کنترل دسترسی در پایگاه داده های مبتنی بر XML توزیع شده

میثم حجازی نیا

چکیده: امروزه با توجه با افزایش حجم اطلاعات ساختاریافته، در قالب XML، و تعاملات تجاری مبتنی بر آن، و نیاز به نگهداری این اطلاعات برای جستجوهای آینده، نیاز به پایگاه های داده های XML افزایش یافته است، بر روی این پایگاه داده ها، روش های کنترل دسترسی متعددی از قبیل روش های مبتنی بر وظیفه، روش کنترل دسترسی مبتنی بر نقش، و کنترل دسترسی تفویضی پیشنهاد شده است. این طرح ها عمدتا بصورت مرکزی می باشد، و لذا برای کاربردهای توزیع شده تنظیم نشده است، در بسترهای کنونی ایجاد شده چون گرید، که در آن سازمان های متعدد، و افراد مختلفی از آنها نیاز به دسترسی به داده های سازمان دیگر می باشند، نیاز به مدل کنترل دسترسی می باشد، که علاوه بر توسعه پذیری که نیازمندی اساسی سیستم های توزیع شده می باشد، بایستی انعطاف پذیری را بدلیل پویایی موجود بر روی این بستر ها ایجاد نماید، در این مقاله به ارائه روشی می پردازیم، که با استفاده از روش های کنترل دسترسی مرکزی، توانمندی استفاده توزیع شده و انعطاف پذیر را در فضاهای توزیع شده فراهم نماید.

1. مقدمه

امروزه پایگاه داده های XML متعددی در سازمان های مختلف وجود دارد که اکثرا تراکنش های سازمان در آنها نگهداری می گردد. سه حالت از کاربردهایی که این پایگاه داده ها دارند عبارتند از:

حالت اول: مستندات XML در انتقال داده بین برنامه های کاربردی تجارت الکترونیک تحت وب را ممکن می سازد. این برنامه های کاربردی نیاز دارند ارتباطاتشان را لاگ نموده و بنابراین با توده ای وسیع از لاگ های مستندات XML روبرو هستند. مدیران ممکن است بخواهند محتواهای این مجموعه XML را تحلیل نمایند.

حالت دوم: وب سایت ها می توانند از مستندات XML برای نگهداری داده ها برای انتشار روی وب استفاده نمایند. رشد وب سایت ها می تواند با رشد تعداد یا سایز مستندات XML روبرو گردد. بعلاوه، برخی از آن مستندات XML بایستی بطور مرتب بروز رسانی گردند.

حالت سوم: مراکز خبری می توانند تمامی مقالات خود را در مستندات XML مقاله خبری نگهداری نمایند. ویرایشگر ممکن است بخواهد تمامی مقالات در مورد انتخابات ریاست جمهوری قبلی را که بیشتر از همه بدان ارجاع شده است بیابد.

که کاربران متعددی در سازمان های مختلف وجود دارند که مدل کنترل دسترسی آنها مدل گراف نقش می تواند باشد، اما مسئله ای که وجود دارد آن است که ممکن است سازمان های متعددی در لایه ای بالاتر کنسرسیومی را تشکیل دهند که این امکان برای آنها فراهم گردد که اطلاعات خود را به اشتراک بگذارند،

این اشتراک اطلاعات جدید از طرفی نبایستی زیرساخت های امنیتی کنترل دسترسی آنها را تحت تاثیر قرار دهد، از سمتی دیگر در بسیاری از مواقع موقتی است، و در درجه سوم نیاز به رویه های کنترل دسترسی جدیدی می باشد، که این دسترسی ها در گراف نقش قبلی تعریف نشده است، با این مسئله بحث کنترل دسترسی توزیع شده مطرح می گردد. راه حل پیشنهادی ما در اینجا به حل این مسائل مطرح در توزیع شدگی پرداخته، و یک گراف نگاشت نقش را بر روی این نقش ها تعریف می نماید،

مطالبی که در ادامه می آید بصورت زیر می باشد، در بخش 2 به توضیح مدل کنترل دسترسی xml مبتنی بر نقش می پردازیم، سپس در بخش 3 مدل نقش وظیفه ای را توضیح می دهیم، در بخش بعدی به ارائه مدل خود که توسعه ای از این مدلهاست می پردازیم.

