Sdílení dat v prostředí s nehomogenními skupinami uživatelů*

Roman Špánek[1,2] and Miroslav Tůma[3]

[1] spanek@cs.cas.cz
Ústav informatiky Akademie věd ČR,
Pod Vodárenskou věží 2, 182 07 Praha 8, Česká Republika
WWW home page: http:www.cs.cas.cz/ spanek/
[2] roman.spanek@vslib.cz
Technická univerzita v Libereci, Hálkova 6, Česká Republika
[3] Ústav informatiky Akademie věd ČR,
Pod Vodárenskou věží 2, 182 07 Praha 8, Česká Republika
WWW home page: http:www.cs.cas.cz/ tuma/

Abstrakt *Článek představuje možné řešení problému bezpečnosti v různých prostředích, která kladou důraz na sdílení zdrojů mezi jednotlivci nebo skupinami uživatelů. Mezi taková prostředí můžeme například zařadit mobilní telekomunikace a sebou nesoucí pojem mobilních databází, superpočítačů tvořených na bázi gridů, peer-to-peer sítí, vizi sémantického webu a v neposlední řadě i technologie počítačových agentů. Všechna tato prostředí mají svá specifika, ale také mají řadu společných jmenovatelů. Naše řešení je založeno na využití virtuálních organizací, které lze definovat jako dynamicky vytvářené skupiny uživatelů a organizací sdílející přístup k počítačům, softwaru, datům a ostatním zdrojům s přesným řízením přístupu a jasnou definicí co, s kým a za jakých podmínek je sdíleno. Náš přístup využívá silného matematického aparátu hypergrafů. Vzhledem s různorodosti a rozsahu cílových prostředí je naším cílem návrh bezpečnostního modelu, který bude mít schopnost samostatného vývoje, bez toho, že by struktura virtuální organizace degenerovala. Cílem článku je podat přesný rozbor jednoho z hlavních problémů v decentralizovaných prostředích a to nalezení konsenzu mezi uživateli virtuální organizace, který je v našem případě představován volbou vedoucího člena virtuální organizace.*

1 Úvod

S příchodem nových technologií umožňující připojení uživatelů k počítačové síti prakticky kdekoli a kdykoli, společně se souvisejícím nárůstem počtu uživatelů, vyvstala i nutnost řešit otázky zabezpečení. Jedním z možných řešení je aplikace velmi odolných šifrovacích algoritmů. Tyto algoritmy nicméně řeší pouze zabezpečení komunikace. Proto je také nutné řešit otázky důvěry mezi skupinami, případně jednotlivými uživateli. Jedno z možných řešení je využití bezpečnostních modelů navržených pro prostředí Virtuálních organizací. Virtuální organizace (VO) jsou dynamicky vytvářené skupiny uživatelů a organizací sdílející přístup k počítačům, softwaru, datům a ostatním zdrojům s přesným řízením přístupu a jasnou definicí co je sdíleno, kým je sdíleno a za jakých podmínek je sdíleno. Model VO je využíván v peer-to-peer sítích, mobilních databázích, sémantickém webu a v neposlední řadě i superpočítačích vytvořených na bázi Gridů, pro které byl model původně navržen. Široké spektrum aplikací poukazuje na použitelnost takového řešení.

Náš příspěvek navazuje na bezpečnostní model navržený v [2], který klade důraz na možnost automatického vývoje a správy virtuální organizace. Takový přístup je velmi vhodný v prostředích, kde počet uživatelů může být velký a navíc mohou být nehomogenní. Jako příklad lze uvést počítačové agenty v prostředí ah hoc sítí nebo sémantického webu. V takovýchto prostředích je nutné mít dostatečně robustní řešení pro správu uživatelů, které bude pracovat co možná nejvíce samostatně bez toho, aby organizace degenerovala nebo ztrácela vlastnost důvěryhodnosti. Degenerací bude myslet vývoj takové VO k jednomu z limitních stavů:

1. jedné VO obsahující všechny uživatele
2. mnoha velmi malých VO

Postup v [2] kombinuje velmi silný matematický model založený na hypergrafech s vhodnou implementací umožňující nasazení v distribuovaném prostředí. Pro ověření navržených postupů byla napsána experimentální aplikace SECGRID v jazyce ANSI C.

