面向社交网络的多方授权模型

霍颖瑜[1] 马 莉[1] 钟 勇[1] 秦小麟[2]
(佛山科学技术学院电子与信息工程学院 佛山 528000)[1]
(南京航空航天大学信息安全研究所 南京 210016)[2]

摘 要 现有的访问控制机制大多局限在用户个人空间内的数据,难以控制个人空间以外的数据,例如用户不能对其在朋友空间中发布的评论进行访问控制,不能对共有的资源进行联合访问控制等。面向社交网络的多方授权模型MRuleSN采用单一所有、多方共有的方法处理所有权问题,采用扩展的w-Datalog规则表达授权,具有更强的灵活性、访问细粒度和表达能力。分析并说明了模型的规则结构、授权语言的语法和语义,最后通过示例说明了该模型的应用和表达能力。

关键词 多方授权,社交网络,授权规则,w-Datalog
中图法分类号 TP309 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.1.027

Multi-party Authorization Model for Social Networks

HUO Ying-yu[1] MA Li[1] ZHONG Yong[1] QIN Xiao-lin[2]
(School of Electronic and Information Engineering, Foshan University, Foshan 528000, China)[1]
(Institute of Information Security, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)[2]

Abstract Most of the current access control mechanisms focus on data of private space of users, which cannot control the data beyond the personal space, such that the remarks published by a user in the space of his friend cannot be controlled by the user and the shared resource cannot be controlled jointly by the sharer. The paper presented the MRuleSN, a multi-party authorization model for social networks. The model processes the problem of ownership by single ownership and multi-party shareholders, and adopts extended w-Datalog rules to express authorization, which owns more powerful flexibility, fine-grained access control and authorization expressiveness. The rule structure, syntax and semantic of authorization language of the model were analyzed and explained. Finally, application and expressiveness of the model were exampled and discussed.

Keywords Multi-party authorization, Social networks, Authorization rules, w-Datalog

1 引言

社交网络成为近年来最为流行的互联网应用之一,如Facebook、Twitter、人人网、微博等都拥有大量的用户。用户在社交网络上分享信息,交流互动,但也带来了社交网络数据被非法使用和个人隐私泄漏等现象。采用访问控制机制对用户个人空间信息进行保护是社交网络的通用做法[1,2]。但现有的访问控制机制多局限在用户个人空间内的数据,难以控制个人空间以外的数据,如用户不能对其在朋友空间中发表的评论进行访问控制,不能对共有的资源进行联合访问控制等[3]。因而,社交网络中的多方访问控制(Multiparty Access Control)问题得到了学术界的广泛关注。

Hu等[3]提出了一种多方授权的模型和机制,同时也提出了多方授权的逻辑表达模型,其方法在授权的表达力灵活性上还略有不足;Thomas等[4]分析了因缺乏多方授权控制而带来的隐私风险,他们以Facebook为例说明了现有的所有者进行访问控制的单方授权方法难以保护用户的隐私;Squicciarini等[5]提出在社交网络中采用所有权共有(co-ownership)的方式处理共享内容,所有权可由内容创建者分配给多个用户,他们的方法中所有权可以由多人同时使用,这带来了管理上的混乱;Amrutha等[6]对多方授权的决策机制和投票机制进行了详细的研究,并提出了一个多方授权模型,该方法未考虑在内容层次上的授权方法,授权的灵活性不足;Shaik等[7]从使用控制的角度对图像隐私的多方保护进行了研究;Yeung等[8]从语义Web的角度对在线图片的多方访问控制进行了研究。现有的方法既缺乏有足够表达力的授权语言来描述多方授权的复杂性,又缺乏社交网络访问控制所需要的灵活性和细粒度性,难以满足社交网络中的多方访问控制需要。

到稿日期:2014-12-23 返修日期:2015-04-09 本文受国家自然科学基金(61373015,11326123),广东省教育厅育苗工程项目(2013LYM0097),佛山市科技发展专项资金项目(2012AA100251),佛山科学技术学院科研项目资助。
霍颖瑜(1979-),女,硕士,讲师,主要研究方向为社交网络访问控制、信息安全,E-mail:fosuhyy@163.com;马 莉(1977-),女,硕士,副教授,主要研究方向为形式化方法、信息安全、访问控制等;钟 勇(1970-),男,博士后,教授,主要研究方向为信息安全、访问控制、数字版权保护技术等;秦小麟(1953-),男,教授,博士生导师,主要研究方向为安全数据库、空间数据库、时空数据库、信息安全等。

针对上述问题，本文在逻辑语言的基础上，提出多方授权模型 MRuleSN，该模型具有如下创新和特色：

(1)采用单一所有、多方共有的方法处理所有权问题，符合社交网络的实际情况。

(2)针对 Datalog 规则无法表达多方授权的可选择性和权重等因素的情况，提出一种扩展的带权重的 w-Datalog 规则，对 w-Datalog 的语法和语义及其评价(evaluation)算法进行了描述，w-Datalog 解决了普通 Datalog 规则难以表达多方决策的问题。

