Question 1: Describe How Social Networks may be represented using Semantic Web Technologies (Lecture 2-1)

Answer: Semantic Web Technologies offers a means to represent social networks in a structured and interconnected way, enhancing the representation of user profiles, relationships, and content. By utilizing machine-readable formats and linking data, Semantic Web technologies enable the establishment of more significant connections and the application of semantic reasoning to social network data. To describe social networks using semantic web technologies, we can use the following Ontologies and Ontology languages like RDF and OWL:

1) Ontology:

Ontologies are used to define the vocabulary and relationships in a domain. In the context of social networks, an ontology can be created to represent concepts such as users, friendships, posts, comments, likes, and other relevant entities. The ontology defines the properties and relationships between these entities, forming a structured representation of the social network domain.

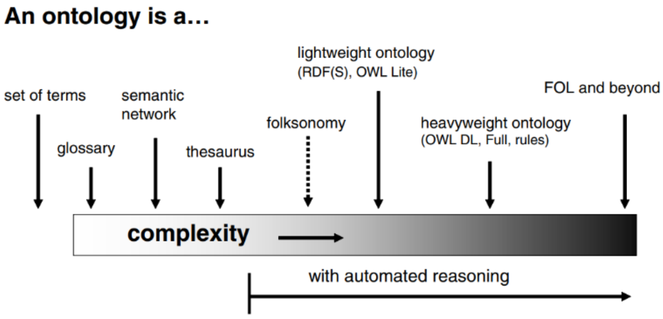


Figure 1.1: Ontologies can be organized according to complexity taken from [1]

Ontologies require a shared understanding within a community, distinguishing them from personal models like database schemas or UML class diagrams. Ontologies can range in complexity, from simple glossaries and controlled vocabularies to semantic networks and thesauri, with varying levels of logic-based reasoning and hierarchical structures. Lightweight ontologies provide minimal descriptions of classes, instances, and properties, while heavyweight ontologies offer precise composition and constraints. At the highest complexity level, knowledge bases employ first-order logic (FOL) to define concepts and their relationships in detail, allowing for advanced reasoning tasks.

Ontology Languages for the Semantic Web:

1. Resource Description Framework (RDF): RDF is a standard data model for representing information in the Semantic Web. It uses subject-predicate-object

triples to express statements about resources. In the case of social networks, user profiles, friendships, and content can be represented as resources, and their attributes and relationships can be expressed using RDF triples.

2. RDF Schema (RDFS) and Web Ontology Language (OWL): RDFS and OWL provide additional expressiveness to RDF by allowing the creation of more complex ontologies and defining constraints and rules. These languages enable the specification of data types, sub-class relationships, property hierarchies, and domain-specific rules, enhancing the representation and reasoning capabilities of the social network data.
3. SPARQL Query Language: SPARQL is a query language for retrieving and manipulating data stored in RDF format. It allows users to query social network data represented using Semantic Web technologies. With SPARQL, users can search for specific user profiles, explore social connections, retrieve posts or comments based on certain criteria, and perform complex graph-based queries.

FOAF:

FOAF (Friend of a Friend) is an RDF vocabulary for describing people and their relationships on the Semantic Web. It enables the representation of social networks and personal information in a machine-readable format. FOAF defines terms and properties to describe individuals, their details, social connections, and relationships. It facilitates the standardized representation of social network profiles, including attributes like names, email addresses, interests, and homepages. FOAF allows for expressing relationships such as friendships, collaborations, and group memberships. By utilizing FOAF, social network data can be interconnected and shared in a standardized manner, fostering integration across platforms and applications.

Example of FOAF:

```
<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
         xmlns:foaf="http://xmlns.com/foaf/0.1/">
  <foaf:Person rdf:about="https://example.com/profile/john">
    <foaf:name>John Doe</foaf:name>
    <foaf:email>john@example.com</foaf:email>
    <foaf:knows rdf:resource="https://example.com/profile/mary"/>
  </foaf:Person>
  <foaf:Person rdf:about="https://example.com/profile/mary">
    <foaf:name>Mary Smith</foaf:name>
    <foaf:email>mary@example.com</foaf:email>
    <foaf:knows rdf:resource="https://example.com/profile/paul"/>
  </foaf:Person>
</rdf:RDF>
```

In this example, we have two individuals, John Doe and Mary Smith. Each person is represented using the ‘foaf:Person’ class. The properties within the ‘foaf:Person’ description provide additional details about each individual. John Doe’s profile includes his name ‘foaf:name’ as "John Doe" and his email address ‘foaf:email’ as "john@example.com". It also states that John knows Mary ‘foaf:knows’, and the relationship is specified by referencing Mary’s profile. Similarly, Mary Smith’s profile includes her name ‘foaf:name’ as "Mary Smith," her email address ‘foaf:email’ as "mary@example.com" and her relationship with Paul ‘foaf:knows’ is indicated by referencing Paul’s profile.

