VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta informačních technologií

Teorie programovacích jazyků

Využití ontologií k řízení přístupu

Roman Petrucha leden 2005

1 Úvod

V současné době je velice obtížné využít „strojů“ k interpretaci a zpracování informací umístěných na WWW. Hlavním problémem je, že sémantika dat je vyjádřena pouze pro potřebu lidí. Lidé vnímají intuitivně jednotlivé pojmy umístěné na WWW a jejich vztahy mezi nimi. Stroje toto v současném prostředí nikdy nemohou dokázat.

Proto vznikl jako rozšíření současného webu web sémantický (SW). Webové stránky v SW jsou okomentovány „anotovány“ koncepty, které jsou formálně definovány v ontologiích (znalostních modelech). Ontologie přináší strukturu do významového obsahu webových stránek.

Po zavedení těchto rozšíření do WWW dojdeme k problému, který zatím není vývojáři kompletně vyřešen. Konkrétně se jedná o problém řízení přístupu k datům. V klasickém webu je problém řízení přístupu řešen na základě centralizovaných schémat. Bezpečnostní administrátor nadefinuje v konkrétním systému uživatele, případně skupiny a jejich práva. Uživatelé se autentizují do systému a na základě této autentizace získají svá práva v rámci tohoto systému.

Pokud se zamyslíme nad koncepcí SW, tak tento způsob centralizovaného přístupu v něm nelze použít. V SW jde o naprosto odlišný přístup, protože potřebujeme uživatelova oprávnění identifikovat i v rámci jiných systémů(v rámci celého webu), než jen v jeho domovském systému. Hlavním rozdílem oproti webu klasickému jsou přímo přístupná metadata, která nejsou nikde uschována pod heslem jako např. v klasických databázových aplikacích.

V následujících kapitolách se seznámíme s několika zajímavými přístupy pro řízení přístupu v sémantickém webu.

2 Řízení přístupu

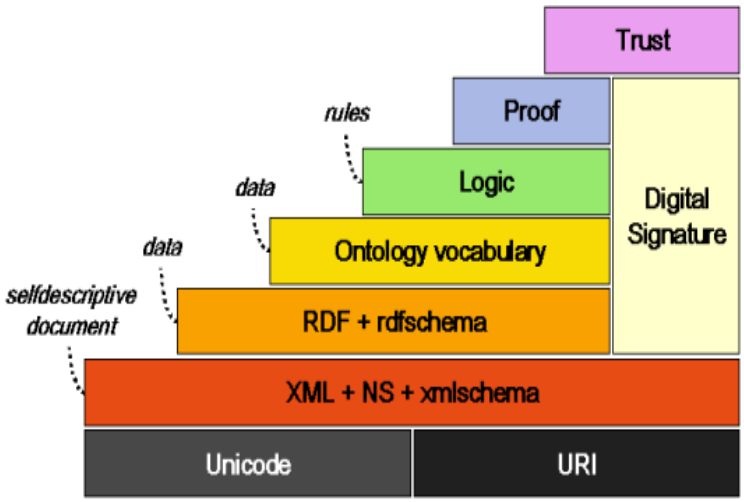
V současné době existuje několik možných přístupů, které by šly využít pro řízení přístupu v sémantickém webu. Přístupy je možné rozdělit do několika proudů podle technologií na kterých jsou založeny.

- XML
- RDF
- Ontologie

2.1 XML přístup

Tento přístup k řešení tohoto problému je inspirován výzkumem v oblastí řízení přístupu v XML dokumentech, kde přístupové oprávnění může být specifikováno v elementech XML dokumentu nebo DTD. XML dokument je považován za instanci jeho DTD schéma. DTD prakticky specifikuje platné elementy a seznamy povolených atributů, jejich strukturu a umístění v dokumentu. Jelikož je SW založen na syntaxi XML(viz. vrstvový model SW na obr č.1.) je logické zkusit použít XML model pro řízení přístupu. Pokud tento model budeme chtít použít na SW je potřeba provést jeho rozšíření (Musíme si uvědomit že SW architektura obsahuje také vyšší vrstvy nad XML jako RDF a ontologie.).

V XML modelu je také možné provádět propagaci(šíření) přístupových oprávnění, což je pojem u klasických systémů nepoužívaný. Specifikace zda je či není propagace umožněna je definována v DTD pro jednotlivé elementy samostatně. Jednotlivá oprávnění se šíří na všechny podelementy nadefinovaného elementu, pro který je propagace vyvolána. Nevýhoda tohoto přístupu spočívá v tom, že se práva šíří dále v XML podstromu nekontrolovaně.



Obr. č. 1 Vrstvový model sémantického webu

2.2 Přístup s využitím RDF

Na této úrovni existuje obrovské množství různorodých přístupů k řízení přístupu k RDF metadatům. Jedním z nich je například Personal Data Access Language [10]. Jiné modely například využívají k autorizaci ke konkrétním RDF metadatům různé certifikáty. Do budoucna s narůstajícím množstvím sémantických metadat jsou tyto přístupy vysoce neefektivní a nepoužitelné.

