

Collusion-resistant multiparty data sharing in social networks

Nisha P. Shetty¹, Balachandra Muniyal¹, Nandini Proothi¹, Bhavya Gopal²

¹Department of Information and Communication Technology, Manipal Institute of Technology, Manipal Academy of Higher Education, Karnataka, India

²Department of Chemical Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Karnataka, India

Article Info

Article history:

Received Sep 27, 2023

Revised Dec 6, 2023

Accepted Dec 26, 2023

Keywords:

Collusion attacks

Identity leakage

Multi-party access control

Policy specification and

management

Security model

Social network computing

Strength of interaction

ABSTRACT

The number of users on online social networks (OSNs) has grown tremendously over the past few years, with sites like Facebook amassing over a billion users. With the popularity of OSNs, the increase in privacy risk from the large volume of sensitive and private data is inevitable. While there are many features for access control for an individual user, most OSNs still need concrete mechanisms to preserve the privacy of data shared between multiple users. The proposed method uses metrics such as identity leakage (IL) and strength of interaction (SoI) to fine-tune the scenarios that use privacy risk and sharing loss to identify and resolve conflicts. In addition to conflict resolution, bot detection is also done to mitigate collusion attacks. The final decision to share the data item is then ascertained based on whether it passes the threshold condition for the above metrics.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Balachandra Muniyal

Department of Information and Communication Technology, Manipal Institute of Technology, Manipal Academy of Higher Education

Manipal, Karnataka, India

Email: bala.chandra@manipal.edu

1. INTRODUCTION

Online social networks (OSNs) are social media platforms that enable users to connect and exchange information with one another. The most popular online social networks for content sharing are Facebook, Twitter, Instagram, Pinterest, and LinkedIn. Online social networks have become essential to today's internet culture because they allow individuals to engage with one another in real-time. As of 2022, social media users have spanned around 4.62 billion [1]. With the vast amount of personal information shared, there is a more significant threat of misuse of this private data and media. Each of these different OSNs has privacy policies outlining the level of control and protection provided to every user.

While these social networks have policies protecting the individual user, most do not have mechanisms that allow multiple users involved in a shared data item to have a considerable say in its distribution in OSNs. While users have control over the visibility of the data they share in their own space, they cannot alter the privacy settings of data posted by another person, even if they were tagged. In most OSNs, removing the tag itself is the maximum extent to which they have control. In collaborative data sharing, each user involved will have different concerns regarding the distribution of their data. The dearth of access control mechanisms for multi-party data poses a considerable privacy risk by leaking sensitive information.

Over the years, numerous models have been implemented for multi-party access control (MPAC) in OSNs. Most MPAC management systems are not collaborative, and the onus is placed on a single user.

Generally, the primary owner of the shared data item aggregates the feedback from other stakeholders and comes to a decision. Utilizing game-theoretic methods has been a prominent strategy for addressing MPAC in OSNs. However, these approaches often oversimplify player behavior and may not accurately capture the intricacies of real-world events. They typically focus on specific objectives, such as maximizing utility or minimizing risk, neglecting certain potential outcomes. In contrast, our proposed methodology considers practical factors like interaction strength and identity leakage (IL), offering a more precise and realistic analysis of policy conflicts.

Unlike game-theoretic approaches that can be complex and challenging to grasp, our suggested method is straightforward and user-friendly. It provides a more accessible framework for a broader audience by relying on easily understandable and explainable measurements. This simplicity does not compromise the analysis's accuracy but enhances its applicability. Our methodology stands out by being adaptable to a variety of policy scenarios, including those involving non-rational actors or non-zero-sum interactions, which may not be suitable for traditional game theory approaches. The versatility of our proposed methodology arises from its ability to analyze a broad spectrum of policy conflicts, coupled with the computational efficiency of the model. This approach, therefore, offers a practical and efficient solution for examining diverse policy scenarios in the context of OSNs.

In response to the escalating privacy challenges within OSNs, our proposed method strategically employs pivotal metrics, notably identity leakage and interaction strength. These metrics serve as instrumental tools to finely tune scenarios, enhancing conflict resolution by precisely identifying and addressing issues related to privacy risk and sharing loss. Recognizing the critical importance of fortifying OSN environments against collusion attacks, our approach integrates a sophisticated bot detection system. The primary research question revolves around how the utilization of identity leakage and interaction strength metrics, coupled with robust bot detection, can effectively preserve user privacy and resolve conflicts in OSNs. This investigation seeks to contribute to a more secure and privacy-conscious online social experience by comprehensively addressing the nuanced challenges posed by the dynamic landscape of OSNs.

Section 2 documents the survey process carried out for the entirety of the paper. Subsection 3.1 introduces the terminologies used throughout the paper, and subsection 3.2 briefs the policy resolution process flow for shared data. Subsection 3.3 covers the standard approach to solving multi-party access control problems using two metrics-privacy risk and sharing loss. Subsection 3.4 talks about introducing a new metric-the strength of interaction. As the name suggests, this metric calculates the strength of the relationship between two users online based on their interactions. It factors in both the quality and quantity of the interactions between users. Identity leakage is introduced in subsection 3.5, and it aims to calculate if a user's identity is being compromised in the data shared online. This is useful in the case of non-consensual sharing of data and can help make the sharing restrictions stricter in case of a conflict. Section 4 covers collusion attacks-often used to exploit automated multi-party access control systems. Specifically, we cover social bot-driven collusion attacks and mitigate them using bot detection techniques. The application of the proposed method is analyzed in various scenarios in section 5. Section 6 concludes the research and suggests appropriate future directions for study in this domain.

2. RELATED WORKS

Hu *et al.* [2] provided an MPAC model to deal with privacy concerns for collaborative data sharing. They also developed a policy specification scheme, a mechanism for evaluation, and a proof-of-concept implementation of the model known as M-Controller. Xu *et al.* [3] proposed a trust-based mechanism to tackle collaborative privacy management where the user can decide whether to post based on the combined opinions of all the stakeholders involved in the shared data item. The parameters for choosing the tradeoff between users are privacy risk and sharing loss-determined through a multi-armed bandit problem which is solved using upper confidence bound policy. Akkuzu *et al.* [4] proposed a system for collaborative privacy management using fuzzy logic decision-making through metrics such as data sensitivity value and confidence value of the targeted group. Users also use trust values to determine reputation value. Based on the socio-technical design paradigm and the social relations model, Ahmad *et al.* [5] developed a model for a personalized multi-party access control mechanism and implemented it using the Facebook application programming interface (API).

Lee *et al.* [6] proposed a fine-grained multi-party access control model so users can control photo-sharing policies on spaces outside of their own. Refinement in the policy was done through a mechanism that could control the face appearance according to users' spatial-temporal information and co-occurrence in photos. They also defined a conflict resolution and policy evaluation scheme to determine the visibility of a face in the photo. Ulusoy and Yolum [7] proposed an agent-based collaborative privacy management system where the standards for privacy are determined on behalf of users whom the agents represent. To achieve

equitable processing of privacy settings and to charge the agents whose privacy settings are selected, the Clarke-Tax method is used for auctioning, and multi-agent simulations are used for evaluation.

Such and Criado [8] suggested various approaches-adaptive, auction-based, and fine-grained-to solve the problem of multi-party access control but highlighted that such tools can never be fully automated and must aim for usability. Cherubini *et al.* [9] proposed a dissuasive and precautionary solution to multi-party privacy conflicts (MPC)-an avenue that aims to avoid MPC before the shared content goes live. Muhammad and Ahmad [10] put forth a joint-sharing approach that uses request evaluations and conflict resolutions resolved by strict, soft, or weak mergers. Some of these solutions rely on votes by the users to accept or deny requests to share information with their initially unintended viewers.

Madeira and Joshi [11] employed machine learning to construct a model for predicting a user's closest friends. The Bayesian network classifier proved to be the most accurate in determining if a group of interactions indicates a connection to a close friend. Krakan *et al.* [12] surveyed to determine the importance of different interaction parameters (likes, comments, chat, tags) and subsequently built a model using Random Forest to quantify relationships in online social networks. They concluded through the model's accuracy that it was feasible to assess friendship intensity through interaction behavior.

Almeny *et al.* [13], [14] introduced two metrics, reachability and audience to combat the privacy risk that arises from the scope of the data item becoming “far-reaching” or “viral” which would increase the likelihood of it to be viewed by silent listeners or invisible audiences. They used centrality metrics to approximate cases in which no information about user activity was found. Domingo-Ferrer *et al.* [15] leveraged game-theoretic concepts to introduce a co-utility framework that relies on cooperation between rational users who aim to help another user achieve their best outcome.