Zbytek příspěvku je organizován: odstavec 2 shrnuje současný stav problematiky bezpečnosti v prostředí virtuálních organizací. Naše konkrétní implemetace je popsána v odstavci 3.1 a volba vedoucího skupiny je v odstavci 3.2. Příspěvek je shrnut závěrem.

* Práce byla podpořena projektem 1ET100300419 programu Informační společnost (Tématického programu II Národního programu výzkumu v ČR: Inteligentní modely, algoritmy, metody a nástroje pro vytváření sémantického webu) a výzkumným záměrem AV0Z10300504 "Informatika pro informační společnost: Modely, algoritmy, aplikace". Práce byla podpořena výkumným centrem: Pokročilé sanační technologie a procesy 1M4674788502, Ministerstva Ministerstvo školství, mládeže a tělovýchovy České Republiky.

2 Současný stav problematiky

Pojem Virtuálních organizací [3] byl zaveden v prostředí Gridů [4]. Gridy jsou rozsáhlé distribuované systémy, tvořené heterogenními výpočetními, datovými a informačními zdroji, propojenými počítačovou sítí, tak aby tyto mohly být využívány jako řešení velmi výpočetně nebo prostorově náročných problémů. Takto propojené zdroje mohou být, a také často jsou, alokovány i velmi daleko od sebe. Velká vzdálenost a také různorodost propojených zdrojů, to jsou hlavní rozdíly mezi gridy a Clustery. V případě gridů je navíc velmi komplikovanou otázkou vyřešení správy přístupu jednotlivých uživatelů. Vzhledem ke geografické různorodosti zdrojů gridu, je velmi moudré předpokládat i stejnou různorodost v případě uživatelů. Tato různorodost bude velmi komplikovat řešení oprávnění přístupu ke gridu, zejména tím, že různé organizace zapojené do gridu mohou mít různá řešení vlastního zabezpečení, různá nastavení přístupových práv a zejména různé způsoby ověřování vlastních uživatelů. Na druhou stranu je nutné, aby uživatel nebyl nucen zadávat stále hesla, případně další osobní data, při připojení k jinému zdroji. Jako další požadavek lze vysledovat možnost delegovat část, případně všechna uživatelova práva na třetí subjekt, tak aby tento mohl provádět úkoly svěřené uživatelem a tedy měl i přístup ke zdrojům na základě uživatelových práv.

Jako jedno z vhodných řešení se ukazuje vytvoření virtuálních organizací. Virtuální organizace je v mnohých aspektech velmi podobná skutečným organizacím. Jedním z hlavní důvodů vytváření VO je poskytovat prostředky pro správu a vytváření důvěry mezi jejími členy. Postupy pro vytváření důvěry v takovém prostředí lze rozdělit na *Policy based* a *Reputation based* přístupy.

Policy based přístup byl navržen pro distribuované architektury služeb [5],[6],[7],[8],[9] a také v kontextu s gridy [10], jako řešení problému autorizace a řízení přístupu. Motivací takového přístupu je vytvořit systém pravidel a postupů pro vytváření a rozhodování o důvěře jednotlivých uživatelů. K tomuto cíli je využíváno jazyků s dobře definovanou sémantikou. Rozhodnutí o důvěře se pak provádí na základě nepřímých atributů uživatele (např. adresa nebo věk), které jsou certifikovány důvěryhodnou třetí stranou.

