mathematics

MDPI

Article

Rule Fusion of Privacy Protection Strategies for Co-Ownership Data Sharing

Tinghuai Ma [1,*], Yuming Su [1], Huan Rong [1], Yurong Qian [2] and Najla Al-Nabhan [3]

[1] School of Computer & Software, Nanjing University of Information Science and Technology, Nanjing 210044, China; suyuming97@nuist.edu.cn (Y.S.); rhuan@nuist.edu.cn (H.R.)
[2] College of Information Science and Engineering, Xinjiang University, Urumqi 830008, China; qyr@xju.edu.cn
[3] Department Computer Science, King Saud University, Riyadh 11362, Saudi Arabia; nalnabhan@ksu.edu.sa
* Correspondence: thma@nuist.edu.cn

Abstract: With the rapid development of social networks, personal privacy leakage has become more and more serious. A social network is a shared platform. Resources in a social network may be shared by multiple owners. In order to prevent privacy leakage, each owner assigns a corresponding privacy protection strategy. For the same shared contents, integrating the privacy protection strategies of all owners is the key problem for sharing. This paper proposes a rule fusion method of privacy protection for the co-ownership of data shared in social networks. First, the content of the protection is defined according to different privacy requirements. Second, this paper uses predicate logic formulas to abstract the natural language-based description of privacy protection and further provides a logical model of privacy protection rules. Third, this paper gives the definition of privacy protection heterogeneous rules and provides a rule fusion algorithm to ensure no conflict exists among these rules. The experimental results show that the proposed rule-based fusion method of privacy protection strategy performs at a higher level than the privacy protection strategy fusion.

Keywords: privacy protection; social network; strategy; integration; rule fusion

Citation: Ma, T.; Su, Y.; Rong, H.; Qian, Y.; Al-Nabhan, N. Rule Fusion of Privacy Protection Strategies for Co-Ownership Data Sharing. *Mathematics* **2022**, *10*, 969. https://doi.org/10.3390/math10060969

Academic Editors: Junjian Huang, Xing He, Huaqing Li and Antanas Cenys

Received: 18 January 2022
Accepted: 12 March 2022
Published: 18 March 2022

1. Introduction

Within information sharing, identity disclosure, link disclosure, and content disclosure may exist. Accordingly, privacy protection in the process of information sharing is considered the protection of shared content and the protection of the sharing owners. The protection methods of shared content mainly include content encryption, generalization and abstraction [1]. The privacy information of the sharing owners mainly includes the identification information, the sensitive attributes and the interpersonal relationships [2]. Compared with traditional privacy protection techniques, such as k-anonymization, generalization, random perturbation, and clustering, it is more challenging to study the privacy protection technology of social networks, which is mainly reflected in [3]. First, due to the complexity of social networks, attackers' background is difficult to simulate. Second, the social interpersonal relationships in social networks should be protected. Third, it is more difficult to measure the amount of information loss in social networks. Fourth, the widespread and immediate nature of information sharing on the Internet has enhanced the timeliness of privacy protection. Therefore, research on the privacy protection of information sharing processes in social networks has significant challenges.

As a sharing platform, the resources shared in social networks may be shared by multiple owners. Usually, privacy protection is carried out for the sharing information involved owners in a social network, so the data co-ownership problem becomes the key to the privacy protection of a social network. The fusion of the privacy control policies of all sharing owners, for the same content, can solve the privacy protection problem of data co-ownership and solve the fusion of multi-owner privacy policies in information sharing. In order to promote the development of social network services, it is necessary

to provide technical support to prevent privacy leakage in social networks, realize the regularization of privacy protection, and provide solutions regarding the fusion of multi-owner privacy policies.

The rest of this paper is organized as follows. Section 2 introduces related works on privacy protection and policy integration. In Section 3, we introduce the rule fusion algorithm of privacy protection strategies. In Section 4, we present the verification and comparison experiments, and the experimental results illustrate its superiority. Section 5 concludes this paper and the future research.

2. Related Works

2.1. Privacy Protection

In the information era, privacy protection is increasingly attractive to researchers, and the research related to privacy protection technology is also becoming more abundant. Anonymity protection, data disturbance, encryption and other technical solutions are the traditional classic privacy protection technologies. These technologies are mainly applied to relational data, but the social network, defined as a graph structure, usually has diverse information such as structures, edges and nodes. Therefore, these classic privacy protection technologies cannot completely transfer to the protection of privacy in social networks. According to the characteristics of the graph structure of social networks, the structure, edge, and node information form privacy in social networks. Therefore, the mainstream research direction of privacy protection of social networks is based on the protection of structures, edges and nodes.

In 2007, Backstrom et al. [4] grouped the protection methods as node anonymity, edge weight disturbance, and edge structure anonymity. Research on the fusion of privacy control strategies of all sharing owners for the same content has become the key to dealing with the issue of data co-ownership. In the interdisciplinary research, the related work on privacy protection also has many merits, such as confidentiality, leakage, anonymity, and intrusion, which are closely related to privacy. Martínez et al. [5], in 2013, and Xia et al. [6], in 2017, proposed the use of privacy protection technology to achieve semantic preservation, so as to prevent the leakage of non-numerical clinical information. In 2016, Wang et al. [7] proposed a lightweight and efficient VANET communication authentication scheme with strong confidentiality that uses self-generated pseudo-identities to ensure privacy protection and condition tracking, which is suitable for real-time emergency reporting applications. In 2016, Li et al. [8] solved the privacy crisis of multi-agent users in smart cities. In addition to privacy protection, other attack schemes use anonymous beacons for information transmission to realize mobile node tracking, such as the nearest neighbor-based probabilistic association data algorithm (NNPDA) proposed by Emara et al. [9] in 2013. With the increasing number of research programs, the launched research focus has shifted from the traditional interpretation of privacy and privacy endangerment to the specific measurement of privacy protection from the perspective of multiple owners.