Such and Rovatsos [16] propose an automated negotiation mechanism that uses the concept of intimacy among agents to determine the utility of proposals. The article presents three heuristics to reduce the complexity of the negotiation mechanism, with Greedy-Branch and Bound (BnB) algorithm performing best overall. The authors suggest that future research should consider the role of disclosing items in shaping user preferences and extend the mechanism to consider the intimacy between negotiating parties. The intimacy metric is used to define relationship strength, like the metric strength of interaction. While the authors assume the values of intimacy are available, the proposed approach defines a way to calculate the same.

There are already plenty of existing works of literature that aim to detect and mitigate Sybil attacks. Jethava and Rao [17] proposed a behavior-based and graph-based approach to detect Sybil attacks in OSNs. AL-Qurishi *et al.* [18] proposed a prediction system consisting of three modules, a data harvesting module, a feature extracting mechanism, and a deep-regression model to evaluate user-profiles and mitigate Sybil attacks on Twitter. Our paper aims to tackle social botnets – an organized group of social bots that collude to carry out malicious attacks in OSNs. Zhang *et al.* [19] illustrated the viability and benefits of using a social botnet for spam distribution and digital influence manipulation using practical Twitter experiments and trace-driven simulations while proposing countermeasures for OSNs to improve their detection systems. There is scant work done in this domain, and our approach attempts to fine-tune privacy protection by considering bot-driven collusion attacks.

3. PROPOSED METHOD

3.1. Terminologies and definitions

Throughout this paper, the following terms and descriptions are used:

- a. Proprietor: The proprietor is the user posting the shared content. In most OSNs, editing and privacy controls remain with the proprietor.
- b. Collaborators: Any user whose data (an image, video, or text) is present in the content shared by the proprietor. For the sake of simplicity, the proprietor is not referred to as a collaborator.
- c. Data: This refers to shared content posted by the user. The data shared is related to the proprietor and one or more collaborators.
- d. Target users: Each collaborator shares content with an intended audience of trusted users T_i .
- e. Untrusted Users: Out of the total intended target, each collaborator has a set of users they do not want to share information with represented as UT_i.
- f. Privacy conflict: For any two users i and j , if $T_i \neq T_j$, there arises a privacy conflict. This is likely to happen in a real-world scenario as most online users interact with different audiences (even if that difference is only of a few users). We must find the total number of conflicting users to resolve a conflict. These conflicting users, referred to as negotiating users (NT_{ij}) We either permit or deny sharing content with these negotiating users to resolve this conflict. In doing so, we encounter one of the following scenarios:

- Privacy risk (PR): An estimated indicator of the controller's (user) privacy risk while disclosing a contradicting piece of data is called privacy risk.
- Sharing loss (SL): An indicator of loss observed by the users due to the decision of not sharing the data with their intended audience.

To understand the concept better, let us take the example of Jane and John. They are friends and have a few mutual users in their friend lists online. If Jane posts a picture of herself and John, it will only be visible to her friends online. For John, this is a conflict; since there are some people that Jane is friends with, and John is not. We could either permit or deny sharing this data to resolve this conflict. The former is an example of privacy risk for John since he will be sharing his data with Jane's friends whom he is not friends with. Moreover, the latter is an example of sharing loss for Jane since we deprive her friends of the ability to view the shared data. To resolve this conflict, we calculate both metrics and give preference to either scenario based on the threshold [20].

3.2. Process flow

A multi-party access-friendly online social network platform should allow all collaborators to have a say in the data viewers or target users of the shared content. Currently, most platforms only allow the proprietor of the content to control who gets to view, comment, and share the same. All collaborators' needs and credibility must be considered to compute the best-case scenario. Figure 1 and Algorithm 1 explain the proposed solution to detect and resolve conflict as it arises in a real-world OSNs. Table 1 provides a comprehensive overview of notations used in the entirety of the paper.

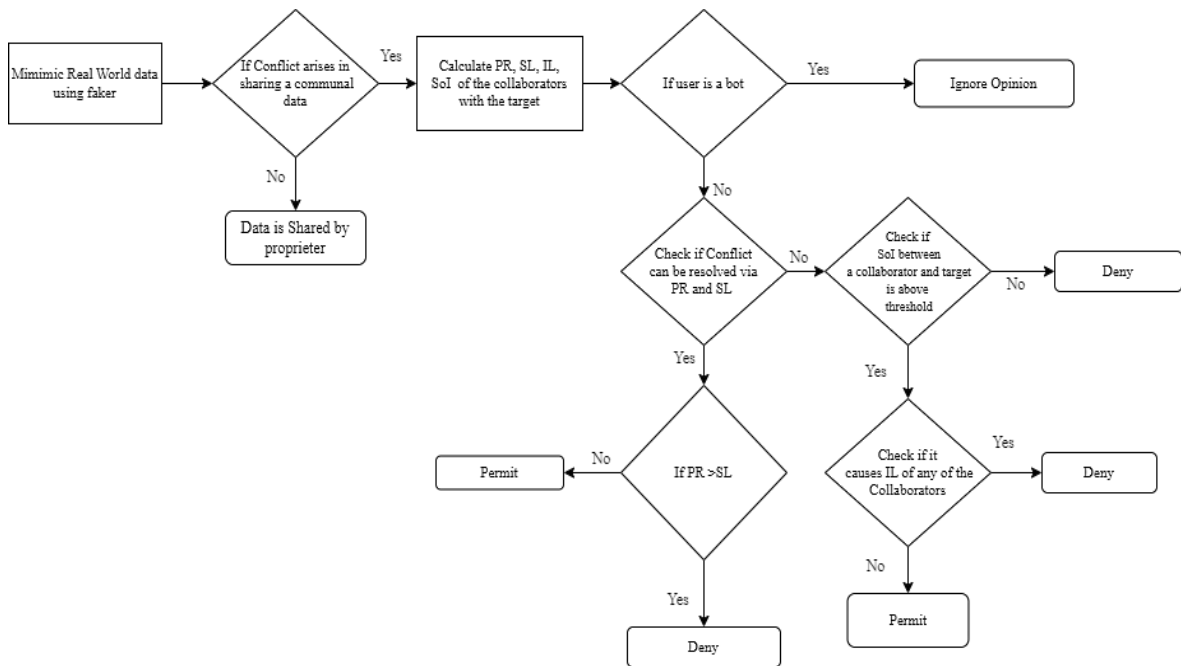


Figure 1. Methodology for identifying and resolving privacy conflicts

Algorithm 1. Proposed method

Input: Users and their friend list

Output: Decision to permit or deny

```

for each shared data item  $i$ 
  for each collaborator  $m$ 
    assess the trusted and untrusted target user for each collaborator  $m$ 
    compute the list of negotiating users  $NU$  for the data item  $i$ 
    check if user  $U \in NU$  is a bot
    if (yes)
      ignore opinion
    else
      compute Privacy risk (PR) and Sharing Loss (SL) between each collaborator  $m$  and every user in the list  $NU$ 
  
```

```

if (PR>SL)
    deny
else
    permit
else
    compute if strength of interaction (S) of each collaborator m and every user in the list NU
    is above the threshold
    if (no)
        deny
    else
        check if data sharing causes identity leakage (I) to collaborator m
        if (yes)
            deny
    else
        permit
return decision permit or deny

```

Table 1. Notations

Symbol	Meaning
T	Target user
UT	Untrusted user
NT	Negotiating user
U	User
F	Frequency of interaction
N	Nature of interaction
W	Weightage
S or SOI	Strength of interaction
PR	Privacy risk
SL	Sharing loss
pc	Privacy concern
sl	Sensitivity level
rep	Reputation
R_u	Recommendations by other users
N_p	Number of posts
$I_{leakage}$	Identity leakage
$G_{wh,wn,wr}$	Associated average group size
ρ	Number of people in the data item
ξ_i	i^{th} activity
Δt_i	time slot of i^{th} activity
γ_i	i^{th} location identifier

3.3. Calculating privacy risk and sharing loss

Let T be the universal set containing the list of target users. Assuming we have three users comprising of a proprietor and two collaborations – U_1, U_2 and U_3 who share data with T . We create another set for each user UT_1, UT_2 and UT_3 representing the users they do not want to share information with, referred to as Negotiating Users. NT_1, NT_2 , and NT_3 represent three sets of negotiating users corresponding to U_1, U_2 and U_3 , respectively. In the case of U_2 and U_3 being collaborators of the shared data and U_1 being the proprietor: PR and SL are calculated based on the negotiating users and sensitivity level-an arbitrary value-of each user. These values are calculated as follows [21], [22]:

- Privacy risk (PR): The privacy risk of collaborative data is said to be the possible harm to the privacy of controllers concerning the shared data item; the higher the privacy risk, the greater the threat to controllers' privacy. The privacy risk is heightened if a greater number of proprietors trust the users in the conflicting segment n_α , the sensitivity of the data is higher and if the data item spreads widely across the network. In the conflicting segment α , the function $PR(\alpha)$ computes the privacy risk associated with the item α , where pc_β and sl_β denote the general privacy concern and sensitivity level chosen by an untrusting proprietor β in (1).

$$PR(\alpha) = \sum_{\beta \in \text{proprietor}_{ut}(\alpha)} (pc_\beta \times sl_\beta) \times n_\alpha \quad (1)$$

- Sharing loss (SL): If the users in the conflicting segment are denied access to view the shared data item, these target users' controllers suffer a loss of data sharing. The overall sharing loss is given by (2).

$$SL(\alpha) = \sum_{\beta \in \text{proprietor}_{ut}(\alpha)} (1 - pc_\beta) \times (1 - sl_\beta) \times n_\alpha \quad (2)$$

3.4. Strength of interaction

In a real-world OSN, the interaction between users takes place by sharing data. Most online platforms offer users the ability to act on shared data, including likes, comments, and shares. Different platforms offer additional interaction methods and call them different names, too (Tweet on Twitter, Post on Instagram). The proposed approach covers standard actions of communication, and this approach can be extended to other forms of actions as well.

The strength of interaction (SoI) is commonly used in ecological research to understand the interactions between different organisms. Garcia *et al.* [23] used the strength of interaction between different researchers to assess the effect of a collaborative network in biomedical research. In this study, the strength of interaction between researchers was given by the number of times they have worked together on a project. In our methodology, the strength of interaction between users depends on the quality and quantity of the interactions. By ignoring either of those metrics, we risk the quality and validity of our results. For example, Denise comments on many of Jane's posts but all of Denise's comments are very negative. If we rely only on the quantity of interaction, we could conclude that Denise and Jane have a high strength of interaction and, in turn, a high trust factor. However, the quality of Denise and Jane's interactions suggests quite the opposite. Similarly, let us take the example of Alfred, who liked Jane's post once. Liking someone's post is positive and indicates that Alfred shows positive feelings toward Jane. However, this was only done once, and it is incorrect to assume that Alfred and Jane share a high level of trust.

While considering this metric of the strength of interaction, it is essential to note that establishing trust and high SoI is a two-way street. Suppose we take the example of Ethel and Jane. Let us say Ethel intermittently likes Jane's posts, tags Jane in comments and posts with a positive implication and even shares Jane's posts online. However, Jane does not reciprocate any of these actions toward Ethel. This does not imply a high strength of interaction between Ethel and Jane. Thus, while calculating the metric, it is seen from the point of view of both users.

While dealing with textual data, we rely on sentiment analysis to understand the nature of the interaction (N). The values can range from $[-1,1]$, with -1 indicating a negative interaction and 1 indicating a positive interaction. The frequency of interaction (F) is a weighted metric representing the quantity of interaction but also relies on the type of interaction, which can be likes, comments, shares, and posts. For ease of calculation, each type of interaction has been assigned a weightage (W) as depicted in Table 2.

Table 2. Weightage of types of interactions in OSNs

Type of interaction	Weightage of interaction (W)
Like	1
Comment	3
Share	5
Post	10

Note that these values are relative and can be changed according to the platform and the significance of these actions. This is not an exhaustive list and does not represent all types of interactions between users. For each engagement, Equation (3) calculates the strength of interaction as:

$$S_i = N * (F * W) \quad (3)$$

For calculating the strength of interaction from one user's end where i represents a singular interaction of any type and n is the total number of interactions, use (4):

$$S_i = \sum_{i=1}^{i=n} N_i * (F_i * W_i) \quad (4)$$

For calculating the cumulative strength of interaction between two users, use (5):

$$S_{ij} = \frac{\sum_{i=1}^{i=n} N_i * (F_i * W_i) + \sum_{j=1}^{j=n} N_j * (F_j * W_j)}{2} \quad (5)$$

While this formula accurately covers most social interactions, an edge case can be understood by this example of Jane and John. Assuming Jane has liked all of John's pictures - indicating overall positive emotions towards John. And say John has only liked one of Jane's many pictures. In this scenario, it is likely that $S_{jane} \rightarrow 1$ and $S_{john} \rightarrow 0$. Hence, $S_{jane,john} \approx 0.5$. In this case, $S_{jane,john}$ is much higher than S_{john} and much lesser compared to S_{jane} and hence does not give us an accurate representation of the strength of their

relationship. To eliminate such a scenario, we are considering the SoI values if the difference in the two values is greater than or equal to 0.5, as shown in (6).

$$S_{ij} = \frac{\sum_{i=1}^n N_i * (F_i * W_i) + \sum_{j=1}^n N_j * (F_j * W_j)}{2} \left. \begin{array}{l} |S_i - S_j| < 0.5 \\ |S_i - S_j| \geq 0.5 \end{array} \right\} \quad (6)$$

3.5. Identity leakage

First introduced by Saini *et al.* [24], identity leakage indicates if any user's identity can be found solely by their participation in the shared data. Leakage of a user's identity could be harmful, especially in the non-consensual sharing of data. The impact of this metric is defined by answering four questions—who the user is, what the user is doing, where the user is, and when the user is there. For simplicity, if we ignore video posts on OSNs, we can remove the question "what the user is doing" as it needs to rely on more than one information frame for a deductive conclusion. As we are left with three questions to answer, we can define identity leakage as a metric between 0 and 1. One way to calculate identity leakage—using facial recognition technology (FRTs), as proposed by Xu *et al.* [25]. However, this approach would require a database of existing scans, which can be hard to gather, as shown in (7) and (8).

$$I_{leakage} = \frac{\rho}{G_{wh,wn,wr}} \quad (7)$$

where,

$$G_{wh,wn,wr} = \frac{1}{n_1 \times n_2 \times n_3} \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \sum_{i_3=1}^{n_3} \mathcal{H}\mathcal{H}(\xi_{i_1}, \Delta t_{i_2}, \gamma_{i_3}) \quad (8)$$

ρ is the no. of people in the data item, and $G_{wh,wn,wr}$ represents the associated average group size and relies on the measure of anonymity and a polymorphic function \mathcal{H} that returns the number of people satisfying the conditions [24]. Here, ξ_i represents the i^{th} activity. In the case of one frame – a photo, there will be only one activity. Moreover, Δt_i represents the time slot in which the i^{th} activity is performed, and γ_i is the i^{th} location identifier.

Introducing this metric strengthens the model's accuracy because it also factors in that in a real-world OSN, there are several instances when a user is tagged in a post, but the shared content is irrelevant to them and is not harming them in any way. In such an instance, prompting the user and informing them of the privacy risk and sharing loss without considering the identity leakage would show a lack of nuance. By measuring identity leakage, the proposed model can ignore scenarios where $I_{leakage} = 0$. Similarly, in cases where the $I_{leakage}$ exceeds the threshold, the proposed model denies sharing of the data even if the user has high values of reputation and strength of interaction.

In several scenarios, especially those that involve non-consensual sharing of data, identity leakage will be high. To derive the overall sensitivity of the situation, we rely on answers to the four questions mentioned above. Identity leakage would only apply to public profiles in a real-world OSN, as a user would be tagged if a publicly visible profile posted a picture and tagged them. This does not apply to private profiles, though some OSNs are implementing features to prevent private profiles from tagging accounts that do not follow them.

4. COLLUSION ATTACKS

Systems composed of dynamic nodes that communicate through a network are networked systems. Different kinds of networked systems exist, such as wireless networks, wireless sensor networks (WSNs), wireless ad hoc networks (WAN), and e-commerce systems. These systems are quite vulnerable to security breaches. A "collusion attack" occurs when a node purposefully enters a covert arrangement with an entity or is compromised by that entity [14].