Reputation based postupy jsou velmi vhodné pro prostředí elektronických komerčních systémů (např. eBay), v peer-to-peer systémech, mobilních databázích a poslední době i pro prostředí sémantického webu [11],[12]. Charakteristikou takového přístupu je odvozování důvěry uživatele na základě jeho chování v minulosti. Důvěra je tedy založena na doporučeních a zkušenostech ostatních členů skupiny [13],[14],[15],[16].

Společným jmenovatelem všech výše zmíněných postupů je skutečnost, že uživatelé jsou do VO vloženy jistou autoritou (např. administrátorem). Toto řešení však nemusí být nejvhodnější v případě, že vezmeme v potaz prostředí s velkým počtem různorodých uživatelů (typicky sémantický web nebo mobilní databáze). Vezmeme-li v potaz takováto prostředí, je vhodné mít nástroj pro automatické vytváření a správu VO.

3 SecGRID

Úkolem SecGRID je umožnit automatické vytváření a správu VO v prostředích s velkým počtem nehomogenních uživatelů. Model SecGRID je založen na matematickém aparátu hypergrafů, který mu poskytuje dostatečně silné protředky pro jeho realizaci a ověření. VO je v SecGRID reprezentována jako ohodnocená hypergrafová struktura. Vztahy mezi členy jsou reprezentovány pomocí váhy ohodnocené hyperhrany. Vyšší ohodnocení hrany implikuje vyšší důvěru mezi členy. Uzly reprezentují jednotlivé uživatele. Ohodnocení uzlu reprezentuje jeho důvěryhodnost, dostupné výpočetní a komunikační prostředky.

3.1 Implementace

Struktura VO v SecGRID je hierarchická. Je tvořena libovolným počtem menších skupin uživatelů, které se dále dělí na menší organizační jednotky. Vzhledem ke skutečnosti, že implementace nerozlišuje mezi VO a jejími dílčími organizacemi, budeme dále používat jen termínu VO. Spodní vrstva hierarchie je tvořena vlastními členy VO. Každá VO má zvoleného vedoucího skupiny, tzv. VO Leader (VOL). Nad touto vrstvou uživatelů je jedna nebo více vrstev tvořených pouze VOL. VOL odpovídá za podřízené jednotky a umožňuje komunikaci mezi jednotlivými VO. Tato struktura zvyšuje důvěryhodnost a bezpečnost SecGRID, neboť právě komunikace členů z jedné VO do jiné představuje největší bezpečnostní riziko pro ostatní členy. Tím že komunikace mezi skupinami je kontrolována VOL je zaručeno, že nedojde k úniku citlivých informací.

Pro potřeby experimentální aplikace byla použito hypergrafů s mohutností incidence hyperhran dvě. Rozšíření aplikace na plnou kardinalitu hyperhran je evidentní.

Vzhledem k požadavku automatického vytváření a správy VO bylo nutné navrhnout řadu pravidel pro ohodnocování hran. Byly vysledovány tři základní varianty, které mohou při vývoji grafové struktury nastat:

1. přidání tranzitivní hrany,
2. vytvoření neorientovaného cyklu,
3. vytvoření orientovaného cyklu.

Nejdůležitější z nich je vytvoření nového orientovaného cyklu (souvislé komponenty v grafu). Tento případ je reprezentován jako vznik uzavřené skupiny uživatelů, kde lze komunikovat mezi všemi členy. Proto je taková souvislá komponenta vhodným kandidátem na vytvoření nové VO. Nová VO je vytvořena na základě ohodnocení hran v komponentě. Dojde-li k vytvoření nové VO, je nutné pro ni zvolit nového vedoucího skupiny (VOL). Ostatní přípustné varianty nejsou pro příspěvek zajímavé, neboť nevyžadují vytvoření nové skupiny (VO) a tedy nevyžadují volbu VOL, která je hlavním tématem příspěvku. Podrobný popis ostatních případů, včetně všech ohodnocovacích pravidel lze nalézt v [2].