In the research of privacy protection, a key issue is to achieve privacy security while ensuring data usability. In 2014 and 2015, Krontiris and Fu et al. [10,11] proposed to minimize the reduction of information or achieve the goal of individual anonymity, which can be achieved by building an integrated architecture system based on actual sharing behavior and combining technical solutions. The well-known k-anonymity privacy protection model proposed by Samarati and Sweeney [12,13] successfully forbade privacy leakage and avoided the possibility for multi-owners to be identified only by a unique identification. The method based on k-anonymity has high universality and low complexity and is widely used in the research of privacy protection. In 2006, Machanavajjhala et al. [14] proposed the *l*-diversity model based on the expansion of the *k*-anonymity method, which prevents intruders from leaking information without recording and identifying operations. In addition, the *t*-intimacy theory proposed by Li et al. [15] in 2007 defined that the distance between two distribution relationships is less than or equal to the specific threshold *t*. The differential privacy model proposed by Zhao et al.[16] in 2019 is another prominent

privacy model. The specific method to measure privacy is to obtain the increased value of a database facing a leakage risk. In most cases, the scheme can provide accurate information regarding relevant databases and ensure a high level of privacy protection.

According to the previous research work, it is easy to find that the research of privacy protection strategies is closely related to the regularization and quantitative of privacy. In 2013, Sankar et al. [17] pointed out that the reasonable balance between privacy protection and data usability can be solved by information theory solutions. In 2013, Lejun et al. [18] defined possibility, accuracy, privacy, and feasibility as four metrics for privacy leakage and combined these four metrics into one. In 2017, Kokolakis et al. [19] proposed a relevant theory of the privacy paradox, which shows that privacy behavior and privacy attitude are not completely consistent. The literature suggests that the research on privacy protection should be based on a comprehensive theoretical model, including the diversity of personal information and the diversity of privacy issues.

2.2. Policy and Privacy Protection Strategies

A privacy protection strategy mode mainly performs authorization management operations on the access request agent, access information, and information access form, as well as the relationship between the three. Due to the diversity of information, all the possible policies in a specific application environment cannot be obtained in advance in general, and with the passage of time, the privacy policies will change dynamically, which requires that the privacy protection policies should have flexible and diverse characteristics. Wang and Ishii et al. [20] obtained dynamic privacy-preserving schemes by carefully designing noise injected to a distributed computing process in 2020. A privacy protection strategy can be regarded as a kind of access authorization management. The strategy defines the specific control scope of the access request agent to perform operations on the access target, and prevents social network members from obtaining services and resources without access authorization. Privacy protection strategies have two authorization strategies: positive and negative access authorization. In social networks, the user generally decides which privacy protection strategy to adopt. In 2013, Young and Quan-Haase et al. [21] researched the privacy protection strategies on Facebook.

In a social network, the access control mechanism is generally used to realize the privacy protection policies, and the implementation of a privacy policy is simply used to answer the access request of a multi-agent. The resource and user composition are critical to access control. The security attributes of resources include access control lists, security labels, etc. There are many security attributes of users, such as identity identification, group labels, security labels, and so on. The access control mechanism matches the user's security attributes with resource security attributes to determine whether the access user is qualified to perform specific operations on the requested resource. Aghili and Sedaghat et al. [22] proposed a kind of access control scheme on the basis of multi-level security attributes in 2022. In general, when the user security level is higher than the security level of the resource, the access mechanism allows the requested resource to be accessed. In addition, attribute review and management performance are also important parts of the access control mechanism.

In the privacy protection process, an access authorization management system is specifically defined by the access control model. The access control model builds a theoretical and conceptual framework for the reasoning of privacy protection strategies, and at the same time, determines the application environment for the access authorization management system. The access control model serves as the bridge between the privacy protection strategy and the access control mechanism. In 2014, Wang and Sun et al. [23] built an access control policy model for privacy protection. From the perspective of social network managers, the access control model is a standardized design and implementation. From the perspective of accessing users, the access control model is an accurate expression of user access needs without ambiguity.

Access authorization management in social networks can be implemented based on a combination of the above-mentioned privacy protection strategy, access control mechanism, and access control model. Firstly, the privacy protection policy and access control model are established to meet the security requirements. Then, the corresponding access control model is designed according to the privacy protection policy and implemented in the application scenario. Finally, the verification step is used to verify whether the privacy protection policy is implemented correctly or not through the access control mechanism.

2.3. Rule Fusion

Fusion technology has been studied in many application fields; for example, Maskell et al. [24] carried out research in the field of target detection and tracking in 2006. Ji et al. [25] carried out research in the field of multi-sensors, Yang et al. [26] conducted research in the field of image fusion, and Kinnunen et al. [27] performed research in the field of speech processing. Information retrieval has also received extensive attention in fusion research, as this type of fusion technology can be used for a variety of research tasks, such as the routing query technology studied by Bigot et al. [28] in 2011, the personnel search proposed by MacDonald et al. [29] in 2009, the blog opinion retrieval technology studied by Wu et al. [30] in 2012, and the automatic ranking of a group of retrieval systems studied by Nurayd et al. [31] in 2006.

According to different retrieval standards, data fusion methods in information retrieval can be divided into different types. According to the required information, we can divide them into two categories: ranking-based and scoring-based. The ranking-based methods mainly include the CombSum method and the CombMNZ method proposed by Fox et al. [32] in 1993, the linear combination method proposed by Vogt [33] in 1998, and the linear correlation method proposed by Wu and McClean et al. [34] in 2006. The scoring-based methods mainly include the Borda counting method proposed by Aslam et al. [35] in 2001, Condorcet fusion proposed by Montague et al. [36] in 2002, and the MAPFuse method proposed by Lillis et al. [37] in 2010. The Bayesian fusion method proposed by Aslam et al. [35] in 2001 can use both scoring and ranking information. On the other hand, in the fusion process, we can give all component systems the same priority, or we can choose to give them different priorities. It is usually possible to obtain a biased view by using some training data to evaluate their performance or the similarity between them. In CombSum, CombMNZ, Borda counting and Condorcet fusion, all systems are treated equally, while in the linear combination method, MAPFuse and probability fusion, different systems are treated in different ways.