This type of attack is not independent of these kinds of networks alone. Social media comprises much sensitive information with its onslaught of connectivity and virtual human interaction. Health conditions, relationship status, family information, residence, political affiliations, and employment details are a few examples. Most security research has revolved around singular attackers. Meanwhile, a collusion attack in OSNs refers to one that consists of numerous malevolent users to improve its outcome (for example, gathering more data on a victim user) by collective action as opposed to separately conducted attacks. One of the most powerful strategies bad users might utilize to avoid existing defenses that primarily deal with lone attackers in networks is the collusion assault, with each attacker exhibiting a different activity pattern or behavior.

Trust is determined based on the reputation of a particular user. The reputation of a certain entity is defined as the perception of that user based on their past actions. This can be discovered either directly, through observation of the user's behavior or recommendations by other users. The user's reputation is then calculated using the aggregate of the two values. However, malicious users can collude to hamper trust relations by decreasing the reputation level of benevolent users or increasing their reputation levels.

To calculate a user's reputation, we can represent an OSN as a social network graph with each user represented as a node and a weighted relationship (SoI) between any two users. Four variables influence the reputation of the user in OSN-recommendation by other users (R_u), number of followers (A), number of following (B), and number of posts made by the user (N_p). By putting a cap on the largest amount of reputation (L) one can receive from a number of followers or following, we get (9):

$$rep = \frac{L}{1+(L-1)e^{-A-B}} - 1 \quad (9)$$

R_u is given more weightage than the other metrics by adding a coefficient [26] (10):

$$rep = 2R_u + N_p + \frac{L}{1+(L-1)e^{-A-B}} - 1 \quad (10)$$

4.1. Social bots

With the ever-growing popularity of OSNs like Twitter and Facebook, there has been a spike in the number of machine accounts that resemble real users. Social bot accounts leave OSNs susceptible to attack. Social bots are computer programs that create material automatically, share it on a specific social network, and engage with its users. While many useful applications of bots help amplify issues of importance and increase the usability and accessibility of online services, we cannot overlook the disservice some others cause. Some bots have heavily impacted the political climate in various ways, from spreading misinformation during crises to distracting users from corrupt government actions through the proliferation of irrelevant information and influencing elections.

Because there are fake social connections, social bots' actions also affect OSNs' social networks. If social bots can access users' accounts, they can collect private data and then use it for spamming and phishing purposes. Additionally, they may compile data from the internet to pass to others, imitate human actions, and persuade people by rating and retweeting. Social bots not only inherently deceive people but also harm the ecosystem by creating fake connections and tainting network content of online social networks.

4.2. Bot-driven collusion attacks

In the proposed approach, we look to resolve bot-driven collusion attacks on OSNs by adopting bot-detection strategies and leveraging the concept of trust in a networked system. To understand trust, we imagine the social media network using graph theory. We use directed and undirected graphs to represent whether the establishment of the trust is one-sided or not. The graph depicts only the friends or associations of each user. While some online social networks allow for unidirectional relationships where user A can follow user B, the reverse might not be true; in this case, it is assumed that friends in a network are both following each other and hence share a bidirectional relationship.

Equation (10) can be exploited by a bot attack by increasing A , B , and N_p . Given that other users are bots, they can easily bump up R_u values for each other. In order to eliminate such an attack, we introduce bot detection. Many tools like Botometer (previously Bot or Not) rely on metrics like the account name, the level of Tweeting, the location in the bio, and the hashtags used. To identify if an account is a bot or not [27], [28]. Thus, we eliminate bot-driven collusion attacks by adding an extra layer of check (11):

$$rep = \begin{cases} 2R_u + N_p + \frac{L}{1+(L-1)e^{-A-B}} - 1, & F_b = 0 \\ 0, & F_b = 1 \end{cases} \quad (11)$$

where F_b states if the user is a bot or not. To determine F_b , we rely on Bebensee *et al.* [29] model that leverages node neighbors and ego-graph topology for bot detection in social graphs [29].

4.3. Applications in real-world OSNs

In a real-world scenario, friend and block lists are constantly being updated. While not fully automated, a multi-party access control system must keep up with the dynamic nature of OSNs and ensure that access policies are updated whenever a change is triggered. Some implementations of multi-party access control can be seen in online social networks: Twitter recently introduced a co-tweet functionality where two

users can co-author a tweet and publish it on their timelines simultaneously [30]. Though it is not a 50-50 partnership between the users as some of the privileges (like pinning the tweet) only lie with the initiator of the collaboration, it is a start towards multi-party access control in online social media networks. Other platforms like Instagram also rolled out Collab, allowing users to collaborate and post content together.

5. RESULT ANALYSIS AND DISCUSSION

5.1. Obtained results

Due to strict data governance laws and restrictions and the lack of accessibility of ethically obtained online social network data, we generated a dataset instead of scraping data from existing platforms. To test our proposed approach, we used Trumania-a scenario-based random dataset generator library in Python3 [31]–[33]. For the following examples, we generated interactions between three users-Matthew, Jennifer, and Nicholas. Trumania allows for defining relationships and attributes for persons in the simulation. By utilizing these features, relationships between friends and interaction types are formed, whereas messages and a user's popularity are considered attributes. Interaction types are picked from the ones listed in Table 2, with some interaction types being more likely than others. The interaction probability as tabulated in Table 3 is provided by observing behaviors of many users across various OSNs.

Table 3. Probability of types of interactions in simulated scenarios

Types of interactions	Probability of interactions (P)
Like	60%
Comment	15%
Share	15%
Post	10%

Since we rely on the quantity and quality of interactions, sentiment analysis is performed on users' textual interactions. We made use of Microsoft's pre-built sentiment analysis model [34]. On running the dataset against the model, we are provided with the following values:

- $Sentiment \in [positive, negative, neutral]$
- $P(positive) \in [0,1]$
- $P(neutral) \in [0,1]$
- $P(negative) \in [0,1]$

where $T_{sentiment}$ represents the overall sentiment of the textual interaction, $P(positive)$, $P(neutral)$, and $P(negative)$ represent the probabilities of the text being positive, neutral, and negative, respectively. An example of the output given by this model is as follows:

Input: Power in Seattle has been restored in 24 hours.

Output: $P(positive) = 0.07, P(neutral) = 0.93, P(negative) = 0.03$

Below we show how the proposed model approaches access control and conflict resolution through these examples:

Example 1: Data sharing between two users - proprietor and collaborator (with PR and SL)

Scenario: Matthew posts a picture with Nicholas

First, we calculate the privacy conflict by evaluating the negotiating users between Matthew and Nicholas; in this case, it is Jennifer - Matthew's friend who is not friends with Nicholas. Privacy risk and sharing loss are calculated by (1) and (2). Here, pc_{β} and sl_{β} denote the general privacy concern and sensitivity level. These can be based on the users' preferences and for this example, we will use the values in Table 4. We calculate the overall PR and SL by summation of the applied formula. Conflict resolution is made by (12), (13):

$$Decision = \begin{cases} \lambda * PR \geq ((1 - \lambda) * SL) & Deny \\ \lambda * PR < ((1 - \lambda) * SL) & Permit \end{cases} \quad (12)$$

where λ denotes the weightage given to PR and SL.

$$\lambda = \begin{cases} 0 \leq \lambda < 0.5 & sharing\ loss\ is\ favoured \\ 0 \leq \lambda \leq 1 & privacy\ risk\ is\ favoured \end{cases} \quad (13)$$

$\lambda = 0.5$ gives equal weightage to PR and SL. If (12) is true, the optimized decision is to permit sharing of data. In this scenario, these are the values obtained are tabulated in Table 5.

Table 4. Metrics and values of sensitivity level for example 1

Metric (User)	Corresponding Value
$PC_{\beta}(\text{Matthew})$	0.66
$PC_{\beta}(\text{Nicholas})$	0.66
$sl_{\beta}(\text{Matthew})$	0.33
$sl_{\beta}(\text{Nicholas})$	0.33

Table 5. Metrics and values of PR and SL for example 1

Metric (User)	Corresponding Value
Privacy Risk	0.43
Sharing Loss	0.45
Decision	Permit

Example 2: Sharing data with target users

Suppose we take the same scenario as example 1. We have learned that Matthew and Nicholas are good friends, whereas Matthew and Jennifer do not have a great relationship. Using the same metrics as Example 1, we calculate PR and SL for both users. In addition, we calculate the strength of interaction using the quantity and quality of past interactions. We find that the strength of interaction between Matthew and Nicholas is greater than that of Matthew and Jennifer, signifying a better relationship between the former. Due to this, Matthew will favor protecting the privacy of Nicholas over the sharing loss incurred from not sharing the data item. Hence, the lambda value will increase. Hence, the optimized decision is to deny sharing. Table 6 shows the interactions between Matthew, Nicholas, and Jennifer in an OSN. Each record corresponds to a different type of interaction which is uniquely identified by the serial number (S_No). The message field is not null for comments and shares (users typically have the option to add a caption). These interactions are a sample of exchanges that take place between the users during their online friendship.