3.2 Volba VOL

Z předchozího odstavce je patrné, jak důležitou roli hrají VOL. Z toho důvodu je nutné mít vhodně vyřešené volení VOL ze členů skupiny a to tak, aby bylo splněno následující:

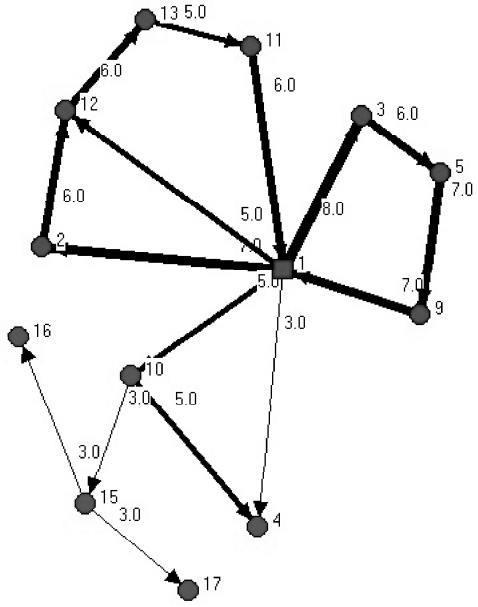
1. nový VOL musí bezpodmínečně být velmi důvěryhodným členem skupiny
2. volba VOL nesmí příliš zatěžovat členy VO
3. volba VOL musí být implementovatelná v distribuovaném prostředí

Choulostivá operace volby nového VOL je řešena v SecGRID pomocí následující procedury:

- jako první VOL je zvolen zakládající člen VO
- v případě nutnosti zvolit nového VOL je do skupiny vyslána RESIGN zpráva původním VOL
- RESIGN zpráva je odeslána sousedovi s nejlepším vztahem (po hraně s největším ohodnocením)
- při přijetí RESIGN zprávy se příjemce rozhodne, zda-li bude novým VOL
- pokud ano, oznámí to skupině pomocí NEWVOLARRIVES zprávy
- v opačném případě předá RESIGN zprávu opět svému sousedu s nímž má nejlepší vztah
- v momentě, kdy kterýkoli člen skupiny, mimo odcházejícího VOL, obdrží RESIGN zprávu podruhé, je tento automaticky zvolen novým VOL

Procedura pro volbu nového VOL má všechny požadované vlastnosti, viz. požadavky výše.

- ad 1. Důvěryhodnost nového VOL je zaručena, neboť k jeho volbě jsou přizváni pouze členové *základní skupiny* (viz. níže) uživatelů
- ad 2. volba nového VOL není náročnou operací, neboť mimo starého VOL žádný ze členů základní skupiny uživatelů nemusí přeposlat RESIGN zprávu vícekrát než jednou. Počet členů základní skupiny je menší než počet všech členů

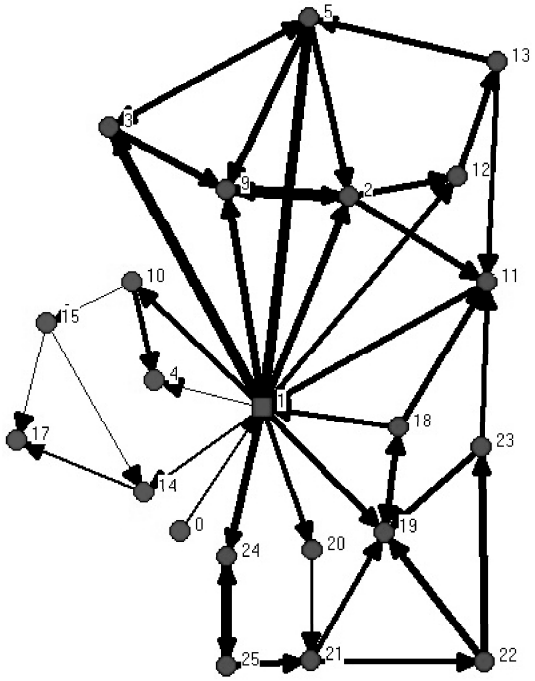


Obrázek 1. Výchozí stav VO.