This above example demonstrates how FOAF can be used to represent individuals, their basic details, and social connections in a structured and machine-readable format.

References:

[1] P. Mika, Semantic Web and Social Networks, Springer, 2008

Question 2: Describe (a) Access Control for Social Networks and (b) the Privacy violations that could occur in Social Networks (Lecture 2-2)
Answer:

(a) Access Control for Social Networks:
Access control for social networks refers to the management and regulation of user access to different resources, features, and information within the social networking platform. It involves implementing measures and policies to ensure the appropriate use, privacy, and security of the platform.
Web-based social networks (WBSNs) are virtual communities on the internet that enable users to connect with others, establish relationships, and share various resources. Over time, many WBSNs have started incorporating Semantic Web technologies like FOAF to represent user data and relationships. This integration facilitates the exchange of information between different WBSNs. However, this adoption has also highlighted the necessity of granting content owners’ greater control over the distribution of their resources. This is essential because their content might be accessible to a much larger community than originally anticipated, emphasizing the importance of ensuring adequate control mechanisms to manage information diffusion.

Access Control Requirements in WBSNs:

Developing an access control model for WBSNs involves addressing challenges in a dynamic environment where users want to share data based on their relationships, which can change over time. The model should support access control based on direct and indirect relationships, considering the length and trust level of the relationship path. Users should be able to define relationship type, depth, and minimum trust for access. Traditional centralized access control can be replaced by making users the administrators of their own data, allowing local authorization based on their policies. To address scalability, a verification approach inspired by Tim Berners-Lee is adopted, shifting the responsibility to the requesting user for proving access control requirements.

The Proposed Access Control Model:

Resource owners define access control policies, specifying conditions for access based on relationship type, depth, and trust level with other users. An access condition is represented as a tuple that includes the node, relationship type, maximum depth, and minimum trust level required for access. Access rules consist of tuples comprising the object identifier and a set of conditions, all of which must be fulfilled. Multiple rules can be assigned to an object to accommodate various access control requirements. Relationships, conditions, and access rules are transformed into logical formulas to simplify the generation of proofs. The verification process entails comparing relationships, conditions, and access rules with existing data until a valid proof is obtained.

Access Condition: *"Given a social network SN, an access condition cond against SN is a tuple (v,rt, Dmax,t min), where v ∈ VSN ∪ {*} is the node with which the requestor must have a relationship, rt ∈ RTSN ∪ {*} is a relationship type, whereas Dmax ∈ N ∪ {*} and t min ∈ TSN ∪ {*} are, respectively, the maximum depth and the minimum trust level that the relationship must have. If v = ∗ and/or rt = ∗, v corresponds to any user in VSN and/or rt corresponds to any relationship in RTSN, whereas if Dmax = ∗ and/or t min = ∗, there is no constraint concerning the depth and/or trust level, respectively."*

Access Rule: "*An access rule rul is a tuple (oid,cset), where oid is the identifier of object obj, whereas cset is a set of conditions {cond1,...,condn}, expressing the requirements a node must satisfy in order to be allowed to access object obj."*

(b) The privacy violations that could occur in Social Networks:

we are going to explore four factors that contribute to privacy violations, supplemented by real-life instances from online social networks when applicable. These four factors encompass: restrictions imposed on users, deficiencies or limitations in design, inadvertent disclosure of information, and conflicts of interest.

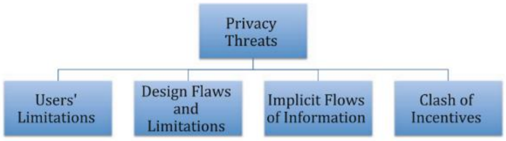


Fig. 2.1 Classification of the causes of uses' privacy leaks

1. Users' Limitations:

Users often make privacy decisions without considering the consequences, due to bounded rationality and limited working memory. Human trust and susceptibility to social engineering attacks also play a role, as users add strangers to their accounts, compromising their intended privacy settings. The text further discusses social bot attacks and targeted friend attacks, demonstrating the willingness of users to accept requests from unknown individuals. The lack of privacy consciousness is not limited to social networks, as users also make irrational decisions in other contexts, such as signing up for loyalty cards or selling their DNA samples for minimal benefits. Overall, these examples highlight the limitations of human decision-making within the bounds of their cognitive abilities.