Table 6. Interactions between users

S_No	Action	Name	Name_2	Message
1	Like	Matthew	Nicholas	-
2	Comment	Matthew	Nicholas	What a lovely picture! Can't wait to see you next week
3	Comment	Nicholas	Matthew	Thank you, Matthew. I can't wait either. It has been such a long time!
4	Like	Nicholas	Matthew	-
5	Share	Nicholas	Matthew	Finally got to meet my best friend!
6	Comment	Matthew	Jennifer	Stop spreading misinformation.
7	Comment	Jennifer	Matthew	I am entitled to my own opinion!
8	Comment	Matthew	Jennifer	Hate speech hurts everyone.

Table 7 depicts the sentiment analysis results for given interactions between the users. Based on (4), we calculate SoI for a user by considering these values of the sentiment of the interactions. SoI is affected by the quality, quantity, and type of interaction. If the sentiment of interaction is negative for a comment, it will reduce the SoI between two users.

Since most of Matthew and Nicholas's interactions are positive, they have a relatively high SoI as compared to Matthew and Jennifer, who have had negative and neutral interactions, thus having a lesser SoI. Conflict resolution is done after accounting for the strength of interaction by modifying (12) to incorporate (14), (15).

$$if \text{ Decision} = \begin{cases} \lambda_{soi} * PR \geq ((1 - \lambda_{soi}) * SL) & \text{Deny} \\ \lambda_{soi} * PR < ((1 - \lambda_{soi}) * SL) & \text{Permit} \end{cases} \quad (14)$$

The lambda value for the (14) is varied depending on S_{ij}

$$\lambda_{soi} = \begin{cases} \lambda + avg(S_{ij}, S_{kj}) & S_{ij} \geq avg(S_{ij}, S_{kj}) \\ \lambda - avg(S_{ij}, S_{kj}) & S_{ij} < avg(S_{ij}, S_{kj}) \end{cases} \quad (15)$$

Here, S_{ij} and S_{kj} represent the strength of interaction between users i and j and users k and j respectively. As per (15), the greater the value of lambda, higher the chances that privacy risk will be favored by the model. Hence, if two users have higher strength of interaction, the chances of them protecting each other's data from the negotiating users is higher as shown in Table 8.

Table 7. Sentiment values for corresponding interactions

S_No	Sentiment	P(Positive)	P(Neutral)	P(Negative)
1	positive	1	0	0
2	positive	0.94	0.05	0.01
3	positive	0.71	0.21	0.08
4	positive	1	0	0
5	positive	0.99	0.01	0
6	negative	0.07	0.38	0.55
7	neutral	0.16	0.69	0.15
8	negative	0	0.01	0.99

Table 8. Metrics and values of SoI for example 2

Metric (User)	Corresponding Value
Privacy Risk	0.43
Sharing Loss	0.45
$S_{\text{Matthew,Nicholas}}$	0.10
$S_{\text{Matthew,Jennifer}}$	0.06
Decision	Deny

Example 3: Non-consensual sharing from a trusted user

Non-consensual sharing of media is prevalent and can often have harmful impacts on the affected persons. It manifests itself in many forms and is a threat to online users' privacy.

a. Matthew shares a picture of Jennifer

- Jennifer's identity leakage is low: Jennifer and Matthew have a low SoI value signifying a weak relationship with each other. Since Jennifer's identity leakage value is low, the risk to her privacy is also low. Due to lower levels of trust between Matthew and Jennifer, the decision will be to permit sharing in Table 9.
- Jennifer's identity leakage is high: Since Jennifer's identity leakage is high, the decision is to deny it to protect her privacy in Table 10.

Table 9. Metrics and values when Jennifer's IL is low

Metric (User)	Corresponding Value
Privacy risk	0.43
Sharing loss	0.45
$S_{\text{Matthew,Jennifer}}$	0.06
$I_{\text{leakage (Low)}}$	0.12
Decision	Permit

Table 10. Metrics and values when Jennifer's IL is high

Metric (User)	Corresponding Value
Privacy risk	0.43
Sharing loss	0.45
$S_{\text{Matthew,Jennifer}}$	0.063
$I_{\text{leakage (High)}}$	0.83
Decision	Deny

b. Matthew shares a picture with Nicholas

- Nicholas' identity leakage is low: Here, Nicholas' identity leakage is low. However, Matthew and Nicholas have a higher SoI, signifying a stronger relationship. In this case, the decision will be to deny since Matthew will favor protecting Nicholas' privacy in Table 11.
- Nicholas' identity leakage is high: Since Nicholas' identity leakage is high, the decision is to deny protecting his privacy in Table 12.

Table 11. Metrics and values when Nicholas' IL is low

Metric (User)	Corresponding Value
Privacy risk	0.43
Sharing loss	0.45
$S_{\text{Matthew,Nicholas}}$	0.10
$I_{\text{leakage (LOW)}}$	0.12
Decision	Deny

Table 12. Metrics and values when Nicholas' IL is high

Metric (User)	Corresponding Value
Privacy Risk	0.43
Sharing Loss	0.45
$S_{\text{Matthew,Nicholas}}$	0.10
$I_{\text{leakage (High)}}$	0.83
Decision	Deny

The above scenarios show how the strength of interaction between two users can affect whether the decision is to permit or deny. Even though both Jennifer and Nicholas had lower values of identity leakage, sharing was still denied in the case of Nicholas because he has a better relationship with Matthew. This is meant to show the intention of favoring privacy when there is more trust between two users.

$$\lambda_{il} = \begin{cases} \lambda_{soi} + I_{leakage} & I_{leakage} \geq 0.5 \\ \lambda_{soi} - I_{leakage} & I_{leakage} < 0.5 \end{cases} \quad (16)$$

Conflict resolution is done after accounting for identity leakage by modifying (14) to incorporate identity leakage as calculated via (16), (17).

$$Decision = \begin{cases} \lambda_{il} * PR \geq ((1 - \lambda_{il}) * SL) & Deny \\ \lambda_{il} * PR < ((1 - \lambda_{il}) * SL) & Permit \end{cases} \quad (17)$$

Example 4: Repeated actions by multiple accounts

Juniper recently started a business, and all their posts have been reported by multiple accounts. Bots can easily alter the variables used to calculate reputation and increase their value. This poses a threat to the privacy of users.

With reputation and bot detection: The expected reputation of bots is zero. However, social bots can sometimes mimic human behavior. Due to collusion-driven attacks where multiple bots can give each other a higher recommendation, the reputation of the bot ends up being a higher number. By using bot detection along with reputation, any bot users are automatically revealed, ensuring that reputation is calculated only for qualified users. By using (10), we calculate the reputation of the users. And we eliminate any bots by using (10), thus precluding a bot-driven collusion attack. In this scenario in Table 13, we calculate the reputation for Juniper and the multiple accounts that are performing the repeated actions. Once we run Bebensee *et al.* [29] model for bot detection, we assign reputation as zero for those users. Upon doing so, we ignore any actions by these accounts while focusing on the privacy of other users [29]. Without bot detection: In the same scenario, if bot detection is not applied, bot users go undetected, and their malicious actions could be considered valid inputs. To avoid this scenario of a collusion attack, bot detection comes in handy.

Table 13. Reputation values for example 4

User	Recommendation	Number of posts	Number of followers	Number of following	Reputation
U_1	0.5	5	50	70	16.00
U_2	0.3	2	20	24	12.60
U_3	0.7	20	200	100	31.40
U_4	0.4	40	60	5	50.80
U_5	0.1	100	5	0	109.98

Example 5: Access control of friends of friends in a scenario with multiple collaborators.

Social media platforms have revolutionized the way people share experiences and interact with friends yet concerns over privacy violations and security breaches persist. Sharing photos and tagging friends creates a personal connection but also poses a risk for privacy attacks, particularly if the friends of friends are

untrustworthy or if the photo contains sensitive information. These attacks can result in the exposure of personal information, harassment, stalking, or even fraud. To mitigate these risks, users should exercise caution when sharing photos, review their privacy settings regularly, and limit their posts' audience to only those they trust. The following scenario iterates how the proposed model prevents privacy risks by using metrics like SoI and IL to control what is being shown to friends of friends. Jane, John, and Juniper are tagged in a post by Jade, who is the proprietor. This example in Table 14 shows how policy resolution is optimized using strength of interaction in a case with multiple collaborators and focuses on what the collaborators' friends get to view.

