Henk Birkholz, Carsten Elfers, Bernd Samjeske, Karsten Sohr

Unternehmensübergreifender Austausch von sicherheitsrelevantem Wissen

Koordinierte verteilte Angriffe im Internet sind ein akutes Problem. Während Angriffe jedoch in kollaborativen Verbünden organisiert werden, verteidigen Unternehmen sich häufig im Alleingang. Sicherheitsmaßnahmen gelten oft als wohl gehütete Geheimnisse und werden ungern preisgegeben. In diesem Beitrag werden Methoden und Voraussetzungen vorgestellt, die einen kooperativen Austausch von sicherheitsrelevantem Wissen ermöglichen, ohne damit etwaige eigene Schwachstellen offen zu legen.

Henk Birkholz

Technologie-Zentrum Informatik und Informationstechnik (TZI) an der Universität Bremen
E-Mail: birkholz@tzi.de

Carsten Elfers

Wissenschaftlicher Mitarbeiter, Technologie-Zentrum Informatik und Informationstechnik (TZI) an der Universität Bremen
E-Mail: elfers@tzi.de

Bernd Samjeske

E-Mail:

Dr. Karsten Sohr

wissenschaftlicher Mitarbeiter am Technologie-Zentrum Informatik und Informationstechnik (TZI) der Universität Bremen
E-Mail: sohr@tzi.de

1 Einleitung

Die Bedrohung durch Angriffe aus dem Internet steigt stetig. In den letzten Jahren erfolgten Angriffe vermehrt durch kollaborative Verbünde [1], z. B. in der Form von internationalen Botnetzen [2]. Während Angriffe durch solche Verbünde erfolgen, erfolgt eine Bekämpfung von Angriffen meist isoliert in einzelnen Unternehmen. In diesem Beitrag wird vorgestellt, wie in einem Netzwerk aus Sicherheitsexperten gemeinsam Angriffe in einem Verbund verhindert werden können. Für diesen Zweck ist der Austausch von sicherheitsrelevantem Wissen zwischen Unternehmen notwendig.

Unter sicherheitsrelevantem Wissen verstehen wir in Bezug auf Angriffe relevante Informationen. Diese Informationen werden für die Korrelation und die darüber hinaus gehende Zusammenführung und Aufbereitung von Sicherheitsvorfällen in entsprechenden *Security Information and Event Management* (SIEM) Systemen (z. B. ArcSight ESM) verwendet. Durch den unternehmensübergreifenden Austausch von Wissen reduziert sich zum einen der Aufwand, dieses zu gewinnen und zu sichern; zum anderen steigt die Reaktionsfähigkeit von Unternehmen in Bezug auf neue Angriffsmuster. Der Austausch von sicherheitsrelevantem Wissen berührt allerdings zwangsläufig immer vertrauliche Informationen der Unternehmen. Eine mögliche Lösung dieser Problematik ist die Bewahrung der Anonymität durch Abstraktion des geteilten Wissens. Dieser Lösungsansatz wird in diesem Beitrag diskutiert.

Um sicherheitsrelevantes Wissen nutzbar zu machen, muss dieses in den eingesetzten SIEM-Systemen integrierbar sein. Erste Versuche, dieses Wissen möglichst generisch mit Hilfe von Ontologien zu beschreiben und für ein SIEM nutzbar zu machen, sind im Projekt FIDeS bereits unternommen worden [3].