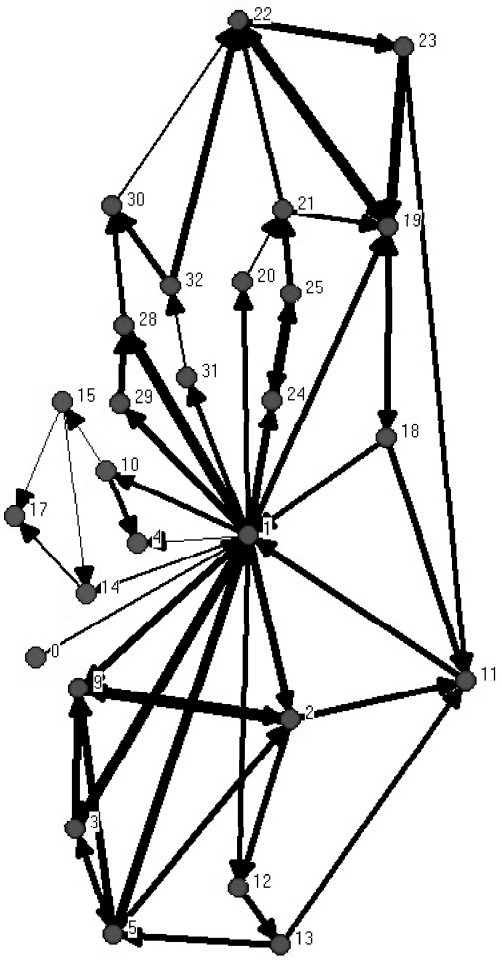
- ad 3. celá procedura využívá systém zpráv, který lze přímo využít v distribuovaném prostředí.

Pojem *základní skupiny* uživatelů poukazuje na skutečnost, že při vytváření struktury VO dochází k vytvoření ustálené skupiny důvěryhodných uživatelů. Na obrázku 1 je stav nově vytvořené VO. Členové VO jsou zobrazeny modře a VOL je naznačen čtverečkem. Síla hran odpovídá ohodnocení. Z obrázku je patrné, že základní skupina je tvořena členy 1,3,5,9. Druhá dobře profilovaná skupina nemá takovou důvěryhodnost. Pokud by došlo k volbě v této konfiguraci, byl by nový VOL zvolen právě ze členů 3,5,9. Uvažujme situaci, kdy byl jako nový VOL zvolen člen 5. Pokud po nějaké době došlo opět k volbě nového VOL, byl by znovu volen pouze ze členů 1,3,9. Je tedy zřejmé, že možní kandidáti na VOL jsou alokováni pouze mezi členy základní skupiny uživatelů, která má menší počet členů, než celá organizace. Například počet členů skupiny na obrázku 1 je osm, nicméně nově volený VOL bude volen pouze ze skupiny tří členů.

Situace po přidání nových členů a provedení přehodnocení je na obrázku 2. Z obrázku je patrné, že nedošlo k dramatickému zvětšení základní skupiny uživatelů. Přestože značně narostl jak počet hran, tak i počet vrcholů, základní skupina uživatelů se rozrostla pouze o jednoho člena na 1,3,5,9,2. Vezmeme-li v po-



Obrázek 2. Stav VO po přidaní hran a uzlů.



Obrázek 3. Vznik *hnízd* ve struktuře VO.

rovnání stávající počet členů, který s zdvojnásobil na šestnáct a počet členů základní skupiny, je tento poměr 16/5. Přitom při původní konfiguraci na obrázku 1 byl tento poměr 8/4. Z toho je jasně patrné, že základní skupina podléhá pomalejšímu růstu. Tento poměr, jak ukazují naše simulace, se bude s přibývajícím počtem členů dále zvětšovat.

