ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 21 за 2013 г.

Коваленко В.Н., Куликов А.Ю.

Определение политики контроля доступа для инфраструктур с массовой интеграцией баз данных

Рекомендуемая форма библиографической ссылки: Коваленко В.Н., Куликов А.Ю. Определение политики контроля доступа для инфраструктур с массовой интеграцией баз данных // Препринты ИПМ им. М.В.Келдыша. 2013. № 21. 20 с. URL: http://library.keldysh.ru/preprint.asp?id=2013-21

О р д е н а Л е н и н а
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В. Келдыша
Р о с с и й с к о й а к а д е м и и н а у к

В.Н. Коваленко, А.Ю. Куликов

Определение политики контроля доступа для инфраструктур с массовой интеграцией баз данных

Москва — 2013

Коваленко В.Н., Куликов А.Ю.

Определение политики контроля доступа для инфраструктур с массовой интеграцией баз данных

Как показывает приведённый в работе обзор, контроль доступа к распределённым сервисам и ресурсам остаётся областью активных исследований. Несмотря на большое количество способов такого контроля, выраженных в формализованных моделях, их применение в конкретных условиях может быть проблематичным. Данная работа посвящена способу описания политики доступа к структурированным данным, которые хранятся в информационной инфраструктуре, образованной путём виртуальной интеграции распределённых гетерогенных баз реляционного типа.

Ключевые слова: интеграция данных, информационная инфраструктура, контроль доступа, факторизация прав доступа, частичные правила

Kovalenko Victor Nikolaevich, Kulikov Alexandr Yurievich

Defining access control policy for infrastructures with mass integration of databases

As shows the review provided in this paper, access control to distributed services and resources remains area of active researches. Despite the existence of a large number of formalized models describing access control, possibility of their application in specific conditions can be problematic. This work is devoted to a way of description of access control policy to the structured data which are stored in information infrastructures formed by virtual integration of distributed heterogeneous relational databases.

Key words: data integration, informational infrastructure, access control, permission factoring, partial rules

Работа выполнена при поддержке Российского фонда фундаментальных исследований, проект 11-07-00147-а и программы фундаментальных исследований Президиума РАН.

1. Введение

В публикации [1] изложен подход, позволяющий создавать масштабные информационные инфраструктуры путём программной интеграции автономных гетерогенных баз данных. Программная (или виртуальная) интеграция отличается от физической тем, что данные интегрируемых баз не перемещаются в общее хранилище, но, тем не менее, система интеграции обеспечивает доступ ко всей их совокупности. В результате интеграции из данных разных баз образуется единое информационное пространство, которое описывается глобальной схемой, аналогичной схеме обычной базы данных. Предложенный подход интеграции характеризуется тем, что элементы глобальной схемы (таблицы в реляционной модели) представляют собой объединение семантически эквивалентных данных из разных баз, которые приведены к унифицированному представлению (рис. 1).

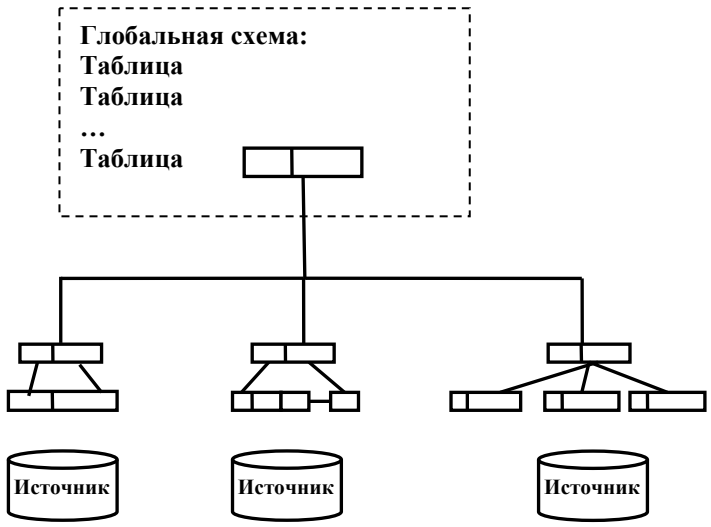

Рис.1. Отображение глобальной схемы на схемы источников. Каждая глобальная таблица является объединением локальных таблиц, приведённых к общему представлению.

Доступ к интегрированным данным осуществляется с помощью поискового запроса SELECT, аналогичного такому запросу в стандарте SQL-92. Отличие от стандартной формы SELECT состоит в том, что данные адресуются в соответствии с глобальной схемой в форме: Имя-группы.Имя-таблицы.Имя-столбца. Под группой понимается программный объект, определяющий совокупность баз, из которых выбираются данные. Группа может содержать как все базы данных, включённые в инфраструктуру, так и какое-то их подмножество. Таким образом, в одном *массовом запросе* можно получать данные из всего интегрированного информационного пространства или из определённой части содержащихся в инфраструктуре баз.

Способ определения области поиска в массовом запросе позволяет не

указывать явно базы-источники, на которых он выполняется, он не зависит от числа баз в инфраструктуре и от физического расположения данных. Однако предусмотрены средства навигации в информационном пространстве: возможность определения области поиска как его подмножества. Отбор баз, включаемых в область поиска, осуществляется на основе метаинформации, описывающей базы данных. В качестве метаинформации может использоваться расширяемый набор метаатрибутов баз данных, например, название и тип организации-владельца базы, местоположение, ведомственная принадлежность.