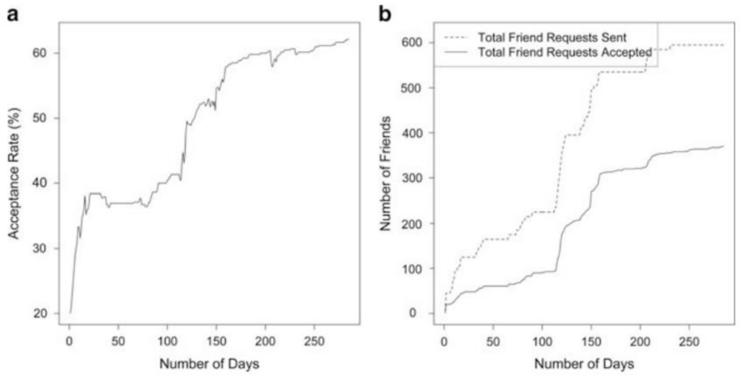


Fig. 2.2 (**a**) Probability of acceptance of our friend requests on Facebook, (**b**) Total number of friends requests sent and total accepted [30]

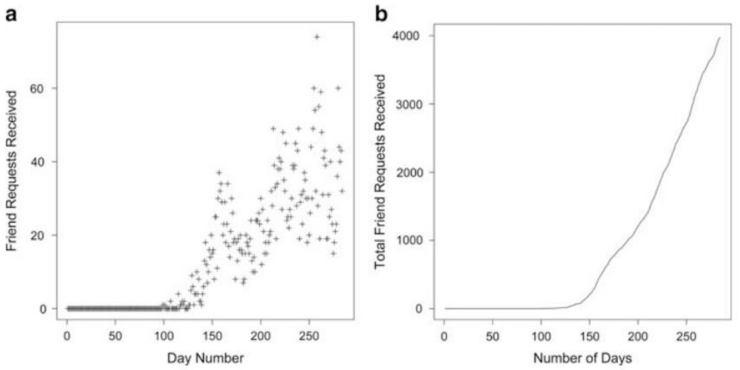


Fig. 2.3 (**a**) Probability of acceptance of our friend requests on Facebook, (**b**) Total number of friends requests sent and total accepted [30]

2. **Design Flaws and Limitations:** Below discussed cases highlight the importance of addressing privacy vulnerabilities and implementing stronger security measures on social media platforms.

   I. Facebook Privacy Settings: Facebook's privacy settings have changed over time, but default settings are rarely modified. The attackers can reconstruct friend-lists using public information. Changes to privacy settings, like Timeline, have limited control over mutual friends and cover photos. Users should have more control over content sharing settings.

   II. Fake Accounts and Cloning Attacks: Social networks lack mechanisms to verify user account authenticity. The attackers can create fake accounts using victims' personal information. The cloning attacks involve creating copies of real accounts. The automated identity theft attacks have been demonstrated.

   III. Permanent Takeover of a Facebook Account: Attackers can permanently take over an account by associating the victim's email address with a new account. Facebook should disallow association of used email addresses with new accounts.

   IV. Facebook's Deactivated Friend Attack: Attackers add victims on Facebook and deactivate their own accounts to remain invisible. They can access private information when reactivating. Facebook made deactivated friends visible on the friend list to address this vulnerability.

   V. Google+ Photo Metadata: Google+ photo metadata, like owner's name and capture details, can impact privacy. Metadata has been used in legal cases to refute claims. The camera make and model can make individuals targets for theft.

   VI. Zuckerberg's Photo Leak and Other Attacks on Facebook: A 2011 security breach resulted in the leak of Mark Zuckerberg's photos. Another vulnerability allowed malicious JavaScript on user profiles. Dhingra and Bonneau found limited hacks to access Facebook photos.

3. **Clash of Interest:** Social networks rely on advertising revenue, which conflicts with user rights. Users want their data protected, while advertisers seek extensive data for targeted ads. The behavioral targeting raises concerns about radicalization. Third-party apps and government requests can complicate privacy and balancing user privacy, revenue, and regulations is challenging for service providers.

4. **Implicit Flows of Information:** Implicit information flows on social networks can unintentionally reveal user details. Analyzing factors like graduation years and

friend connections can accurately estimate users' ages. Examining "likes" and video preferences can expose personal interests and potentially sensitive information. Even seemingly harmless actions, such as brand preferences, can disclose personal characteristics or health conditions. Defending against implicit information flows is difficult as they introduce numerous unknown correlations, making them attractive to adversaries.

References:

Answer 1: https://link.springer.com/chapter/10.1007/119153072_80

Answer2: https://link.springer.com/book/10.1007/978-3-7091-0894-9

Question 3: Describe aspects of Social Media Analytics (Lecture 2-3).
Answer: Aspects of Social Media Analytics: Location Mining. We are going to discuss Tweeque approach based on social graphs for Location Mining.