The linear combination method is a general form of CombSum in which scoring is available. For how to determine the appropriate weight of the linear combination method, several different methods have been studied. In 2006, Wu and McClean et al. [34] studied the weighted model considering both system performance and system differences. In 2009, Wu et al. [38] studied a series of performance power functions. In 2012, Wu et al. [39] studied a method based on multiple linear regression. Some optimization methods, such as the conjugate gradient method and the golden section search method, are also used to mine applicable weights. Some parts of these weighting patterns may be useful for weighted Condorcet fusion. In addition, M.D. Ruiz [40] and others also made corresponding research on meta-association rules and information fusion in 2017.

3. Rule Fusion of Privacy Protection Strategies

A social network is a network between people. It connects people through the carrier of the network, so as to form groups with certain characteristics. Private information is information that is relevant to individuals or groups of individuals and can reveal details about their lives or other characteristics that may affect them. In social networks, the privacy information contained in user data is diverse and complex. Therefore, service providers often develop various privacy protection strategies to reduce the risk of user privacy leakage. However, due to the complexity of privacy, the privacy rules for the

privacy protection of different users in the system are also very complicated and often cross each other. Since social networks are a sharing platform, the resources shared in social networks may be involved with multiple sharing owners, so dealing with the problem of data co-ownership has become the key to the privacy protection of social networks. The fusion of the privacy control policies of all sharing owners for the same content can reasonably solve the privacy protection problem of data co-ownership. However, once these heterogeneous rules for different privacy information are abnormal in the fusion process, it will bring users data leakage and serious security problems. Therefore, it is necessary to study the rule fusion of privacy protection policies. In this section, the abstract model of rule fusion for privacy protection is presented. The rule fusion scheme for privacy protection is abstracted into predicate logic formula, and the rule fusion of the privacy protection policy (RFPP) is realized.

3.1. Privacy Protection Rules

Users' personal information and behavior information may involve user privacy, and some control strategies customized by social network service providers also need to be protected. In the data interaction of social networks in a cloud environment, access control is the basic component unit of these interactions. Users and their privacy information are important elements in user privacy protection. This section first gives the basic definition of these elements, and then gives the privacy protection strategy in the form of an abstract model based on predicate logic.

3.1.1. Privacy Protection Elements

In social networks, privacy protection mainly involves the protected target user, other users, and the latter's various authority for the former's private information. It is worth noting that the meaning of other users in this context includes not only the persons who may be exposed to private information, but also other social network service providers.

Definition 1. *In this paper, T is used to represent the set of target users in the social network. These users are the targets of the social network service providers to formulate privacy protection rules. Use U to represent a collection of other users, and these other users may pose a potential threat to T's privacy. A represents the collection of permissions of T corresponding to different U. In this part, the meaning of permissions mainly refers to the access and modification permissions of private information.*

Definition 2. *The privacy permissions is defined as a triple containing the elements in Definition 1.*

$$(t, a, u), (t \in T) \wedge (a \in A) \wedge (u \in U) \tag{1}$$

This triple defines the permissions of other users for the target user. The permissions can generally be divided into three categories, namely no permissions, partial permissions, and full permissions. No permission at all means that the other user has no permission to access or modify target users' private information. Partial permissions indicate that other users have certain access or can even modify permissions for private information. Full authority means that other users have the authority to access and modify any private information.

3.1.2. Privacy Protection Predicate Logic Formula

Privacy protection rules in social networks, to a certain extent, can be considered as a kind of access control policy, which is used to describe the permission relationship of users to a certain protection object, users or their privacy resources. Among them, users, objects and permissions are the basic elements of privacy protection rules. Combined with other elements, such as the environment, a complete set of unified rules description is formed. A complete rule description is expressed by a language or multiple languages with no ambiguity, and these rules must be expressed by predicate logic. Specifically, the so-called predicate formula is a formula formed by connecting some predicates with

predicate connection symbols. In the predicate logic formula, the atomic proposition, that is, the smallest proposition, is composed of individual words and predicates. Individual words represent things or things that can exist independently. Predicates are words used to describe the nature of individual words, that is, words that describe a certain connection between things.

Definition 3. *The privacy protection rules are defined as the authority A of other users U against the target user T. We use the IF-THEN structure to describe the status of these persons in a specified environment:*

IF condition (environment) THEN (T,A,U).

In this paper, the basic structure of privacy protection rules is defined in Backus–Naur form:

Definition (1) *<privacy protection rules>::=<conditions>→<permissions>*

Definition (2) *<condition>::= <subcondition> {< [and|or] ><sub-condition>}*

Definition (3) *<permission>::= <behavior> {< [and|or] ><behavior>}*

Where ::= means "can be positioned as", and → is a logical connector that means "implies" or "necessary and sufficient conditions"; that is, when other users meet all the conditions defined in the privacy protection rules, they have the authority to access this private information. In terms of conditions and permissions, the conditions in rules can be expressed as logical expressions of a group of sub-conditions, and the permissions obtained by users when they get permission are also a set of logical expressions of their legitimate behaviors.

3.1.3. Logic Model of Privacy Protection Rules

When the predicate logic formula for privacy protection is given, the privacy protection rule model can be established accordingly. The specific permissions, behaviors, and conditions in the model are different in different situations. In this section, predicate logic formulas are still used to describe the relationship between these objects. Among them, the symbol ∧ is used to represent "and", which means that the condition in the rule is reached by a number of sub-conditions and represents a Boolean expression of various constraints.

For example, the school may only allow the students in the school to connect to the online library through the campus network at a specific time. In this case, the campus online library becomes the privacy that the school needs to protect, and the students are the users who request access rights, so they can access the library data only under certain conditions. The rules for protecting the privacy of users can be described in Table 1.

Table 1. When the time is 9:00–21:00, students can access the online library of the school through the campus network.

	Rule Elements: Object S; User s; Time t.
	(1) person(s) ∧ object(S)
	(2) student(s) ∧school(S)
	(3) statusIn(s,S) ∧ inCampusNetwork(s, S)
	(4) timeAfter(9) ∧ timeBefore(21)
	(5) school(S) ∧ openLibarayAcess(S) requester(s)

The above five items represent the execution conditions of privacy rules. If privacy protection rules are activated, the users can meet the conditions of privacy policies and obtain permissions only when all conditions are fully met. In this way, we can clearly observe the requirements of privacy policy and regularize the logic of the privacy protection policy in the way of predicate logic. In this example, the privacy protection rule will only be activated when the target user is a school, the requesting user is a student, and the requested content is access to the campus network. When we use natural language to describe these rules to stimulate conditions, it is more redundant, so further, we can use the Backus–Naur Form paradigm to define the set of these conditions.

