Kann Big Data Security unsere IT-Sicherheitssituation verbessern?

Die Schäden durch Angriffe im Internet zeigen, dass wir uns zurzeit nicht angemessen schützen.

Das Internet mit seinen vielfältigen innovativen Möglichkeiten hat eine hohe Relevanz in unserer modernen Gesellschaft erreicht, die noch weiter steigen wird.
Die Werte der Unternehmen werden fast ausschließlich mit Hilfe von IT-Systemen verwaltet. Wir beobachten, dass die Angriffsflächen der IT- und Internet-Technologie durch komplexere Software und kompliziertere Zusammenhänge zwischen Protokollen, Diensten und Infrastrukturen vielfältiger und deutlich größer werden. Die Angriffe auf unsere immer höheren Werte auf den IT-Systemen und deren Verfügbarkeit werden verteilter, raffinierter, professioneller und sehr erfolgreich ausgeführt. Die IT-Kriminalität erfährt eine zunehmende Industrialisierung und damit eine nicht zu unterschätzende und nie dagewesen professionalisierte Nachhaltigkeit.
Eine kritische Beurteilung der aktuellen IT-Sicherheitssituation des Internets zeigt, dass wir uns zurzeit nicht angemessen schützen. Welches sind die größten Sicherheitsprobleme im Internet?

„Einfallstor Software“: Die Software stellt in allen Branchen einen immer größeren Wertschöpfungsanteil dar. Wir nutzen Software in PCs, Notebooks, Smartphones, in sehr großen Rechnerzentren, aber auch immer mehr in Autos, in Industrieanlagen, im Haus und vielen weiteren Lebensbereichen. Ein großes Sicherheitsproblem ist, dass in der aktuell genutzten Software zu viele Schwachstellen vorhanden sind. Die Software-Qualität der Betriebssysteme, Anwendungen und Dienste reicht bei der heutigen Bedrohungslage nicht mehr aus. So liegt die Fehlerdichte, also die Anzahl an Softwarefehlern pro 1.000 Zeilen Code, in qualitativ hochwertiger Software heute im Schnitt bei 0,3. Da gängige Betriebssysteme zehn Millionen Zeilen Code und mehr haben, sind danach durchschnittlich mehr als 3.000 Software-Fehler zu finden. Teile von diesen Softwarefehlern sind Ziele für erfolgreiche Angriffe. Bei den großen Betriebssystemen, Anwendungen und Diensten ist in den nächsten zehn Jahren auch mit keiner sprunghaften Verbesserung der Software-Qualität zu rechnen und selbst wenn: Auch bei verbesserter Software-Qualität werden die Angreifer noch vorhandene Software-Schwachstellen professioneller ausnutzen.

“Schlechter Schutz vor Malware“: Malware ist der Oberbegriff für "Schadsoftware" wie Viren, Würmer, Trojanische Pferde und andere. Angreifer - wie kriminelle Organisationen, Spione oder Terroristen - nutzen Software-Schwachstellen aus, um Malware auf IT-Endgeräten zu installieren. Hauptsächlich über E-Mail-Anhänge oder unsichere Webseiten mit Hilfe von sogenannten Drive-by Downloads wird Malware in IT-Endgeräte unbemerkt eingeschleust. Das Institut für Internet-Sicherheit geht zurzeit davon aus, dass auf jedem 25. IT-Endgerät in Deutschland ungewollte Malware vorhanden ist, die über ein Botnetz gesteuert wird /Pohl13/. Ein Botnetz ist eine Gruppe von IT-Endgeräten, die unter zentraler Kontrolle eines Angreifers steht und von ihm für Angriffe genutzt wird. Dadurch können Angreifer Informationen von IT-Endgeräten auslesen (Keylogger, Trojaner), IT-Endgeräte für die Spam-Verteilung und DDoS-Angriffe nutzen und zum Beispiel Daten verschlüsseln und Lösegeld für die Entschlüsselung verlangen. Bei Lösegeldforderungen verschlüsseln die Angreifer mit Hilfe der Malware wichtige Daten auf dem IT-Endgerät und verlangen vom