2. مدل کنترل دسترسی مبتنی بر نقش برای پایگاه داده های xml

پایه این مدل دو مدل گراف نقش و مدل کنترل دسترسی پایگاه داده های شیء گرا می باشد. مدل گراف نقش مبتنی بر سه گراف می باشد، یکی گراف نقش هاست، دوم گراف گروه هاست، و سوم گراف افراد است. در گراف نقش هر نود نقش می باشد، و گراف بصورت جهت دار می باشد و هر یال ها از نقش های ارشد تر شروع می گردد. این گراف مشخص می نماید، که برای مثال نقش مدیر مالی در یک سازمان اجازه مشاهده اطلاعاتی را دارد که کارشناس مالی اجازه دسترسی به آنها را دارد، در این گراف در صورتی که مسیری از نقشی به نقشی دیگر وجود داشته باشد، این نقش مجاز به تمامی دسترسی هایی است که نقش دیگر دارد. مولفه دوم گراف گروه می باشد، در این گراف ممکن است افراد مختلف از به گروه های مختلفی نگاشت شوند، این امر نسبت دادن دسترسی ها را ساده می سازد، به این شکل که با یک نگاشت یک گروه به یک مجوز، افراد مختلف با نقش های متعدد در داخل آن گروه می توانند آن مجوز را داشته باشند. و گراف سوم هم گراف نگاشت افراد به نقش هاست، که به این شکل هر فرد می تواند به نقش های متعدد نگاشت گردد، لذا اگر نقشی مجوزی را داشته باشد، تمامی افرادی که آن نقش را دارند می توانند به آن نقش دسترسی پیدا کنند. پایه دوم این مدل، مدل کنترل دسترسی پایگاه داده شیء گراست، که در این مدل برای دسترسی به صفات مختلف عملگرهایی تعریف می گردد، که این عملگرها با هم یک گراف می سازند، که این گراف جهت دار، اگر از نودی به نود دیگر یالی وجود داشته باشد، یعنی فردی که مجوز دارد عملیات متناظر نود اول را انجام دهد، می تواند عملیات متناظر نود دوم را نیز به انجام رساند، همچنین در کلاس ها گرافی تعریف می گردد، که ارگ در این گراف یالی از نودی که متناظر کلاس الف است، به نودی که متناظر کلاس ب است، وجود داشته باشد، پس اگر فردی مجوز مشاهده کلاس الف را داشته باشد، مجوز مشاهده کلاس ب را نیز خواهد داشت.

در مدل کنترل دسترسی مبتنی بر نقش برای پایگاه داده های xml ما بجای کلاس ها، مسیرها را در پایگاه داده xml تعریف می نمائیم، به این طریق ما در صورتی که روی مولفه ای دسترسی را تعریف نموده باشیم، این دسترسی تا وقتی که دسترسی آشکاری روی زیر مولفه آن داده یا رد شده باشد، به نقش زیر مولفه ها منتقل می شود، که البته انتشار این دسترسی ها مطابق مدل دسترسی شیء گرا می باشد، نمونه ای از این دسترسی ها در زیر نشان داده شده است:

$C_{AcademicCounceller} = \{(// completedCourse, update)\}$

که این بیانگر آن است که فردی که نقش academicCounceller را دارد، مجوز بروز رسانی درس های کامل شده را دارا می باشد.

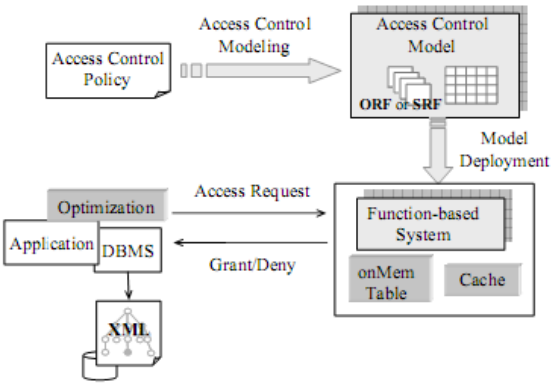
مشکلی که در این مدل وجود دارد، یکی این نکته است که این دسترسی ها صرفا

روی مولفه ها و زیر مولفه ها تعریف شده است، و به attribute هایی که در فایل xml تعریف می گردد توجهی نشده است، ثانیا برای کوئری هایی که در آن // استفاده شده است، پیش بینی صورت نگرفته است.