Tuto skutečnost lze vysvětlit vznikem izolovaných *hnízd*. Jako hnízdo budeme myslet vznik skupiny uživatelů, mající k sobě navzájem velmi dobré vztahy a navíc tvořící souvislou komponentu s vysokým ohodnocením hran. Na obrázku 3 je taková struktura dobře patrná mezi členy 19, 22 a 23. Simulace ukazují, že s přibývajícím počtem členů takovýchto struktur uvnitř VO přibývá a navíc bývají alokována dále od VOL (vzhledem k délce orientované cesty v grafu). Skutečnost, že hnízda jsou izolována a nejsou v blízkosti VOL je snadné vysvětlit, neboť pokud by byla blízko VOL, stala by se součástí základní skupiny uživatelů. Vzhledem k patrné izolovanosti hnízd a k ohodnocení jeho hran je dobré se zamyslet, zda-li by nebylo lépe takové struktury úplně od VO oddělit a vytvořit z nich vlastní nové menší VO. Odpověď na tuto otázku nelze položit jednoznačně, neboť oddělení od zbytku VO by mělo za následek ztrátu spojení s dalšími členy skupiny a tedy izolovanost, která by ovšem mohla být na škodu členům a to jak hnízda tak i zbytku VO. Na druhou stranu je nutné podotknout, že jistá míra izolovanosti je již zachycena ve vlastním ohodnocení hran uvnitř hnízda. Sdílení informací je tedy daleko snazší mezi členy hnízda, než mezi zbytkem VO. Dalším silným argumentem pro nevytváření nové VO je skutečnost, že každá nově vytvořená organizace musí uchovávat informace o okolních strukturách (ve smyslu uložení dat o okolních VOL). Okolní skupiny pak musí ukládat informace o nově vznikajících skupinách. Pokud by tedy byly nové VO vytvářeny příliš rychle a s malým počtem členů, znamenalo by to značnou zátěž pro všechny zainteresované VOL. Druhou stranou mince je

skutečnost, že příliš velké VO se špatně udržují. Proto náš model počítá s rozdělením VO na dílčí menší v momentě, kdy bude splněna podmínka $d(k) > \delta$, kde $d(k)$ je průměr VO a δ je celé kladné číslo.

Dalším zajímavým aspektem procedury pro volbu nového VOL je skutečnost, že kandidátní VOL jsou alokovány vždy v blízkosti původních VOL. Tato skutečnost je důsledkem podmínky pro přeposlání RESIGN zprávy, aby příjemce měl s odesílatelem nejdůvěrnější vztah. Tedy aby hrana mezi odesílatelem a příjemcem měla nejvyšší ohodnocení a tedy byla mezi členy základní skupiny.

Algoritmus volby VOL zvolí nového VOL při složitosti $O(n)$, kde n je délka orientovaného cyklu v hypergrafu představujícího VO. Vzhledem ke skutečnosti, že vytvoření nové VO je podmíněno vznikem orientovaného cyklu je tedy zaručeno, že algoritmus vždy skončí zvolením nového VOL. Složitost je ve skutečnosti nižší vzhledem k vytvoření základní skupiny uživatelů.

4 Závěr

Cílem příspěvku bylo navrhnout a experimentálně ověřit metodu pro dosažení konsenzu mezi členy virtuálních organizací. Celý příspěvek je začleněn do širšího problému návrhu bezpečnostního modelu pro prostředí virtuálních organizací, který bude použitelný v prostředích majících velký počet různorodých uživatelů. V takovýchto prostředích je nutné, aby model měl schopnost samostatného vývoje. Jedním z klíčových momentů v životě VO je volba vedoucího člena (VOL), který je zodpovědný za její správu a také za komunikaci mezi ostatními skupinami. Proto byl navržen postup jak dosáhnou shody mezi členy VO bez toho, aby tato volba měla dopad na efektivitu a použitelnost našeho modelu v reálném prostředí. Hlavní limitující faktory jsou požadavky, aby algoritmus volby byl přímo implementovatelný v distribuovaném prostředí a zachovával důvěru mezi členy VO. Naše experimenty ukazují, že navržený postup volby splňuje všechny požadavky na něj kladené. Pro ověření korektnosti byla použita experimentální aplikace SecGRID. Na základě provedených experimentů, byla dále ukázána celá řada zajímavých momentů ve vývoji VO, které mají klíčový dopad především na reálné využití našeho modelu.