(3)访问模型建立在社交网络内容的结构层次上，具有更强的灵活性和更细的访问粒度。

2 社交网络访问控制模型 RuleSN

基于规则的社交网络访问控制模型(Rule Based Social Network Access Control Model，RuleSN)是我们提出的一种基于 Datalog 逻辑规则的社交网络访问控制模型。该模型通过用户之间的关系、用户与资源的关系、资源与资源之间的关系以及实体属性进行访问控制。在该模型中，具有属性的实体包括主体、客体和关系。

主体包括用户、用户群和系统群。用户可加入用户群成为该群的成员，将社交网络自身看作一个系统群，默认所有用户和用户群均是该系统群的成员，用户群之间不再有包含关系，图 1 所示是以 qq 系统为例的主体层次。

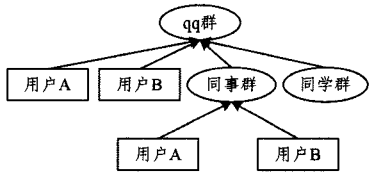

图 1 主体层次示例

客体包括内容和空间。内容是独立存在的客体，如网络链接、文章、博客、评论、注释以及各类图片、视频、音频等都属于内容客体。通过从属关系，内容之间可形成如图 2 所示的内容层次，对文章 A 的评论与文章 A 之间存在直接从属关系，该评论与对该评论的回复也存在直接从属关系，形成如图 2 所示的内容层次。空间用于存放内容，空间里的内容为该空间所有，所有内容均应属于某空间，空间可包含子空间，形成如图 2 所示的空间层次。

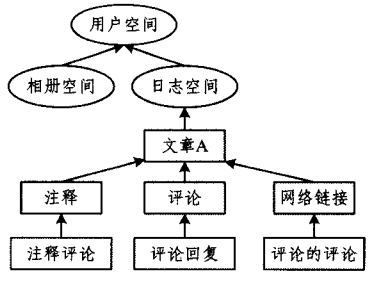

图 2 客体层次示例

系统、主体和客体均具有访问控制策略，当主体访问客体时，需要通过系统策略、客体拥有者的访问控制策略、客体自身的访问控制策略 3 方的验证。

关系作为一种实体存在，分为两类：一类是用户与用户之间的关系，如朋友关系等；另一类是主体和客体之间的关系，如拥有关系、共有关系、原创关系和传播等。RuleSN 模型的关系如图 3 所示。

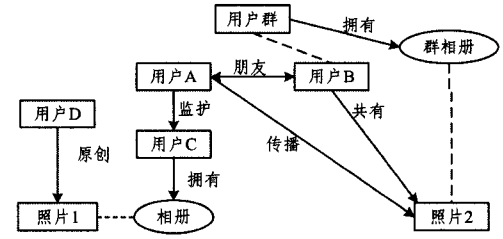

图 3 关系网络图示例

3 多方授权模型 MRuleSN

3.1 RuleSN 模型中的多方授权问题

RuleSN 模型能完整地表达用户之间的关系、用户与资源的关系、资源与资源之间的关系以及实体属性，具有较强的授权表达力和灵活性，但 RuleSN 模型未考虑多方授权问题，主要原因在于 Datalog 规则无法表达多方授权的可选择性和权重等因素。

社交网络控制中的多方授权问题包括：

(1)共有客体授权问题。如对提交到朋友空间的客体，提交者与空间所有者对该客体的联合授权问题。

(2)第三方隐私控制问题。对牵涉到第三方的客体如图片中的第三方(提交者与所有者之外)的隐私控制问题。

(3)多关系客体的联合授权问题。主客体之间包括拥有、标注、共有等关系，具有多关系的主体的联合授权问题。

(4)多层次客体的授权冲突问题。属于不同层次的客体导致的授权冲突问题。

(5)客体复制移动导致的多方授权问题。

3.2 主客体关系性质

主客体关系包括下列类型：

(1)拥有关系。主体是客体的所有者(owner)。我们规定客体只有唯一的所有者。

(2)共有关系。主体是客体的共有者(sharer)，共有者拥有所有者的部分权限。

(3)原创关系。主体是客体的原创者(creator)，原创者是客体的最初所有者。

(4)传播关系。主体是客体的传播者(disseminator)，主体将客体传播张贴到自己或其它主体空间。

对主客体关系中的主体有如下定义。

定义 1(客体层次) 客体层次 CSH 是二元组($C \cup SP$，$\leqslant_{CS}$)，其中 C 是内容集，SP 是空间集，$\leqslant_{CS}$ 是偏序关系，其中：

(1)对任意的 $sp \in SP, sp' \in SP$，$sp \leqslant_{CS} sp'$ 当且仅当 sp 是 sp' 的子空间。

(2)对任意的 $c \in C, c' \in C$，$c \leqslant_{CS} c'$ 当且仅当 c 是 c' 的直接从属内容。

(3)对任意的 $c \in C, sp \in SP$，$c \leqslant_{CS} sp$ 当且仅当内容 c 存放在空间 sp 中且 c 是其所在内容层次的最高点。