3. مدل کنترل دسترسی مبتنی بر وظیفه برای پایگاه داده های xml

مشکلی که در این مدل بدان پرداخته شده است، بحث عدم بهینه بودن مدل های کنونی و حجم بالای آنها می باشد. در این مدل سه نوع مجوز دسترسی تعریف می گردد، یکی r+ که اجاز دسترسی فقط روی یک مولفه و نه زیر مولفه ها را می دهد، دوم R+ که اجازه دسترسی به تمامی زیر مولفه ها را می دهد، و سوم R– می باشد، که مجوز دسترسی روی مولفه ای را می گیرد. در این روش برای حل مشکل ACL و مدل صلاحیت، و حل مشکلی که در مدل مبتنی بر نقش وجود دارد، سه نوع وظیفه قانون تعریف می گردد، یکی وظیفه قانون شیء گرا (ORF)، یکی وظیفه قانون کاربر گرا (SRF) و سوم وظیفه قانون عمومی (GRF) می باشد، که بترتیب در اولی وظایف قانون نسبت به اشیاء شاخص شده، در دومی نسبت به کاربر، و سومی متعلق به وظایف قانونی می باشد که در هیچکدام از دو مورد قبل نیامده است. این شامل مواردی است که از // استفاده می نمایند. سپس مطرح شده است که برای ترکیب این موارد بایستی به چه شکل عمل نماییم، تا وظایف قانون به زبان جاوا نوشته شود. این روش همچنین Caching ای را برای افزایش performance ارائه نموده است.

معماری این سیستم بصورت زیر می باشد:



در مدلسازي كنترل دسترسي سياست كنترل دسترسي به كلاس هاي جاوا تبديل مي گردد كه هر كلاس نمايش دهنده قوانين كنترل دسترسي براي تعداد خاصي از اشياء و كاربرهاست. در حين اين مدلسازي جدول نگاشت متناظر نيز توليد مي گردد. بعلاوه قطعه كد جاوا در كدبايتي بنحوي كامپايل مي گردد كه قابليت اجرا داشته باشد. فرآيند ديگر تحويل مدل مي باشد، كه ابتدائا جدول نگاشت جاوا را به حافظه اصلي بار نام دارد كه در حافظه اصلي نمونه هاي شيء InMemTableنموده و سپس جدول عمومي خالي سيستم ديگري را آماده مي سازد كه جاوايي را كه براي پردازش درخواست دسترسي لازم مي باشد، نگهداري مي نمايد. با توجه به آنكه هر گروه وظايف داراي نام يكساني است از نام گروهي به عنوان كليد و ارتباط آن با اشياء نمونه استفاده مي نمايد.

4. راه حل پیشنهادی

مشکل اصلی که در دو روش فوق بدان توجه نشده است، آن است که مدل های فوق برای سیستم های محلی مناسب می باشند، و انعطاف پذیری کافی برای سیستم های توزیع شده ای را که افراد و سازمان ها در آن پویایی بالاتری را دارند، را ندارند، و لذا بدلیل نیاز به مقیاس پذیری (Scalability) در این فضا جدول دسترسی مبتنی بر نقش نمی تواند بطور مرکزی ایجاد گردد، و بایستی پویایی لازم را داشته باشد، لذا در مدل پیشنهادی ما نگاشتی از نقش های مرتبط یک سیستم به سیستم دیگر برای هر جفتی از سیستم ها انجام می گیرد، برای این نگاشت ها

بایستی مشخص گردد، که انتشار اختیار به چه شکل، و اضافه کردن و حذف کردن از آنها به چه شکل باشد، که در این بخش به توضیح هر کدام از این موارد می پردازیم. با فرض اینکه هر کدام از سایت های درگیر دارای مدل کنترل دسترسی مبتنی بر نقش می باشند، وقتی یک فرد از سایتی می خواهد به سایتی دیگر دسترسی داشته باشد، طبق مدل دسترسی آن سایت این فرد بایستی نقشی را داشته باشد که به او assign شده باشد، اما این سایت فقط می داند که فرد از سایت دیگری آمده است، افرادی که از سایت دیگر دسترسی به این سایت دارند لزوما دارای نقش های یکسانی نیستند، از سمت دیگر این افراد نه تنها متغیر هستند، و این متغیر بودن، در سازمان اول مشخص می شود، و ممکن است در صورتی که در مدل مبتنی بر نقش سایت جدید وارد شود، برای از بین بردن آن اطلاع از سمت سایت اول داده نشود، لذا در صورتی که این افراد را در مدل مبتنی بر نقش سازمان دوم درج نماییم، بحث درج، و حذف آنها دارای پیچیدگی بوده، و مدل مناسب نمی باشد. از سمتی این افراد در سایت خود برای کاهش حجم جداول ACL و Capability، به نقش هایی نگاشت شده اند، که دسترسی ها بر اساس این نقش ها صورت می گیرد، پس در صورتی که گراف نقش جدیدی را تعریف کرده و افراد را بدان نگاشت نمائیم، بدلیل متغیر بودن، و حجم بالای کار در زمانی که سازمانی به این شبکه می پیوندد و از آن جدا می گردد، این روش مناسبی نمی باشد.