2.3 Přístup s využitím ontologií

V praxi je důležité nejen si vynutit řízení přístupu v elementech, dokumentu, nebo na DTD úrovni, ale existuje zde také potřeba řídit přístup na konceptové úrovni. Například je pochopitelné, že někdo bude chtít omezit přístup k webovým datům, které obsahují informace o biologických zbraní lidem z jistých zemí. Dalším klasickým případem je zamezit lidem mladším 18 let přístupu k webovým datům „xxx“. Místo specifikace oprávnění nad každým elementem v každém souvisejícím dokumentu nebo DTD je daleko efektivnější specifikovat řízení přístupu nad koncepty jako „biologické zbraně“, „xxx“ a vynutit si je nad všemi jejich datovými instancemi. Tyto koncepty jsou definovány v ontologiích, které jsou důležitou součástí SW.

Základní rozdíly a možné problémy oproti předcházejícím modelům jsou následující:

- Pokud budeme uvažovat o každém konceptu samostatně a ignorovat vztahy mezi koncepty pak může dojít k porušení bezpečnosti. Příkladem může být nadefinování práv určitému konceptu(který na ně nemá nárok) a pak může dojít pomocí inference k přenesení této bezpečnostní chyby dál.
- Informace může být nepřístupná(autorizovanému uživateli) pokud neuvažujeme vztahy mezi koncepty. Příkladem je přístup ke dvěma ekvivalentním konceptům, přičemž na jeden práva k přístupu máme a na druhý ne.
- Redukce počtu zadávání explicitních přístupových oprávnění, díky možnosti odvození dalších pravidel.
- Ontologie se může vyvíjet(přidání konceptu, změna vztahů, atd.). Například v případě přidání ekvivalentního konceptu je možné mu ihned automaticky nastavit korektní oprávnění.

Tato oblast řízení přístupu na úrovni ontologií není zatím moc prozkoumána. Na této úrovni jsem zatím objevil základy pouze jednoho modelu, kterému se budu věnovat dále.

3 Concept-level based model

Tento model je založen na přístupu s využitím ontologií popsaném v předchozí kapitole. Nejdříve se musíme seznámit aspoň se základními pojmy.

3.1 Ontologie

Ontologie je termín vypůjčený z filozofie, který náleží vědnímu oboru, který se zabývá popisem jednotlivých druhů entit ve světě a jejich vzájemných vztahů. V literatuře zabývající se umělou inteligencí je ontologie definována jako formální explicitní specifikace konceptualizace.

Obecně ontologie definuje běžnou slovní zásobu pro subjekty, kteří potřebují sdílet informace v určité oblasti (doméně). Obsahuje strojově interpretované definice základních konceptů v dané doméně a relace mezi nimi.

Pro vyjádření ontologie potřebujeme znát speciální jazyk. Současným nejpoužívanějším standardem jsou jazyky RDFS, DAML+OIL(podporovaný DARPA) a OWL(podporovaný W3C). Do budoucna se počítá pouze s vývojem na platformě jazyka OWL.

3.1.1 Role automatické inference

Webové ontologie byly od počátku chápány nikoliv jako pasivní soubory platných vztahů, ale jako znalostní báze, nad kterými lze strojově odvozovat. Kombinací různých ontologií(např. propojením ekvivalentních konceptů) lze získat nové vztahy mezi různými ontologiemi a pomocí nich odvodit nové doposud neznámá data. Inferenci mají na starost různé typy inference engine.

3.1.2 Příklad ontologie

Jak vypadá část ontologie v grafické podobě ve formě orientovaného označeného grafu je vidět na následujícím obrázku č.2. Jednotlivé ovály nám specifikují konkrétní koncepty. Pojmenované orientované hrany definují vztahy-relace mezi koncepty.

Na tomto příkladu budeme dále vysvětlovat další principy.