Table 14. Users and their friend lists for example 5

User	Friend List
Jane	Jade, Alex, Mark
John	Jade, Amanda, Mark, Eva
Juniper	Jade, Anna, Ava, Mark, Eva
Jade	Jane, John, Juniper, Eva, Mark, Kate

In the following examples, $PR=SL$, hence if we use the traditional formula for policy resolution, we will not get an optimized result. The following scenarios show how SoI and IL help optimize the result. In this scenario, Jane and Jade have a high SoI and Jane and Alex also have a high SoI. Since, Keeping Jade's identity leakage in mind, the decision is to permit Alex to view the shared data item in Table 15. Here, John and Jade have a high SoI whereas John and Amanda have a low SoI. Keeping Jade's identity leakage in mind, the decision is to deny Amanda to view the shared data item in Table 16.

Table 15. Metrics and values for example 5; Alex

Metric (User)	Corresponding Value
Privacy Risk	0.50
Sharing Loss	0.50
$S_{\{Jane,Jade\}}$	0.78
$S_{\{Jane,Alex\}}$	0.83
$I_{leakage} (Jade)$	0.42
Decision	Permit

Table 16. Metrics and values for example 5; Amanda

Metric (User)	Corresponding Value
Privacy Risk	0.50
Sharing Loss	0.50
$S_{\{John,Jade\}}$	0.67
$S_{\{John,Amanda\}}$	0.35
$I_{leakage} (Jade)$	0.42
Decision	Deny

Here, Juniper and Jade have a low SoI and Jade and Kate have a low SoI too. Even though Juniper's identity leakage is low, the decision is to deny Anna to view the shared data item in Table 17. Here, Juniper and Jade have a low SoI and Juniper and Ava have a high SoI. Keeping Jade's identity leakage in mind, the decision is to deny Ava to view the shared data item in Table 18. Here, Jade and Jade have a high SoI and Jade and Eva have a low SoI. Additionally, because Jane's identity leakage is low, the decision is to permit Ava to view the shared data item in Table 19.

Table 17. Metrics and values for example 5; Anna

Metric (User)	Corresponding Value
Privacy Risk	0.50
Sharing Loss	0.50
$S_{\{Juniper,Jade\}}$	0.23
$S_{\{Jade,Kate\}}$	0.12
$I_{leakage} (Juniper)$	0.21
Decision	Deny

Table 18. Metrics and values for example 5; Ava

Metric (User)	Corresponding Value
Privacy Risk	0.50
Sharing Loss	0.50
$S_{\{Juniper,Jade\}}$	0.23
$S_{\{Juniper,Ava\}}$	0.57
$I_{leakage} (Jade)$	0.42
Decision	Deny

Table 19. Metrics and values for example 5; Eva

Metric (User)	Corresponding Value
Privacy Risk	0.50
Sharing Loss	0.50
$S_{\{Jane,Jade\}}$	0.78
$S_{\{Jade,Eva\}}$	0.92
$I_{leakage}(Jane)$	0.13
Decision	Permit

5.2. Ramifications of the study findings

The proposed methodology, which incorporates metrics such as SoI, IL, user reputation, and bot detection, along with established parameters like PR and SL, offers several benefits to the policy resolution of shared data in OSNs.

- Enhanced threat detection: The focus on assessing a user's reputation based on potential bot status introduces a novel element in threat detection. This deliberate consideration of social bots enhances the ability to identify and mitigate collusion attacks initiated by automated entities, providing a more comprehensive approach to security.
- Comprehensive understanding of interactions: By considering multiple metrics and parameters, the methodology offers a more intricate understanding of online user interactions. This goes beyond traditional approaches that may focus on a limited set of factors. The inclusion of SoI, IL, user reputation, bot detection, PR, and SL provides decision-makers with a holistic view, allowing them to assess the quality and quantity of interactions more comprehensively.
- Informed decision-making: The multi-faceted approach empowers stakeholders and decision-makers with a broader set of considerations. This, in turn, facilitates more informed and educated decisions regarding policy disputes. Understanding the nuanced nature of online interactions allows for the implementation of policies that are not only effective but also considerate of the intricacies involved.
- Balancing quantity and quality: The methodology's objective is to assess both the quality and quantity of interactions addresses a common challenge in OSNs. Instead of focusing solely on the volume of interactions, the approach acknowledges the strength and nature of these interactions. This balance is crucial in ensuring that policy resolutions not only restrict unwanted activities but also promote positive and meaningful user engagements.
- Narrowing the gap between theory and practice: While fully automated policy resolution remains a challenge, the proposed methodology represents a step towards bridging the gap between theoretical concepts and practical application. It acknowledges the complexity of online interactions, making it more realistic and applicable in real-world OSN scenarios.
- Foundation for continuous improvement: The inclusion of various metrics and parameters creates a flexible framework that can be adapted and improved over time. As the landscape of online interactions evolves, the methodology provides a foundation for continuous enhancement, ensuring that policy resolution strategies remain relevant and effective.

5.3. Comparison with related works

By taking a data-driven approach to policy analysis, the proposed method can help identify and resolve conflicts in a way that balances the interests of different stakeholders and achieves the desired outcomes. Table 20 provides a brief overview of how the proposed method is advantageous over notable existing methods in the domain. The following justifications support the efficacy of this method: i) Identity leakage is a critical metric for privacy risk analysis in online social networks. By measuring the potential for personal information to be exposed or linked to an individual, identity leakage can help identify scenarios where privacy risks are high and where additional privacy protections may be necessary; ii) The strength of interaction metric can help identify scenarios where collusion attacks are more likely to occur. By analyzing the strength of connections between different actors in the network, the method can identify scenarios where collusion could occur and develop policies to prevent it; iii) Using metrics such as privacy risk and sharing loss to identify and resolve conflicts can help balance the interests of different stakeholders. By quantifying the potential privacy risks and sharing losses associated with different policy options, the method can help identify policies that minimize negative impacts on users while still achieving the desired outcome; iv) Fine-tuning scenarios based on specific metrics can help ensure that the analysis is tailored to the specific context of the online social network being studied. This can improve the accuracy of the analysis and increase the likelihood that the resulting policies will be effective; and v) Consideration of real-world factors: The proposed method takes into account real-world factors, such as identity leakage and the strength of

interaction between users. The proposed method is relatively easy to interpret and communicate to stakeholders, particularly compared to game theory approaches.