定义 2(拥有关系) 拥有关系 $R_{own} \subseteq S \times O$，其中 S 是主体集，O 是客体集，按如下确定拥有关系。

(1)主体与其所属或创建的空间客体具有拥有关系；

(2)主体与其所拥有的空间客体内存放的内容客体具有拥有关系，即 $\forall s\in S, c\in C, R_{own}(s,c)\Rightarrow \exists sp\in SP, c\leqslant_{CS} sp \wedge R_{own}(s,sp)$。

拥有关系具有如下性质。

性质 1(拥有关系的唯一性) 主客体间的拥有关系具有唯一性，单个客体不能同时与多个主体具有拥有关系，即 $\forall s\in S, o\in O, R_{own}(s,o)\Rightarrow !\exists s'\in S, s\neq s'\wedge R_{own}(s',o)$。

性质 2(拥有关系的层次传递性) 在客体层次中，拥有关系具有传递性，即 $\forall s\in S, o\in O, o'\in O, R_{own}(s,o)\wedge o'\leqslant_{CS} o)\Rightarrow R_{own}(s,o')$。

由性质 1 和性质 2，易证如下定理。

定理 1(客体层次的同一主体拥有关系定理) 同一客体层次中的所有客体只能与同一主体具有拥有关系。

定义 3(共有关系) 共有关系 $R_{share}\subseteq S\times O$，其中 S 是主体集，O 是客体集，共有关系指客体对主体之间具有标记(tagging)，也即：$\forall s\in S, c\in C, R_{share}(s,c)\Rightarrow c$ 中具有标记指向 s。

定义 4(原创关系) 原创关系 $R_{create}\subseteq S\times O$，其中 S 是主体集，O 是客体集，原创关系 $R_{create}(s,c)$ 指主体 s 是客体 c 的最初创造者。

定义 5(传播关系) 传播关系 $R_{disseminate}\subseteq S\times O$，其中 S 是主体集，O 是客体集，传播关系 $R_{disseminate}(s,c)$ 指主体 s 是客体 c 的当前传播者，即 $\forall s\in S, s'\in S, c\in C, R_{disseminate}(s,c)\wedge R_{own}(s',c)\Rightarrow s$ 将客体 c 传播张贴到 s' 的空间内。

3.3 授权策略

逻辑许可证[9]是我们提出的一种主要基于 Datalog 规则的权利管理语言。在该权利语言中，授权策略以逻辑许可证的形式表达，由于多方授权决策和策略表达的需要，我们采用带权重的事务控制逻辑方式表达。

3.3.1 Datalog 语言

Datalog 是一种有限形式的逻辑编程，它拥有变量、谓词和常数，但却缺乏函数符号。Datalog 规则是具有以下格式的规则：

$$R_0(t_{0,1},\cdots,t_{0,k0})\leftarrow R_1(t_{1,1},\cdots,t_{1,k1}),\cdots,R_n(t_{n,1},\cdots,t_{n,kn}) \tag{1}$$

其中，$R_0,\cdots,R_n$ 是关系谓词符号，每一项 $t_{i,j}(0\leqslant i\leqslant n, 1\leqslant j\leqslant k_i)$ 既可以为常量也可以为变量。公式 $R_0(t_{0,1},\cdots,t_{0,k0})$ 被称作规则头，序列 $R_1(t_{1,1},\cdots,t_{1,k1}),\cdots,R_n(t_{n,1},\cdots,t_{n,kn})$ 为规则体。如果 $n=0$，则规则体为空且该规则被称作一个事实。一个规则是安全的仅当所有变量发生在规则头并出现在规则体中。一个 Datalog 程序就是 Datalog 规则的有限集合。

定义单个常量或单个变量为项，让 $p(x_1,\cdots,x_n)$ 作为任意谓词符号，$x_1,\cdots,x_n$ 是相应取值域上的项，称 $p(x_1,\cdots,x_n)$ 是一个原子公式(简称原子)，称 $p(x_1,\cdots,x_n)$ 或其求反 $\neg p(x_1,\cdots,x_n)$ 为一个文字。

Datalog 规则是如下形式的规则：

$$P\leftarrow L_1,\cdots,L_n \tag{2}$$

其中，P 是内涵谓词原子，$L_1,\cdots,L_m$ 是任意原子。

称 Datalog 规则中的规则体为事务，事务具有下列形式：

$$L_1,\cdots,L_m \tag{3}$$

$L_i(0\leqslant i\leqslant m)$ 是任意文字。

3.3.2 w-Datalog 程序语义和评价

定义 6(w-Datalog 规则) w-Datalog 规则是具有如下形式的规则：

$$w_0:P\leftarrow w_1:L_1,\cdots,w_k:L_k,[w_{k+1}:L_{k+1}],\cdots,[w_n:L_n],L_{n+1},\cdots,L_m \tag{4}$$