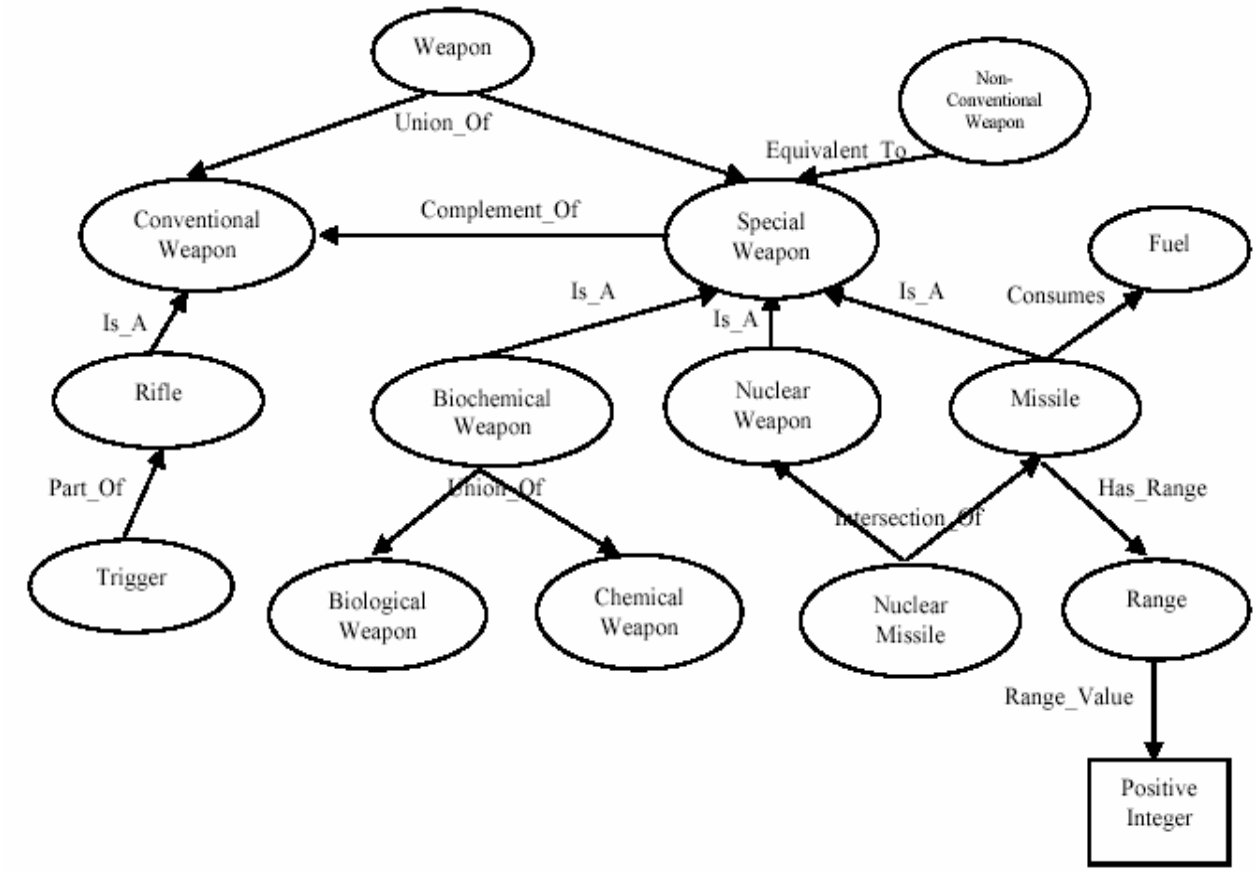


Figure 1. Part of an Ontology for Weapons.

Obr. č.2. Část zbraňové ontologie

3.1.3 Definice konceptu

Koncept c je definován jako n-tice
c = (O, T, (P, R, V), I)

O je ontologie, ve které je koncept c definován
T je množina taxonomií(vztahy organizující hierarchickou strukturu)
P je množina vlastností
R je množina omezení nad P
V je množina hodnot, kterých může nabývat P (primitivních datových typů)
I je množina instancí konceptu c

Koncept ‘střela’ je definován jako druh speciální zbraně, spotřebovávající benzín, který je definován v ontologii http://x.edu/ontology/fuel a je charakterizována číslem, které udává její dosah v ontologii http://x.edu/ontology/weapon. Instance ‘střely’ může být například Irácká střela. V následujícím příkladu je tato věta zapsáno dle předcházející definice.

missile = (http://x.edu/ontology/weapon, (subclassOf, special weapon),
(Has_Range, allValuesFrom, range),
(Consumes, someValuesFrom, http://x.edu/ontology/fuel#fuel), Iraq’s missile)

V další fázi narazíme na pojem *instance konceptu*, což jsou vlastně RDF data, které jsou anotována konceptem.

3.2 Základní vztahy mezi koncepty

V této kapitole se seznámíme se základními vztahy, které lze identifikovat mezi koncepty. Nemusíme se pouze omezovat na oblast dané ontologie, ale vztahy mohou existovat i mezi koncepty více různých ontologií. Příklad ontologie zbraní, kterou jsme si zde uvedli na obrázku č. 2 patří do kategorie doménových ontologií, které nám popisují jen určitou oblast(v tomto případě zbraňovou ontologii). Zajímá nás identifikace pouze doménově nezávislých vztahů. To znamená, že hledáme obecné vztahy mezi koncepty, které mohou existovat ve „všech“ ontologiích. Toto je ale v současné době téměř nemožné, dokud nebude v oblasti sémantického webu jasně definován standard, který se bude používat.

Nejvíce záleží na tom v jakém jazyce je daná ontologie vytvořena. Jelikož existují různé ontologické jazyky, které umožňují specifikovat vztahy mezi koncepty na různých úrovních rozlišení. I v rámci jazyka jako je OWL je možno narazit na rozdíly v existujících vztazích v různých jeho verzích. Díky tomu má LITE verze nejmenší sílu inference oproti FULL verzi.

Pokud by se při vytváření modelu pro řízení přístupu jednalo o konkrétní ontologii, měla by být umožněna i specifikace jiných ne standardních vztahů mezi koncepty. Následuje seznam několika doménově nezávislých vztahů, které se vyskytují a jsou používány v naší příkladové ontologii.