Table 20. Advantages of proposed method over existing methods

Authors	Parameters incorporated	Access control	Advantages of proposed method over existing technique
Abid and Daud [35]	Mutual interests of users and their relationships	Text-based dynamic and fine-grained	It only considers the relationship between resources and user clusters based on common interests. It fails to consider the trust and interaction parameters that makes it susceptible to bot related collusion attacks from mimicked human behavior.
Xu <i>et al.</i> [3]	Aggregated trust-based decisions by users	Trust-based	This method uses the multi-armed bandit approach which is computationally intensive. Additionally, they only consider privacy risk and sharing loss while ignoring the complex relationships between online users – which is better determined by factors like strength of interaction.
Gao <i>et al.</i> [36]	Blockchain, Ciphertext-policy attribute-based encryption (CP-ABE), and InterPlanetary File System (IPFS)	Fine-grained	This paper considers user privacy when uploading files to a platform that is governed by a third party's privacy policy. While blockchain ensures the decentralization of data, humans – the weakest link in security – are not addressed. We introduce parameters like Identity Leakage and an optimized model that concludes which access policy is optimal for the user based on their past interactions on social media.
Hu <i>et al.</i> [37]	Privacy risk, sharing loss	Dynamic trust-based	The proposed method uses a more comprehensive analysis of policy conflicts by considering Nash Equilibrium. Game theory approaches often focus on a single objective, such as maximizing utility or minimizing risk, and may not consider the full range of outcomes. By considering multiple objectives, the proposed method provides a more nuanced analysis of policy conflicts and enables stakeholders to make more informed decisions.
Jasmine and Hymavathi [38]	Aggregated trust-based decisions by users	Trust-based	This method considers only factors like privacy risk and sharing loss while applying a multi-armed bandit approach. Our approach adds nuance and considers factors like quality of interaction, strength of interaction between users online, and identity leakage of the collaborators. These real-world considerations help make the model more nuanced.
Mouhsein and Madhavi [32]	User approval and security inclination	Fine-grained	The approach considers user approval and security inclination to co-control mutual information. Such a model is prone to collusion attacks. In our proposed method, we are considering bot detection using reputation and in turn stopping collusion attacks.
Niksirat <i>et al.</i> [28]	User input and explanation	Fine-grained	By using a mediation bot, they aim to resolve multi-party conflict on social media platforms. This requires user interaction and input. Our model proposes an automated system where decisions are suggested based on the users' previous interactions online.
Ramteke and Talmale [39]	Sensitivity and trust levels	Trust-based	This method uses a voting scheme for detecting conflicts between users with a shared data item. It also introduces a proof-of-concept social network for providing secure access but does not exemplify how conflict resolution would be done. While the method considers the trust levels, it does not specify the exact parameters based on which trust is calculated and incorporated in the computation.
Hu <i>et al.</i> [40]	Privacy Risk, Sharing Loss	Trust-based	The paper proposes a new method to detect and resolve privacy conflicts in collaborative data sharing on OSNs. It considers both the privacy risk and sharing loss and a proof-of-concept implementation named "Retinue" is also presented with extensive evaluation. PR and SL are useful in balancing the interests of different stakeholders. Our approach goes one-step further as to fine-tune these scenarios based on metrics like Identity Leakage and Strength of Interaction to tailor the analysis to specific contexts and improve accuracy.
Ali <i>et al.</i> [31]	Cryptographic-based techniques	Fine-grained	This method proposes a framework for preserving user privacy OSNs by incorporating collaborative content sharing and cryptographic-based techniques. The proposed framework includes an access management server that acts as middleware between the OSN server and users. Cryptographic techniques often complex to implement and have the potential for decreased usability. The proposed method, on the other hand, relies on intuitive metrics that are easy to understand and explain, making it more accessible to a wider range of stakeholders.
Lee <i>et al.</i> [6]	Relationship type, spatial temporal information, and co-occurrence users	Fine-grained	This method utilizes a fine-grained approach to tackle access control at a face-level instead of a photo-level. Our approach also considers interaction parameters and attempts to cover privacy risks posed by collusion attacks from bots.
Iliia <i>et al.</i> [41]	Relationship level and trust	Threshold-based Secret Sharing	The paper proposes a socially-aware privacy-preserving system for protecting users' privacy from other users and the service provider. The system includes a collaborative multi-party access control model that enables all users associated with a resource to participate in defining the access control policy as well as enforcing it. Since the mechanism is based on trust, there could be potential threats from bot collusion attacks which are addressed in our approach through bot detection by reputation.

6. CONCLUSION

In this paper, we have modelled an approach that relies on sophisticated factors like strength of interaction, reputation, and identity leakage to assign trust and resolve privacy conflicts in a multi-party sharing environment. These metrics consider real-world scenarios that include but are not limited to: non-consensual sharing of data, bot-driven collusion attacks, and trust-based policy resolution of shared data. The strength of interaction metric can help identify scenarios where collusion attacks are more likely. By analyzing the strength of connections between different actors in the network, the method can identify scenarios where collusion could occur and develop policies to prevent it. Using metrics such as privacy risk and sharing loss to identify and resolve conflicts can help balance the interests of different stakeholders. By quantifying the potential privacy risks and sharing losses associated with different policy options, the method can help identify policies that minimize negative impacts on users while still achieving the desired outcome. Identity leakage is a critical metric for privacy risk analysis in online social networks. By measuring the potential for personal information to be exposed or linked to an individual, identity leakage can help identify scenarios where privacy risks are high and where additional privacy protections may be necessary. Fine-tuning scenarios based on specific metrics can help ensure that the analysis is tailored to the specific context of the online social network being studied. This can improve the accuracy of the analysis and increase the likelihood that the resulting policies will be effective. By taking a data-driven approach to policy analysis, the method can help identify and resolve conflicts in a way that balances the interests of different stakeholders and achieves the desired outcomes.

In the future, we intend to make the SoI metric more realistic by factoring in superficial interactions on the internet. We also plan to extend this approach to increase its usability by performing user testing and enhancing our model to be better accustomed to a real-world scenario. Finally, we plan to collate these results into a functioning application that respects users' privacy and ensures that multi-party privacy does not diminish the user experience with a special focus on access control for historical data, keeping in mind the dynamic nature of online interactions and relationships.

ACKNOWLEDGEMENTS

We acknowledge the contribution of Mrs. Sharada P. Shetty for her insights during the course of the study.





REFERENCES

- [1] S. Kemp, "Digital 2022: Global overview report," *Datareportal*, 2022. <https://datareportal.com/reports/digital-2022-global-overview-report> (accessed Nov. 01, 2023).
- [2] H. Hu, G.-J. Ahn, and J. Jorgensen, "Multiparty access control for online social networks: Model and mechanisms," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 7, pp. 1614–1627, Jul. 2013, doi: 10.1109/TKDE.2012.97.
- [3] L. Xu, C. Jiang, N. He, Z. Han, and A. Benslimane, "Trust-based collaborative privacy management in online social networks," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 1, pp. 48–60, Jan. 2019, doi: 10.1109/TIFS.2018.2840488.
- [4] G. Akkuzu, B. Aziz, and M. Adda, "Fuzzy logic decision based collaborative privacy management framework for online social networks," in *Proceedings of the 5th International Conference on Information Systems Security and Privacy*, 2019, pp. 674–684, doi: 10.5220/0007702206740684.
- [5] A. Ahmad *et al.*, "A relation-aware multiparty access control," *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 1, pp. 227–239, Jul. 2019, doi: 10.3233/JIFS-179080.
- [6] C. Lee, W. Wang, and Y. Guo, "A Fine-grained multiparty access control model for photo sharing in OSNs," in *2016 IEEE First International Conference on Data Science in Cyberspace (DSC)*, Jun. 2016, pp. 440–445, doi: 10.1109/DSC.2016.113.
- [7] O. Ulusoy and P. Yolum, "Collaborative privacy management with auctioning mechanisms," in *Studies in Computational Intelligence*, 2021, pp. 45–62.
- [8] J. M. Such and N. Criado, "Multiparty privacy in social media," *Communications of the ACM*, vol. 61, no. 8, pp. 74–81, Jul. 2018, doi: 10.1145/3208039.
- [9] M. Cherubini, K. Salehzadeh Niksirat, M.-O. Boldi, H. Keopraseuth, J. M. Such, and K. Huguenin, "When forcing collaboration is the most sensible choice: desirability of precautionary and dissuasive mechanisms to manage multiparty privacy conflicts," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW1, pp. 1–36, Apr. 2021, doi: 10.1145/3449127.
- [10] T. Muhammad and A. Ahmad, "A joint sharing approach for online privacy preservation," *World Wide Web*, vol. 24, no. 3, pp. 895–924, May 2021, doi: 10.1007/s11280-021-00876-5.
- [11] M. Madeira and A. Joshi, "Analyzing close friend interactions in social media," in *2013 International Conference on Social Computing*, Sep. 2013, pp. 932–935, doi: 10.1109/SocialCom.2013.145.
- [12] S. Krakan, L. Humski, and Z. Skočir, "Determination of friendship intensity between online social network users based on their interaction," *Tehnicki vjesnik - Technical Gazette*, vol. 25, no. 3, Jun. 2018, doi: 10.17559/TV-20170124144723.
- [13] J. Alemany, E. Del Val, J. M. Alberola, and A. García-Fornes, "Metrics for privacy assessment when sharing information in online social networks," *IEEE Access*, vol. 7, pp. 143631–143645, 2019, doi: 10.1109/ACCESS.2019.2944723.
- [14] J. Alemany, E. del Val, J. Alberola, and A. García-Fornes, "Estimation of privacy risk through centrality metrics," *Future Generation Computer Systems*, vol. 82, pp. 63–76, May 2018, doi: 10.1016/j.future.2017.12.030.
- [15] J. Domingo-Ferrer, S. Martínez, D. Sánchez, and J. Soria-Comas, "Co-Utility: Self-enforcing protocols for the mutual benefit of participants," *Engineering Applications of Artificial Intelligence*, vol. 59, pp. 148–158, Mar. 2017, doi: 10.1016/j.engappai.2016.12.023.