Детальное описание массового запроса SELECT дано в [1]. Здесь мы ограничимся примером, иллюстрирующим способ адресации данных и возможность определения пространства поиска:

SELECT SP.Persons.name FROM SP.Persons WHERE SP.Persons.income>40000, SP.region=”Москва”

Запрос выдаёт значения поля name из таблицы Persons для лиц, чей доход (income) превышает определённую величину. Запрос выполняется в поисковой области SP, состав которой специфицируется в конструкции WHERE: в группу SP включаются базы данных организаций, которые расположены в Москве (метаатрибут region).

Одной из важнейших задач, которая должна решаться системой интеграции, является обеспечение безопасности данных. В условиях информационных инфраструктур эта задача осложняется тем, что в них содержатся данные разной тематики, принадлежащие разным учреждениям, с ними работают люди разных специальностей и разного уровня компетенции. В связи с этим рассматривается вопрос о создании таких механизмов защиты данных, которые бы позволяли ограничить доступ определёнными срезами информационного пространства.

Основная тема данной статьи – описание политики контроля доступа в информационных инфраструктурах, используемых коллективно распределёнными пользователями. Предлагаемый способ излагается в разделе 4. В разделе 2 формулируются базовые положения систем безопасности и вводятся соответствующие понятия. В разделе 3 даётся обзор современных подходов определения политики контроля доступа.

2. Основные принципы обеспечения безопасности распределённых систем

Проблема безопасности решалась для различных компьютерных систем. На современном этапе большое внимание уделяется защите систем, которые дистанционно предоставляют сервисы и различные типы ресурсов (компьютерные, информационные) распределённым коллективам пользователей.

Задача системы безопасности – предотвращение раскрытия хранящейся в защищаемой компьютерной системе информации (обеспечение секретности) или модификации её данных и других ресурсов (обеспечение сохранности) [2].

Одним из направлений решения этой задачи является *контроль доступа*, направленный на то, чтобы исключить возможность выполнения неправомочных действий и в то же время обеспечить постоянную доступность защищаемой системы и её ресурсов для выполнения легитимных операций.

Помимо наличия большого количества конкретных результатов, можно, по-видимому, говорить об утверждении общих принципов организации систем безопасности и, в частности, о сложившейся архитектуре контроля дистанционного доступа, которая обладает свойствами общности, универсальности и модульности. Под общностью понимается возможность применения системы контроля доступа для защиты ресурсов разных типов, например компьютерных ресурсов, что характерно для вычислительных грид-инфраструктур, или ресурсов баз данных, как в рассматриваемом нами случае интегрированных информационных инфраструктур. Универсальность заключается в том, что система контроля доступа применима для многих защищаемых систем, давая общие механизмы безопасности, которые могут быть встроены в различные приложения. Модульность позволяет использовать разные методы для реализации отдельных механизмов, комбинируя их в рамках целостной схемы контроля. Как показано в обзоре [3], в современных системах контроля доступа выражена тенденция к стандартизации как в аспекте взаимодействия их компонент (Web- и грид-службы), так и в дескриптивном аспекте (SAML – Security Assertion Markup Language [4], XACML – eXtensible Access Control Markup Language [5]).

В архитектуре безопасности контроль доступа осуществляется двумя последовательно выполняющимися компонентами.

- Все обращения к защищаемой системе поступают, прежде всего, в компоненту аутентификации. Её функция - верификация идентичности пользователя, от имени которого подан запрос. Запросы пользователей, не имеющих права доступа к системе, отсеиваются, их обработка прекращается.
- Компонента авторизации выполняет анализ запроса, проверяя, все ли запрашиваемые действия являются правомочными. Проверка заключается в определении того, может ли аутентифицированный пользователь выполнять заданные в запросе операции над указанными ресурсами. Результатом анализа является либо отказ в обслуживании запроса, либо присвоение ему статуса авторизованного.

Авторизованные запросы могут передаваться в дополнительные компоненты системы безопасности, например в компоненту *аудита*, или непосредственно в защищаемую систему для выполнения.

Исходной информацией для выполняемой компонентой авторизации проверки служит то, какие из защищаемых ресурсов доступны (или недоступны) каждому из пользователей системы. Формально соответствие между пользователями и доступными ресурсами задаётся набором правил, устанавливающих так называемые привилегии, совокупность которых определяет *права доступа* пользователей. Форма и способ интерпретации

правил зависит от используемой в той или иной системе *модели контроля доступа*. Конкретный набор правил составляет *политику контроля доступа*, которая разрабатывается для каждой установки компьютерной системы. Эта задача возлагается на администратора системы безопасности.

Простейшей формой правил является список контроля доступа (Access Control List – ACL). Правила модели ACL задают частичное отображение пар {имя пользователя ($u \in U$), операция ($o \in O$)} на множество ресурсов: $[U \times O] \rightarrow R$. Правило означает, что пользователь u может выполнять операцию o на некотором подмножестве ресурсов $r \in R$. Основная проблема ACL, также как и более развитых моделей DAC и MAC [6], связана с администрированием политики контроля доступа. Все эти модели предполагают определение прав доступа индивидуально для каждого пользователя. Такое допустимо в рамках отдельных предприятий, но в условиях большого количества и распределённости ресурсов и пользователей администрирование политики контроля доступа становится трудно разрешимой проблемой.