- Superclass/Subclass
- Ekvivalence
- Část/Celek (Part/Whole)
- Overlap/Intersection (“přesah/průnik”)
- Sub-koncept/Sjednocení(Sub-concept/Union)
- Komplement

3.3 Klasifikace vztahů

Vztahy – relace se kterými jsme se seznámili v předchozí kapitole je možné klasifikovat na základě následující vlastnosti, která je velice podstatná pro potřeby sémantického web. Konkrétně se jedná o to zda a jak je možné provést inferenci. Na tomto základě lze vztahy klasifikovat do třech následujících skupin.

- *Inferovatelné vztahy* *(Inferable Relationship)* **IR**
- *Částečně inferovatelné vztahy* *(Partially Inferable Relationship)* **PIR**
- *Ne-inferovatelné vztahy* *(Non-Inferable Relationship)* **NIR**

Instance konceptu c_j mohou být inferovány z instancí konceptu c_i. Značíme: $c_i => c_j$
Neproveditelná inference, značíme: $c_i !=> c_j$

3.3.1 Inferovatelné vztahy(IR)

Vztahy mezi koncepty c_i a c_j můžeme považovat za IR jestliže platí $c_i => c_j$. IR vztahy lze dále rozdělit na slabé a silné vztahy.

IR je slabé, $c_i =>_w c_j$, jestliže instance z c_i nemohou odhalit všechny vlastnosti z c_j, jinak je IR silné $c_i =>_s c_j$

IR mohou obsahovat následující vztahy:

1) Ekvivalenci mezi dvěma koncepty: $c_i \equiv c_j$, pak $c_i =>_s c_j$ a naopak $c_j =>_s c_i$

2) Vztah mezi konceptem a jeho součástí, jestliže $c_j \in \{c_i\}$ pak $c_i =>_s c_j$

3) Vztah podtřída nebo nadtřída konceptu (subclass, superclass). Jestliže c_i □ c_j pak $c_i =>_s c_j$
Dále obsahuje:

vztahy sub-konceptu k sjednocení konceptů
jestliže $c_i = c_1 \cup c_2 \cup \ldots \cup c_k$ pak $c_i =>_s c_j$, $i = 1, \ldots, k$

vztah průniku konceptu k jakémukoliv z přesahujících konceptů
jestliže $c_i = c_1 \cap c_2 \cap \ldots \cap c_k$ pak $c_i =>_s c_j$, $j = 1, \ldots, k$

Příklad(silné): ‘nukleární střela‘ je průnikem ‘nukleární zbraně‘ a ‘střely‘ . Instance konceptu ‘nukleární zbraně‘ a ‘střely‘ mohou být inferovány z instance ‘nukleární střely‘.

Příklad(slabé): Jestliže instance subclass konceptu může být inferována z instancí superclass konceptu.

3.3.2 Částečně inferovatelné vztahy(PIR)

Vztah mezi koncepty c_i a c_j je PIR:
právě když $c_i \wedge c_{k1} \wedge \ldots \wedge c_{kn} => c_j$, kde $k_1, \ldots k_n \neq i, j$ a $c_{k1}, \ldots, c_{kn} \neq \varnothing$

PIR obsahují vztahy například z každého z přesahujících konceptu k průnikovému konceptu, to znamená

jestliže $c_i = c_1 \cap c_2 \cap \ldots \cap c_k$ pak $c_1 \wedge c_2 \wedge \ldots \wedge c_k => c_i$

3.3.3 Ne-inferovatelné vztahy(NIR)

Koncepty c_i a c_j jsou ve vztahu NIR:
jestliže $c_i \wedge c_{k1} \wedge \ldots \wedge c_{kn}$!=> c_j a $c_j \wedge c_{k1} \wedge \ldots \wedge c_{kn}$!=> c_i, kde $k_1, \ldots k_n \neq i, j$

Příkladem NIR je třeba vztah konceptu k jeho komplementu. V obrázku námi uvedené ontologie se jedná o vztah mezi koncepty ‘konvenční zbraň‘ a ‘speciální zbraň‘, jelikož neobsahují žádné společné instance.

3.4 Model řízení přístupu v SW

Tento model představuje pouze několik základních vlastností, které by šly do budoucna určitě dále rozšířit. Je založen na přístupu založeném na využití ontologií jehož základy jsou popsány v kapitole 2.3.

Hlavním myšlenkou je explicitní specifikace několika přístupových oprávnění v ontologii. Přístupová oprávnění k ostatním konceptům se vygenerují pomocí propagace, založené na vztazích mezi jednotlivými koncepty. Na tomto základě pak aplikace, která přistupuje k RDF datům zobrazí subjektu jen data, které jsou pro něj v ontologii povolené.