Besitzer eine Summe für die Informationen, mit denen die Daten wieder entschlüsselt werden können.
Wir müssen kritisch feststellen, dass die Anti-Malware-Produkte heute mit 75 bis 95 Prozent eine zu schwache Erkennungsrate haben. Bei direkten Angriffen auf ein IT-System liegt die Erkennungsrate im Schnitt sogar nur bei 27 Prozent.
Advanced Persistent Threat (APT) ist die Begrifflichkeit, die sich für intelligente Malware wie Stuxnet und deren „Verteilung“ international etabliert hat. Unter dem Namen Stuxnet wird ein Botnet mit einer qualitativ sehr hochwertigen Malware verstanden, die speziell für Produkte zur Überwachung und Steuerung technischer Prozesse (SCADA-System) der Firma Siemens entwickelt wurde. Stuxnet wurde mit dem Ziel geschrieben, die Leittechnik einer Anlage zur Uran-Anreicherung im Iran zu sabotieren. Stuxnet hat eine neue Qualität an Malware eingeleitet, die sehr viel intelligenter ist, viel gezielter vorgeht und vor allem einen sehr viel größeren Schaden anrichten kann. Stuxnet markiert den Startpunkt der Entwicklung von qualitativen Cyberwaffen, die Industrien und Infrastrukturen ganzer Länder lahmlegen können.
Allgemein wird der Begriff Advanced Persistent Threat (APT) in der Regel als ein gezielter Angriff mit komplexen Angriffstechnologien und -taktiken sowie aufwendigen Hintergrundinformationen eines Opfer-IT-Systems und dessen Umgebung verstanden. Dabei nimmt der professionelle Angreifer einen großen Aufwand auf sich (Advanced), um erfolgreich auf ein Opfer-IT-System zuzugreifen und dabei möglichst lange (Persistent) unentdeckt zu bleiben. So ist es möglich, über einen längeren Zeitraum Informationen auszuspähen oder Schaden anzurichten. Gegen diese Art von hochentwickelten und professionellen Angriffen mit intelligenter Malware in Verbindung mit Social Engineering haben wir im Prinzip heute keine passenden Abwehrtechnologien im Einsatz!

Weitere Sicherheitsprobleme: Passworte werden immer noch für die Authentikation im Internet genutzt, unsere Webseiten sind zu unsicher, mobile Geräte bringen neue Angriffsvektoren, Verschlüsselungen werden zu wenig genutzt, Cloud Computing stellt eine weitere Herausforderung dar, usw.

Kritische Bewertung der aktuellen IT-Sicherheitssituation des Internets

Professionelle Hacker greifen alles und weltweit erfolgreich an! Professionelle Hacker haben in den letzten Jahren sogar amerikanische IT-Firmen, wie z.B. Google und RSA Security erfolgreich gehackt. Auch die großen US-Zeitschriften wie New York Times und Washington Post waren für die professionellen Hacker leichte Beute, und fast alle US-Behörden waren für die Eindringlinge wie ein Schweizer Käse (siehe auch http://www.heise.de/security/artikel/NSA-Affaere-Die-Krise-als-Chance-begreifen-2037053.html). Wir müssen realisieren, dass unsere heutigen IT-Sicherheitslösungen weder Geheimdienste, noch professionelle Hacker stoppen können. Wenn die professionellen Hacker dieser Welt das richtige Wissen haben und über genug Geld verfügen, können sie im Sinne von Advanced Persistent Threats jede Organisation erfolgreich hacken.

Welche IT-Sicherheitsstrategien stehen zur Verfügung, um uns zu schützen?

Im Folgenden werden drei prinzipielle IT-Sicherheitsstrategien beschrieben, die uns helfen können, uns zu schützen.

1. Vermeiden von Angriffen

Die einfachste IT-Sicherheitsstrategie ist die, dass Angriffe erfolgreich vermieden werden. Beispiele von Vermeidungsstrategien sind:

- Nicht alle IT-Systeme ans Internet anschließen
- So wenig wie möglich Daten generieren, die angegriffen werden können
- Keine Technologie mit Schwachstellen verwenden
- Usw.

Abb: Vermeiden von Angriffen

Bewertung: Vermeiden von Angriffen ist die beste IT-Sicherheitsstrategie, aber praktisch nur begrenzt umsetzbar, da wir mit unseren IT-Systemen ans Internet angeschlossen sein wollen, um die Vorteile nutzen zu können.