First of all, the relationship reasons in privacy protection rules are used to describe each object. T represents the collection of target objects, and U represents the collection of other user objects. These user objects in the collection usually issue various access requests to the objects T and represent the collection of rules; A represents a collection of rules. In addition to the above abstract description, for specific objects, we use $t_i \in T$ to represent the specific protection user object individual, such as the school in the above example. Use $u_i \in U$ to indicate a specific requesting user object individual; t_i and u_i have different specific connotations in different situations. Similarly, we use $c_i \in C$ to represent the object of the privacy content and the object that needs to be protected in the social network. In privacy protection rules, the fundamental protection object is the privacy content of the target object t_i. c_i represents the specific privacy contents of t_i, and C represents a collection of privacy contents.

In the above example, a predicate formula is used to express the relationship between two objects. For example, statusIn(*,*) judges whether the student belongs to the school by judging whether the student's status is in the school. The relationship between the target object (T) and other users (U) are expressed in this form. Due to the presence of multiple relationships, multiple predicate formulas are often used. Similarly, predicates such as student(*) are used to express the identity or attributes of an object. On the basis of the above content, we use E to represent the environment of policy execution.

Definition 4. *The execution environment in privacy protection rules can be written as follows:*

$$\begin{gathered} E ::= \exists t_i \in T, \exists u_i \in U, \exists c_i \in C \\ [relation(t_i, u_i), \ldots | object(t_i), \ldots | behave(u_i), \ldots | c_i, \ldots] \end{gathered} \tag{2}$$

The execution environment of privacy protection rules is given in the form of the Backus–Naur Form paradigm in formula (2). Then, we will define standard privacy rules based on policy primitives.

Definition 5. *The standard privacy rules based on policy primitives can be written as follows:*

$$\begin{gathered} P ::= \forall t_i \in T, \forall u_i \in U, \forall c_i \in C, \forall a_i \in A \\ [relation(t_i, u_i) \wedge object(t_i) \wedge behave(u_i) \rightarrow access(u_i, c_i, a_i)] \end{gathered} \tag{3}$$

According to formula (3), the privacy protection rules are described as the following model.

Definition 6. *The privacy protection rule model can be written as follows:*

$$\begin{gathered} \forall k_1 \in K_1, \forall k_2 \in K_2, \forall k_3 \in K_3, \ldots \forall k_n \in K_n \\ [p_1(k_1, k_2, \ldots k_n) \wedge p_2(k_1, k_2, \ldots k_n) \wedge \ldots p_m(k_1, k_2, \ldots kc_n) \rightarrow access(u_i, c_i, a_i)] \end{gathered} \tag{4}$$

We can simplify formula (4) to:

$$\begin{gathered} \forall k_1 \in K_1, \forall k_2 \in K_2, \forall k_3 \in K_3, \ldots \forall k_n \in K_n \\ [f \rightarrow access(u_i, c_i, a_i)] \end{gathered} \tag{5}$$

Among them, K_1 , K_2 , … K_n represent different categories of first-order logic polynomials, and k_1 , k_2 , … k_n represent specific items under the corresponding categories, and p_1 , p_2 , … p_n represent atomic formulas about these items. In this part, we use f to represent the first-order logic formula.

According to the above privacy protection rule model structure, we can clearly describe the specific connotation of the rule and the multiple objects it contains. Rules are expressed as a set of constraints in the model. When and only when the conditions of these rules are satisfied, the privacy information will be sent to the requesting user, so as to achieve the function of protecting privacy data in the social network. In the previous example, "When

the time is 9:00–21:00, student users can access the school's online library through the campus network", the privacy rule described by the natural language rule can be modeled into the following first-order multi-category logic formula:

$$\begin{array}{c} \forall t_i \forall u_i \forall a_i (statusIn(t_i, u_i)) \wedge inCampusNetwork(t_i, u_i) \wedge timeIn(9, 21) \\ \rightarrow access(u_i, openLibraryAccess(a_i), a_i)] \end{array} \quad (6)$$

Simple natural language to express rules is very flawed, so in this section, we give the rules for expressing privacy protection rules in the form of a first-order multi-category logic formula. In specific situations, the constraints in the rules should be extracted according to the different requirements, and these rules should be re-expressed in the language of mathematics and logic.

3.2. Privacy Protection Rule Integration

3.2.1. Unified Description of Heterogeneous Rules

According to the description of privacy components in Section 3.1.1, the definition of privacy data is more complicated, and in many cases, privacy rules themselves may be part of privacy. For the convenience of description, in this section, we uniformly describe the user's request as access; that is, the behavior of other users can be expressed as a request for access to the target user's private data.

Definition 7. *Privacy request attributes. We use 'privacy request attributes' to define the user's specific access request content for private data, and we will denote it as a privacy attribute* (PA)*. Specifically, the value of PA represents the response content of the target object to the user's request. For the example in Section 3.1, the school provides students with access rights to the online library for a specific period of time. When a student wants to access the library, he first needs to make this request to the server. In this case, the value of PA is the student's access for the school's online library.*

Even in the same environment, privacy protection rules allow different target objects, users, and privacy content to be protected, as well as different privacy request attributes. Regarding the attributes of the privacy request, the user's access request may involve multiple privacy contents. In many cases, these contents usually have unequal levels, so they can be described in the form of a tree. As an example shown in Section 3.1, we can draw the access control hierarchy as the tree in Figure 1. In order to achieve a more general and abstract definition, a unified description is given in the following form:

$$Rules \rightarrow RULES = \{MA, SA, BA\} \quad (7)$$

Among them, MA represents a main privacy request attribute (main attribute), SA represents a source attribute, and BA represents a behavior attribute. In Figure 1, requests for login and registration can be expressed as behavior request attributes, and download requests for resources can be regarded as resource request attributes. We construct the user's access attributes into a privacy attribute tree (PA-T), which is composed of a limited number of nodes. Among them, the privacy request attribute MA is the root node, and we define the child nodes of MA as $PA - T_0$, $PA - T_1$, $PA - T_2$, ... $PA - T_n$. According to these child nodes, we can decompose a privacy attribute tree into different subtrees, each of which has $PA - T_i$ $(0 \leq i \leq n)$ as the root node. In the example in Figure 1, the primary request is the root node of the PA tree, which contains three child nodes. In addition, the child node login also includes student login and teacher login. Similarly, other subtrees also contain one or more child nodes.