3. Ролевая модель контроля доступа и её развитие

Потребности разных областей практики привели к появлению ряда моделей контроля доступа. Наибольшую известность и распространение получила модель RBAC (Role Based Access Control), которая предложена в работе [7] и доведена до уровня стандарта [8]. Востребованность RBAC подтверждается тем, что она в той или иной форме реализована в нескольких СУБД: INFORMIX Online Dynamic Server Version 7.2, Sybase Adaptive Server release 11.5, Oracle DBMS, Microsoft SQL Server, PostgreSQL 8.1 и других.

Основная мотивация RBAC – повышение эффективности администрирования при большом числе пользователей. Для этого присваивание прав разбивается на две части: во-первых, пользователям присваиваются роли, и, во-вторых, ролям присваиваются привилегии (рис. 2). Эффективность достигается за счёт того, что такой подход отражает организационную структуру предприятий: даже при большом числе персонала должностей не бывает очень много, и именно должность определяет права доступа. Поскольку права доступа отнесены к ролям, регистрация нового пользователя сводится к присваиванию ему роли (или ролей), а изменение должности – к присваиванию новой роли.

Рис.2. Схема присваивания привилегий в модели RBAC.

В спецификации стандарта предложены три уровня модели RBAC: базовая (Core), иерархическая (Hierarchical) и с ограничениями (Constrained).

Базовая модель предписывает, что отображение пользователи-привилегии доступа разбивается на две части: пользователи-роли и роли-привилегии, причём оба эти отношения имеют кардинальность многие-ко-многим. Кроме того, определяется аппарат селективной активизации/деактивизации ролей в рамках пользовательской сессии.

Hierarchical RBAC вводит иерархию ролей: частичный порядок, определяемый отношением родитель-ребёнок между ролями. Транзитивное замыкание этого отношения порождает отношение предок-потомок, в соответствии с чем роли-потомки наследуют привилегии своих предков.

Constrained RBAC накладывает ограничения на использование ролей. Может быть ограничено число одновременно работающих пользователей, имеющих одну роль. Кроме того, определён способ, направленный на исключение конфликтов ролей, которые могут возникать из-за того, что пользователь в результате авторизации получает права от несовместимых ролей. Для этого применяется разбиение ролей на взаимоисключающие множества.

Дальнейшее развитие проблематики контроля доступа во многом связано с развитием модели RBAC. Перечислим наиболее важные направления развития.

Атрибутный RBAC. Модель RBAC не определяет конкретный способ приписывания ролей пользователям. Простое решение состоит в том, что роли присваиваются вручную, исходя из обязанностей пользователя на предприятии. Однако в распределённых системах, когда администратор безопасности не имеет непосредственного контакта с пользователями, а ресурсы и пользователи принадлежат различным доменам, присваивание ролей составляет самостоятельную проблему.

Предложенная в работе [9] модель вводит, в дополнение к авторизационным правилам, правила порождения ролей (рис. 3). Для идентификации пользователя применяется унифицированный набор атрибутов, которые его содержательно характеризуют. Такими атрибутами могут служить место работы, специальность, возраст. Правила порождения ролей определяют отображение значений атрибутов на роли. Модель получила название Rule-Based RBAC (RB-RBAC).

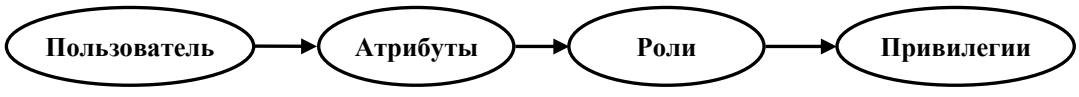

Рис. 3. Схема присваивания привилегий в модели RB-RBAC.

Учёт условий среды. Практический опыт показал, что в ходе авторизации бывает нужно учитывать условия, в которых она происходит. Так, например, иногда требуется, чтобы права, которые получает пользователь в

рабочее время, были шире, чем в нерабочее. Также права могут ограничиваться, если пользователь осуществляет доступ не с рабочего места, а удалённо по мобильному устройству. Совокупность факторов, определяющих применимость правил контроля доступа, образует *контекст авторизации*. Собственно, два таких условия – ограничение на число пользователей в одной роли и исключение возможности использования взаимоисключающих ролей – введены уже в модели RBAC. В дополнение разработаны модели, учитывающие факторы времени Temporal-RBAC [10] и местоположения GEO-RBAC [11].

Делегирование прав доступа. Важным для практики рабочим приёмом является временная передача прав доступа от одного лица другому. Это необходимо в случаях, когда один из пользователей подменяет другого, или для выполнения своих обязанностей он должен получить доступ к защищаемым ресурсам, которые принадлежат обслуживаемому лицу. По этой теме выполнены многочисленные исследования. Различные варианты делегирования полномочий рассмотрены в работе [12].

Параметризация правил. Несмотря на те улучшения, которые даёт модель RBAC, проблема эффективности администрирования остаётся актуальной. Примером может служить информационная система крупного университета, которая обслуживает несколько тысяч студентов. Персональные данные каждого из них или по крайней мере их часть должны быть доступны только владельцу. Если следовать модели RBAC, потребуется ввести количество ролей, пропорциональное числу студентов. В ряде работ исследованы способы параметризации правил RBAC (см. обзор применяемых подходов в [13]). В соответствующих моделях роли и права доступа снабжаются параметрами, в качестве которых обычно выступают атрибуты пользователей и ресурсов. Так, роль Student может быть параметризована персональным идентификатором студента (pid). При администрировании группы пользователей Student эффект от параметризации достигается следующим образом. Во-первых, определяется роль (AnyStudent), которую могут получить все члены группы, и для неё определяются общие для всех права. Во-вторых, для множества ролей Student (pid) определяются персональные правила, в которых ролям с заданными значениями атрибутов ставятся в соответствие права доступа к личным данным.