- **Tweeque: Spatio-Temporal Analysis of Social Networks for Location Mining Using Graph Partitioning:** This text discusses the reluctance of social media users to disclose their locations due to privacy and security concerns. The paper that we have taken the **reference from**, introduces a novel algorithm called Tweeque, which predicts the current location of users based solely on their social network data. Unlike previous approaches, Tweeque considers geospatial proximity, friendship connections, and migration patterns. The algorithm utilizes graph partitioning to identify social groups and incorporates time as a factor in predicting the user's most recent city location. Extensive experiments were conducted to demonstrate the accuracy and efficiency of the system.

  INTRODUCTION: Social networks are popular and widely researched. While location connects the physical and online worlds, facilitating social connections and revealing insights into user behavior, protecting location privacy is important to prevent unwanted interest and information sharing. Trustworthiness is crucial in social media, as demonstrated during protests like the "Arab Spring." Social media impacts marketing by enabling communication and quick feedback. However, accessing user location faces challenges due to security and privacy concerns. The authors' contributions involve exploring geospatial proximity and friendship on Twitter, analyzing migration's impact on location prediction, and proposing algorithms for social group identification. The Tweeque approach surpasses existing methods with implicit temporal analysis.

  GEOSPATIAL PROXIMITY AND FRIENDSHIP: The research investigates the impact of geographical proximity on the probability of friendship in Twitter. Despite the internet's global reach, users tend to form connections with individuals from their offline social circles. The study aims to comprehend the distinct characteristics of Twitter relationships in comparison to platforms like Facebook and LiveJournal. The directionality of relationships, presence of celebrities and spammers, and the requirement for mutual following redefine the concept of friendship on Twitter.

Experimental analyses are conducted to examine the association probability and geographic distance between pairs of users. The findings demonstrate similarities to previous studies on LiveJournal and Facebook, revealing a power law distribution of friendship probability based on distance.

EFFECT OF MIGRATION: In this section, the paper explores experiments and studies that highlight the significant movement of people and the need for temporal analysis to accurately predict a user's current location. It collected demographic information from over 300,000 public Facebook profiles of users in the United States, focusing on their age, hometown, and current location. The data showed that most users, ranging from 63% to 72%, no longer reside in their hometowns. It also analyzed migration rates among Twitter users, dividing age groups into 10-year buckets. Notably, there was a high migration rate, exceeding 9%, among users aged 20 to 29, aligning with the common tendency to relocate after completing education. Furthermore, an examination of Twitter's demographics revealed that individuals aged 25 to 34 constituted a significant portion of the user population. Based on these observations, we conclude that Twitter users have a propensity to migrate. Therefore, while accurately predicting a user's current location, migration should be considered as an important factor by any algorithm.

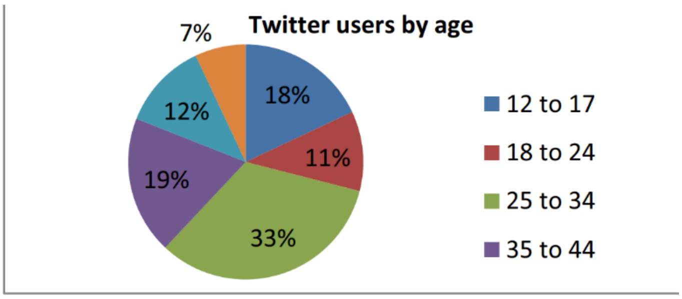


Figure 4. Distribution of Twitter users according to age

TEMPORAL DATA MINING: The paper presents a theory in social science based on two observations: the formation of friendships in cliques and the migration of individuals over time. According to this theory, by identifying and studying social cliques, it is possible to ascertain a user's current location. The study explains the process of dividing the social graph into separate cliques and introduces the concept of purity to determine the current location within each clique. It also introduces the Shi-Malik algorithm for partitioning the graph and a voting algorithm based on purity to determine the final location. It highlights the importance of aggregating locations and boosting concepts at different levels of specificity, such as city, state, and country.

EXPERIMENTS AND RESULTS: We now discuss the quality of algorithms evaluated in the paper. The paper describes the evaluation of the algorithm for forming social cliques and the accuracy of the location prediction algorithm. The evaluation involved handpicking a group of known Twitter users, forming social cliques through graph partitioning, and verifying the correctness of each group. The results showed promising accuracy for the graph partitioning algorithm, indicating its usefulness in obtaining social cliques. The location prediction algorithm demonstrated an accuracy of 76.3% in correctly predicting the current city of a user, surpassing the accuracies of Tweethood and Tweecalization. The average size of a city group was found to be 1.82, and the algorithm achieved an accuracy of 84.9% at the country level, outperforming the content-based approach and the other algorithms.

CONCLUSION: The paper presents an innovative spatial-temporal method called Tweeque, which integrates sociology and data mining to better understand the relationship between geospatial proximity and friendship on online social networks. The approach surpasses previous methods by highlighting the significance of user migration as a hidden factor for temporal analysis. Through experiments, it demonstrates the effectiveness of Tweeque, achieving an accuracy of 76.3% for city-level prediction and 84.9% for country-level prediction. The outcomes of the approach surpass the performance of conventional content-based methods and past social graph-based approaches like Tweethood and Tweecalization.