In social networks, service providers may generally use different languages to describe privacy protection rules. However, these rules based on different languages are difficult to integrate. Therefore, a unified method is needed to express different rules. However, different rules include target objects, other users, and possible access requests. Due to the limited kind of access requests, we describe them in the form of a tree. Therefore, on

the basis of constructing the privacy request attribute tree, we can match and search the attributes in different rules. When different rules match successfully, we can extract the corresponding attribute value contained in the node and use it as the element in the set of formula (7). In the process of rule fusion, we still use the form of tree description in order to ensure the hierarchy and logic of different rules. As Algorithm 1 shows, we uniformly describe the privacy protection rules of multiple subjects in social networks and use the same form of privacy attributes to represent different rules.

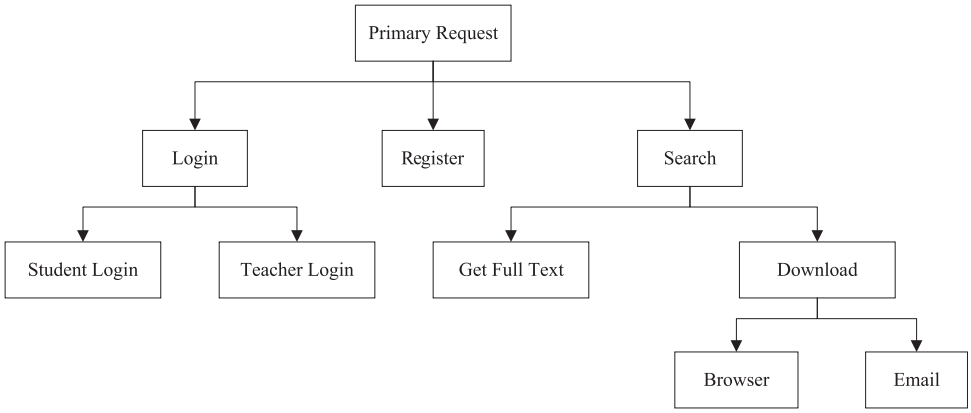

Figure 1. Examples of privacy request attributes.

Algorithm 1 Uniform description algorithm of privacy protection rules

Input: Different forms of privacy protection rules $rule_i$

Output: Uniformly described privacy protection rules $RULE$

1: **for** $rule_i$ **of** Service provider I **do**
2: **for** User request attributes **in** rules **do**
3: Grab the user request attribute described in the rule and assign it to the corresponding attribute node;
4: Establish a private request attribute instance tree T_i;
5: **end for**
6: According to the description of the object of the access request, T_i add to $rule_i$
7: **end for**

From the above algorithm, different rules of multiple users can be merged with each other. The algorithm in this section can unify these heterogeneous rules into the same description. Due to the fact that heterogeneous rules under social networks have different formalities and different contents, the requests faced by different target objects are also diverse. Therefore, the privacy content is also very different. Due to the above reasons, privacy protection rules still face some logic and priority problems in the process of fusion. Once this fusion is wrong, privacy will still be leaked. Therefore, we also propose a heterogeneous rule fusion algorithm based on an attribute tree.

3.2.2. Fusion of Heterogeneous Rules

In the process of fusion of different privacy protection rules, the first problem is to detect whether these rules have conflicts. Once the conflicts in these rules are eliminated, the remaining rules can be well organized into fusion rules. In actual situations, different privacy protection rules may be originated from different user requests.

Specifically, when the privacy protection rules of multiple social network subjects have the same privacy tree structure, the value set of a group of privacy tree node attributes has the following expression:

$$\begin{gathered} RULES_1 = \{MA_1, SA_1, BA_1\} \\ RULES_2 = \{MA_2, SA_2, BA_2\} \\ RULES_3 = \{MA_3, SA_3, BA_3\} \\ MA_1 = MA_2 = MA_3, SA_1 = SA_2 = SA_3, BA_1 = BA_2 = BA_3 \end{gathered} \tag{8}$$

When the privacy protection rules of multiple social network subjects have different privacy tree structures, the set of values of a group of privacy tree node attributes has the following expression:

$$\begin{gathered} RULES_4 = \{MA_4, SA_4, BA_4\} \\ RULES_5 = \{MA_5, SA_5, BA_5\} \\ RULES_6 = \{MA_6, SA_6, BA_6\} \\ MA_1 \cap MA_2 \cap MA_3 = MA_1||MA_2||MA_3 \\ SA_1 \cap SA_2 \cap SA_3 = SA_1||SA_2||SA_3 \\ BA_1 \cap BA_2 \cap BA_3 = BA_1||BA_2||BA_3 \end{gathered} \tag{9}$$

The structure of the privacy tree is the same, so we need to compare the attributes of root nodes, resources and behaviors. If the above attributes are consistent, there may be a set inclusion relationship between different privacy rules. If the utility of different rule nodes is different, then the conflict type between rules is utility conflict. In the case of the same privacy tree structure and inconsistent node attributes, we only need to compare the hierarchical relationship between resource attributes and the inclusion relationship of other attributes. The detail method is described in Algorithm 2.

Algorithm 2 Privacy protection rule conflict detection fusion algorithm

Input: Privacy rules $rule_i$

Output: Fusion rules $RULE_F$, Conflict rules $RULE_C$

for $rule_i$ **of** Service provider I **do**
    Read each property
    Compare the property values between rules to see if they are the same
    if Property values are the same **then**
        Only one of them will be added to $RULE_F$
        Continue to the next property
    else if Property conflicts with different property values **then**
        The conflicting attributes are stored in $RULE_C$
        for All rules in $RULE_C$ **do**
            Attribute union between rules
            Add to $RULE_F$
        end for
    end if
end for

For privacy protection rules in different situations, Algorithm 2 first fuses the same items of attribute values in these rules, leaving only one item in the fusion result rules. For conflict rules, the conflict attributes are stored in the conflict rule set, and the confliction attributes are merged to reduce the conflict rules. For the fusion rule set, the algorithm of rule fusion forms the same description of rules by establishing an attribute tree. Once these

conflicting rules are resolved, the risk of user privacy disclosure will be greatly reduced, which can further protect the privacy data security of the target object. For the new non-conflict rule attribute tree, a new consistency rule with a hierarchical structure is provided. The privacy protection rule consistency description and fusion method proposed in this paper can ensure user privacy security with low computational cost and time complexity.