其中，P 是内涵谓词原子，$L_1,\cdots,L_n$ 是任意带权重的文字，称为权重文字，$L_{n+1},\cdots,L_m$ 是任意文字，$w_i(0\leqslant i\leqslant n)$ 是大于零的实数，称为权重。

w-Datalog 规则包括 3 类文字：固定权重文字、可选权重文字(带中括号的权重)和不带权重的文字。为简单起见，我们考虑如下 3 类基本规则：

$$w_0:P\leftarrow w_1:L_1,\cdots,w_n:L_n \quad \text{(I 型基本规则)}$$

$$w_0:P\leftarrow [w_1:L_1],\cdots,[w_1:L_n] \quad \text{(II 型基本规则)}$$

$$P\leftarrow L_1,\cdots,L_n \quad \text{(III 型基本规则)}$$

其余的 w-Datalog 规则可看作是这 3 类规则的组成。

以下为简单起见，不考虑负文字的情况，III 型基本规则就是传统的 Datalog 规则，在此不再重复，先考虑 I 型基本规则的评价。

定义 7(操作语义) 设 DB 是 w-Datalog 程序，定义 DB 的操作语义 $O(DB)$ 为：

$$O(DB)=\{p(\tilde{X})\mid DB\rightarrow^{\times} p(\tilde{X})\} \tag{5}$$

其中，$p(\tilde{X})$ 是任意原子，$DB\rightarrow^{\times} p(\tilde{X})$ 表示 $p(\tilde{X})$ 能从 DB 中推导(derivations)。

定义 8(固定点[10]语义) 设 DB 是 w-Datalog 程序，定义 DB 的固定点语义 $Fix(DB)=T_{DB}\uparrow\omega$。

其中 T_{DB} 是直接操作算子，$T_{DB}\uparrow\omega$ 定义直接操作算子 T_{DB} 在 DB 中反复执行直到固定点(fixpoint)。w-Datalog 的 Herbrand 基和 Datalog 的 Herbrand 基是有一样的形式，即 $\beta=\{p_1,\cdots,p_n\}$，其中 $p_i(0\leqslant i\leqslant n)$ 是任意原子，类似于文献[11]，下面定义 I 型直接操作算子。

定义 9(I 型直接操作算子) 设 DB 是由 I 型基本规则组成的 w-Datalog 程序，I 是一个解释(interpretation)，则按如下定义 T_1 算子($2^{\beta}\rightarrow 2^{\beta}$)：

$$\begin{aligned}T_1(I)=\{\ & p(\tilde{X})\mid \exists \text{重命名规则}\\ & w_0:p(\tilde{t})\leftarrow w_1:L_1(\tilde{Y}_1),\cdots,w_n:L_n(\tilde{Y}_n)\in DB\\ & w=(\sum w_i\mid \forall L_i(\tilde{X}_i)\in I\wedge \exists\theta\,\theta L_i(\tilde{Y}_i)=\\ & L_i(\tilde{X}_i))\wedge w\geqslant w_0\\ & \text{且在不共享变量的情况下有 } p(\tilde{X})=\theta p(\tilde{t})\\ & \}\end{aligned} \tag{6}$$

直接算子 $T_1(I)$ 说明规则的成立要求任意规则体中的文字 $L_i(\tilde{Y}_i)$ 如果经变量通代(θ)后即 $\theta L_i(\tilde{Y}_i)$ 属于解释 I，那么累加该文字的权重 w_i 到 w 中，最后这些文字的权重之和 w 应大于等于 w_0。

定理 2(I 型操作语义和固定语义的同一性) 设 DB 是由 I 型基本规则组成的 w-Datalog 程序，则有 $O(DB)=Fix(DB)$。

基本 Datalog(pure Datalog)中该定理是 Datalog 评价(evaluation)的基础，假定 $T_p(I)$ 是基本 Datalog 中的直接操作算子，在基本 Datalog 中，有：

$$I_1\subseteq I_2\Rightarrow T_p(I_1)\subseteq T_p(I_2) \tag{7}$$

$T_1(I)$是在$T_p(I)$基础上增加了对权重的限制条件，用函数$F(T_p(I))$表示该限制条件，我们只需要证明：

$$I_1 \subseteq I_2 \Rightarrow F(T_p(I_1)) \subseteq F(T_p(I_2)) \tag{8}$$

对于任意事实$p(\tilde{t}) \in F(T_p(I_1))$，假定该事实由规则$w_0 : p(\tilde{t}) \leftarrow w_1 : L_1(\tilde{Y}_1), \cdots, w_n : L_n(\tilde{Y}_n)$推出，那么对任意$L_i(\tilde{X}_i), \theta L_i(\tilde{X}_i) \in I_1 \Rightarrow \theta L_i(\tilde{X}_i) \in I_2$，这是由于$I_1 \subseteq I_2$，在$I_2$中$w$的值不会小于$I_1$中$w$的值，也即$w_{I2} \geqslant w_{I1}$，因此$p(\tilde{t}) \in F(T_p(I_2))$，即下列规则成立：

$$p(\tilde{t}) \in F(T_p(I_1)) \Rightarrow p(\tilde{t}) \in F(T_p(I_2)) \tag{9}$$