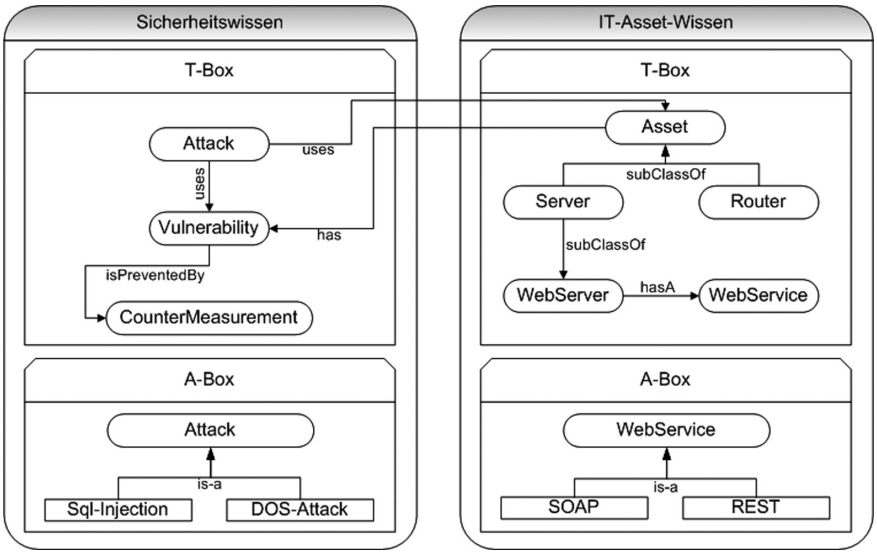
2 Ontologien und OWL

Eine Ontologie dient der strukturellen Abbildung von Entitäten (z. B. Webservices, Router) und der Beschreibung der Relationen zwischen diesen Entitäten (z. B. ‚ein Webserver ist ein Computer'). Die *Web Ontology Language* (OWL) [4] ist ein Formalismus zur Repräsentation dieses strukturellen Wissens. Sie wird genutzt, um Wissen in digitalisierter, maschinenlesbarer Form unter Anwendungsprogrammen auszutauschen.

Eine Ontologie ist separiert in zwei Bereiche. Der erste Bereich ist die T-Box, welche das terminologische Wissen beschreibt. Sie besteht aus den Konzepten (Classes) und deren Relationen (Properties) zueinander, die in diesem Fall die gewünschte Domäne Sicherheit bzw. IT-Assets repräsentiert (Abbildung 1). Diese bildet das Vokabular, durch das das Wissen über die Domäne beschrieben wird.

Der zweite Bereich ist die A-Box, die aus Individuen dieser Konzepte besteht. Die-

Abbildung 1 | Struktur der ausgetauschten Ontologie



ser Bereich repräsentiert die speziellste Form des Wissens und entspricht konkreten Instanzen der Konzepte aus der T-Box.

Die Konzepte und Relationen lassen sich in einer Baumstruktur von allgemeinen zu spezialisierten Konzepten und Relationen anordnen. OWL bietet dafür die Sprachelemente „subClassOf" für die Anordnung der Konzepte und „subPropertyOf" für die Anordnung der Relationen. OWL wird in XML gespeichert und basiert dabei auf dem *Resource Description Framework* (RDF) [5]. Das folgende Beispiel zeigt die Beschreibung zweier Konzepte *Attack*, und *DenialOfService* als Unterkonzept von *Attack* in *OWL*:

OWL Beispiel - Relationen

```
<owl:ObjectProperty rdf:ID="uses">
 <rdfs:domain rdf:resource="#Attack" />
 <rdfs:range rdf:resource="#Vulnerability" />
</owl:ObjectProperty>
```

Beziehungen zwischen Individuen lassen sich durch Relationen (*Object Properties*) repräsentieren. Das folgende OWL-Beispiel zeigt die Relation „uses", die zwischen Individuen des Konzeptes „Attack" und Individuen des Konzeptes „Vulnerability" bestehen kann:

OWL Beispiel - Konzepte

```
<owl:Class rdf:ID="Attack">
 <rdfs:subClassOf rdf:resource="#Thing" />
  ...
</owl:Class>
<owl:Class rdf:ID="DenialofService">
  <rdfs:subClassOf rdf:resource="#Attack" />
  ...
</owl:Class>
```

Dies modelliert die Ausnutzung einer Verwundbarkeit durch einen Angriff. In dieser Form können so heterogene sicherheitsrelevante Informationen zentralisiert zugänglich gemacht werden.

2.1 Wissensaustausch und dessen Nutzen

Sicherheitsrelevante Informationen können sinnvoll mit Hilfe von zwei Ontologien beschrieben und zwischen Unternehmen ausgetauscht werden (Abbildung 1):

- Die erste Ontologie umfasst IT-Asset-Wissen. Dieses besteht aus Informationen über die IT-Infrastruktur des Unternehmens. In diesen Bereich fallen u. a. die Inhalte klassischer Inventar-Datenbanken, Netz-Topologien, Rechte-Infrastrukturen in Form von Nutzer- und Rollenkonzepten oder auch *Configuration Management Databases* (CMDB).
- Die zweite Ontologie umfasst das eigentliche Sicherheitswissen. Dieses beinhaltet im Einzelnen das Wissen über die Struktur von Angriffen, Verwundbarkeiten, den Umsetzungsgrad von Sicherheitsmaßnahmen und Schutzbedarfsdefinitionen. Dieses Wissen basiert auch auf externen Quellen wie CERT-Meldungen oder strukturierten Verwundbarkeitsdatenbanken wie der öffentlich zugänglichen *National Vulnerability Database* (NVD) [6].