4. Experiments

In the experiment, we will focus on the analysis of the substitutability and time complexity of the proposed algorithm. We randomly generate a large number of privacy protection rules that may be used in social networks. These rules and policies are similar to some rules published by famous service providers in logical and goal aspects.

4.1. Substitutability

In order to verify the usability of the rule-based fusion of privacy protection policies, the rule substitution test experiment is designed. In the experiment, the proposed method is verified by calculating the substitutability between the privacy fusion rules and the original rules. Rule replaceability mainly includes two aspects: user attribute and resource attribute. The calculation formula is as function 10.

$$\begin{aligned} A(p_1, p_2) &= A_M(p_1, p_2) * W_M + A_S(p_1, p_2) * W_S \\ A_M &= 1 - \frac{Path_{\min}(M_1, M_2)}{Path_{\max}(M_1, M_2)} \\ A_S &= 1 - \frac{Num(S_1 \cap S_2)}{Num(S_1 \cup S_2)} \end{aligned} \tag{10}$$

Among them, A_M and A_S are the substitution of users' attributes and resource attributes, p_1, p_2 are two privacy rules, M_1, M_2, S_1, S_2 are two users' attributes and source attributes, respectively, as defined in Equation (7), *Path* represents the distance of the users' attributes in the hierarchical relationship of the attribute tree, and W represents the weights. In the experiment, W_M and W_S are assigned a value of 0.7 and a value of 0.3, respectively.

In the experiment, we randomly selected 20, 40, 60, 80 and 100 rules from 100 regularized privacy protection policies to form 5 groups. After fusing the privacy protection rules in the group, a new unified description privacy rule is obtained. We compare the fusion rules of each group with all the rules in the group and obtain several alternative results. We use the average of all the substitutability results as the final fusion rule substitutability of the group. In order to verify the superiority of the method proposed in this chapter, the XACML general strategy fusion method is compared. The comparison results of the fusion substitutability of each group are shown in Figure 2.

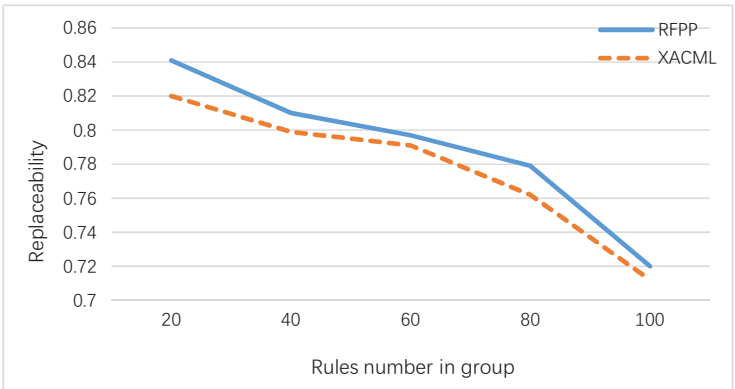

Figure 2. Alternative results of privacy rule fusion.

The experimental results show that the privacy-preserving rule fusion method proposed in this chapter has better substitutability; however, with an increased number of rules, the substitutability of the fusion results will decline.

4.2. Time Complexity

In this section, two experiments are designed to verify the conflict detection. In the first experiment, we compared the time consumption of fusion according to the number of heterogeneous rules in the proposed scheme from 100 to 1000, with an interval of 100, as shown in Figure 3, where the time unit is milliseconds. It is not difficult to find that with an increasing number of heterogeneous rules, the consumption of privacy protection rule fusion time also increases. The rule number and time consumption almost maintain a linear relationship, and the increase of time consumption is acceptable compared with the method of conflict-free detection. At the same time, the proposed fusion algorithm can avoid conflict among the rules and has good practical value.

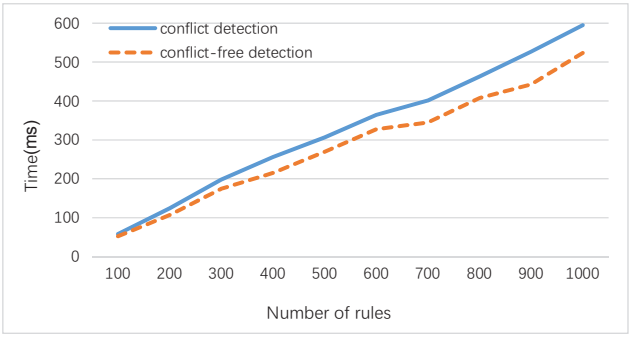

Figure 3. Time consumption of different heterogeneous rules.

In order to further study the characteristics of the time complexity of the proposed scheme, the experiments in this section not only analyze the influence of the number of heterogeneous rules on the fusion time, but also consider the influence of the number of attributes in the heterogeneous rules on the fusion time. As shown in Figure 4, the time consumption of fusion is compared with the number of attributes in the heterogeneous rules in the scheme proposed, from 1000 to 10,000, with an interval of 1000, and the time unit is also in milliseconds. It can be found that the fusion time also increases with the number of attributes, and the increase in time consumption is also within the acceptable range compared with the conflict-free detection method. With the increase in the number of attributes, the fusion time of this scheme will increase significantly.