3.4.1 Přístupová oprávnění v SW

Přístupová oprávnění nám specifikují zdali daný subjekt může provádět určité akce na konkrétním objektu. Je definováno jak čtveřice:

ca = {obj, a, s, sub}, kde

obj	může být buď ontologie nebo koncept nebo množina konceptů v ontologii, identifikovatelná pomocí URI ontologie rozšířené o path expression výraz (např. #fuel)
a	privilegium – select, update, create, delete
s	sign ∈{+, -} kladný pro povolení, záporný pro zákaz
sub	je subjekt, kterému je oprávnění přiděleno

Například: Subjekt Roman má select oprávnění ke konceptu ‘speciální zbraň’ v ontologii ….
ca = {<http://x.edu/ontology/weapon.owl#special_weapon>, select, +, Roman}

Tento typ oprávnění je specifikován explicitně security administrátorem. Značí se jako *explicitní autorizační báze*. Oproti tomu existuje další přístup, který spočívá v odvozování a šíření už existujících pravidel na základě vztahů mezi koncepty. Tento přístup se nazývá propagace.

3.4.2 Propagace

Propagační autorizační báze – množina všech oprávnění odvozených pomocí propagace z explicitně daných nebo existujících přístupových oprávnění.

Základní symbolika:

→	udává směr propagace pokud je povolena a použita
!→	není možné použít propagaci

Zdrojový koncept – koncept ve kterém začíná propagace

Cílový koncept – koncept ve kterém propagace končí

negativní propagace – brání jakémukoliv možnému přímému nebo nepřímému neoprávněnému přístupu

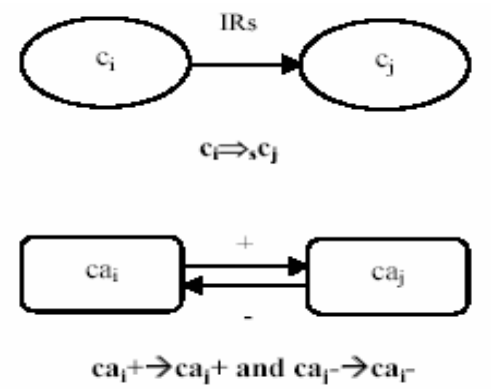
pozitivní propagace – zajistí že subjekty mají přístup pouze k informacím, ke kterým jsou oprávněny

3.4.2.1 Propagační politika pro IR

Není nutné kontrolovat autorizaci žádného jiného konceptu než je zdrojový koncept. Propagační politika mezi IR závisí, ale také na tom zda je vztah mezi koncepty silný nebo slabý.

Pokud IR z konceptu c_i do c_j je **silné** $c_i \Rightarrow_s c_j$ pak

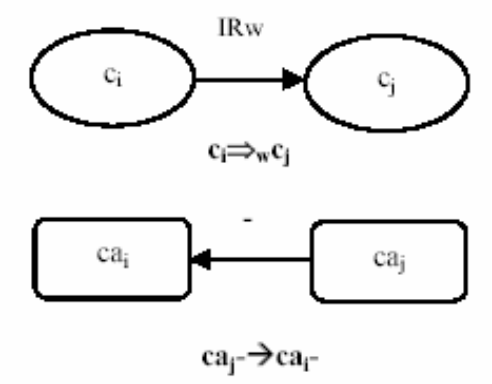
- Máme dáno přístupové oprávnění $ca_i = \{c_i, a, +, sub\}$ pak ca_i může být propagováno z konceptu c_i do c_j. Výsledkem bude odvozené nové oprávnění $ca_j = \{c_j, a, +, sub\}$ Toto propagační pravidlo značíme takto: $ca_i + \rightarrow ca_j +$
- Máme dáno přístupové oprávnění $ca_j = \{c_j, a, -, sub\}$ pak ca_j může být propagováno z konceptu c_j do c_i. Výsledkem bude odvozené nové oprávnění $ca_i = \{c_i, a, -, sub\}$ Toto propagační pravidlo značíme takto: $ca_j - \rightarrow ca_i -$



V horním obrázku je znázorněna silná IR vazba mezi koncepty, zatímco dolní nám ukazuje proveditelná propagační pravidla. Podobně je to znázorněno v dalším obrázku.

Pokud se jedná o **slabé** IR, $c_i \Rightarrow_w c_j$, pak existuje pouze jeden směr propagace

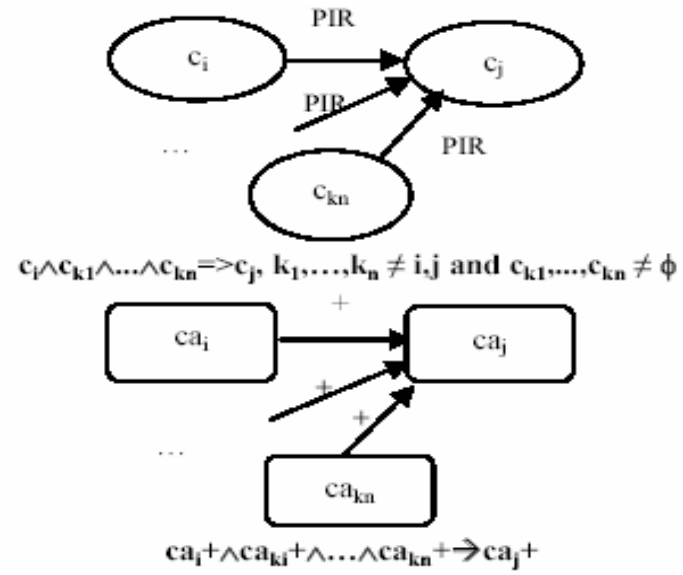
- Máme dáno přístupové oprávnění $ca_j = \{c_j, a, -, sub\}$ pak ca_j může být propagováno z konceptu c_j do c_i. Výsledkem bude odvozené nové oprávnění $ca_i = \{c_i, a, -, sub\}$ Toto propagační pravidlo značíme takto: $ca_j - \rightarrow ca_i -$