Sicherheitsrelevantes Wissen hingegen umfasst zudem jegliches Wissen, das mit Hilfe dieser beiden Ontologien beschrieben wird. Die Unterteilung der Wissensbeschreibung durch zwei Ontologien ist erforderlich, da jedes Unternehmen seine eigene IT-Infrastruktur aufnehmen und diese mit dem geteilten Wissen in Beziehung setzen muss. Regeln zur Angriffskorrelation in intelligenten SIEM-Systemen wiederum bauen auf dem jeweiligen Gesamtbild von IT-Asset- und Sicherheitswissen auf.

2.2 Ereignisverarbeitung in SIEM-Systemen

Das mit Hilfe von Ontologien ausgetauschte Wissen wird in SIEM-Systemen für die Verarbeitung von Echtzeit-Informationen verwendet. Unter Echtzeit-Informationen fallen einfache System-Ereignisse wie die Veränderungen von Verfügbarkeit und Systemgesundheit, der Erfolgsgrad von Authentifikationsversuchen und komplexe Ereignisse, die bereits eine aufwendige Vorverarbeitung erfahren haben. Komplexe Ereignisse sind beispielsweise erkannte Malware in Dateien oder böswillige Protokollverletzungen in Paketströmen, die mit Hilfe von Intrusion Detection Systemen (IDS) detektiert werden können.

In der SIEM-Domäne ist es verbreitet, Reaktionen auf bestimmte Echtzeitinformationen in Form von Regeln auszudrücken. Ontologien können zur Modellierung von SIEM-Regeln verwendet werden, indem sicherheitsrelevantes Wissen in Form von logischen Beziehungen zwischen Tripeln ausgedrückt wird. Jedes Tripel besteht aus zwei Entitäten (z. B. Konzepte oder Individuen) und einer Relation zwischen diesen.

Beispielsweise kann so beschrieben werden, dass eine Sicherheitslücke vorliegt, wenn bei einem Host eine bestimmte Version des Apache Webservers installiert und dieser vom Internet erreichbar ist. Regeln dieser Form können durch die Anfragesprache SPARQL [7] für Ontologien formuliert und dann durch einen ontologischen Schlussfolgerungsmechanismus (wie z. B. Pellet [8]) überprüft werden. Für die korrekte Schlussfolgerung ist es jedoch nötig, dass die dem SIEM-System vorliegenden Informationen korrekt und vollständig sind. In dem genannten Beispiel muss daher sichergestellt werden, dass die installierten Apache-Webserver als Individuen der unternehmenseigenen IT-Asset-Ontologie bekannt sind.

2.3 Automatische Erhebung von IT-Asset-Wissen

Die Erhebung spezifischer IT-Asset-Individuen kann durch den Einsatz automatischer Methoden flankiert werden. Zwar ist es häufig der Fall, dass eine automatische Erfassung der manuellen Überprüfung auf Korrektheit bedarf – in Teilbereichen ist eine vollautomatische Erfassung jedoch möglich.

Die Topologie und Konfiguration von Datennetzen ist dafür ein anschauliches Beispiel: Datennetze in Unternehmen zeichnen sich typischerweise durch eine strikte Trennung interner, externer und demilitarisierter Zonen (DMZ) aus. Die Adressen und Kommunikationseinschränkungen einzelner IT-Systeme, Dienste und Zonen sind klar definiert. Die konzeptuellen Strukturen von Unternehmensnetzen sind dabei vergleichbar und lassen sich so in einer generellen Netz-Ontologie (als Teil der IT-Asset-Ontologie) modellieren. Die individuelle Ausprägung dieses ontologischen Modells pro Unternehmen lässt sich dann automatisch erheben, wenn die Konfigurationen der involvierten Komponenten – Switches, Router und Firewalls – durch entsprechende Kollektoren zugreifbar sind oder die Konfiguration bereits mit Hilfe einer CMDB zentralisiert ist.

Nicht nur, dass eine gut gepflegte Netz-Ontologie als Teil der Infrastrukturinformation notwendig für die Bewertung von sicherheitsrelevantem Wissen ist, sie bietet darüber hinaus einen direkten Mehrwert bei der Gestaltung von Unternehmensnetzen. Inkonsistenzen in der Netz-Konfiguration können automatisch erkannt werden. Ein einfaches Beispiel dafür ist die Erreichbarkeit von Diensten über Zonengrenzen hinweg: Ist die Nutzung eines Dienstes für den Betrieb vorgesehen, wird dieser aber durch die aktuelle Konfiguration des Netzes unterbunden, so kann dies automatisch erkannt und gemeldet werden.