2. Entgegenwirken von Angriffen

Das Entgegenwirken von Angriffen ist die meist verwendete IT-Sicherheitsstrategie, um Schäden durch Angriffe zu vermeiden. Hier werden IT-Sicherheitstechnologien verwendet, die eine hohe Wirkung gegen bekannte Angriffe zur Verfügung stellen und damit die Werte angemessen schützen.
Beispiele für IT-Sicherheitstechnologien sind Verschlüsselungssicherheitssysteme (Datei-, Festplatten-, E-Mail-Verschlüsselung, VPN-Systeme, SSL, ...), Authentikationsverfahren (Passworte, Biometrie, Challenge-Response, …), Firewall-Systeme, Rechteverwaltung, usw.

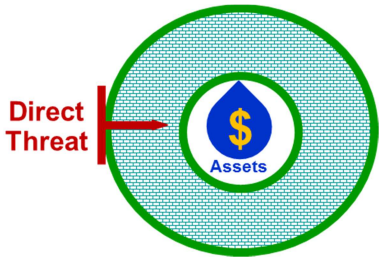

Abb: Entgegenwirken von Angriffen

Bewertung: Das Entgegenwirken von Angriffen ist eine naheliegende IT-Sicherheitsstrategie, aber leider stehen zurzeit nicht genug wirkungsvolle IT-Sicherheitstechnologien, -Lösungen und -Produkte zur Verfügung oder werden nicht angemessen eingesetzt, was die erfolgreichen professionellen Hacker und diverse Geheimdienste uns jeden Tag vor Augen halten.

Im Schnitt sind es nur ca. 5 % aller vorhandenen Daten in Unternehmen, die besonders schützenswert sind. Aber welche Daten besonders schützenswert sind, ist nicht im jedem Unternehmen durch die Stakeholder definiert.

Innovationen auf dem Gebiet des Entgegenwirkens von Angriffen mit Hilfe von IT-Sicherheitstechnologien liegen im Bereich von proaktiven IT-Sicherheitssystemen.
Die zunehmende Vielfalt und Komplexität unserer IT-Endgeräte und IT-Infrastrukturen brauchen deutlich verlässlichere, robustere und wirkungsvollere IT-Sicherheitskonzepte. Wir müssen weg von ausschließlich reaktiven hin zu modernen, proaktiven IT-Sicherheitssystemen, die eine Ausführung von intelligenter Malware, eines der größten Probleme zurzeit, verhindern können. Solche proaktiven IT-Sicherheitssysteme arbeiten mit einem kleinen Sicherheitskern und Virtualisierung (weniger als 100.000 Zeilen Code), können Software messbar, vertrauenswürdig machen und mit einer starken Isolation, Anwendungen mit ihren Daten separieren und nachhaltige und angemessene IT-Sicherheit bieten. Für proaktive IT-Sicherheitssysteme muss die Softwarearchitektur der IT-Endgeräte allerdings grundlegend anders aufgebaut sein als bisher. Außerdem müssen Sicherheits-Infrastrukturkomponenten gemeinsam umgesetzt werden, damit diese IT-Sicherheits- und Vertrauenstechnologien organisationsübergreifend genutzt werden können. Auf der Forschungsebene wurden die Vorteile der proaktiven IT-Sicherheitssysteme schon längst dargestellt und nachgewiesen. Die ersten IT-Sicherheitsunternehmen bieten heute bereits ausgereifte Lösungen. Jetzt wird es Zeit, dass diese von der Industrie und den Behörden eingeführt werden, damit eine notwendige höhere Sicherheit und Vertrauenswürdigkeit der IT-Endgeräte und IT-Infrastrukturen erzielt werden kann /PoSp2013/.

3. Erkennen von Angriffen, denen nicht erfolgreich entgegengewirkt werden kann.

Wenn Angriffen mit Hilfe von IT-Sicherheitsmechanismen nicht angemessen entgegengewirkt werden kann, dann bleibt nur noch die Strategie, diese zu erkennen und zu versuchen, den Schaden so schnell wie möglich zu minimieren.

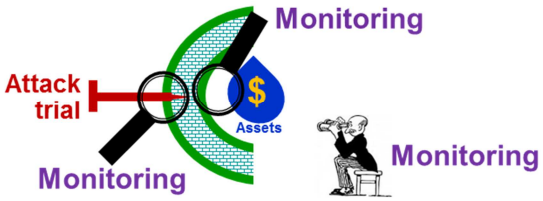

Abb: Erkennen von Angriffen

In diesem Bereich gibt es z.B. ein Intrusion Detection System oder generell IT-Sicherheitssysteme, die Warnungen erzeugen, wenn Angriffe erkannt werden. Hier ist die Idee, dass in einem definierten Bereich (Kommunikationsinfrastruktur, Endgeräte, …) nach Angriffssignaturen oder Anomalien gesucht wird, um dann entsprechend reagieren zu können, um Schaden zu verhindern oder zu reduzieren.