References:

1. Tweeque: Spatio-Temporal Analysis of Social Networks for Location Mining Using Graph Partitioning

Question 4: Describe Web Services and Cloud-based Infrastructures for Social Networks (Lecture 2-4)

Answer:

Web Services: "*A Web Service is a software system designed to support interoperable machine-to-machine interaction over a network. It has an interface described in a machine-processable format (specifically WSDL). Other systems interact with the Web service in a manner prescribed by its description using SOAP messages, typically conveyed using HTTP with an XML serialization in conjunction with other Web-related standards.*"

A web service is implemented by a concrete agent that sends and receives messages, while the service itself represents the abstract set of functionalities provided.

The requesters are entities that wish to use a web service, while providers are entities that offer the service through their agents. The success of message exchange between requesters and providers relies on their agreement on semantics and mechanics.

The mechanics of the message exchange are documented in a web service description (WSD), which specifies the interface, message formats, transport protocols, and network locations. The semantics refer to the shared expectations and behaviors of the web service, representing the contract between requesters and providers.

There are different types of web services, including SOAP-based web services and RESTful web services. SOAP (Simple Object Access Protocol) uses XML for message formatting and relies on protocols such as HTTP, SMTP, or others for transport. It provides a strict and structured approach to web service communication. On the other hand, REST (Representational State Transfer) is an architectural style that uses lightweight protocols like HTTP for communication. RESTful web services are more flexible and loosely coupled, making them popular for building scalable and efficient APIs.

Cloud-based Infrastructures for Social Networks:

Cloud-based infrastructures for social networks refer to the use of cloud computing technologies and services to support the storage, processing, and delivery of social networking platforms. These infrastructures leverage the scalability, flexibility, and cost-efficiency of cloud computing to meet the growing demands of social network users.

Cloud computing providers offer their services according to three fundamental models Infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS) where IaaS is the most basic and each higher model abstracts from the details of the lower models.

There are four types of cloud deployment models that can be utilized for social media networks:

Public Cloud: In a public cloud deployment, the social media network infrastructure is hosted and operated by a third-party cloud service provider. The infrastructure and resources are shared among multiple organizations and users. This model offers scalability, cost-efficiency, and ease of management since the cloud provider handles infrastructure maintenance and upgrades.

Private Cloud: In a private cloud deployment, the social media network infrastructure is dedicated to a single organization. The infrastructure can be managed internally by the organization or hosted by a third-party service provider.

Hybrid Cloud: A hybrid cloud deployment combines elements of both public and private clouds. It involves the integration of on-premises infrastructure with public and/or private cloud resources. In this model, certain components of the social media network infrastructure may be hosted on a public cloud, while others are maintained on a private cloud or on-premises. Hybrid clouds offer flexibility, allowing organizations to leverage the benefits of both public and private clouds. This model is suitable for social media networks that require a combination of scalability, control, and data privacy.

Community Cloud: A community cloud is a type of cloud deployment model that is specifically designed for a particular community or group of organizations with shared interests, requirements, or goals. In the context of a social network, a community cloud can be utilized to create a specialized and collaborative platform for a specific community or group of users. It can be operated and managed by the community members themselves or by a third-party service provider.

There are some key elements of cloud-based infrastructures for social networks which includes data storage and management, computing power, scalability and elasticity, content delivery and distribution, security and privacy and integration with third-party services.