Реализация современных моделей контроля доступа. С наибольшей полнотой описанные выше расширения модели RBAC реализованы в системах OASIS [14] и PERMIS [15]. Рассмотрим основные положения модели OASIS (Open Architecture for Secure Interworking Services).

Политика контроля доступа в OASIS задаётся правилами в форме дизъюнктов Хорна [16]. Существуют два сорта правил. Правила авторизации отображают роли на права доступа. Правила активации ролей отображают пользователей на роли, а также роли на другие роли.

Оба сорта правил состоят из двух частей – предпосылки и цели. Правила

авторизации имеют вид:

r, e_1, ..., e_{n_e} ⊢ p

Предпосылка (левая часть) в этих правилах всегда содержит одну роль r. Правило присваивает права доступа (p) этой роли. И роли, и права доступа могут быть параметризованы. В предпосылке также могут быть указаны охарактеризованные выше контекстные условия в виде предикатов e_k (k ∈ $[0..n_e]$). При вычислении значений предикатов может использоваться информация, внешняя по отношению к системе OASIS, что позволяет учитывать произвольные факторы.

Вид правил активизации ролей следующий:

r_1, r_2, ..., r_{n_r} , ac_1, ..., $ac_{n_{ac}}$, e_1, ..., e_{n_e} ⊢ r

В результате интерпретации такого правила пользователь получает новую целевую роль r, если удовлетворяются все предварительные условия, перечисленные в предпосылке. Условия r_k (k ∈ $[0..n_r]$) представляют собой список ролей, которые должны быть активизированы пользователем предварительно. Элементы ac_k (k ∈ $[0..n_{ac}]$) обозначают так называемые *назначения*, которые реализуют механизм делегирования прав доступа. Для срабатывания правила все назначения должны быть сделаны. Элементы e_k (k ∈ $[0..n_e]$), как и в правилах авторизации, – это предикаты, вычисляемые с привлечением внешних данных. Любой из списков: предварительно активизированных ролей, назначений и контекстных условий, может быть опущен.

4. Политика контроля доступа для интегрированных информационных инфраструктур

Программные средства, основанные на описанных выше моделях, находят всё большее применение при создании распределённых систем: вычислительных грид-инфраструктур [17], информационных и управляющих систем [18]. В то же время, информационные инфраструктуры имеют специфику, которую необходимо учесть в модели контроля доступа. Предлагаемый способ основывается на следующих положениях.

Атрибутное описание пользователей. Права доступа определяются исходя из формализованного описания пользователей с помощью набора значений характеризующих их атрибутов. Регистрация пользователей и присваивание им значений атрибутов распределяется между организациями–владельцами баз данных: каждая организация регистрирует своих сотрудников. Атрибуты предъявляются в электронном документе безопасности, например, в расширенном сертификате X.509 или в атрибутном сертификате RFC 5755.

Централизованный контроль доступа. Политика доступа описывается набором правил, которые определяются централизованно. Рассматриваются только операции чтения данных, поскольку предполагается, что наполнение баз выполняется организациями–владельцами. Контроль доступа производится службой безопасности информационной инфраструктуры, от имени которой происходит выполнение поисковых запросов в базах данных. Каждая база предоставляет доступ системе управления запросами ко всем своим данным, включаемым в инфраструктуру. Таким образом, политика безопасности на локальном уровне не зависит от состава и числа пользователей информационной инфраструктуры.

Условия контроля доступа. Защищаемыми ресурсами в информационной инфраструктуре являются данные, интегрированные в единое пространство. Структура этого пространства описывается глобальной схемой, в соответствии с которой выполняются поисковые запросы. В этом плане информационная инфраструктура схожа с одной базой данных, и для контроля доступа можно было бы использовать имеющиеся методы и средства. Отличия, которые представляется необходимым отразить в модели контроля доступа, заключаются в следующем.

1. Информационная инфраструктура строится путём объединения автономных, принадлежащих разным учреждениям баз данных. Массовый поисковый запрос, сформулированный в терминах глобальной схемы, позволяет получить данные из всех баз данных инфраструктуры. Если для защиты использовать также элементы глобальной схемы, то контроль доступа будет весьма грубым: определять права доступа можно только для совокупности одинаковых по семантике (однотипных) данных из всех баз. Необходим дополнительный контроль, позволяющий определять права по отношению к определённым подмножествам баз данных. Таким способом можно выразить связь пользователя с некоторым ограниченным кругом учреждений по признакам местоположения, ведомственной принадлежности, виду деятельности и пр.

2. Предполагается, что в информационных инфраструктурах будет существенно большее, чем в обычных базах, тематическое разнообразие данных: финансовые, кадровые, регистрационные, производственные, медицинские и т.д. Имеет смысл ограничивать состав данных, доступных пользователю, соответственно его специальности: финансовый контролёр, врач, специалист по планированию производства должны получать доступ к разным типам данных.