由规则(9)可推出$F(T_p(I_1)) \subseteq F(T_p(I_2))$，规则(8)得证，也即定理1得证。

下面定义II型直接操作算子。

定义 10(II型直接操作算子)　设DB是由II型基本规则组成的w-Datalog程序，I是一个解释，则下面定义T_2算子$(2^\beta \to 2^\beta)$：

$$\begin{aligned} T_2(I) = \{\ & p(\tilde{X}) \mid \text{对规则} \\ & w_0 : p(\tilde{t}) \leftarrow w_1 : L_1(\tilde{Y}_1), \cdots, w_n : L_n(\tilde{Y}_n) \\ & w = \sum_{i=1}^{n} \sum w_i \mid \forall L_i(\tilde{X}_i) \in I \wedge \exists \theta\, \theta L_i(\tilde{Y}_i) = \\ & L_i(\tilde{X}_i) \wedge w \geqslant w_0 \\ & \text{且在不共享变量的情况下有 } p(\tilde{X}) = \theta p(\tilde{t}) \\ & \} \end{aligned} \tag{10}$$

直接算子$T_2(I)$说明解释I中的任何文字$L_i(\tilde{X}_i)$经变量通代(θ)后能够等于规则中的任意文字$L_i(\tilde{Y}_i)$，则累积该文字的权重w_i到w中，最终需要w大于等于w_0。

例如，假设解释$I = \{p(1), q(2), o(1), o(2)\}$，对如下的w-datalog规则：

$$4 : q(x) \leftarrow 1 : p(x), [2 : (y)] \tag{11}$$

让规则(11)中的x绑定1后，成为如下规则：

$$4 : q(1) \leftarrow 1 : p(1), [2 : o(y)] \tag{12}$$

解释I中$o(1)$和$o(2)$让可选文字成立两次，因而有$w = 5$，规则(12)成立，因而$T(I) = \{p(1), q(1), q(2), o(1), o(2)\}$。

定理 3(II型操作语义和固定语义的同一性)　设DB是由II型基本规则组成的w-Datalog程序，则有$O(DB) = Fix(DB)$。

类似定理1的证明，在此不再赘述。

w-Datalog规则由3类基本规则组合而成，因而其评价方法也由3类基本规则的评价方法组合实现。

3.3.3　授权规则

系统的取值域如下：

主体集S，用户集U，用户群集G，系统群SYS，$S = U \cup G \cup SYS$；

客体集O，内容集C，空间集SP，$O = C \cup SP$；

关系集R，用户关系集RU，主客体关系集RO，$R = RU \cup RO$；

主客体关系集RO包括拥有关系集R_{own}、共有关系集R_{share}、原创关系集R_{create}、传播关系集$R_{disseminate}$，$RO = R_{own} \cup R_{share} \cup R_{create} \cup R_{disseminate}$；

实体集$E = S \cup O \cup R$，实体属性集$ATTR$；

操作集P，授权类型集SIGN$= \{+, -\}$；

整数集Z，自然数集N。

在定义授权规则前，我们首先规定授权规则必须是安全的。安全性指在规则体中出现的任何变量都必须出现在某个非求反的子目标谓词中，保证安全性的目的是防止不受限制的变量产生不受输入数据库控制的规则或事实集。授权规则规定，对任意的无限制域变量，规则体中必存在包含该变量的非负原子L_i，而其它变量均取值于离散的有限集，受其取值域限制。

定义 11(主体授权规则，*AuthS* 规则)

AuthS 规则是如下形式的规则：

$$AuthS(s_1, s_2, o, \langle sign \rangle p) \leftarrow L_1, \cdots, L_n \tag{13}$$

其中，$s \in S, o \in O, sign \in SIGN, p \in P$，$L_i (1 \leqslant i \leqslant n)$是代表主客体双方属性的任意文字。

$AuthS(s_1, s_2, o, \langle sign \rangle p)$，主体$s_1$授权$s_2$对客体$o$进行$\langle sign \rangle p$操作。规则如下：

$$AuthS(s_1, s_2, o, +read) \leftarrow R_{own}(s_1, o), date(d), d \leqslant \text{“2014/09/01”} \tag{14}$$

规则(14)说明客体o中的所有者s_1授权s_2在2014/09/01前拥有阅读o的权限。

定义 12(多方授权规则，*AuthD* 规则)

AuthD 规则是如下形式的w-Datalog规则：

$$\begin{aligned} & w : AuthD(s, o, \langle sign \rangle p) \Leftarrow w_1 : AuthS(s_1, s, o, +p), \cdots, \\ & w_k : AuthS(s_n, s, o, p), [w_{k+1} : AuthS(s_1, s, o, +p)], \cdots, [w_n : \\ & AuthS(s_n, s, o, p)], L_1, \cdots, L_n \end{aligned} \tag{15}$$

其中，$s \in S, o \in O, sign \in SIGN, p \in P$，$w_i : AuthS(s_i, s, o, p)$ $(1 \leqslant i \leqslant k)$是固定权重文字，$w_j : AuthS(s_j, s, o, p)$ $(k+1 \leqslant i \leqslant n)$是可选权重文字，$L_i (1 \leqslant i \leqslant n)$是代表主客体双方属性任意文字。