3.4.2.2 Propagační politika pro PIR

Oproti IR zde existuje zásadní rozdíl. Propagace záleží na autorizaci jak konceptu c_{kn} tak c_i.

- Máme dána přístupová oprávnění $ca_i = \{c_i, a, +, sub\}$ a $ca_{ki} = \{c_{ki}, a, +, sub\}$, kde i = 1, …, n pak ca_i a ca_{ki} může být propagováno z konceptu c_i a c_{ki} do c_j. Výsledkem bude nově odvozené oprávnění $ca_j = \{c_j, a, +, sub\}$ Toto propagační pravidlo značíme takto: $ca_i + \wedge ca_{ki} + \wedge \ldots \wedge ca_{kn} + \rightarrow ca_j +$. Graficky je znázorněno na následujícím obrázku.



3.4.2.3 Propagační politika pro NIR

Pokud jsou koncepty c_i a c_j ve vztahu NIR. To znamená že

$c_i \wedge c_{k1} \wedge \ldots \wedge c_{kn}$!=> c_j a $c_j \wedge c_{k1} \wedge \ldots \wedge c_{kn}$!=> c_i, kde $k_1, \ldots k_n \neq i, j$

Pak je celkem zřejmé že nelze provést inferenci. Z toho vyplývá že pak nelze provést ani propagaci.

3.5 Náměty na další výzkum

- Zajímavé by bylo prozkoumat možnosti jak integrovat způsob ontologického modelu řízení přístupu přímo do ontologického jazyka, pomocí rozšíření jeho slovníku.

- Kombinace tohoto modelu s již zavedenými a používanými modely v oblasti klasických nesémantických systémů. Zajímavá by mohla být kombinace s Role Based Access Control Model, používaným v OODBMS.

- Prostudování všech existujících a používaných ontologických jazyků a objektových modelů k identifikaci nových univerzálních doménově nezávislých vztahů – relací mezi koncepty.

- Studium možných konfliktů a jejich řešení, které mohou vzniknout při složitějších situacích během propagace.

- Sestavení ontologie pro privilegia (select, update, create, delete). V případě že mám nejvyšší privilegium automaticky získávám ostatní práva….

- Vyřešení problému důvěryhodnosti v RDF data, ontologie v sémantickém webu.

- Identifikace uživatele v rámci sémantického webu. Je zapotřebí, aby se uživatel někde autorizoval svým (loginem, IP, …) a na základě vztahů k jeho autorizované osobě se v rámci SW identifikují jeho práva.

4 Řízení přístupu a integrita dat pro DAML+OIL 2002

4.1 Bezpečnostní ontologie pro DAML+OIL

Jednou ze specifických aplikací sémantického webu je specifikace přístupových omezení k webovým stránkám a integrita obsahu webových stránek. K tomuto účelu vytvoříme specializovanou bezpečnostní ontologii. Tato definovaná ontologie bude obsahovat notaci pro značení webových zdrojů s přihlédnutím k řízení přístupu a integritě dat. Musíme vzít v úvahu následující fakta.

- K webovému zdroji může být omezen přístup. Aby k němu určitý subjekt získal přístup musí vyžadující agent nebo uživatel úspěšně projít autorizačním testem(využití tradičních autentizačních technik).
- Data mohou být neúmyslně nebo záměrně modifikována. Zde je možné použít standardních kryptografických technik, které zajišťují integritu uložených nebo přijatých dat.

Definujeme bezpečnostní ontologii v jazyce DAML+OIL, která umožní anotovat webové zdroje s přihlédnutím k základním principům řízení přístupu a integrity dat. Tato ontologie zasahuje do současného standardu pro XML signature syntax.

4.1.1 Řízení přístupu

Proces pro určení a ověření identity(autentizace) vyžadující strany ve webové aplikaci je často základem rozhodnutí zdali je přístup povolen nebo ne. Ověření identity je založeno na tokenu, který se také nazývá credential, který vyžadující strana má nebo zná(login-password, veřejné a soukromé klíče, certifikáty). Našim cílem je specifikovat omezení přístupu k webovým stránkám nebo službám, které využívají autentizaci jako podmínku pro autorizovaný přístup.