Bewertung: Die IT- Sicherheitsstrategie, Erkennen von Angriffen, ist sehr hilfreich, hat aber definierte Grenzen. Advanced Evasion Techniques (AET) ist z.B. eine Technologie, die dafür sorgt, dass Angriffe unentdeckt an Sensoren vorbei gehen

können. Dies kann z.B. durch die Provozierung von IP-Fragmentierung erreicht werden, bei der IP-Pakete in kleine Teile aufgeteilt und damit Signaturen verschleiert werden.

Big Data Security

Big Data Security soll Advanced Persistent Threats erkennen, um Schaden zu verhindern oder zu reduzieren. Die Themen „Big Data“ und „Security“ werden kombiniert, um mehr IT-Sicherheit erzielen zu können.
Die grundsätzliche Idee von Big Data ist, große Datenmengen aus vielfältigen Quellen mit neu entwickelten Methoden und Technologien zu erfassen und zu analysieren, um gewünschte Ergebnisse zu erzielen.
Der Bereich Big Data Security soll mit so viel Input von sicherheitsrelevanten Informationen und Kontextinformationen wie möglich, die Beurteilung der IT-Sicherheitssituation ermöglichen, damit Advanced Persistent Threat (APT) erkannt werden können.

Datenquellen für Big Data Security

Als Datenquellen werden sicherheitsrelevante und Kontext Informationen in dem zu überwachenden IT-Bereich generiert, gesammelt und verarbeitet.

Logdaten und weitere sicherheitsrelevante Informationen von

- Netzkomponenten, wie Switche, Route, Gateways, …
- Servern, wie Webservern, E-Mail-Servern, SIP-Servern, …
- Sicherheitskomponenten, wie Firewall, Intrusion Detection, …
- Endsystemen (Betriebssystemen, Dateisystemen, Anwendungen, CPU-Auslastung, Speicherauslastung, Stromversorgung, usw.)
- Speziellen Sensoren von der Kommunikationsinfrastruktur und Endsystemen
- Usw.

Kontextinformationen
Im Bereich Kontextinformationen werden Daten gesammelt, die geeignet sind, ein dynamisches Lagebild der IT-Sicherheit zu erstellen. Hierzu zählen insbesondere:

- Informationen über bekannte Bugs, Schwachstellen, Sicherheitsvorfälle von den eingesetzten Technologien, Lösungen und Produkten.
- Lagebilder anderer Organisationen, die zur Verfügung gestellt werden.
- Sicherheitsinformationen nationaler und internationaler CERTs
- Genutzte Technologien
- Usw.

Methoden zum Finden von relevanten Angriffsdaten

Im Umfeld von Big Data ist die Auswertung von sehr großen Mengen an sicherheitsrelevanten und Kontext Informationen erst mit den heutigen leistungsfähigen Computersystemen in einen praktisch umsetzbaren Rahmen gerückt.
Augenscheinlich ungeordnete Datenstrukturen erschweren den Weg für ein zielgerichtetes Erkennungssystem von Angriffen aus einer Vielzahl von sicherheitsrelevanten Informationen. Mit den Möglichkeiten durch Algorithmen des maschinellen Lernens lassen sich viele Strukturen in den sicherheitsrelevanten