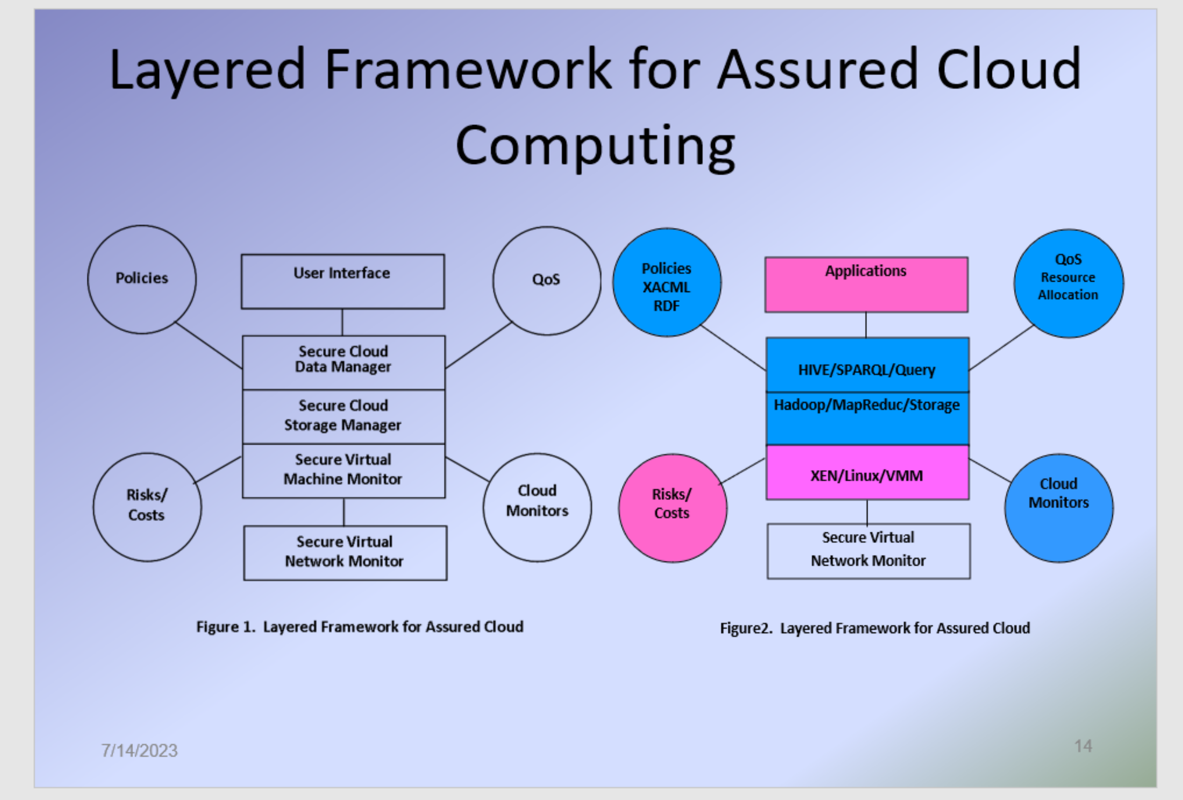
References: Lecture Slides from Lecture 2-4

Question 5: Describe Cloud-based Assured Information Sharing in Social Networks (Lecture 2-5)

Answer:

Cloud-based assured information sharing in social networks pertains to the utilization of cloud computing technology for the purpose of enabling secure and dependable exchange of information within social network platforms. This approach involves harnessing the capabilities provided by cloud infrastructure and services to guarantee the confidentiality, integrity, availability, and privacy of shared data among users of social networks.

The Assured Information Sharing Approach is a method that aims to improve trust and collaboration among various entities through policy and incentive-based information sharing. It involves integrating and analyzing Medicaid claims data, enforcing policies to measure information loss, and identifying reliable partners for sharing. Additionally, it includes evaluating the incentives and risks associated with sharing data, considering factors like data privacy, security, and the potential benefits for all parties involved. Implementing this approach enables organizations to create a secure and efficient framework for information sharing, while also addressing concerns regarding privacy, trust, and compliance.



Layered Framework for Assured Cloud Computing

Figure 1. Layered Framework for Assured Cloud

Figure2. Layered Framework for Assured Cloud

7/14/2023

14

Figure: Layered framework for assured cloud

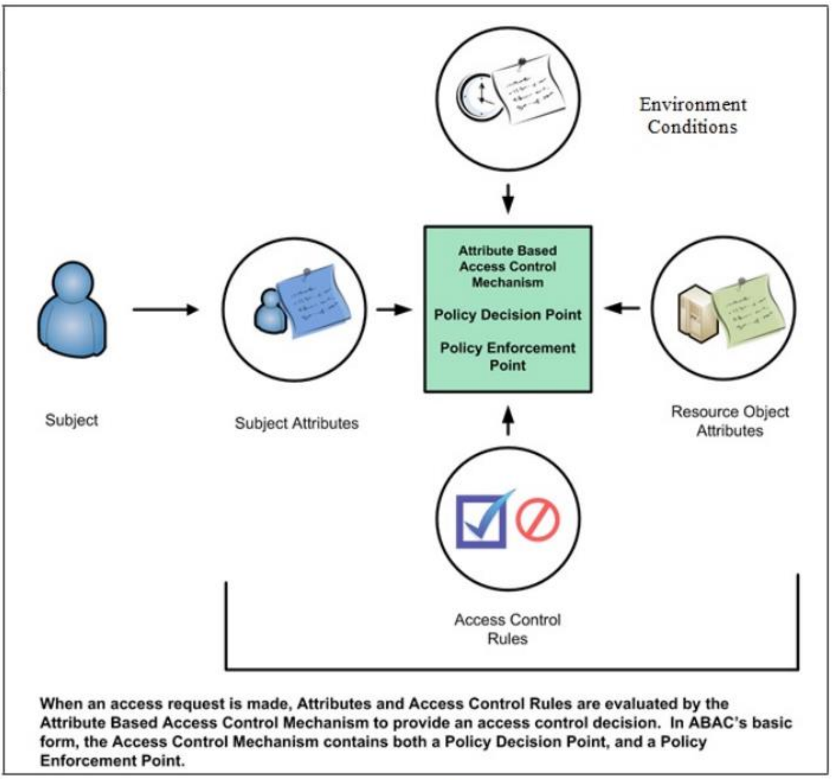
[from lecture slides 2-5, slide 14]

A layered framework for assured cloud computing in social networks is a structured approach that provides a systematic and comprehensive way to ensure the security, reliability, and privacy of cloud-based services within social networking environments. The framework consists of different layers, each addressing specific aspects of assurance in cloud computing for social networks.

