that allows individuals to work together in a group to effectively achieve a common purpose.

⇒ Social capital can also be thought of as the potential ability to obtain resources, favors, or information from one's personal connections.

⇒ Social capital allows a group of people to work together effectively to achieve a common purpose or goal.

⇒ It allows a society or organization, such as corporation or a nonprofit, to function together as a whole through trust and shared identity, norms, values and mutual relationship.

Personal Networks.

structural Dimension :-

Relat structural dimension of social capital refers of patterns of relationships or positions that provide in terms of accessing large, important parts of network.

Relational Dimension :-

Relational dimensions of social capital which concern the kind of personal relationships that people have developed with each other through a history of interactions.

Cognitive dimension :-

Cognitive dimensions of social capital refers to those resources providing shared Representation, Interpretations and system of meaning.

Structural Hole :-

It occurs in the space that exists between closely clustered communities.

1.8 Historical overview of Privacy and Security :-

Early Days :-

Back when computers were just starting, privacy and security weren't big worries. There weren't many connections between computers. So things were simpler.

Internet Boom :-

When the Internet got in the 1990s, people started caring more about privacy and security. They came up with basic ways to keep information safe when sending it over the web.

Social Media Takes off :-

As social media like Myspace and Facebook got popular, people started sharing a lot more personal stuff online. This made everyone think more about how their info was being used and who could see it.

Big Data Breaches :-

Some big companies had major data breaches, where hackers got into their systems and stole lots of personal info.

This made government make new rules, like GDPR in Europe and CCPA in California, to protect peoples' privacy better.

5. Better Security Tools :-

Over time, new tech made it easier to keep data safe online. Things like encryption and two factor authentication became common ways to protect sensitive info.

6. Privacy - Friendly options :

Recently, more services focused on privacy have popped up. These are alternatives to big companies that promise to keep your data safer, like using apps that don't track your every move.

7. Ongoing Debates :-

Even with all these changes, there are still arguments about how much privacy we should have online and whether it's worth giving up some privacy for convenience.

UNIT II SECURITY ISSUES IN SOCIAL NETWORKS

The evolution of privacy and security concerns with networked technologies, Contextual influences on privacy attitudes and behaviors, Anonymity in a networked world.

2.1 The evolution of privacy and security Concerns with networked technologies :-

Early Concerns:

In the early days of social networks, users were primarily concerned about setting privacy controls and managing who could see their post and personal Information.

Emergence of Cyber Threats:-

⇒ As social networks gained popularity, Cyber Threats such as phishing scams, malware distribution, and spam became prevalent, posting risks to user security.

Phishing :-

⇒ When attackers send scam emails (or text messages) that contain links to malicious websites.

Identity Theft and Hacking :

With the increase in user engagement on social platforms, incidents of identity theft, hacking accounts, and impersonation became significant security concerns.

Privacy Issues :-

Social media companies faced security over privacy practices, including the collection and use of personal data for targeted advertising, leading to concerns about user privacy and data protection.

Data Breaches :-

Instances of data breaches on social networks exposed sensitive user information to unauthorized access, raising questions about the security measures employed by these platforms.

Misinformation and Fake news :-

The spread of misinformation and fake news on social media platforms emerged as a security issues, impacting public trust, influencing opinions, and even posting risks to democratic processes.

Regularity Responses :-

Governments and regularity bodies responded to these concerns by introducing legislation and regulations aimed at safeguarding user privacy, promoting transparency in data practices, and addressing cyber threats on social networks.