- [16] J. M. Such and M. Rovatsos, "Privacy policy negotiation in social media," *ACM Transactions on Autonomous and Adaptive Systems*, vol. 11, no. 1, pp. 1–29, Apr. 2016, doi: 10.1145/2821512.
- [17] G. Jethava and U. P. Rao, "User behavior-based and graph-based hybrid approach for detection of Sybil Attack in online social networks," *Computers and Electrical Engineering*, vol. 99, Apr. 2022, doi: 10.1016/j.compeleceng.2022.107753.
- [18] M. Al-Qurishi, M. Alrubaian, S. M. M. Rahman, A. Alamri, and M. M. Hassan, "A prediction system of Sybil attack in social network using deep-regression model," *Future Generation Computer Systems*, vol. 87, pp. 743–753, Oct. 2018, doi: 10.1016/j.future.2017.08.030.
- [19] J. Zhang, R. Zhang, Y. Zhang, and G. Yan, "The rise of social botnets: attacks and countermeasures," *IEEE Transactions on Dependable and Secure Computing*, vol. 15, no. 6, pp. 1068–1082, Nov. 2018, doi: 10.1109/TDSC.2016.2641441.
- [20] N. P. Shetty, B. Muniyal, and S. Mowla, "Policy resolution of shared data in online social networks," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 4, pp. 3767–3776, Aug. 2020, doi: 10.11591/ijece.v10i4.pp3767-3776.
- [21] N. P. Shetty *et al.*, "Trust based resolving of conflicts for collaborative data sharing in online social networks," In: *Dutta, P., Chakrabarti, S., Bhattacharya, A., Dutta, S., Shahnaz, C. (eds) Emerging Technologies in Data Mining and Information Security. Lecture Notes in Networks and Systems*, vol. 490. Springer, Singapore, doi: 10.1007/978-981-19-4052-1_5.
- [22] H. Hu and G.-J. Ahn, "Multiparty authorization framework for data sharing in online social networks," in *IFIP Annual Conference on Data and Applications Security and Privacy*, 2011, pp. 29–43.
- [23] J. M. Siqueiros-García, R. García-Herrera, E. Hernández-Lemus, and S. Alcalá-Corona, "A game-theory modeling approach to utility and strength of interactions dynamics in biomedical research social networks," *Complex Adaptive Systems Modeling*, vol. 5, no. 1, Dec. 2017, doi: 10.1186/s40294-017-0044-0.
- [24] M. K. Saini, P. K. Atrey, S. Mehrotra, and M. S. Kankanhalli, "Privacy aware publication of surveillance video," *International Journal of Trust Management in Computing and Communications*, vol. 1, no. 1, 2013, doi: 10.1504/IJTMCC.2013.052523.
- [25] K. Xu, Y. Guo, L. Guo, Y. Fang, and X. Li, "My privacy my decision: Control of photo sharing on online social networks," *IEEE Transactions on Dependable and Secure Computing*, vol. 14, no. 2, pp. 199–210, Mar. 2017, doi: 10.1109/TDSC.2015.2443795.
- [26] "How to calculate users' reputation," *stackexchange*. <https://math.stackexchange.com/q/2252166> (accessed Nov. 01, 2023).
- [27] Y. Roth and N. Pickles, "Bot or not? The facts about platform manipulation on Twitter," X blog, 2020. <https://blog.twitter.com/en/topics/company/2020/bot-or-not> (accessed Nov. 01, 2023).
- [28] K. Salehzadeh Niksirat, D. Korka, H. Harkous, K. Huguenin, and M. Cherubini, "On the potential of mediation chatbots for mitigating multiparty privacy conflicts - a wizard-of-Oz study," *Proceedings of the ACM on Human-Computer Interaction*, vol. 7, no. CSCW1, pp. 1–33, Apr. 2023, doi: 10.1145/3579618.
- [29] B. Bebensee, N. Nazarov, and B.-T. Zhang, "Leveraging node neighborhoods and egograph topology for better bot detection in social graphs," *Social Network Analysis and Mining*, vol. 11, no. 1, Dec. 2021, doi: 10.1007/s13278-020-00713-z.
- [30] "About CoTweets," X Help Center. <https://help.twitter.com/en/using-twitter/cotweets> (accessed Nov. 01, 2023).
- [31] S. Ali, A. Rauf, N. Islam, and H. Farman, "A framework for secure and privacy protected collaborative contents sharing using public OSN," *Cluster Computing*, vol. 22, no. S3, pp. 7275–7286, May 2019, doi: 10.1007/s10586-017-1236-2.
- [32] S. H. Mouhsein and V. Madhavi, "Access control among multi-user for online social networks: Model and mechanisms," *International Journal of Scientific Engineering and Technology Research*, pp. 0923–0926, 2018.
- [33] J. Sicart, "ReallImpactAnalytics/trumania," *GitHub*. <https://github.com/ReallImpactAnalytics/trumania> (accessed Nov. 01, 2023).
- [34] "Sentiment analysis prebuilt model," *Microsoft*. <https://learn.microsoft.com/en-us/ai-builder/prebuilt-sentiment-analysis> (accessed Nov. 01, 2023).
- [35] S. Abid and M. I. Daud, "A dynamic and automated access control management system for social networks," *Security and Communication Networks*, vol. 2022, pp. 1–11, Oct. 2022, doi: 10.1155/2022/1929339.
- [36] H. Gao, Z. Ma, S. Luo, Y. Xu, and Z. Wu, "BSSPD: A blockchain-based security sharing scheme for personal data with fine-grained access control," *Wireless Communications and Mobile Computing*, vol. 2021, pp. 1–20, Feb. 2021, doi: 10.1155/2021/6658920.
- [37] H. Hu, G.-J. Ahn, Z. Zhao, and D. Yang, "Game theoretic analysis of multiparty access control in online social networks," in *Proceedings of the 19th ACM symposium on Access control models and technologies*, Jun. 2014, pp. 93–102, doi: 10.1145/2613087.2613097.
- [38] S. Jasmine and M. Hymavathi, "Implementation of collaborative privacy control scheme in OSN," *International Journal of Scientific Engineering and Technology Research*, vol. 8, pp. 545–548, 2019.
- [39] A. Ramteke and G. Talmale, "Authorization mechanism for multiparty data sharing in social network," *International Journal of Research in Engineering and Technology*, vol. 03, no. 04, pp. 311–316, Apr. 2014, doi: 10.15623/ijret.2014.0304056.
- [40] H. Hu, G.-J. Ahn, and J. Jorgensen, "Detecting and resolving privacy conflicts for collaborative data sharing in online social networks," in *Proceedings of the 27th Annual Computer Security Applications Conference*, Dec. 2011, pp. 103–112, doi: 10.1145/2076732.2076747.
- [41] P. Ilija, B. Carminati, E. Ferrari, P. Fragopoulou, and S. Ioannidis, "SAMPAC: Socially-aware collaborative multi-party access control," in *Proceedings of the Seventh ACM on Conference on Data and Application Security and Privacy*, Mar. 2017, pp. 71–82, doi: 10.1145/3029806.3029834.

BIOGRAPHIES OF AUTHORS







Nisha P. Shetty     has published in the areas network security and machine learning. Currently she is working in the area of social network security. She serves as assistant professor in the Department of Information and Communication Technology, Manipal Institute of Technology, Manipal. She has completed her bachelors and master's in computer science and engineering from Visvesvaraya Technological University, Karnataka, India. She is now pursuing her PhD. at Manipal Institute of Technology, MAHE, Manipal. She can be contacted at email: nisha.pshetty@manipal.edu.







Balachandra Muniyal     research area includes network security, algorithms, and operating systems. He has more than 30 publications in national and international conferences/journals. Currently he is working as the Professor in the Department of Information and Communication Technology, Manipal Institute of Technology, Manipal. He has 25 years of teaching experience in various Institutes. He also serves as a Coordinator for the Centre of Excellence-Cybersecurity at Manipal Institute of Technology, Manipal. He can be contacted at email: bala.chandra@manipal.edu.



Nandini Proothi     completed her bachelor's in information technology from Manipal Institute of Technology, India. Her current research interests include social media security, accessibility of secure online systems, and the intersection of human-computer interaction and privacy. She can be contacted at email: nandini.proothi1@learner.manipal.edu. Her profile can be found at <https://www.linkedin.com/in/nandini-proothi/>.



Bhavya Gopal     received a Bachelor of Technology in Chemical Engineering from Manipal Institute of Technology, Karnataka, India, in 2022. Her current research interests include artificial intelligence, social media security and privacy. He can be contacted at email: bhavya.gopal1@learner.manipal.edu. Her profile can be found at <https://www.linkedin.com/in/bhavyagopal>.