Question 6: Describe Attribute-based Access Control and XACML

Answer:

ABAC, or Attribute-Based Access Control, is a method of access control that uses attributes or characteristics of users, objects, and environmental conditions to make authorization decisions. In ABAC, access control policies are defined based on attributes such as user roles, job titles, location, time of access, and other relevant factors. These attributes are evaluated during the access request process to determine whether the requested operation should be allowed or denied. ABAC provides a flexible and granular approach to access control, allowing organizations to define fine-grained policies based on specific attributes and conditions. It is commonly used in complex systems and environments where access control requirements are dynamic and diverse.



When an access request is made, Attributes and Access Control Rules are evaluated by the Attribute Based Access Control Mechanism to provide an access control decision. In ABAC's basic form, the Access Control Mechanism contains both a Policy Decision Point, and a Policy Enforcement Point.

Figure 3: Core ABAC Mechanisms

Figure: ABAC [1]

ABAC mechanism involves policy definition based on attributes, access request generation, attribute evaluation against policies, policy decision making, access enforcement, and audit/logging for accountability and security purposes.

XACML:

XACML, also known as the eXtensible Access Control Markup Language, is an established language and framework used in information systems to establish and enforce access control policies. It offers organizations a versatile and scalable method for defining detailed policies that dictate the who, what, and when of resource access.

At its core, XACML operates on a policy-based access control model, where access control decisions are driven by policies expressed in XML format. These policies outline the specific rules and conditions that determine whether an individual should be granted or denied access to a particular resource.

The key components of XACML include:

Policy Decision Point (PDP): The PDP evaluates access requests and makes decisions based on the defined policies. It acts as the central authority responsible for enforcing access control.

Policy Enforcement Point (PEP): The PEP is the component that intercepts access requests and forwards them to the PDP for evaluation. It enforces the decisions made by the PDP by either granting or denying access.

Policy Information Point (PIP): The PIP provides additional information that may be required during the evaluation of access requests. It acts as a data source for the PDP, supplying attributes and context information.

Policy Administration Point (PAP): The PAP is responsible for managing and defining the access control policies. It allows administrators to create, update, and delete policies according to the organization's requirements.

XACML policies consist of a set of rules that define the conditions under which access is allowed or denied. These rules consider various attributes, such as user roles, resource attributes, environmental factors, and time of access. XACML also supports the concept of obligations and advice, which allow for additional actions to be performed after access decisions are made.

Overall, XACML provides a standardized approach for implementing fine-grained access control in diverse environments, enabling organizations to enforce their access control policies consistently across different systems and applications.

References: Using XACML for access control in Social Networks

Question 12: Describe Inference/Privacy Control for Social Networks

ANS:

Privacy control in social networks refers to the mechanisms and strategies implemented to protect the privacy of users' personal information and control the access and visibility of their data within the social network platform. It involves giving users the ability to manage their privacy settings, decide who can view their profile and posts, and control the sharing of their personal information with others.

The Inference problem occurs when unauthorized information can be deduced or concluded from available information. If an attacker can infer or obtain highly sensitive private information, it raises privacy control concerns. For instance, an insurance agency deducing that a customer has cancer would be an example of such a problem.

To address this issue, multiple Inference Controllers have been developed over time. The initial Inference Controller compared the responses it generated for a specific query with a history of responses. This comparison aimed to determine whether classified information could be inferred. The user query was protected by policies, and these policies transformed the user query into an input. The Inference Controller then processed the query and generated an output. If the response was clear and did not reveal any information, it was released; otherwise, it was withheld. However, a challenge arose in determining the memory capacity required for the Inference Controller to compare responses effectively. A depiction of this Inference Controller is provided.