Nejjednodušším řešením jak vytvořit bezpečnostní ontologii by bylo definovat tyto tři způsoby k řízení přístupu(přístup bez oprávnění, přístup pomocí loginu, přístup pomocí certifikátu).

```
<AccessControl rdf:ID="None">
  <requiredCredential>NoCredential</requiredCredential>
</AccessControl>

<AccessControl rdf:ID="AccessByLogin">
  <requiredCredential>Login</requiredCredential>
</AccessControl>

<AccessControl rdf:ID="AccessByX509v3Certificate">
  <requiredCredential>X509v3Certificate</requiredCredential>
</AccessControl>
```

Není však možné jen definovat několik základních tříd v naší bezpečnostní ontologii pro řízení přístupu. Námi definovaná ontologie by měla obsahovat hierarchickou strukturu tříd oprávnění. To znamená, že na nejvyšší úrovni je třída *Credential*, která má několik podtříd *NoCredential, Certificate, Key*, které mohou mít další podtřídy, jenž specifikují další konkrétní nebo obecné typy (např. X509 certifikát).

Tato struktura nám v budoucnosti umožní rozšíření o nové druhy oprávnění(credential) založených na nových technologiích jako hlas nebo biometrické údaje(otisky prstů, sítnice, …).

Nejdůležitější třídou je třída *AccessControl*, která obsahuje minimálně jeden specifikovaný requiredCredential použitý pro autentizaci. Jelikož třída *Credential* obsahuje i „NoCredential“ pak umožňuje anotovat webové zdroje tak, aby neměly žádná přístupová omezení. Pokud je nadefinováno více použitelných credential, pak je umožněno zvolit si jeden z nich pro úspěšný přístup k webovým zdrojům.

```
<daml:Class rdf:ID="AccessControl">
  <rdfs:subClassOf>
    <daml:Restriction daml:minCardinality="1">
      <daml:onProperty rdf:resource="#requiredCredential"/>
    </daml:Restriction>
  </rdfs:subClassOf>
</daml:Class>
```

Některé credential třídy také přicházejí se speciálními vlastnostmi. Na příklad, login credential vždy vyžaduje uživatelské jméno a heslo. Proto je nutné je dodefinovat dle následující ukázky.

```
<daml:DatatypeProperty rdf:ID="passphrase">
  <rdfs:comment>
    passphrase is a DataTypeProperty whose range is xsd:string.
  </rdfs:comment>
  <rdf:range rdf:resource="http://www.w3.org/2000/10/XMLSchema#string"/>
</daml:DatatypeProperty>
```

Obdobně by byla nadefinována vlastnost uživatelské jméno „name“.

Pak bude credential třída pro login vypadat následovně.

```
<ClassOfCredentials rdf:ID="Login">
  <rdfs:comment>
    A login has exactly one user-name and one passphrase.
  </rdfs:comment>
  <rdfs:subClassOf rdf:resource="#Credential"/>
  <rdfs:subClassOf>
    <daml:Restriction daml:Cardinality="1">
      <daml:onProperty rdf:resource="#name"/>
    </daml:Restriction>
    <daml:Restriction daml:Cardinality="1">
      <daml:onProperty rdf:resource="#passphrase"/>
    </daml:Restriction>
  </rdfs:subClassOf>
</ClassOfCredentials>
```

Abychom mohli propojit dokument nebo jiný zdroj s definicí jeho bezpečnostní politiky(„s bezpečnostní ontologií“) je nutné definovat následující objektovou vlastnost *usesAccessControl* pro RDF metadata.

```
<rdf:ObjectProperty rdf:ID="usesAccessControl">
  <rdfs:comment>
  A Resource uses an access control technique, such as Login or Certificates.
  <rdfs:domain rdf:resource="rdfs:Resource"/>
  <rdfs:range rdf:resource="#AccessControl"/>
</rdf:ObjectProperty>
```

4.1.2 Využití XML Signature

Doposud jsme prezentovali elementy ontologie, které popisují řízení přístupových práv na velice abstraktní úrovni. Na názorném příkladu si ukážeme jak se situace může zkomplikovat.

Například webová stránka, která používá certifikáty akceptuje jen X509 verzi 3 certifikátů nebo jen akceptuje certifikáty vydané konkrétní certifikační autoritou(např. dále používané CA VeriSign). Takováto informace je velice užitečná pro směrování software agentů k webovým zdrojům, které jsou jim dostupné. Software agent, který vyhledává konkrétní webovou službu může být vybaven kolekcí certifikátů. Kdykoli pak agent narazí na službu, která splňuje jiné funkční aspekty požadované služby, může provést jednoduché výpočty aby vyvodil zda služba bude dostupná nebo ne.

Aby bylo možno použít takové detaily při popisu přístupových omezení bylo rozhodnuto využít již standardizovaných prostředků XML Signature Syntax a Processing Rules. Tyto standardy byly navrženy organizacemi IETF a W3C. Definují nám reprezentaci podpisu jakéhokoliv digitálního obsahu ve formátu XML. Konkrétně, XMK Signature syntax framework obsahuje odkazy na dobře známé kryptografické algoritmy a klíčové management systémy.