规则(16)的授权决策要求多数共有者同意才能进行授权。

$$\begin{aligned} & n/2 : AuthD(s, o, +p) \Leftarrow [1 : AuthS(s_1, s_2, o, +p)], \\ & R_{share}(s_1, o), sumof(o, SHARER, n) \end{aligned} \tag{16}$$

其中谓词*sumof*获得客体o共有者的数量。

定义 13(操作授权规则，*cando* 规则)

cando 规则是如下形式的规则：

$$cando(s, o, p) \leftarrow L_1, \cdots, L_n \tag{17}$$

其中，$s \in S, o \in O, p \in P$，$L_i (1 \leqslant i \leqslant n)$是代表主体授权、多方授权和客体双方属性的任意文字。

cando 规则是最后的授权结果规则，与主体授权或多方授权规则配合使用，能够实现各类多方授权决策。下列规则(18)说明最终授权以客体所有者的决策优先：

$$\begin{aligned} & cando(s, o, p) \leftarrow AuthS(s_1, s_2, o, +p), R_{own}(s_1, o) \\ & cando(s, o, p) \leftarrow AuthD(s, o, +p), \neg AuthS(s_1, s, \\ & o, -p), R_{own}(s_1, o) \end{aligned} \tag{18}$$

规则(19)说明要求所有共有者完全同意才能授权：

$$\begin{aligned} & 1 : AuthD(s_2, o, -p) \Leftarrow 1 : AuthS(s_1, s_2, o, -p), R_{share}(s_1, \\ & o) \\ & cando(s, o, p) \leftarrow \neg AuthD(s, o, -p) \end{aligned} \tag{19}$$

3.3.4　多方授权逻辑许可证

定义 14(许可证)　许可证lic是五元组$\{D, AuthPolicy, DecionPolicy, CandoPolicy, Attr\}$，其中$D$是客体的唯一标识符，$AuthPolicy$是主体授权策略和授权逻辑规则集，*De-*

cionPolicy 是多方决策策略集，*CandoPolicy* 是授权规则集，*Attr* 是属性集。

为简单起见，我们只考虑属性(谓词)及其单个取值的情况，并将其表示为 *name*≡*value* 的属性绑定形式。

图 4 是逻辑许可证示例。

```
1. License (//许可证示例
2.   imagespace //相册空间
3.   { AuthS (s1, s2, o, +read) ←Rown (s1, o), relation (s1, s2,
       "friend"), trust(s1, s2, "friend", n), n≥0.5;
4.     AuthS (s1, s2, flower. jpg, +read) ←Rdisseminate (s1, o), relation
       (s1, s2, "collegaue"), depth(s1, s2, "collegaue", m), m<2, di-
       rin(flower. jpg, imagespace);
5.     AuthS (s1, s2, o, +p)←Rown (s1, o), INPUT(s1, s2, o, +p,
       ss), ss="permit";
6.     AuthS (s1, s2, o, +p)←Rshare (s1, o), INPUT(s1, s2, o, +p,
       ss), ss="permit";
7.     AuthS (s1, s2, o, +p)←Rcreate (s1, o), INPUT(s1, s2, o, +p,
       ss), ss="permit";
8.     AuthS (s1, s2, o, +p) ←Rdisseminate (s1, o), INPUT (s1, s2,
       o, +p, ss), ss="permit";
   }, // AuthPolicy
9. { n/2: AuthD (s, o, +p)⇐[1: AuthS(s1, s, o, +p)], Rshare (s1, o),
     sumof (o, SHARER, n);
10.    1: AuthD (s, o, +p)⇐[0.2: AuthS(s1, s, o, +p)], Rshare (s1,
       o), 0.7: AuthS(s2, s, o, +p), Rcreate (s2, o), [0.5: AuthS(s3, s,
       o, +p)], Rdisseminate (s3, o);
   }, // DesionPolicy
11.  { cando(s, o, p)←AuthS(s1, s, o, +p), Rown (s1, o)
12.    cando(s, o, p)←AuthD(s, o, +p), ¬AuthS(s1, s, o, −p),
       Rown (s1, o)
   }, // CandoPolicy
13.  { owner≡'李华'; id≡'123'; version≡'1.0'; expire≡'15/12/
       31' }, //Attr
14. )
```

图 4　逻辑许可证示例

图 4 中语句 3 说明李华允许可信度不低于 0.5 的好友访问属于自己的相册空间；语句 4 说明图片 flower. jpg 的传播者允许自己的直接同事访问该图片，该图片直接存储在李华的相册空间内，谓词 depth 用于计算关系之间的深度[12]，dirin 用于说明客体之间的直接从属关系。语句 5－8 分别说明主客体 4 类关系都可以进行主体授权，谓词 INPUT 用于输入主体的意见；语句 9 假定共有者的重要性相同，平均分配给每位共有者的权重为 1，但共有者必须一半以上同意($\geq n/2$)才能获得决策授权；语句 10 按照关系的重要程度设置权值，最重要的原创者权重为 0.7，次重要的传播者权重为 0.5，共有者权重为 0.2，形成阈值度的授权(关系主体权重和必须超过 1 的阈值度)；语句 11 说明如果 s 获得的是客体所有者的主体授权，则 s 直接获得操作授权；语句 12 说明如果 s 获得客体的多方授权且没有客体所有者的否认主体授权，s 才能获得操作授权，也即客体所有者对授权有一票否认权。