3. Понятие роли, являющееся ключевым в модели RBAC, отражая естественную организацию производственной деятельности, сохраняет значение и в информационных инфраструктурах. Однако такая простая классификация пользователей может быть достаточной для спецификации прав доступа в рамках отдельных учреждений, но выход за их границы требует учёта и других характеристик. Это обстоятельство отмечается и в связи с различными

применениями RBAC: приходится вводить “составные” роли типа Врач_Терапевт, Инженер_Электронщик. В отличие от модели RBAC, мы будем рассматривать роль как важную, но равноправную с другими характеристику пользователя.

4.1. Частичные правила контроля доступа

Вариант присваивания прав по характеристикам пользователей представлен в модели атрибутного RBAC в виде двухэтапной схемы: набор пользовательских атрибутов вначале отображается в роли, которым в правилах контроля доступа присваиваются привилегии. В рассматриваемых условиях, однако, роль не является исключительной характеристикой, определяющей привилегии. Как показано в предыдущем разделе, права доступа в информационных инфраструктурах зависят от совокупности факторов, отражаемых в атрибутах. Прямое решение, позволяющее выразить эту зависимость, – задать в явной форме отображение всех допустимых наборов значений атрибутов на множество привилегий. Этому соответствуют правила вида:

$$A_1, A_2, \ldots, A_N \Rightarrow \mathcal{P} \qquad (1)$$

A_i – представляют пары {имя, значение} атрибутов, $\mathcal{P}$ – множество привилегий.

Недостаток этой формы правил, которые будем называть полными, в том, что в них смешано влияние всех факторов. Предлагаемый подход заключается в использовании частичных правил. Частичными мы называем правила, которые определяют права на неполном наборе атрибутов и выражают влияние отдельных факторов на множество привилегий. Частичные правила имеют вид:

$$A_{i1}, A_{i2}, \ldots, A_{ik} \Rightarrow \mathcal{P}' \qquad 0 \le k \le N. \quad (2)$$

Политика контроля доступа в целом описывается системой частичных правил. Получая набор допустимых значений атрибутов пользователя, механизм контроля доступа действует следующим способом.

- В первую очередь образуется множество применимых правил. Правило считается применимым, если значения всех представленных в нём атрибутов совпадают со значениями соответствующих атрибутов пользователя.
- Вычисляются результирующие права путём композиции прав, заданных в применимых правилах.

Подход, в котором права доступа определяются исходя из произвольного набора атрибутов, был, по-видимому, впервые предложен в модели TCM (Tees Confidentiality Model) [19] и успешно применяется в медицинской отрасли.

Далее описывается система частичных правил для информационных инфраструктур, в которой используется:

- двухуровневая классификация пользователей;
- наследование прав доступа от общих к детализирующим правилам;

- раздельное определение привилегий для четырёх слоёв организации данных.

Отметим, что предлагаемая система не является единственно возможной и в отношении выбранных атрибутов, и в отношении состава правил.

4.2. Состав атрибутов и описание прав доступа

Прежде всего, перечислим состав атрибутов пользователя, которые выбраны с ориентацией на профессиональное использование информационных инфраструктур:

- СПЕЦ – специальность;
- РОЛЬ – должностное положение;
- СФД – сфера деятельности.

Два первых атрибута – классифицирующие. Атрибут СПЕЦ разбивает множество пользователей на группы, в которых права определяются независимо. При необходимости в отдельных группах может быть проведено детализирующее ранжирование с помощью атрибута РОЛЬ, отражающего должностные обязанности.

Достаточно типична ситуация, когда пользователи одной специальности (и роли) должны получать доступ к различающимся наборам источников. Различия могут быть связаны с географическими (источники, содержащие данные по определённому региону), тематическими (базы данных учреждений, специализирующиеся в некоторой области) или иными ограничениями. Чтобы учесть влияние этого фактора, вводится атрибут СФД - сфера деятельности.

Исходя из этого набора атрибутов, приведём общий вид полных правил, дополнив определение (1) способом задания прав доступа:

$$\text{СПЕЦ, РОЛЬ, СФД} \Rightarrow (\textit{Таблица}\ ((\textit{Столбец}\ldots)\ (\textit{Строка}\ldots)\ (\textit{БД}\ldots))\ldots) \qquad (3)$$

Права доступа задаются вложенной конструкцией, на первом уровне которой указывается список имён глобальных таблиц (*Таблица*…), доступных пользователям с указанными в левой части значениями атрибутов. К каждой таблице могут быть приписаны три списка, в которых задаются ограничения доступа к ней: доступные столбцы (*Столбец*…), строки (*Строка*…) и базы данных (*БД*…). Все эти списки не являются обязательными и могут быть опушены. При опущенных списках столбцов или строк считаются доступными все соответствующие элементы таблицы. Если не указан список баз данных, таблица доступна на всех базах инфраструктуры. Способы задания ограничивающих списков описаны в п. 4.4.

4.3. Способ определения политики контроля доступа

Предлагаемая система частичных правил основывается на следующих общих положениях. Во-первых, защита информационных инфраструктур, описываемых глобальной схемой, осуществляется на четырёх слоях организации – таблица, её столбцы и строки, а также базы данных, в которых находятся данные таблицы. В предлагаемом подходе этим слоям ставятся в

соответствие разные компоненты прав доступа, а правила строятся так, чтобы определять каждую компоненту отдельно.

Во-вторых, результирующие права получаются в результате композиции прав, заданных в отдельных правилах. С учётом того, что используются атрибуты, разбивающие пользователей на группы с общими свойствами, а ограничения на доступность столбцов, строк и источников связаны с конкретными таблицами, в качестве способа композиции в системе правил используется наследование прав доступа – от более общих правил к детальным.