برای حل این مشکل بایستی نگاشتی بین نقش ها تعریف گردد، ما علاوه بر گراف های نقش که در هر کدام از سایت ها تعریف می گردد، یک گراف نگاشت بین نقش ها (act-as) تعریف می نمائیم، این گراف گرافی است که نگاشت نقش های مربوط سایت الف را به سایت ب نگاشت می نماید. از طرف دیگر بحث دومی که مطرح می گردد، وجود مواردی است، که دسترسی به مولفه های یک سایت، نبایستی توسط نقش نگاشت شده در سایت دیگر مورد انجام شود، برای حل این مطلب دو سناریو مطرح می گردد، سناریو اول حالتی است که مثلا administrator یک سایت به administrator یک سایت دیگر، نگاشت گردد، این نگاشت ذاتا در گراف act-as دارای مشکل بوده است. سناریوی دوم می تواند این باشد که مثلا دانش آموزان یک سایت اجازه دسترسی به عنوان دانش آموزان سایت دیگر را داشته باشند، در این حالت مشکلی که وجود دارد، آن است که ممکن است دانش آموزان سایت الف دسترسی به موارد خاصی را داشته باشند، که این دسترسی برای دانش آموزان سایت دوم نبایستی وجود داشته باشد، برای این بحث ما گرافی نفی را تعریف می نمائیم، در این گراف که برای افراد یک سایت جهت دسترسی به سایتی دیگر مطرح می گردد، ما نفی نقش هایی را تعریف می کنیم، مثلا برای دانش آموز، نقشی به نام آنتی دانش آموز پیشنهاد می نمائیم، برای این آنتی دانش آموز، اجازه دسترسی به موارد اختصاصی به وضوح نفی می گردد، اضافه کردن این نقش به گراف نقش سایت الف، مشکلی ایجاد نمی کند، چرا که پویایی در آن وجود ندارد. عدم پویایی بدین معناست، بدلیل آنکه این موارد عمومی برای هر سایتی است که به سایت الف می خواهد متصل گردد، لذا یک دفعه تعریف آن کفایت می کند، از سوی دیگر در صورتی که برای سایت خاصی مثلا، سایت مربوط به مردم ایران بخواهیم این دسترسی را منع نماییم، باز می توان از تعریف یک نقش جدید روی این گراف استفاده

نمود. لذا برای هر نقش ما نگاشتی از نقش های نفی و نقش های اصلی را داریم، برای آنکه دسترسی به کاربری از سایت دیگر داده شود، دسترسی تنها در صورتی داده می شود، که مجموعه نقش هایی که به یک فرد داده می شود، روی آن آیتم داده نفی نشده باشند. نکته دیگری که مطرح می گردد، آن است که دسترسی ها به چه صورت انتشار می یابد، این مسئله آنجا مطرح می شود که از نتیجه انتشار نقش ها در سایت الف، ما مجوزهایی برای نقش خاصی داریم، برای سایت ب نیز ما مسئله مشابهی را داریم، لذا بایستی به نحوی اختیارات منتج از انتشار سایت ب را به سایت الف برای دسترسی منتقل نمائیم. برای این انتقال پس از تشکیل گراف اولیه نگاشت بین نقش های سایت ها، ابتدا گراف بستار را در سایت ب ساخته، و سپس بر اساس آن نگاشت را انجام می دهیم.