V následující ukázce je ukázáno rozšíření o XML signature syntax. Nejdříve je definována nová podtřída *XMLSignatureX509v3Certificate* třídy *Certificate* v bezpečnostní ontologii. Podobně by šly nadefinovat i jiné standardy jako PGP nebo SKPI veřejné klíče.

```
<ClassOfCredentials rdf:ID="XMLSignatureX509v3Certificate">
  <rdfs:subClassOf rdf:resource="#Certificate">
  <daml:Restriction daml:cardinality="1">
    <daml:onProperty rdf:resource="#associatedData"/>
    <daml:hasClass rdf:resource="http://www.w3.org/2000/09/xmldsig#X509Data"/>
  </daml:Restriction>
</ClassOfCredentials>

<!-- similar for PGPData or SPKIData elements -->
```

Teprve teď můžeme anotovat webovou stránku, která poskytuje přístup uživateli, která má X509v3 certifikát, podepsaný od VeriSign.

```
<AccessControl
rdf:ID="AccessByXMLSignatureX509v3CertificateIssuedByVerisign">
  <requiredCredential>
     <ClassOfCredentials rdf:resource="#XMLSignatureX509v3Certificate">
       <associatedData>
         <dsig:X509Data>
           <dsig:X509IssuerSerial>
             <dsig:X509IssuerName>VeriSign</dsig:X509IssuerName>
           </dsig:X509IssuerSerial>
         </dsig:X509Data>
       </associatedData>
     </ClassOfCredentials>
  </requiredCredential>
</AccessControl>
```

4.1.3 Integrita dat

Ve spoustě komerčních aplikací je nezbytně nutné vědět, že informaci kterou jsme získali můžeme opravdu věřit. To znamená, že vyžadující klient by rád získal jistotu, že obsah webové stránky, kterou získal je stejný, tak jak byla jejím autorem publikována. Je třeba zamezit jak chybě při přenosu tak úmyslné modifikaci dat. Existuje několik technik jak zajistit tuto integritu dat (checksums, MACs, digital signatures on hash values). Podobně jako tomu bylo v předchozí části je možné dodefinovat do této bezpečnostní ontologie tyto možnosti s využitím XML Signatures.

4.1.4 Webové služby

Pro popis sémantických webových služeb existuje speciálně definovaný jazyk DAML-S. Z tohoto pohledu je nutné odlišit přístup ke standardním webovým zdrojům a webovým službám. V definované bezpečnostní ontologii bude pro ně třeba nadefinovat vlastní třídy. Konstrukce bude principielně obdobná jako v předchozích případech.

5 Závěr

V této práci představuji základy několika modelů a motivaci pro řízení přístupu v prostředí sémantického webu. Dle mého názoru je řízení přístupu za pomocí ontologií unikátní přístup, který má s budoucím rozšířením sémantického webu obrovské možnosti. Tato oblast je minimálně prozkoumaná, většina výzkumů v oblasti řízení přístupu k webovému obsahu probíhá pouze na úrovni RDF[7], kde je možné vytvořit již finální funkční přístupy, které jsou však omezeny jen na oblast určité domény. V budoucnu bude množství dat na sémantickém webu exponenciálně narůstat a dojde k potřebě mít model, který bude použitelný na takové to obrovské množství dat.

Finální řešení tohoto problému řízení přístupu je zatím neřešitelné, do doby než budou standardizovány všechny důležité technologie, na kterých je řízení přístupu závislé (logic, proof, trust vrstvy). Tento nový zabezpečený sémantický web by pak měl umožnit uživatelům bezpečné využití inteligentních agentů a sémantických webových služeb.

Z hlediska dalšího výzkumu v této oblasti by mohla být zajímavá kombinace zde popsaných modelů, který by umožnil automatickou propagaci přístupových práv.

Literatura

[1] RDF, http://www.w3.org/RDF/

[2] Semantic Web, http://www.w3.org/2001/sw/

[3] Petrucha, R. .: Semantic Web, Proceedings of the 10th Conference STUDENT EEICT 2003.

[4] The Semantic Web Community Portal, http://www.semanticweb.org/

[5] Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. Scientific American, 2001.

[6] Svátek V.: Ontologie a WWW, DATAKON, 2002

[7] Semantic Web Trust and Security Resource Guide, http://www.wiwiss.fu-berlin.de/suhl/bizer/SWTSGuide/

[8] Qin, L., Alturi, V.: Concept-level access control for the Semantic Web, Proceedings of the 2003 ACM workshop on XML security, 2003.

[9] An Introduction To Role-Based Access Control, http://csrc.nist.gov/rbac/NIST-ITL-RBAC-bulletin.html

[10] Personal Data Access Language, http://www.w3.org/2002/01/pedal/thesis.html

[11] Web-Ontology (WebOnt) Working Group, http://www.w3.org/2001/sw/WebOnt/

[12] Jena2, http://jena.sourceforge.net/

[13] Sesame RDF Framework, http://aduna.biz/products/sesame/index.html