



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Entwurf vom
11. Oktober 2017

Masterarbeit

**Datenschutzfreundliche Speicherung
unternehmensinterner Überwachungsdaten mittels
Pseudonymisierung und kryptographischer
Schwellwertschemata**

vorgelegt von

Tom Petersen

geb. am 13. Dezember 1990 in Hannover

Matrikelnummer 3659640

Studiengang Informatik

eingereicht am 11. Oktober 2017

Betreuer: Dipl.-Inf. Ephraim Zimmer

Erstgutachter: Prof. Dr. Hannes Federrath

Zweitgutachter: Prof. Dr. Mathias Fischer

Aufgabenstellung

Die technologiegestützte Bekämpfung von Insider-Angriffen im Unternehmenskontext basiert aktuell häufig auf der Analyse des Nutzerverhaltens einzelner Mitarbeiter und der Erkennung von Abweichungen zum erwarteten Normalverhalten. Diese sogenannte Anomalieerkennung benötigt umfassende Überwachungsdaten aller digitalen Endgeräte und Datenkommunikationssysteme zur Erstellung und eindeutigen Zuordnung von Nutzerprofilen. Dabei entsteht ein Konflikt mit dem Datenschutz der Mitarbeiter, da die Erhebung, Verarbeitung und Speicherung von Überwachungsdaten einen schweren Eingriff in die Privatsphäre und die informationelle Selbstbestimmung der Mitarbeiter darstellt. Um diesen Konflikt zu lösen, können auf der einen Seite Datenschutztechniken eingesetzt werden, die den unmittelbaren Personenbezug gesammelter Daten entfernen. Auf der anderen Seite kann mithilfe von Kryptographie die Rückgewinnung des Personenbezugs im Verdachtsfall und unter der Voraussetzung einer mehrseitigen Kollaboration ermöglicht werden.

Das Ziel der Masterarbeit ist die konzeptionelle Erarbeitung einer solchen datenschutzfreundlichen und mehrseitig sicheren Erhebung, Verarbeitung und Speicherung von Überwachungsdaten sowie die prototypische Implementierung auf Basis eines Security Information Event Management Systems. Dabei sollen insbesondere die folgenden Punkte bearbeitet werden:

- Wie ist der aktuelle Stand sowohl der Technik als auch der Wissenschaft im Bereich der Pseudonymisierung und der kryptographischen Schwellwertschemata?
- An welcher Stelle des konzipierten Systems können die Überwachungsdaten entsprechend des Datenschutzes und der späteren möglicherweise erforderlichen Rückgewinnung des Personenbezugs verarbeitet werden und welche Auswirkungen können entstehen?
- Wie und an welcher Stelle muss das Schlüsselmanagement der benötigten kryptographischen Funktionen erfolgen?
- Welche Alternativen gibt es neben der Pseudonymisierung und den kryptographischen Schwellwertschemata zur Lösung des genannten Zielkonflikts und wie können diese in das Konzept und die prototypische Implementierung integriert werden?

Todo list

■ TBW	4
■ Wo passen Abschnitte zu folgenden Stichworten hin? - Verschiedene Datenarten: Identifizierend, Traffic, nicht relevant, ... - Grundlegende Definition Insiderangriff .	6
■ Schadenshöhe (siehe Antrag)?	6
■ EZ: Warum nicht? Einige werben damit, Innentaeter erkennen zu koennen, z.B. IBM QRadar	6
■ EZ: Warum? Welche?	6
■ EZ: Was genau ist das Szenario? Liegt dein Fokus nun auf den zusaetzlichen Daten- quellen und Erkennungslogiken oder auf den datenschutzrechtlichen Bedenken? . .	7
■ Zu ergaenzen: Aufbau der Arbeit, Literatur	7
■ Evtl. auch Telemediengesetz erwahnen	14
■ Erweiterte Beschreibung	16
■ EZ zu Schritt 3: $u_i = (i, \dots)$???	16
■ Hier ergaenzen.	17
■ Details	19
■ TBW	19
■ TBW	22