مسئله دیگر آن است که واقعا این گراف نگاشت نقش بایستی در کجا قرار گیرد، این گراف نگاشت در دو جا می تواند قرار گیرد، یکی در سایت الف می باشد، که به این شکل بایستی در یک ارتباط امن تمامی نقش های متناظر برای کاربر، به سایت ب فرستاده شود، روش دیگر آن است که این گراف در سایت الف موجود باشد، به این روش با فرستادن نقش فرد از طرف سایت ب، در یک ارتباط امن ممکن می باشد، که سایت الف نقش های متناظر او را یافته و دسترسی ها را بررسی و به تناسب نتیجه را باز می گرداند. برای جلوگیری از مشکل پیچیدگی کنترل دسترسی بدلیل وجود گراف های متعدد، در حالتیکه مثلا توپولوژی اتصال بین سایت ها بصورت ستاره باشد، روش اول مناسب تر است، به این روش، برای هر دسترسی سایت مبدا براحتی نقش ها را نگاشت نموده و برای سایت دوم ارسال می نماید، به این روش پیچیدگی حجم محاسبات در همان سایت متصل شونده، متمرکز شده، و لذا حجم محاسبات در سایتی که بدان قرار است دسترسی پیدا گردد کاهش می یابد.

5. پاسخ به تراکنش های توزیع شده

جهت پاسخ به تراکنش توزیع شده، نیاز به هماهنگی بین سایت های مختلف وجود دارد، برای این مطلب تنها کافی است پروتکل 2PC تغییر نماید، لذا در حالتی که هر سایت پاسخ می دهد که آیا توانایی انجام این تراکنش را دارد یا خیر، این درخواست را با توجه به نقش های فرستاده شده از سایت فرستنده پاسخ می دهد، لذا پس از پاسخ به سایت درخواست کننده، این پاسخ ها طبق پروتکل 2PC بررسی می گردند، در این روش پاسخ ها در پروتکل 2PC متفاوت می گردد، این پاسخ می تواند به دو شکل باشد، یکی پاسخ رد امنیتی می باشد، که به این شکل سایت درخواست کننده متوجه می شود که دسترسی به اطلاعات آن سایت مجاز نیست، لذا پاسخ منع دسترسی به کاربر باز گردانده می شود، پاسخ دوم نیزم می تواند پاسخ عدم دسترسی بدلیل Lock باشد، که در این حالت کاربر متوجه می شود که پس از مدتی می تواند مجددا درخواست اجرای کوئری را بنماید.

البته در اجاری پروتکل توزیع شده، دسترسی به اطلاعات دو روش مطرح می گردد، روش اول آن است که اجرای تراکنش در همان سایت انجام گیرد، که در فوق پاسخ آن توزیع داده شد، روش دوم آن است که تراکنش با انتقال داده ها به سایت اصلی اتفاق افتد، که در این حالت، مسئله جریان اطلاعات مطرح می گردد،

چرا که ممکن است اطلاعاتی که برچسب امنیتی خورده باشند، به جایی انتقال پیدا کنند، که در آنجا اطلاعات فاش گردد، برای جلوگیری از این مسئله در کپی کردن اطلاعات برچسبی کوچکتر مساوی برچسب اطلاعات اولیه را می خورند، اما مسئله ای که مطرح می شود آن است، که ممکن است پس از انتقال اطلاعات به سایت ثانویه با حمله ای، این اطلاعات سطحشان تغییر کند، و این منجر می گردد که اطلاعات سایت الف فاش گردد، برای جلوگیری از این مطلب، بایستی اطلاعات بنوعی کد گردد، با توجه به آنکه اطلاعات xml برای دسترسی توسط برنامه های کاربردی است، لذا فاش شدن، در دسترسی در بین کامپونت های نرم افزاری رخ می دهد، پس برای هر اطلاعاتی که کپی می گردد، پیشنهاد می گردد، که کد محافظی تعبیه گردد، اطلاعات فرستاده شده در داخل این کد تعبیه می گردند، و با توجه به نقش دریافت شده توسط آن فاش می گردد، لذا برای هر کدام اطلاعات کد جاوایی نوشته می شود، که در آن دو نوع وظایف قانونی بکار برده می شود، این وظایف قانونی یکی وظایف قانونی شیء گرا، و یکی وظایف قانونی عمومی می باشد، که این وظایف قانونی عمومی بحث دسترسی های بصورت '//' را بر طرف می نماید. از طرفی برای نگهداری اطلاعات در این مدل ما سطوح دسترسی را بصورت ریز دانه تعریف می نماییم، لذا سطوح دسترسی برای هر فایل xml در داخل فایل xsd آن از قبل تعریف شده است. برچسب فایل xsd در ابتدای ایجاد معادل برچسب ایجاد کننده آن می باشد، اما دسترسی ریز دانه روی زیر مولفه های یک مولفه همان طور که در مدل نقش وظیفه ارائه شده است، قابل تعریف است، نکته ای که در مقالات متعدد به آن توجه نشده است، بحث تعریف سطوح دسترسی روی صفات است، که پیشنهاد ما آن است، که بر روی فایل xsd برای صفات هر کدام از المنت ها تعریف گردد.