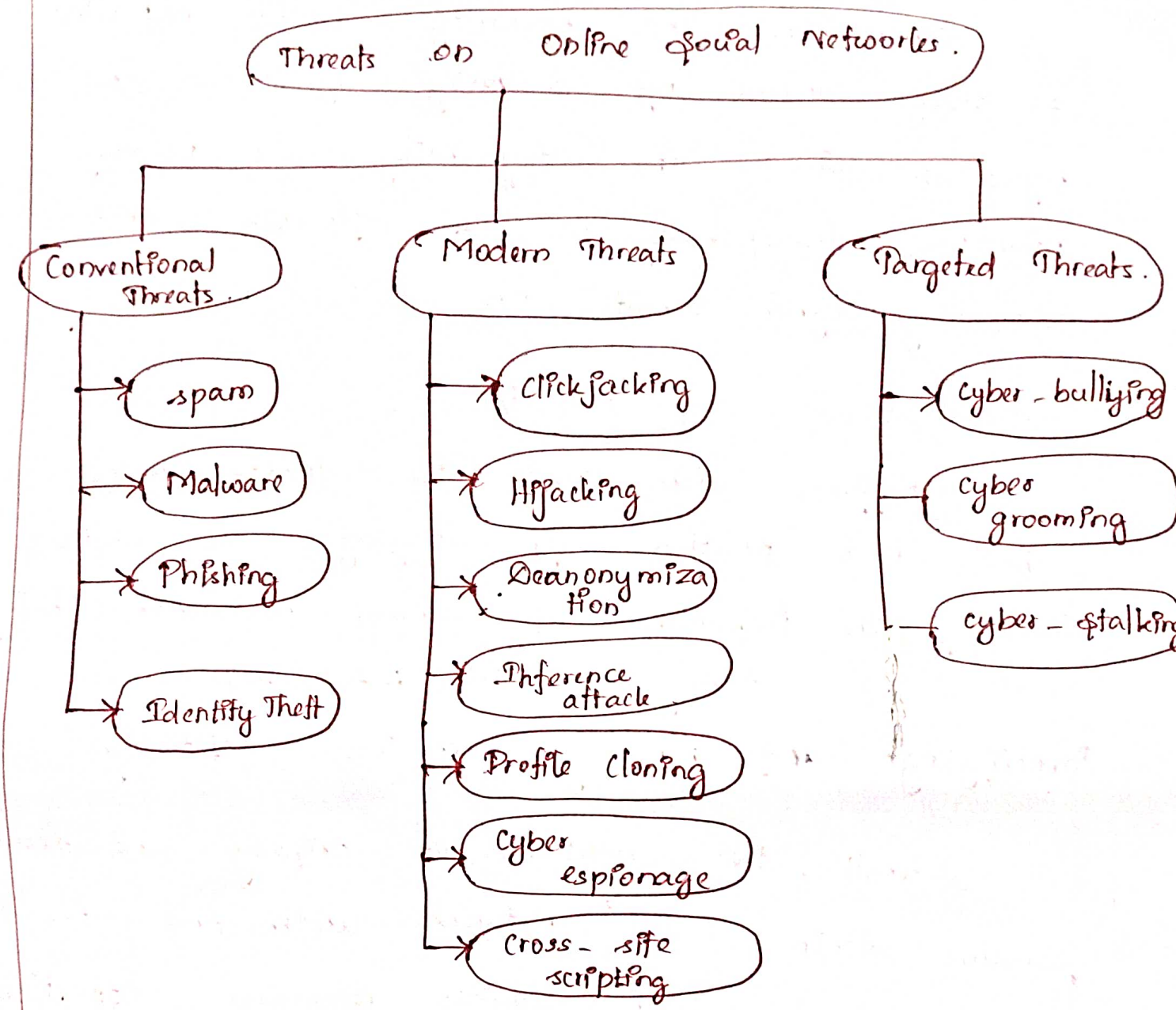
User awarness and Education :-

⇒ Efforts were made to educate users about best practices for staying safe online, including the importance of strong passwords, being cautions of sharing personal information, and identifying and reporting suspicious activity.

⇒ Overall, the evolution of privacy and security concerns within social networks underscores the importance of robust security measures, effective regulation, and user education to mitigate risks and promote a safer online environment.

36

Various Threats on online social network and media.



Conventional Threats :-

Threats that users have been experiencing from the beginning of the social network.

Modern Threats :- are attacks that use advanced techniques to compromise accounts of use.