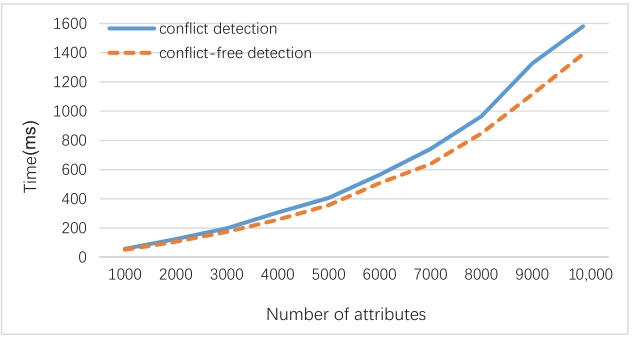

Figure 4. Time consumption of different attribute numbers in heterogeneous rules.

5. Conclusions

This paper proposes a rule fusion method of privacy protection strategies to solve the privacy security problem of co-ownership data sharing in social networks. First, according to the different connotations of private data, the content of the protection information is defined. Then, a predicate logic formula is used to abstract the natural language description of privacy protection, and a logical model of privacy protection rules is constructed. Finally, this paper gives the definition of heterogeneous privacy protection rules and gives the algorithm for rule fusion, which avoids the conflict of privacy protection requirements in

different application scenarios in social networks and the leakage of privacy data. The experimental results verify that the rule-based fusion method of privacy protection strategy proposed in this paper can perform well regarding rule substitution, and the time consumption of the algorithm is also acceptable.

In the complex social network environment, different access requests, services, and privacy data increase the difficulty of formulating privacy rules. Although the regularized privacy description and rule logic model proposed in this article provide a basis for rule fusion to a certain extent, the unified description of privacy protection rules in this article is still relatively simple, so the design of the fusion algorithm is relatively elementary. In future work, we will study and design a more comprehensive unified expression method for heterogeneous rules to further improve the accuracy and security of the privacy rule fusion mechanism.

Author Contributions: Conceptualization, T.M.; algorithm and methodology, Y.S.; writing—review and editing, H.R. and Y.Q.; funding acquisition, T.M. and N.A.-N. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Key Research and Development Program of China (No.2021YFE014400). This work was supported in part by the National Natural Science Foundation of China (No.61877034). The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding this work through research group no. RGP-1441-33.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yuan, M.; Chen, L.; Yu, P.S.; Yu, T. Protecting Sensitive Labels in Social Network Data Anonymization. *IEEE Trans. Knowl. Data Eng.* **2013**, *25*, 633–647. [CrossRef]
2. Michota, A.K.; Katsikas, S.K. Designing a seamless privacy policy for social networks. In Proceedings of the 19th Panhellenic Conference on Informatics (PCI 2015), Athens, Greece, 1–3 October 2015; Karanikolas, N.N., Akoumianakis, D., Nikolaidou, M., Vergados, D.D., Xenos, M., Giaglis, G.M., Gritzalis, S., Merakos, L.F., Tsanakas, P., Sgouropoulou, C., Eds.; ACM: New York, NY, USA, 2015; pp. 139–143.
3. Zhang, L.; Yang, S.; Li, J.; Yu, L. A Particle Swarm Optimization Clustering-Based Attribute Generalization Privacy Protection Scheme. *J. Circuits Syst. Comput.* **2018**, *27*, 1850179:1–1850179:21. [CrossRef]
4. Backstrom, L.; Dwork, C.; Kleinberg, J.M. Wherefore art thou R3579X?: Anonymized social networks, hidden patterns, and structural steganography. *Commun. ACM* **2011**, *54*, 133–141. [CrossRef]
5. Martínez, S.; Sánchez, D.; Valls, A. A semantic framework to protect the privacy of electronic health records with non-numerical attributes. *J. Biomed. Inform.* **2013**, *46*, 294–303. [CrossRef]
6. Xia, Z.; Wang, X.; Zhang, L.; Qin, Z.; Sun, X.; Ren, K. A Privacy-Preserving and Copy-Deterrence Content-Based Image Retrieval Scheme in Cloud Computing. *IEEE Trans. Inf. Forensics Secur.* **2016**, *11*, 2594–2608. [CrossRef]
7. Wang, M.; Liu, D.; Zhu, L.; Xu, Y.; Wang, F. LESPP: Lightweight and efficient strong privacy preserving authentication scheme for secure VANET communication. *Computing* **2016**, *98*, 685–708. [CrossRef]
8. Li, Y.; Dai, W.; Ming, Z.; Qiu, M. Privacy Protection for Preventing Data Over-Collection in Smart City. *IEEE Trans. Comput.* **2016**, *65*, 1339–1350. [CrossRef]
9. Emara, K.; Woerndl, W.; Schlichter, J.H. Vehicle tracking using vehicular network beacons. In Proceedings of the IEEE 14th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM 2013), Madrid, Spain, 4–7 June 2013; IEEE Computer Society: Washington, DC, USA, 2013; pp. 1–6.
10. Krontiris, I.; Langheinrich, M.; Shilton, K. Trust and privacy in mobile experience sharing: Future challenges and avenues for research. *IEEE Commun. Mag.* **2014**, *52*, 50–55. [CrossRef]
11. Fu, Z.; Sun, X.; Liu, Q.; Zhou, L.; Shu, J. Achieving Efficient Cloud Search Services: Multi-Keyword Ranked Search over Encrypted Cloud Data Supporting Parallel Computing. *IEICE Trans. Commun.* **2015**, *98-B*, 190–200. [CrossRef]
12. Samarati, P. Protecting Respondents' Identities in Microdata Release. *IEEE Trans. Knowl. Data Eng.* **2001**, *13*, 1010–1027. [CrossRef]
13. Sweeney, L. k-Anonymity: A Model for Protecting Privacy. *Int. J. Uncertain. Fuzziness Knowl. Based Syst.* **2002**, *10*, 557–570. [CrossRef]