Reference

1. Clarke, F., Ekeland, I.: Nonlinear oscillations and boundary-value problems for Hamiltonian systems. Arch. Rat. Mech. Anal. **78** (1982) 315–333
2. Špánek,R., Tůma, M.: Secure Grid-based Computing with Social-Network Based Trust Management in the Semantic Web. submitted to NNW. (2006) 15str.
3. Foster, I., Kesselman, C. and S. Tuecke: The Anatomy of the Grid. Enabling Scalable Virtual Organizations. International Journal of Supercomputer Applications, 2001.
4. Foster,I., Kasselman, C.: The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufmann, 1999.
5. P. Bonatti and P. Samarati, "Regulating service access and information release on the web" in *CCS '00: Proceedings of the 7th ACM conference on computer and communications security*, pages 134–143. ACM Press, 2000.
6. N. Li and J. Mitchell, *A Role-based Trust-management Framework.*, In DARPA Information Survivability Conference and Exposition (DISCEX), Washington, D.C., Apr. 2003.
7. R. Gavriloaie, W. Nejdl, D. Olmedilla, K. E. Seamons, and M. Winslett, *No registration needed: How to use declarative policies and negotiation to access sensitive resources on the semantic web.*, In 1st European Semantic Web Symposium (ESWS 2004), volume 3053 of Lecture Notes in Computer Science, pages 342–356, Heraklion, Crete, Greece, may 2004. Springer.
8. M. Y. Becker and P. Sewell, *Cassandra: distributed access control policies with tunable expressiveness.*, In 5th IEEE International Workshop on Policies for Distributed Systems and Networks, Yorktown Heights, June 2004.
9. P. A. Bonatti and D. Olmedilla, *Driving and monitoring provisional trust negotiation with metapolicies.*, In 6th IEEE International Workshop on Policies for Distributed Systems and Networks (POLICY 2005), pages 14–23, Stockholm, Sweden, jun 2005. IEEE Computer Society.
10. J. Basney, W. Nejdl, D. Olmedilla, V. Welch, and M. Winslett, *Negotiating trust on the grid.*, In 2nd WWW Workshop on Semantics in P2P and Grid Computing, New York, USA, may 2004.
11. L. Kagal, T. Finin, and A. Joshi, *A policy based approach to security for the semantic web.*, In Proceedings of the 2nd International Semantic Web Conference, Sanibel Island, Florida, USA, Oct. 2003.
12. G. Tonti, J. M. Bradshaw, R. Jeffers, R. Montanari, N. Suri, and A. Uszok, *Semantic web languages for policy representation and reasoning: A comparison of KAoS, Rei and Ponder.*, In Proceedings of the 2nd International Semantic Web Conference, Sanibel Island, Florida, USA, Oct. 2003.
13. K. Aberer and Z. Despotovic, *Managing trust in a peer-2-peer information system.*, In Proceedings of 10th International Conference on Information and Knowledge Management, pages 310–317, 2001.
14. E. Damiani, S. D. C. di Vimercati, S. Paraboschi, P. Samarati, and F. Violante, *A reputation-based approach for choosing reliable resources in peer-to-peer networks.*, In Proceedings of ACM Conference on Computer and Communications Security, pages 202–216, 2002.
15. S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, *Eigenrep: Reputation management in p2p ne-*