Targeted Threats :- are attacks that are targeted on some particular user which can be committed by any user for varied personal vendattas.

Spam :-

⇒ Term is used for unsolicited bulk electronic msgs.

⇒ The communication details of legitimate users can easily be obtained from company websites, blogs and newsgroup.

Malware :- → Fake accounts / profile, Social n/w API, Driven by download attack, shortened and Hidden IPAs, Cross-site scripting attacks

⇒ Malware is noxious programming.

⇒ It may get installed by clicking a malicious URL,

⇒ An attacker can inject some malicious script in URLs and clicking on that URLs can make that script run on a system that collect sensitive information from that system.

Phishing :-

It is a kind of social engineering attacks where the aggressor can acquire sensitive and confidential information like username, password and credit card details of a user through fake websites and emails which appear to be real.

Identity theft :-

⇒ The assailant utilizes someone's else's identity like social number, mobile number & address, without their permission to commit attackers.

⇒ With the help of these details, the attacker can easily gain access to a victim's friend list.

2.2 Contextual Influences on privacy attitudes and Behaviors:-

1. Social norms.
2. Perceived Risks.
3. Trust In Platforms.
4. Peer Influence
5. Regulatory Environment
6. Technological Factors
7. Informational Content

Social Norms:-

Description:-

Social norms encompass societal expectations and behaviors regarding privacy with social networks For eg: In some cultures, it may be common for individuals to share information openly, while in others, privacy may be highly valued, leading to more cautious sharing behaviors.

Eg)

In a social group where it's customary to share personal photos and updates regularly, individuals may feel pressured to conform to this norm and share more than they are comfortable with.

2. Perceived Risks: -

Description : -

⇒ Perceived risks refers to individuals subjective assessments of potential negative consequences associated with sharing personal information on social networks. These risks can include identity theft, online harassment, or data breaches.

Eg)

⇒ After hearing about a friends account being hacked on a social network, individuals may become more cautious about the information they share and adjust their privacy settings accordingly

3. Trust In Platforms : -

Description :-

⇒ Trust In platforms reflects users confidence in social network platforms. ability to safeguard their social ~~net~~ personal information and maintain their privacy.

⇒ Factors influencing trust include platform reputation, transparency in data handling practices, and responsiveness to security concerns.

Eg) Users may be more willing to share sensitive information on a social network with a

strong track record of privacy protection and clear privacy policies.

4. Peer Influences: -

Description: -

⇒ Peer Influences refers to the Impact of social connections on individuals. privacy attitudes and behaviors. People may be influenced by the behaviors and sharing practices of their friends, family and social n/w contact.

Eg) If a user observes their friends sharing personal experiences openly on a social network, they may feel more comfortable doing the same, even if they were initially hesitant.

5. Regulatory Environment: -

Description: -

⇒ The regulatory environment encompasses laws, regulations, and policies governing data protection and privacy in social networks. These regulations vary by jurisdiction and can impact platform policies, user rights, and data handling practices.

Eg)

⇒ Compliances with the General Data Protection Regulation (GDPR) in European Union requires social

networks to provide users with greater control over their personal data and transparent data processing practices.

6. Technological Factors: -

Description: -

⇒ Technological factors refers to the design, features, and security measures implemented by social n/w platform to protect privacy.

⇒ These includes encryption protocols, privacy controls, and data encryption methods.

Eg) Social n/w may introduce end-to-end encryption for private messaging to enhance user privacy and security.