Sollte die Sicherheitsstrategie eines Unternehmens dies erlauben, ließen sich darüber hinaus auch Zugriffsbeschränkungen in Bezug auf einzelne Nutzer modellieren, wenn sie z. B. nach 802.1x auf Link-Layer-Ebene implementiert sind. Auch die Ergebnisse einer regelmäßigen Überprüfung durch Schwachstellen-Scanner (wie z. B. Nessus) können automatisch in entsprechende Ontologie-Teile integriert werden.

Diese Zusatzinformationen sind kein Ersatz für die manuelle Pflege der IT-Asset-Informationen, sie stellen allerdings eine erhebliche Verringerung des Erhebungsaufwands dar und bieten einen entscheidenden Vorteil bei der zeitnahen Erkennung von Konfigurationsproblemen oder der Vorbeugung von Inkonsistenzen.

3 Austausch von sicherheitsrelevantem Wissen

Sicherheitsrelevantes Wissen über die Grenze eines Unternehmens hinweg zu publizieren ist bei vielen Unternehmen aufgrund des Aufzeigens von möglichen Sicherheitslücken undenkbar. Beim Austausch dieses Wissens ist es daher notwendig, einen Rückschluss auf das austauschende Unternehmen zu vermeiden und Anonymität zu gewährleisten.

Dies ist keine einfache Aufgabe, wenn das relevante Wissen vollständig erhalten bleiben soll, aber dennoch auf spezifische Informationen verzichtet werden muss, um einen Rückschluss auszuschließen. Die Verwendung einer Sicherheits- und IT-Asset-Ontologie kann dazu sehr nützlich sein. Durch die Ontologien können sehr spezifische Beschreibungen in dem auszutauschenden Wissen abstrahiert werden, z. B. durch die Ersetzung einer exakten Produktbezeichnung (eines Individuums) „CheckPoint UTM-1 Appliance 136 R71“, die ggf. Rückschlüsse auf bestimmte Unternehmen erlauben würde, durch das abstraktere Konzept „CheckPoint UTM-1 Appliance“. Die Erforderlichkeit zu einer solchen Abstraktion des Wissens kann z. B. mit Hilfe des k-ANONYMITY-Verfahrens überprüft werden [9]. Bei diesem Verfahren muss eine zu spezifizierende Mindestanzahl von Nutzern eine Beschreibung verwendet haben, um die Anonymität nicht zu gefährden. Beispielsweise muss die Firewall-Bezeichnung nicht abstrahiert werden, wenn Sie bei vielen Unternehmen in dieser Version bereits zur Beschreibung ähnlichen Wissens verwendet wurde.

Konsequenz einer solchen Abstraktion ist jedoch, dass im Falle des Austauschs von Angriffswissen die *false positive*-Rate bei der Erkennung von Angriffen kurzfristig steigt. Sobald unterschiedliche Quellen das ausgetauschte Angriffswissen bestätigen, kann die Abstraktion sukzessive wieder aufgehoben werden. Dies erfordert jedoch eine gewisse Redundanz der Modellierung in den Unternehmen, die allerdings geringer ausfällt als im Falle einer je Unternehmen getrennten Modellierung.