Informationen allerdings ordnen. Bei der Vorbereitung werden gleichartige Eingabeparameter erzeugt und die sicherheitsrelevanten Informationen normalisiert. Die Dimension der sicherheitsrelevanten Informationen lässt sich ebenfalls reduzieren, um den Rechenaufwand zu minimieren und dadurch den Gesamtablauf zu optimieren. Maschinelles Lernen unterscheidet dabei zwei Ansätze, überwachtes und unüberwachtes Lernen. Zum Aufgabenfeld des überwachten Lernens gehören das Regressions- und Klassifizierungsproblem. Ziel der Regression ist es, Daten vorherzusagen und eine klare Zahlaussage zu generieren. So ließen sich beispielsweise Zahlen zur Darstellung von kritischen Sicherheitsrisiken hervorbringen. Aufgaben der Klassifikation befassen sich damit, Daten in verschiedene Klassen einzuteilen, möglich wäre z.B. dadurch etwa eine Einteilung von Software in bösartig oder harmlos.
Die Stärke im unüberwachten Ansatz liegt darin, nach Mustern auch in vorher unklassifizierten Daten zu suchen, um sie nach vorheriger Aufbereitung besser beschreiben zu können. Mittels Clustering werden ähnliche Datengruppen miteinander in Verbindung gesetzt. Messwerte, die z.B. eine aufgenommene Leistung an Servernetzteilen dokumentieren, lassen sich beispielsweise so gruppieren, dass anschließend festgestellt werden kann, ob ein Server sich im Leerlauf befindet oder auffällig viel Aktivität verzeichnet.
Die Erwartungshaltung an diesen Erkennungsansatz liegt unter anderem darin, Dinge zu erkennen, die vorher anders nicht greifbar waren. Da der Algorithmus selbstständig lernt, werden klassische Fehler in diesem Sinne also nicht produziert, was dann aber zu einem anderen Problem führen kann. Es ist wichtige zu erkennen, ob der Algorithmus auch in die gewünschte Richtung lernt. Für die erfolgreiche Erkennung von Angriffen ließen sich diese Ansätze kombinieren, um eine Erweiterung in der Systemsicherheit zu entwickeln. Mit unüberwachtem Lernen müssen alle relevanten Gegebenheiten miteinander abgeglichen und dadurch Korrelationen gefunden werden. Aufgrund der Ergebnisse des vorherigen Ansatzes ließe sich dann mittels überwachten Lernens ein Klassifizierungsproblem ableiten, um die Angriffsvektoren zu identifizieren. Aufgrund der individuellen Gegebenheiten eines Unternehmens ist der Einsatz von Lernalgorithmen an die vorhandenen IT-Infrastrukturen anzupassen.

Technologien, die die Umsetzung von Big Data Security unterstützen

Big Data Security setzt nicht allein auf neue Ansätze zur Erkennung von Angriffen, sondern nutzt bereits schon teilweise Jahrzehnte alte Konzepte, die bislang jedoch nicht praktikabel umgesetzt werden konnten und in der Regel auch nicht speziell für diesen Zweck entwickelt wurden. Die enormen Fortschritte im Bereich der Computer-Hardware ermöglicht es, diese Ansätze nun praktisch zu umzusetzen. Heutige, marktübliche Computer verfügen über genügend Computerleistung, um viele dieser umfangreichen Prozesse von maschinellem Lernen in akzeptabler Zeit durchzuführen. Die Möglichkeiten der Parallelisierung aufgrund leistungsfähigerer Hardware steigern diese Leistung noch einmal.
Dazu kommen noch hohe Geschwindigkeiten in der Datenübertragung, die ein Auslagern verschiedener Prozesse von maschinellem Lernen auf dafür bereitgestellte Server erlauben. Im Bereich der Server ist die Leistung der Hardware dann noch einmal um ein vielfaches höher. So können riesige Mengen von sicherheitsrelevanten und Kontext Informationen dauerhaft auf Festplatten gesichert, aber auch für die Analyse in sehr großen Arbeitsspeichern gehalten und verarbeitet werden. Big Data Security ist aufgrund dieses Leistungszuwachses erst heute praktisch umsetzbar.