6. تحقیقات آینده

تمامی بحث هایی که گفته شد در مورد کنترل دسترسی بر روی خواندن بود، و دسترسی های نوشتن در هر سایت تنها توسط همان سایت انجام می گیرد. برای دسترسی نوشتن، ما دو تغییر در اینجا داریم، یکی تغییر روی فایل های xsd می باشد، و یکی دسترسی روی خود فایل های xml، می باشد، در بحث نوشتن مطرح می گردد، که آیا فردی از یک سایت دیگر مجاز به تغییر اطلاعات دیگر سایت ها می باشد یا خیر، که این مطلب خود در دو بعد مورد توجه قرار می گیرد، یکی بحث فایل هایی است که یک فرد در جای دیگر ایجاد نموده است، و مورد دوم نیز فایل هایی می باشد، که در سایت بصورت محلی ایجاد شده، که خود این فایل ها نیز به دو دسته فایل های اختصاصی و فایل های عمومی تقسیم می گردد، که می تواند برای تحقیقات بعدی مورد استفاده قرار گیرد.

7. نتایج

در این مقاله به بررسی روش های کنترل دسترسی xml قالب پرداختیم و مزایا و معایب آنها را توضیح دادیم. سپس روش کنترل دسترسی را معرفی نمودیم که پاسخگوی نیاز امروز در مورد تراکنش های توزیع شده، و یا دسترسی به منابع سایت دیگر، در محیط هایی چون گرید یا شرکت های شریک و سازمان های انتزاعی بود. این روش می تواند در مدیریت کنترل دسترسی replica ها مورد استفاده قرار گیرد، این روش دارای خاصیت انعطاف پذیری و مقیاس پذیری بود، گراف نگاشت نقش در این مقاله تعریف گردید، که با قرار دادن آن روی سایت

های درخواست کننده، ترافیک سرور های سرویس دهنده را کاهش دادیم، در نهایت بیان شد که این مدل کنترل دسترسی صرفا برای عملیات خواندن می باشد، و برای عملیات نوشتن نیاز به توسعه های بیشتری نیز دارد.

8. مراجع

[1] A Role-Based Approach to Access Control for XML Database, Jingzhu Wang, Sylvia L. Osborn, Department of Computer Science,The University of Western Ontario,London, Ontario, Canada ,N6A-5B7,[ACM2003].

[2] A Function-Based Access Control Model for XML Databases, Naizhen Qi, Michiharu Kudo, Jussi Myllymaki, Hamid Pirahesh, IBM Research, Tokyo Research Lab,[ACM2005].

[3] A Flexible Mandatory Access Control Policy for XML Databases, Hong Zhu, Renchao Jin, Kevin Lü, Huazhong University of Science and Technology[ACM2007]

[4] An Access Control Method Based on the Prefix Labeling Scheme for XML Repositories, hohei Yokoyama, Manabu Ohta, Kaoru Katayama and Hiroshi Ishikawa, Graduate School of Engineering .Tokyo Metropolitan University[ACM2005]

[5] Relevancy Based Access Control of Versioned XML Documents, Mizuho Iwaihara, Somchai Chatvichienchai, Chutiporn Anutariya, Vilas Wuwongse, Dept. Social Informatics, Kyoto University, Kyoto, 606-8501 Japan[ACM2005]

[6] A Bitmap-Based Access Control For Restricted Views Of XML Documents, Abhilash Gummadi, Jong P. Yoon, Biren Shah, Vijay Raghavan, Center For Advanced Computer Studies, University of Louisiana at Lafaye□e[ACM2003]

[7] Access Control of XML Documents Considering Update Operations, Chung-Hwan Lim, Seog Park, Sang H. Son. Dept of Computer Science, University of Virginia[ACM2003].