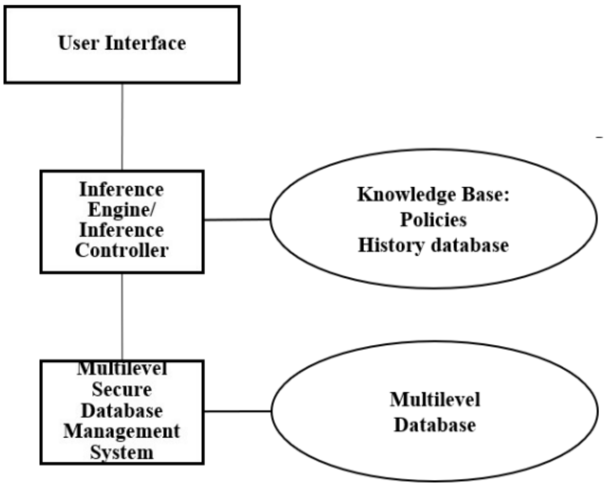

User Interface
Inference Engine/ Inference Controller
Knowledge Base: Policies History database
Multilevel Secure Database Management System
Multilevel Database

Figure: from [1]

As the inference controller had the capability of policy-based information sharing, the CAISS++ architecture, including CAISS [3], could be expanded to accommodate social network systems. This extension involves specifying policies in formats such as XACML and RDF, and utilizing translators to convert these policies. By using a policy engine, role-based access control can be provided. The architecture of CAISS is illustrated below.

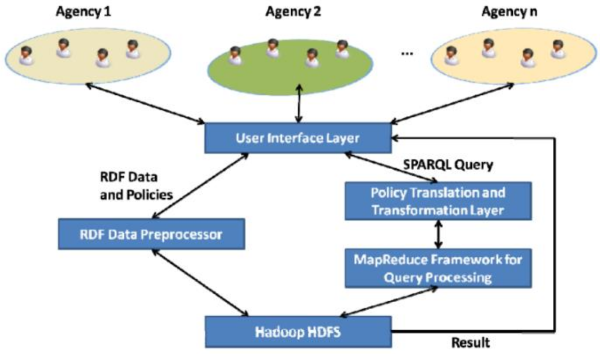


Figure: Taken from [1]

References:

[1]CS6301 Lecture #3-7 Course Material, Dr. Bhavani Thuraisingham, The University of Texas at Dallas

[2] Bhavani M. Thuraisingham, William R. Ford, Marie Collins, J. O'Keeffe: Design and Implementation of a Database Inference Controller. Data Know. Eng. 11(3): 271-297 (1993)

[3] Tyrone Cadenhead, Murat Kantarcioglu, Vaibhav Khadilkar, Bhavani M. Thuraisingham: Design and Implementation of a Cloud-Based Assured Information Sharing System. MMM-ACNS 2012: 36-50

Question 13: Describe three attacks to Social Networks

Answer:

Attacks on the Social Media Platform:

In these attacks, the attacker exploits social media to gather information from users. For example, in a De-Anonymization attack, the attacker aims to obtain personal information by leveraging publicly available group data. By analyzing a user's browser history, the attacker can determine their social media group memberships and ultimately identify the user. Another attack, known as XSS (Cross Site Scripting), involves the insertion of malicious code into comments. When a user views a page containing such a comment, the code is executed on their browser, potentially granting the attacker access to sensitive information stored in cookies or session tokens. Major social networking sites like Twitter and Facebook have been targeted by such attacks, as evidenced by the "Over the Rainbow" attack on Twitter, where an executable JavaScript code was embedded and spread through retweeting functionality. Additionally, Evil Twin Attacks occur when an attacker obtains information about legitimate users and then poses as the user by posting information on the website.

Attacks on Computing Systems and Infrastructures:

Databases and high-performance server units are primary targets for these attacks due to the valuable user information they store. Despite password protection, databases can still be cracked through brute force attacks or by exploiting vulnerabilities in software patches. Implementing stronger passwords and keeping software up to date with the latest stable versions can mitigate these attacks. Software infrastructures are also vulnerable to insider attacks, where employees with limited access can tamper with sensitive data. The architecture of the system itself can also create vulnerabilities, with potential exploitation of hooks between the operating system and the database.

Attacks on Social Media Systems:

These attacks include Zero Day Attacks, DDoS (Distributed Denial of Service), and Man-in-the-Middle Attacks. Zero Day Attacks exploit vulnerabilities in software that are unknown to the public, taking advantage of users who haven't updated to the latest version. In DDoS attacks, the server is bombarded with multiple requests, causing it to become unresponsive to legitimate users. These attacks often serve as distractions while the actual targeted attack focuses on system vulnerabilities. Man-in-the-Middle Attacks involve intercepting data between the user and the system without modifying system functionality, making them difficult to detect and diagnose.

These various attacks pose significant threats to the security and privacy of social media users, necessitating robust security measures and countermeasures to protect against them.

References:

[1] G. Wondracek, T. Holz, E. Kirda and C. Kruegel, "A Practical Attack to De-anonymize Social Network Users," 2010 IEEE Symposium on Security and Privacy, Oakland, CA, USA, 2010, pp. 223-238, doi: 10.1109/SP.2010.21.