Herausforderungen von Big Data Security

Es gibt zahlreiche Herausforderungen bei der Umsetzung eines Big Data Security Systems, von denen hier drei Beispiele aufgezeigt werden:
Obwohl die Computerleistung der Hardware in den letzten Jahren kontinuierlich zunimmt, können einige Berechnungen, abhängig von der Anzahl der Eingabeparameter und des Lernmodells, auch diese Grenzen erreichen. Berechnungen eines komplexen neuronalen Netzes erreichen schon bei handelsüblichen Computern schnell ihre Grenzen.
Der Umgang mit maschinellem Lernen ist oftmals mit Ausprobieren verbunden. Gerade im Bereich Big Data Security lässt sich nicht eindeutig vorab bestimmen, welcher Ansatz der geeignetste für eine bestimmte Aufgabenstellung ist. Gerade im unüberwachten Ansatz besteht die reale Möglichkeit, dass Korrelationen in den Daten gefunden werden, die in die Irre zu führen scheinen. So sollte beispielsweise Ende der 90er eine Abteilung einer Supermarktkette die Einkaufsgewohnheiten der Konsumenten über einen bestimmten Zeitraum analysieren. Hierbei wurde überraschend festgestellt, dass junge Väter, wenn sie abends Windeln für ihren Nachwuchs einkaufen, oft gleichzeitig Bier in ihren Einkaufwagen legen. Dieses Beispiel zeigt, dass durchaus ungewöhnliche Korrelationen vorhanden sein können, über deren Nutzbarkeit letztlich ein Spezialist zu entscheiden hat. Diese Entscheidungen müssen jedoch mit großer Vorsicht getroffen werden, denn auch diese Korrelationen können eine entscheidende Relevanz aufweisen, auch wenn diese nicht direkt erkennbar ist.
Wenn sicherheitsrelevanten Informationen zur Absicherung einer IT-Infrastruktur herangezogen werden, ist der Datenschutzaspekt zwingend zu berücksichtigen, denn personenbezogene Daten dürfen nicht einfach verwendet oder abgespeichert werden. Relevant wird dies insbesondere dann, wenn beispielsweise Deep-Packet-Inspection, das intensive Untersuchen einzelner Datenpakete einer Netzwerkverbindung, als Teil eines Intrusion-Detection Systems eingesetzt wird. Aber auch vergleichbare Sicherheitstechnologien, wie LogDaten-Systeme müssen den Datenschutzaspekt zwingend berücksichtigen. Diese Aspekte des Datenschutzes zeigen, dass es neben den im klassischen Big Data Bereich längst nicht eindeutigen Datenschutzrichtlinien noch weitere juristische Besonderheiten zu beachten gilt, auf die an dieser Stelle aber nicht weiter eingegangen werden kann.

Ausblick

Big Data Security könnte als Ergebnis einen allumfassenden IT-Sicherheitslagebericht über die aktuelle IT-Infrastruktur signifikant ergänzen und damit das Sicherheitsniveau langfristig steigern. Gefahren wie Advanced Persistent Threats sollen erkannt und im Idealfall direkt verhindert werden können. Sinnvollerweise sollten Anomalien identifiziert werden, ohne dass ähnliches Verhalten vorher aufgezeichnet wurde. Auf der anderen Seite sollten so wenig wie möglich Fehlalarme entstehen, die den Arbeitsablauf der Sicherheitsabteilung erschwert. Wenn Big Data Security sich als starke Ergänzung der IT-Sicherheit etabliert, sollten sich auch europäische Softwarehersteller intensiv mit dieser IT-Sicherheitsthematik auseinandersetzen.

Firmen, die Big Data Security anbieten

Dieses IT-Sicherheitsthema wird auch von der Softwareindustrie beachtet. So bieten einige, hauptsächlich US-Amerikanische Unternehmen, wie beispielsweise IBM und

RSA Security unterschiedliche Lösungen im Bereich Big Data Security an. Vor allem RSA Security erhöht seine Bemühung, seit sie selbst zum Opfer eines Advanced Persistent Threat geworden sind.

Bewertung: Big Data Security
Ob Big Date Security das intelligente Finden der Nadel (APT) im Heuhäufen darstellt oder eher das Finden der Nadel durch das Verstreuen von zusätzlichem Heu noch schwerer macht, muss sich im Praxistest erst noch herausstellen.

Zusammenfassung
Klar ist, dass wir uns zurzeit nicht angemessen gegen die professionellen Hacker schützen können.
Big Data Security ist ein Ansatz, der uns helfen kann, komplexe Sicherheitsprobleme zu identifizieren und damit Schäden zu vermeiden oder zu reduzieren. Es wird sich herausstellen müssen, ob der große Aufwand von Big Data Security durch die Ergebnisse gerechtfertigt werden kann. Der Versuch lohnt sich!

Wichtig zu erkennen ist aber auch, dass wir immer noch IT- Sicherheitsmechanismen benötigen, um z.B. mit proaktiven IT-Sicherheitssystemen, die 5 % der Kronjuwelen der Unternehmen angemessen zu schützen.

Literatur:

/Pohl13/ N. Pohlmann: „Daten gegen Diebstahl sichern“, Wirtschaftsspiegel, IHK Münster, 2/2013

/PoSp2013/ N. Pohlmann, A. Speier: „Eine Diskussion über Trusted Computing“. IT-Sicherheit – Management und Praxis, DATAKONTEXT-Fachverlag, 5/2013