Перечислим типы правил предлагаемой системы (рис. 4):

СПЕЦ ⇒ *Таблица* … (4)

СПЕЦ, РОЛЬ ⇒ (*Таблица* ((*Столбец*…) (*Строка*...))…) (5)

СПЕЦ, СФД ⇒ (*Таблица (БД*…)…) (6)

СПЕЦ, РОЛЬ, СФД ⇒ (*Таблица (БД*…)…) (7)

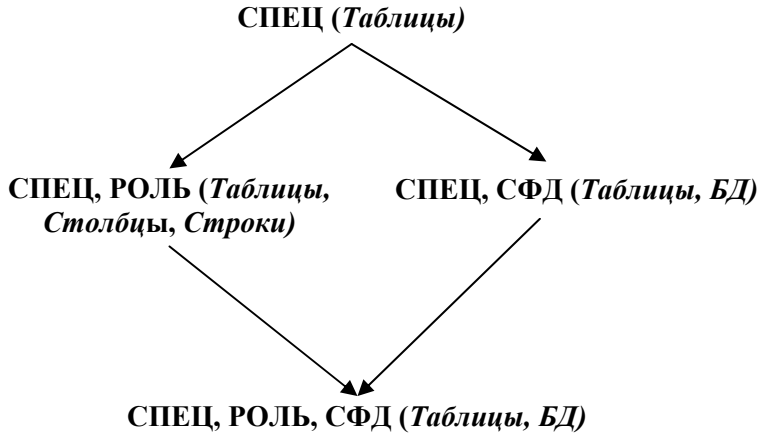

Рис. 4. Система частичных правил. Показаны состав атрибутов пользователя, определяемые компоненты прав доступа, а также наследование прав между правилами.

Первое и самое общее правило (4) задаёт одну компоненту прав доступа – список таблиц. Список задаётся, исходя из значений атрибута СПЕЦ, и интерпретируется как максимальный набор таблиц, который доступен пользователям, работающим по тематике данной специальности. Эта компонента прав наследуется более полными правилами (5, 6, 7), но может быть изменена в правилах (5).

В правилах (5) для допустимых в рамках данной специальности значений атрибута РОЛЬ добавляются ограничения на строки и столбцы указанных в (4) таблиц. В правилах достаточно указать лишь те наследуемые от правил (4) таблицы, для которых такие ограничения существуют. Остальные определённые в (4) таблицы остаются доступными в полном объёме. Кроме того, доступ к таблице может быть закрыт полностью, для этого указывается пустые списки

ограничений. Если для какого-то значения роли ограничений этого типа нет вообще, то соответствующее правило не определяется.

Ограничению по фактору сферы деятельности соответствует атрибут СФД, который определяет списки доступных источников для данной специальности в зависимости от значения этого атрибута. Правила (6), наследуя, как и правила (5), от правил (4) набор таблиц, позволяют определить для каждой таблицы ограничение на набор источников. Правило определяется только для тех значений атрибута СФД, для которых такое ограничение имеется.

Совокупность правил (4-6) задаёт права для полного набора значений атрибутов (СПЕЦ, РОЛЬ, СФД): пользователю доступны таблицы, определённые в правиле типа (4) с учётом дополнительных ограничений на столбцы, строки и источники данных, если они заданы в (5) и (6). Однако в правилах (6) предполагается, что все роли данной специальности имеют одинаковую сферу деятельности. Это представляется типовым случаем, но, при необходимости, сфера деятельности может быть определена с учётом не только специальности, но и роли в правилах (7). Они наследуют ограничения на списки баз данных из правил (6), а их действие заключается во введении списка баз для таблиц, которые не указаны в (6), или замене этих списков.

При вычислении прав для данного набора значений атрибутов механизм контроля доступа прежде всего формирует множество применимых правил. В это множество всегда входит одно из правил (4), которое определяет набор доступных таблиц – частичные права. Если имеются применимые правила и других типов, производится последовательная модификация частичных прав, причём существенно, что правило (7) применяется после правила (6).

Заметим, что рассмотренная система правил не является единственно возможной. Например, к ней может быть добавлено правило с пустым набором атрибутов. Заданные в таком правиле права наследуются всеми пользователями. Ещё один вариант – возможность задания в правилах (4) не только таблиц, но также и каких-либо ограничений, если они являются общими для пользователей одной специальности.

4.4. Определение доступности данных

Доступные данные описываются в полных правилах контроля доступа в виде:

(*Таблица* ((*Столбец*…) (*Строка*…) (*БД*…))…)

В частичных правилах указываются отдельные составляющие этой конструкции. Таблицы и их столбцы задаются своими именами в соответствии с глобальной схемой информационной инфраструктуры. В уточнении нуждаются способы задания списков баз данных и доступных строк.

Для задания списка баз данных (*БД*…) предлагается более мощный способ, чем перечисление конкретных баз. В используемом методе интеграции источники – базы данных описываются расширяемым набором метаатрибутов

(см. раздел 1). При составлении массовых поисковых запросов для выделения подмножества баз используется предикатное выражение, в котором в качестве переменных выступают метаатрибуты, сравниваемые с константными значениями. Этот же способ применяется и для задания списка баз данных в правилах контроля доступа. Например, выделить базы данных, которые принадлежат организациям некоторого региона и специализируются в области машиностроения, можно с помощью выражения из двух предикатов:

region=”Московская область”, sphere=”машиностроение”

Здесь подразумевается оператор конъюнкции, но в более общем случае могут быть использованы операторы дизъюнкции и отрицания. Возможны ситуации, когда сформированный таким регулярным способом список необходимо немного поправить, добавив или исключив из него какие-то базы. Для этого предусматривается прямая адресация посредством метаатрибута name – имя базы данных.