7. Information content: -

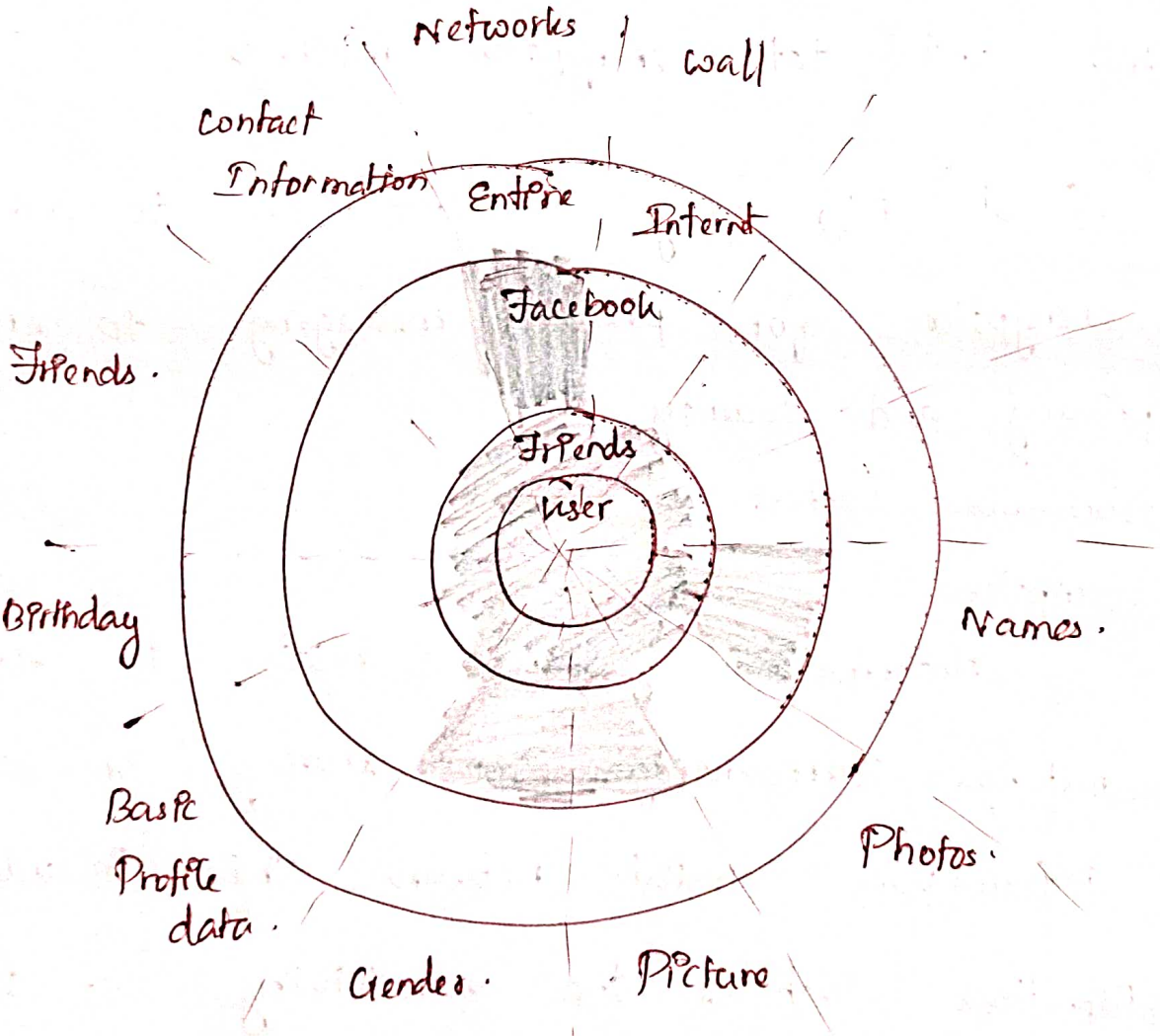
Description: -

⇒ Informational content relates to the circumstances surrounding the sharing or accessing of information social network. This includes considerations such as the sensitivity of the information, the intended audience, and the users over this shared contantent.