14. Machanavajjhala, A.; Kifer, D.; Gehrke, J.; Venkitasubramaniam, M. *L*-diversity: Privacy beyond *k*-anonymity. *ACM Trans. Knowl. Discov. Data* **2007**, *1*, 3. [CrossRef]
15. Li, N.; Li, T.; Venkatasubramanian, S. t-Closeness: Privacy Beyond k-Anonymity and l-Diversity. In Proceedings of the 23rd International Conference on Data Engineering (ICDE 2007), The Marmara Hotel, Istanbul, Turkey, 15–20 April 2007; Chirkova, R., Dogac, A., Özsu, M.T., Sellis, T.K., Eds.; IEEE Computer Society: Washington, DC, USA, 2007; pp. 106–115.
16. Zhao, J.; Chen, Y.; Zhang, W. Differential Privacy Preservation in Deep Learning: Challenges, Opportunities and Solutions. *IEEE Access* **2019**, *7*, 48901–48911. [CrossRef]
17. Sankar, L.; Rajagopalan, S.R.; Poor, H.V. Utility-Privacy Tradeoffs in Databases: An Information-Theoretic Approach. *IEEE Trans. Inf. Forensics Secur.* **2013**, *8*, 838–852. [CrossRef]
18. Lejun, F.; Yuanzhuo, W.; Xiaolong, J.; Jingyuan, L.; Xueqi, C.; Shuyuan, J.; Francesco, P. Comprehensive Quantitative Analysis on Privacy Leak Behavior. *PLoS ONE* **2013**, *8*, e73410.
19. Kokolakis, S. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Comput. Secur.* **2017**, *64*, 122–134. [CrossRef]
20. Wang, X.; Ishii, H.; He, J.; Cheng, P. Dynamic Privacy-preserving Collaborative Schemes for Average Computation. *IFAC-PapersOnLine* **2020**, *53*, 2963–2968. [CrossRef]
21. Young, A.L.; Quan-Haase, A. Privacy protection strategies on facebook. *Inf. Commun. Soc.* **2013**, *16*, 479–500. [CrossRef]
22. Aghili, S.F.; Sedaghat, M.; Singelée, D.; Gupta, M. MLS-ABAC: Efficient Multi-Level Security Attribute-Based Access Control scheme. *Future Gener. Comput. Syst.* **2022**, *131*, 75–90. [CrossRef]
23. Wang, H.; Sun, L.; Bertino, E. Building access control policy model for privacy preserving and testing policy conflicting problems. *J. Comput. Syst. Sci.* **2014**, *80*, 1493–1503. [CrossRef]
24. Maskell, S.; Everitt, R.G.; Wright, R.; Briers, M. Multi-target out-of-sequence data association: Tracking using graphical models. *Inf. Fusion* **2006**, *7*, 434–447. [CrossRef]
25. Ji, Z.; Wu, Q.M.J. An improved artificial immune algorithm with application to multiple sensor systems. *Inf. Fusion* **2010**, *11*, 174–182. [CrossRef]
26. Yang, S.; Wang, M.; Jiao, L.; Wu, R.; Wang, Z. Image fusion based on a new contourlet packet. *Inf. Fusion* **2010**, *11*, 78–84. [CrossRef]
27. Kinnunen, T.; Li, H. An overview of text-independent speaker recognition: From features to supervectors. *Speech Commun.* **2010**, *52*, 12–40. [CrossRef]
28. Bigot, A.; Chrisment, C.; Dkaki, T.; Hubert, G.; Mothe, J. Fusing different information retrieval systems according to query-topics: A study based on correlation in information retrieval systems and TREC topics. *Inf. Retr.* **2011**, *14*, 617–648. [CrossRef]
29. Macdonald, C. The voting model for people search. *SIGIR Forum* **2009**, *43*, 73. [CrossRef]
30. Wu, S. Applying the data fusion technique to blog opinion retrieval. *Expert Syst. Appl.* **2012**, *39*, 1346–1353. [CrossRef]
31. Nuray, R.; Can, F. Automatic ranking of information retrieval systems using data fusion. *Inf. Process. Manag.* **2006**, *42*, 595–614. [CrossRef]
32. Fox, E.A.; Koushik, M.P.; Shaw, J.A.; Modlin, R.; Rao, D. Combining Evidence from Multiple Searches. In Proceedings of the First Text REtrieval Conference (TREC 1992), Gaithersburg, MD, USA, 4–6 November 1992; Harman, D.K., Ed.; National Institute of Standards and Technology (NIST): Gaithersburg, MD, USA, 1992; Volume 500-207, pp. 319–328.
33. Vogt, C.C.; Cottrell, G.W. Predicting the Performance of Linearly Combined IR Systems. In Proceedings of the SIGIR '98: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Melbourne, Australia, 24–28 August 1998; Croft, W.B., Moffat, A., van Rijsbergen, C.J., Wilkinson, R., Zobel, J., Eds.; ACM: New York, NY, USA, 1998; pp. 190–196.
34. Wu, S.; McClean, S.I. Performance prediction of data fusion for information retrieval. *Inf. Process. Manag.* **2006**, *42*, 899–915. [CrossRef]
35. Aslam, J.A.; Montague, M.H. Models for Metasearch. In Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, New Orleans, LA, USA, 9–13 September 2001; Croft, W.B., Harper, D.J., Kraft, D.H., Zobel, J., Eds.; ACM: New York, NY, USA, 2001; pp. 275–284.
36. Montague, M.H.; Aslam, J.A. Condorcet fusion for improved retrieval. In Proceedings of the 2002 ACM CIKM International Conference on Information and Knowledge Management, McLean, VA, USA, 4–9 November 2002; ACM: New York, NY, USA, 2002; pp. 538–548.
37. Lillis, D.; Zhang, L.; Toolan, F.; Collier, R.W.; Leonard, D.; Dunnion, J. Estimating probabilities for effective data fusion. In Proceeding of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2010), Geneva, Switzerland, 19–23 July 2010; Crestani, F., Marchand-Maillet, S., Chen, H., Efthimiadis, E.N., Savoy, J., Eds.; ACM: New York, NY, USA, 2010; pp. 347–354.
38. Wu, S.; Bi, Y.; Zeng, X.; Han, L. Assigning appropriate weights for the linear combination data fusion method in information retrieval. *Inf. Process. Manag.* **2009**, *45*, 413–426. [CrossRef]
39. Wu, S. Linear combination of component results in information retrieval. *Data Knowl. Eng.* **2012**, *71*, 114–126. [CrossRef]
40. Ruiz, M.D.; Gómez-Romero, J.; Molina-Solana, M.; Ros, M.; Martín-Bautista, M.J. Information fusion from multiple databases using meta-association rules. *Int. J. Approx. Reason.* **2017**, *80*, 185–198. [CrossRef]