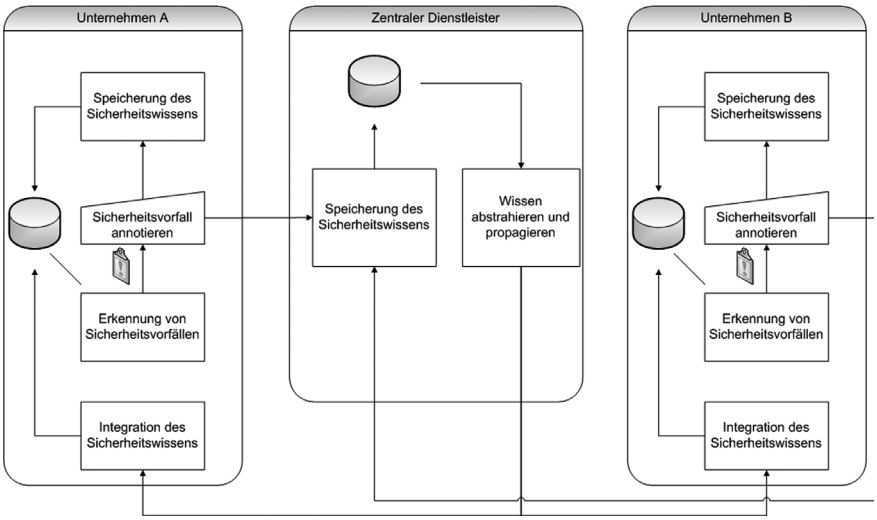
Das Aufheben einer Abstraktion kann mit der gleichen Bewertungsmethode erfolgen, die auch die Abstraktion auslöst. Dies erfordert jedoch eine große Beteiligung an dem Austausch von Wissen durch Teilnehmer, die eine ähnliche Sicherheitsinfrastruktur verwenden. Der Vorteil der Abstraktion liegt jedoch im frühzeitigen Austausch von sicherheitsrelevantem Wissen für die zeitnahe Reaktion auf Sicherheitslücken bei gleichzeitiger Gewährleistung der Anonymität des Verfassers.

Das Auftreten von falsch-positiven Meldungen aufgrund von abstraktem Wissen kann als erneuter Vorfall behandelt werden. Dieser wird durch den gleichen Prozess annotiert und mit der zentralen Wissensbasis ausgetauscht und dient damit der Spezialisierung des Wissens, das Ursache für das Auftreten dieser falsch-positiven Meldung ist. Dies führt zur mittelfristigen Vermeidung dieser Falschmeldungen für alle Unternehmen. Die beteiligten Unternehmen erweitern das Vokabular (T-Box), dessen Konkretisierungen (A-Box) und die sich darauf beziehenden Regeln, wodurch in einem iterativen Prozess eine gemeinsame Wissensbasis geschaffen wird.

Abhängig von der unternehmensspezifischen IT-Infrastruktur identifiziert und integriert ein SIEM-System relevantes propagiertes Wissen von der zentralen, gemeinsamen Wissensbasis. Dies Umfasst neben den Ontologien auch Wissen, das mit Hilfe des in den Ontologien beschriebenen Vokabulars erstellt wurde, insbesondere auch SIEM-Regeln zur Korrelation von sicherheitsrelevanten Echtzeit-Informationen.

Abbildung 2 zeigt eine Skizze des Prozesses eines Wissensaustausches zwischen zwei Unternehmen A und B über eine zentrale Stelle. Für den Fall, dass ein Unternehmen A einen Sicherheitsvorfall erkennt, muss dieser durch einen Experten annotiert werden. Die standardisierte Beschreibung des Vorfalls kann dann sowohl für die unternehmenseigene Wissensbasis genutzt werden als auch für eine zentrale Wissensbasis außerhalb des Unternehmens.

Abbildung 2 | Skizze des Austauschprozesses von Sicherheitswissen



Andere Unternehmen haben keinen direkten Zugriff auf dieses Wissen, sondern nur über die zentrale Stelle nach einem möglichen Abstraktionsschritt, der den Rückschluss auf Unternehmen A (mit dem Vorfall) verhindert. Basierend auf diesem Wissen kann dann in einem anderen Unternehmen B ggf. eine neue Form von Vorfällen detektiert werden. Diese können dann ebenfalls wieder für die unternehmenseigene Wissensbasis beschrieben werden und auch zur Verfeinerung des Wissens für andere Unternehmen dienen.

Daraus entsteht eine Schleife, die das bestehende Wissen unternehmensübergreifend vervollständigt bzw. spezialisiert. Dieser Prozess kann selbstverständlich durch ein Unternehmen ergänzt werden, das speziell sicherheitsrelevantes Wissen recherchiert und dem Verbund zur Verfügung stellt, z. B. Hersteller von kompatiblen SIEM-Korrelationssystemen. Bei dieser Form würde die Abstraktion des Wissens natürlich entfallen. Das Wissen kann jedoch unabhängig davon durch die unterschiedlichen Unternehmen ergänzt werden.

Zum Austausch von Wissen, das direkt in Unternehmen mit möglicherweise heterogenen Infrastrukturen und unterschiedlichen IDS-Erkennungsmethoden verwendet werden soll, ist ein Standard für die wesentlichen Konzepte zur Beschreibung von sicherheitsrelevantem Wissen notwendig. Dies umfasst einen Standard der beiden bereits genannten Ontologien für Sicherheitswissen und IT-Assets.