Защита данных на уровне строк является наиболее сложной частью определения политики контроля доступа. Покажем, однако, что её можно реализовать и в рамках предлагаемого подхода.

4.5. Ограничение выбора строк

В конструкциях описания прав элемент (*Строка…*) представляет собой список ограничений, которые выделяют из таблицы строки, объявляемые в соответствующем правиле доступными. В качестве ограничения выступает оператор SELECT, результат которого – множество доступных строк. Фактически это пользовательское представление VIEW, но его определение в правилах контроля доступа имеет важное достоинство: все приложения работают с глобальной схемой данных. Такой способ реализован в Oracle Virtual Private Database (VPD) [20], однако для условий большого числа пользователей предлагается его расширение на основе параметризации: в ограничивающем операторе SELECT вместо константных значений полей могут использоваться параметры. Например, оператор:

*SELECT * FROM Salary WHERE employee-id = $user-id*

позволяет получить строки таблицы Salary, в которых поле employee-id имеет значение, совпадающее со значением параметра user-id (символ $ используется для обозначения параметров).

В качестве параметров могут выступать описанные ранее атрибуты пользователя, однако, кроме того, могут быть введены дополнительные параметры, которые не участвуют в отборе правил контроля доступа, но поставляются в идентификационных документах. Для профессиональной деятельности могут быть полезными атрибуты места работы (ограничение данными, относящимися к организации) и персональный идентификатор пользователя (врач может получать данные, касающиеся только его пациентов).

В приведённом выше примере доступность строк вычисляется по данным самой защищаемой таблицы (Salary). Более типична ситуация, когда данные, соответствующие идентифицирующим атрибутам пользователя, включаются в состав одной таблицы, а требуется защитить строки другой. Эта можно сделать, если существует цепочка ссылок, реализуемая последовательностью операторов JOIN, ведущая от первой таблице к защищаемой. В этом варианте можно говорить о наследовании прав доступа.

Приведём пример из медицинской области, который показывает, как ограничить доступ врачей к личным данным только тех пациентов, которых они обслуживают. Упрощая ситуацию, предположим, что связь врачей с обслуживаемыми пациентами отражена в таблице Doctor со схемой (iddoc, idpat), где iddoc – идентификатор врача, idpat – идентификатор пациента. Вторая таблица Patient имеет схему (idpat, pdata), pdata – набор полей личных данных. Ограничение выбора строк может задано так:

SELECT Patient.pdata FROM Doctor, Patient WHERE Doctor.iddoc=$user-id, Doctor.idpat= Patient.idpat

Параметр $user-id в этом операторе обозначает идентифицирующий атрибут пользователя – врача.

Если известно, что данные глобальной таблицы Doctor хранятся в базах по месту работы врачей, пространство поиска может быть ограничено путём введения группы (SP).

SELECT Patient.pdata FROM SP.Doctor, Patient WHERE Doctor.iddoc=$user-id, Doctor.idpat= Patient.idpat, SP.name=$wplace

В этом запросе используется ещё один параметр – атрибут пользователя wplace, представляющий место работы врача. Предикат SP.name=$wplace отбирает в группу SP базы с метаатрибутами, равными значению этого параметра.

Существует несколько разных подходов к интерпретации детального ограничения рассматриваемого типа. Модель Oracle VPD предполагает автоматическую модификацию исходных запросов пользователя, которые составляются в соответствии со схемой базы данных. Модификация заключается в том, что ограничение на выбор строк вводится в часть WHERE исходного запроса. Как показано в [21], такой подход имеет недостатки. Во-первых, увеличивается время обработки запросов. Во-вторых, автоматическая модификация в некоторых случаях приводит к неверным результатам (не тем, которые ожидает пользователь). В связи с этим в работах [21], [22] предложены подходы к оценке корректности запроса путём его текстуального сравнения с правилами контроля доступа. Эти подходы также не лишены недостатков: соответствующие алгоритмы используют эвристики и имеют высокую вычислительную сложность. Ориентируясь в первую очередь на модификацию

запросов, мы считаем необходимым проведение дополнительных исследований вопроса о способе интерпретации ограничений на выбор строк.

5. Заключение

Обеспечение надёжного контроля доступа при условии приемлемых затрат на администрирование остаётся одним из проблемных вопросов при разработке распределённых систем. Наиболее востребованные современные модели контроля доступа основаны на ролевом подходе, ключевая идея которого – определение прав доступа для групп пользователей. Однако в имеющихся вариантах ролевых моделей в малой степени затрагивается вопрос о способе компактного описания защищаемых ресурсов, что особенно существенно для ресурсов информационных инфраструктур.

В статье представлен способ определения политики контроля доступа к структурированным данным, которые хранятся в распределённых базах и интегрированы в единое информационное пространство. Основная направленность предлагаемого способа – представление в явном виде зависимости между существенными характеристиками пользователей и доступными им данными. Для этого предложено расширение состава характеристик – атрибутов пользователя, которые наряду с ролью включают специальность, сферу деятельности, место работы. Исходя из этого состава, разработана система частичных, то есть содержащих неполный набор атрибутов, правил. Предлагаемая система частичных правил обладает следующими свойствами.