4　模型实现

首先说明 MRuleSN 授权程序是分层的(stratified)[13]和安全的。

由于 MRuleSN 授权规则必须符合安全性要求，我们只说明 MRuleSN 程序的分层性，MRuleSN 程序可进行如下分层，如表 1 所列。

表 1　MRuleSN 语言的分层

层	谓词	规则体包含谓词
1	外延谓词 P_{ext}	主客体属性等谓词
2	主体授权规则，AuthS	外延谓词 P_{ext}
3	多方授权规则，AuthD	外延谓词 P_{ext} 和 AuthS 文字
4	操作授权规则，cando	外延谓词 P_{ext}、AuthS 文字和 AuthD 文字

因而逻辑许可证程序是在 w-Datalog 语义基础上增加了分层负文字，w-Datalog 语言的评价在 3.3.2 节已做说明，分层负 Datalog 的评价可参考文献[13]。

5　表达力示例

我们的模型具有较强的表达力，能表达文献[3]提出的所有模型。

(1)决策投票(Decision Voting)模型

所有多方参与者一人一票，超过半数通过：

$$n/2::AuthD\ (s,o,+p)\Leftarrow 1:AuthS(s_1,s,o,+p), R_{create}(s_1,o), [1:AuthS(s_2,s,o,+p)], R_{share}(s_2,o), 1:AuthS(s_3,s,o,+p), R_{create}(s_3,o), [1:AuthS(s_4,s,o,+p)], R_{disseminate}(s_4,o), sumof\ (o,ALL,n) \tag{20}$$

如果多方参与者重要性不同，可以设定每位参与者具有不同权重，实现带权重的决策投票。

(2) 敏感度投票(Sensitivity Voting)

敏感度(Sensitivity Level)是参与者对共享数据隐私保护敏感程度的衡量，比如一般中国人可能对年龄信息的泄露不太敏感，而部分公众人物可能会很敏感。敏感度一般是 0 到 1 之间的有理数，下列规则根据参与者对客体的敏感度计算投票结果：

$$w:AuthD\ (s,o,+p)\Leftarrow[1/l:AuthS(s_1,s,o,+p)], SL(s_1,o,l) \tag{21}$$

其中谓词 SL 说明 s_1 对 O 的敏感度为 l，规则(21)说明，数据敏感度越高，其权重越小，因而数据的敏感度越高，授权的可能性越小。

(3)客体所有者优先

见规则(18)。

(4)所有参与者必须都同意

即否认优先，见规则(19)。

(5)多数参与者同意

下列规则说明超过半数以上参与者同意才能授权。

$$n/2:AuthD\ (s,o,+p)\Leftarrow[1:AuthS(s_1,s,o,+p)], sumof\ (o,ALL,n) \tag{22}$$

结束语　本文针对社交网络中的多方授权问题，提出面向社交网络的多方授权模型 MRuleSN。该模型采用单一所有、多方共有的方法处理所有权问题，采用扩展的 w-Datalog 规则表达多方授权，解决了普通 Datalog 规则难以表达多方决策的问题。MRuleSN 模型具有较强的多方授权灵活性和授权表达能力。

下一步，将对社交网络中客体的移动和复杂的具体访问控制方法做深入研究。

参 考 文 献

[1] Park J, Sandhu R, Cheng Y. Acon: Activity-centric access control for social computing[C]// 2011 Sixth International Conference on Proc. of Availability, Reliability and Security (ARES). IEEE, 2011: 242-247

[2] Mahmood S. Online Social Networks: Privacy Threats and Defenses[M]// Security and Privacy Preserving in Social Networks. Springer Vienna, 2013: 47-71

[3] Hu Hong-xin, Gail-Joon Ahn, Jan Jorgensen. Multiparty Access Control for Online Social Networks: Model and Mechanisms[J]. Proc. of IEEE Transactions on Knowledge and Data Engineering, 2013, 25(7): 1614- 1627

[4] Thomas K, Grier C, Nicol D M. unfriendly: Multi-party privacy risks in social networks[C]// Proc. of Privacy Enhancing Technologies. Springer Berlin Heidelberg, 2010: 236-252

[5] Squicciarini A C, Shehab M, Wede J. Privacy policies for shared content in social network sites[J]. The VLDB Journal-The International Journal on Very Large Data Bases, 2010, 19(6): 777-796

[6] Amrutha P, Sathiyaraj R. Privacy Management of Multi User Environment in Online Social Networks (OSNs)[J]. GJCST-E: Network, Web & Security, 2013, 13(10): 01-07