Ein weiteres Problem ist die Vertrauenswürdigkeit des ausgetauschten Wissens. Es ist nicht auszuschließen, dass Angreifer bewusst falsches Wissen beisteuern, um einen Angriff zu verbergen. Um dies dennoch zu verhindern, ist es erforderlich, dass eine zentrale, vertrauenswürdige Instanz die Teilnehmer hinsichtlich ihrer Vertrauenswürdigkeit überprüft. Es ist nicht erwünscht, dass z. B. ein Konkurrenzunternehmen falsches Wissen verbreiten kann, um ein anderes Unternehmen zu schädigen.

Eine mögliche Lösung stellt dabei die Überprüfung des Wissens bezüglich der Konsistenz zu dem Wissen unterschiedlicher, voneinander unabhängiger Unternehmen dar. Diese Unabhängigkeit festzustellen ist jedoch nicht zwingend möglich. Eine weitere Möglichkeit besteht in der Überprüfung des Wissens hinsichtlich des Widerspruchs zu bereits länger bekanntem und validiertem Wissen, z. B. durch Abstraktion und Überlappung mit diesem. Ein weiterer Sicherheitsmechanismus ist die Validierung durch die zentrale Stelle anhand der dort vorliegenden Informationen über die Infrastruktur der teilnehmenden Unternehmen. Meldet beispielsweise ein Unternehmen einen Sicherheitsvorfall, der nicht zu der eingesetzten Infrastruktur passt (z. B. für Server, die nicht bei diesem Unternehmen im Einsatz sind), sollte dies zu einer Verringerung des Vertrauensgrades dieses Wissens führen.

Für das Unternehmen, das Wissen anderer Unternehmen nutzen möchte, sollte zudem die Möglichkeit gegeben werden, eine Gruppe von Teilnehmern zu definieren, die ein besonderes Vertrauen genießt. Dies kann z. B. den Abstraktionsgrad des Wissens für diese Gruppe und den Vertrauensgrad hinsichtlich des von dieser Gruppe gemeldeten Wissens beeinflussen.

4 Fazit

In diesem Beitrag haben wir ein Verfahren beschrieben, das es ermöglicht, sicherheitsrelevantes Wissen zwischen Unternehmen anonym auszutauschen und dieses zur Bekämpfung von Angriffen zu verwenden. Dabei spielen Ontologien eine zentrale Rolle für die Beschreibung dieses Wissens, u. a. in Form von Korrelationsregeln und deren Abstraktion zur Gewährleistung der Anonymität bei gleichzeitiger Reaktionsschnelligkeit. Die wesentlichen Probleme dieses Ansatzes sind sowohl das Fehlen eines Standards für die Beschreibung der Regeln und der Ontologien als auch das Problem, einen zentralen Dienstleister zu finden, dem alle Unternehmen vertrauen.

Dieser Ansatz sollte im nächsten Schritt zwischen einigen nicht konkurrierenden Unternehmen evaluiert werden, um eine erste Einschätzung der Übertragbarkeit auf größere Unternehmensverbünde zu ermöglichen.

Dank

Diese Arbeit wurde durch das Bundesministerium für Bildung und Forschung unter dem Förderkennzeichen 01|S08022 (Projekt FIDeS) gefördert.

Literatur

[1] Mirkovic, J.; Reiher, P. (2004). *A Taxonomy of DDoS Attack and DDoS Defense Mechanisms*. ACM SIGCOMM Computer Communication Review.

[2] Lee, W.; Wang, C. & Dragon, D. (2008). *Botnet Detection: Countering the Largest Security Threat*.

[3] *FIDeS*. http://www.fides-security.org/

[4] *OWL Web Ontology Language*. http://www.w3.org/TR/2004/REC-owl-features-20040210/

[5] *Resource Description Framework (RDF)*. http://www.w3.org/RDF/

[6] Stoddard, M; Bodeau, D.; Carlson, R; et al. (2005). *Process Control System Security Metrics – State of Practice*.

[7] *SPARQL Query Language for RDF*. http://www.w3.org/TR/rdf-sparql-query/

[8] *Pellet: OWL 2 Reasoner for Java*. http://clarkparsia.com/pellet

[9] Sweeny, L. (2002). *k-ANONYMITY: A Model for Protecting Privacy*. International Journal on Uncertainty, Fuzziness and Knowledge-based Systems.