- Состав атрибутов позволяет факторизовать правила контроля доступа таким образом, чтобы частичные наборы атрибутов определяли права для разных слоёв структуризации данных.
- В способе интерпретации системы частичных правил использован механизм наследования прав от одного или нескольких правил. Наследуемые права могут быть модифицированы в детализирующих правилах, либо, если модификация не нужна, детализирующие правила могут быть опущены.
- В правилах могут быть заданы ограничения на доступ к отдельным строкам таблиц глобальной схемы. Эти ограничения могут быть параметризованы с использованием атрибутов пользователя.

Перспективы развития работ мы связываем с созданием инструментов конструирования политики контроля доступа в конкретных применениях, а также с обобщением подхода, направленным на поддержку расширяемого состава атрибутов пользователя.

6. Литература

[1]. Коваленко В.Н., Куликов А.Ю. Интеграция данных и язык запросов в масштабных информационных инфраструктурах // Программные продукты и системы. № 3, 2012, с. 124-130.

[2]. Samarati P., De Capitani di Vimercati S. Access control: Policies, models, and mechanisms // Foundations of Security Analysis and Design, Springer-Verlag, New York, 2001.
[3]. Ghiselli A., Federico S., Zappi R. Review of Security Models Applied to Distributed Data Access // Lecture Notes in Computer Science, Volume 4375/2007, pp. 34-48, 2007.
[4]. Security Assertion Markup Language (SAML) V2.0 Technical Overview. URL: http://www.oasis-open.org/committees/download.php/27819/sstc-saml-tech-overview-2.0-cd-02.pdf
[5]. OASIS eXtensible Access Control Markup Language (XACML). URL: http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xacml
[6]. Sandhu R. S., Samarati P. Access Control: Principles and Practice // Communications Magazine, IEEE, Vol. 32, No. 9. (1994), pp. 40-48.
[7]. Ferraiolo D.F., Kuhn D.R. Role-Based Access Control // 15th National Computer Security Conference, October 1992, pp. 554–563. URL: http://csrc.nist.gov/groups/SNS/rbac/documents/ferraiolo-kuhn-92.pdf
[8]. Proposed NIST Standard for Role-based Access Control / Ferraiolo D.F. [etc.] // ACM Transactions on Information and Systems Security, Volume 4, Number 3, August 2001.
[9]. Al-Kahtani M., Sandhu R. A model for attribute-based user-role assignment // Proceedings of the 18th Annual Computer Security Applications Conference, Las Vegas NV, December 2002.
[10]. A Generalized Temporal Role-Based Access Control Model / Joshi James B.D. [etc.] // IEEE Transactions on Knowledge and Data Engineering, p. 4-23, January, 2005.
[11]. GEO-RBAC: A spatially aware RBAC / Bertino E. [etc.] //. ACM Transactions on Information and System Security (TISSEC), Volume 10, Issue 1, February 2007. URL: http://www.cs.purdue.edu/homes/bertino/geo-rbac.pdf
[12]. Barka E., Sandhu R. A role-based delegation model and some extensions // In Twenty-third National Information Systems Security Conference, Baltimore, MD, October 2000.
[13]. Abdallah Ali E., Khayat Etienne J. A formal model for parameterized role-based access control // IFIP International Federation for Information Processing, Volume 173, 2005, pp. 233-246.
[14]. Bacon J., Moody K., Yao W. A model of OASIS role-based access control and its support for active security // ACM Transactions on Information and System Security, 2002, Volume 5, Issue 4, pp. 492–540.
[15]. Chadwick D., Otenko A. The PERMIS X.509 Role Based Privilege Management Infrastructure // Future Generation Computer System, 19(2):277-289, 2003.
[16]. Alfred Horn. On sentences which are true of direct unions of algebras // Journal of Symbolic Logic, 16, 1951, 14-21.

[17]. A Flexible Attribute Based Access Control Method for Grid Computing / Lang B. [etc.] // J. Grid Computing (2009) 7:169–180.
[18]. Designing access control model and enforcing security policies using PERMIS for a smart item e-health scenario / Hasan Mahmud [etc.] // International Journal of Engineering Science and Technology, Vol. 2(8), 2010, pp. 3777-3787.
[19]. Howitt A. The formal specification of the Tees confidentiality model // University of Teesside, 2008. URL: http://tees.openrepository.com/tees/handle/10149/112674.
[20]. The Virtual Private Database in Oracle9iR2: An Oracle Technical White Paper // Oracle Corporation, January 2002.
[21]. Extending query rewriting techniques for fine-grained access control / Rizvi S. [etc.] // Proceedings of the 2004 ACM SIGMOD international conference on Management of data (SIGMOD '04), ACM, New York, NY, USA, pp. 551-562. URL: http://www.cse.iitb.ac.in/~sudarsha/Pubs-dir/nontruman-sigmod04.pdf
[22]. Assessing query privileges via safe and efficient permission composition. Sabrina De Capitani di Vimercati [etc.] // Proceedings of the 15th ACM conference on Computer and communications security (CCS '08). ACM, New York, NY, USA, pp. 311-322. URL: http://spdp.dti.unimi.it/papers/ccs2008.pdf

Оглавление