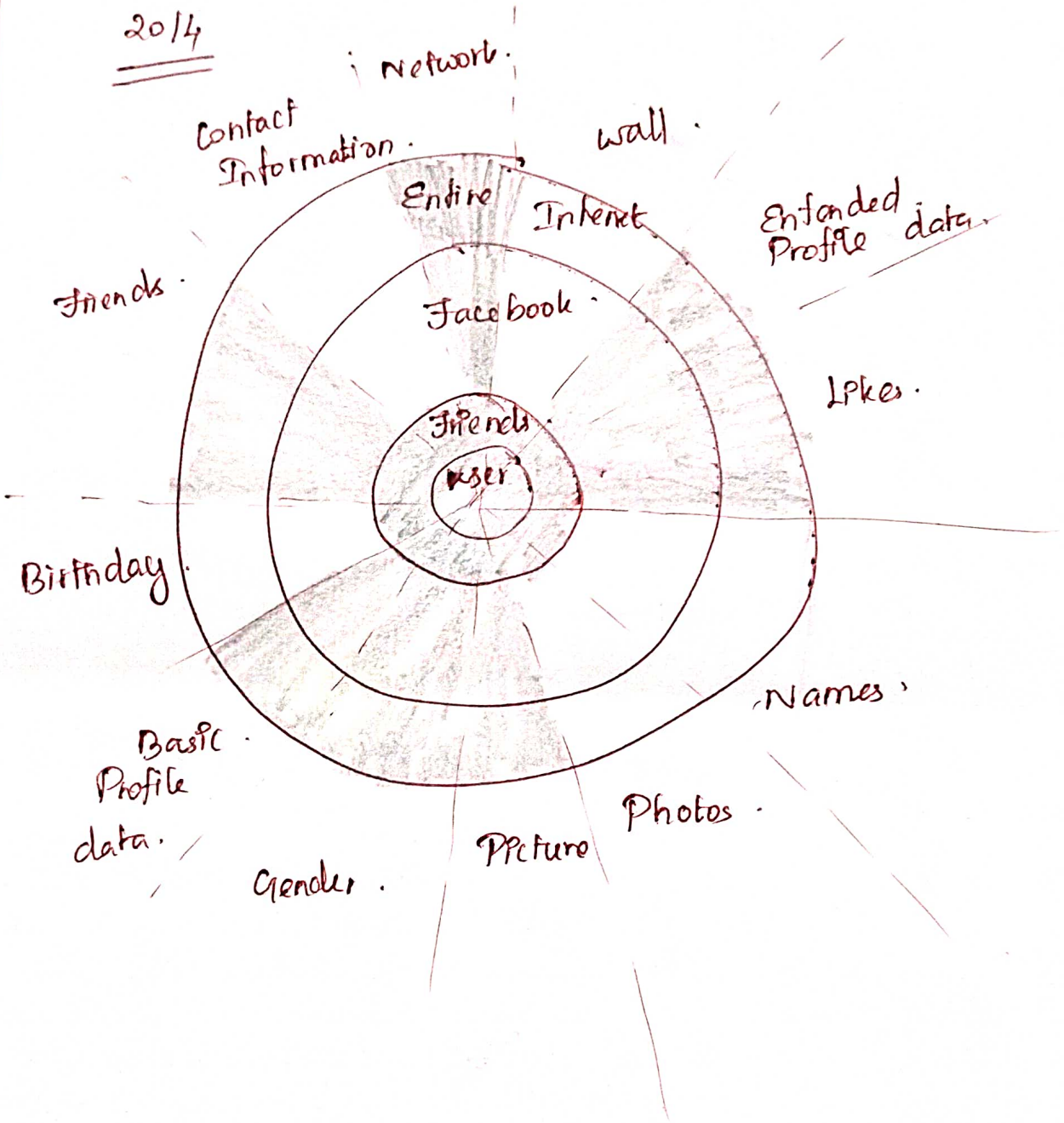
Eg) Users may choose to share personal experience publicly on their profiles but limit access to more sensitive information, such as financial details, to a select group of ~~tra~~ trusted contacts.

Default visibility settings for social media over time :-

☑ Visible (default settings) ☐ Not visible.



2014


Network
Contact Information
Wall
Entire Internet
Extended Profile data
Friends
Facebook
Likes
Friends
User
Birthday
Names
Basic Profile data
Photos
Gender
Picture