[7] Subhani S, Rajasekhar M. A photo privacy for tagged images using rule-based access control in social networks[J]. International Journal of Research Sciences and Advanced Engineering, 2012, 2(5): 45-49

[8] Yeung C A, Kagal L, Gibbins N, et al. Providing Access Control to Online Photo Albums Based on Tags and Linked Data[C]// Proc. of AAAI Spring Symposium: Social Semantic Web: Where Web 2. 0 Meets Web 3. 0. 2009: 9-14

[9] Zhong Yong, Zhang Hong, Liu Feng-yu, et al. A Digital Rights Management Mechanism and Implementation Based on Logic Framework[J]. Journal of Computer Research and Development, 2010, 47(2): 223-230(in Chinese)
钟勇, 张宏, 刘凤玉, 等. 一种基于逻辑框架的数字版权管理机制和实现[J]. 计算机研究与发展, 2010, 47(2): 223-230

[10] Bertino E, Catania B, Gori R, et al. Active-U-Datalog: integrating active rules in a logical update language [C]// Proc. of International Seminar on Logic Databases and the Meaning of Change, LNCS 1472. Berlin: Springer, 1998: 107-133

[11] Montesi D, Bertino E, Martelli M. Transactions and updates in deductive databases[J]. IEEE Trans. Knowl. Data Eng., 1997, 9(5): 784-797

[12] Carminati B, Ferrari E, Perego A. Rule-based access control for social networks[C]// Proc. of on the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops. Spring-Verlag Lecture Notes in Computer Science, LNCS 4278, 2006: 1734-1744

[13] Jajodia S, Samarati P, Sapino M L, et al. Flexible support for multiple access control policies[J]. ACM Transaction on Database System, 2001, 26(2): 214-260

(上接第 93 页)

[4] Chan E Y, Ching W K, Ng M K, et al. An optimizationalgorithm for clustering using weighted dissimilaity measure[J]. Pattern Recoginzation, 2004, 37(5): 943-952

[5] Bai Liang, Liang Ji-ye, Dang Chuang-yin, et al. The impact of cluster representatives on the convergence of the K-modes type clustering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(6): 1509-1522

[6] Yang Yi-ming. An evaluation of statistical approaches to text categorization[J]. Information Retrieval, 1999, 1(1/2): 69-90

[7] Information G M. Uncertainty and the utility of categories[C]// Proc. of the Seventh Annual Conf. on Cognitive Science Society. Lawrence Erlbaum, 1985: 283-287

[8] Barbará D, Li Yi, Couto J. COOLCAT: an entropy-based algorithm for categorical clustering[C]// Proceedings of the Eleventh International Conference on Information and Knowledge Management. ACM, 2002: 582-589

[9] Aggarwal C C, Procopiuc C, Yu P S. Finding localized associations in market basket data[J]. IEEE Transactions on Knowledge and Data Engineering, 2002, 14(1): 51-62

[10] Cao Fu-yuan, Liang Ji-ye, Bai Liang, et al. A framework for clustering categorical time-evolving data[J]. IEEE Transactions on Fuzzy Systems, 2010, 18(5): 872-882

[11] Wrigley N. Categorical data analysis for geographers and environmental scientists[M]. Blackburn Press, 2012

[12] Chmielewski M R, Grzymala-Busse J W. Global discretization of continuous attributes as preprocessing for machine learning[J]. International Journal of Approximate Reasoning, 1996, 15(4): 319-331

[13] Dash M, Liu Huan. Consistency-based search in feature selection [J]. Artificial Intelligence, 2003, 151(1): 155-176

[14] Guyon I, Elisseeff A. An introduction to variable and feature selection[J]. The Journal of Machine Learning Research, 2003, 3: 1157-1182

[15] Zhou Zhi-hua. Three perspectives of data mining[J]. Artificial Intelligence, 2003, 143(1): 139-146

[16] Huang Zhe-xue. Extensions to the k-means algorithm for clustering large data sets with categorical values[J]. Data Mining and Knowledge Discovery, 1998, 2(3): 283-304

[17] Lee M, Pedrycz W. The fuzzy C-means algorithm with fuzzy P-mode prototypes for clustering objects having mixed features [J]. Fuzzy Sets and Systems, 2009, 160(24): 3590-3600

[18] Yu Jian. General C-means clustering model[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(8): 1197-1211

[19] Alamuri M, Surampudi B R, Negi A. A survey of distance/similarity measures for categorical data[C]// 2014 International Joint Conference on Neural Networks (IJCNN). IEEE, 2014: 1907-1914

[20] Andritsos P, Tsaparas P, Miller R J, et al. LIMBO: Scalable clustering of categorical data[M]// Advances in Database Technology-EDBT 2004. Springer Berlin Heidelberg, 2004: 123-146

[21] Chan E Y, Ching W K, Ng M K, et al. An optimization algorithm for clustering using weighted dissimilarity measures[J]. Pattern Recognition, 2004, 37(5): 943-952

[22] Comaniciu D, Meer P. Mean shift: A robust approach toward feature space analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 603-619