Zusammenfassung

TBW

Inhaltsverzeichnis

1	Einführung	6
1.1	Related work	7
2	Grundlagen	8
2.1	Definitionen und Notationen	8
2.2	Arbeitnehmerdatenschutz	8
2.3	SIEM-Systeme	13
2.4	Pseudonymisierung	14
2.5	Schwellwertschemata	15
3	Stand der Wissenschaft und Auswahl von Verfahren	18
3.1	SIEM-Systeme	18
3.1.1	OSSIM	18
3.2	Pseudonymisierung	18
3.3	Schwellwertschemata	19
4	Implementierung	20
4.1	Anforderungen	20
4.2	Architektur und Umsetzung	20
4.3	Einbindung in OSSIM	20
4.4	Umsetzung der Pseudonymisierung	20
4.5	Implementierung und Integration des Schwellwertschemas	20
4.6	Evaluation	20
5	Alternativen	21
5.1	Alternative1	21
5.2	21
5.3	Vorgehen zur Integration	21
6	Fazit	22
	Literatur	23

1 Einführung

Liest man von erfolgreichen Angriffen auf Unternehmensnetzwerke, so ist die implizite Annahme von außenstehenden, unternehmensfremden Angreifern weit verbreitet. Doch häufig sind die Angreifer bereits im Netzwerk ansässig. Es handelt sich um (ehemalige) Mitarbeiter oder zumindest Personen mit legitimem Zugriff auf das Netzwerk, wie Geschäftspartnern oder Kunden. Hierbei geht es keineswegs lediglich um Einzelfälle.

Wo passen Abschnitte zu folgenden Stichworten hin? - Verschiedene Datenarten: Identifizierend, Traffic, nicht relevant, ... - Grundlegende Definition Insiderangriff

In dem *IBM Cyber Security Intelligence Report* von 2015 werden 55% der Angriffe als aus dem internen Netz stammend angegeben [Bra+15]. Zu beachten ist allerdings, dass nicht nur mit Absicht ausgeführte Angriffe hierunter erfasst wurden, sondern auch unbeabsichtigte wie das versehentliche Veröffentlichen schützenswerter Kundendaten.

Auch der Branchenverband bitkom führt in seiner *Spezialstudie Wirtschaftsschutz* aus dem Jahr 2016 nach einer Befragung von über 1000 Unternehmen aus, dass etwa 60% der erfolgten Handlungen aus dem Bereich Datendiebstahl, Industriespionage oder Sabotage durch (ehemalige) Mitarbeiter erfolgten [bit16].

Schadenshöhe (siehe Antrag?)

Auch wenn die genauen Zahlen aufgrund von unterschiedlichen Annahmen und der in diesem Bereich nicht zu vernachlässigenden Dunkelziffer¹ mit Vorsicht zu betrachten sind, so geben sie doch Hinweise darauf, dass Angriffe von Innentätern weit verbreitet sind und ein hohes Schadenspotenzial aufweisen. Die Erkennung und Verhinderung solcher Angriffe sollte daher ein wichtiger Teil des IT-Sicherheitskonzepts eines Unternehmens sein.

Zur Erkennung von Angriffen in Netzwerken können SIEM-Systeme eingesetzt werden (siehe Abschnitt ??). Diese sind jedoch in erster Linie auf das Erkennen von externen Angriffen ausgelegt und in ihrer derzeitigen Form kaum sinnvoll für das Erkennen von Innentätern zu nutzen.

Hierfür würden zusätzliche Datenquellen und Erkennungslogiken nötig sein. Zusätzlich spielen auch datenschutzrechtliche Bedenken im Bezug auf das Sammeln von großen Datenmengen über Mitarbeiter des eigenen Unternehmens hier eine entscheidende Rolle. Details hierzu sind im folgenden Kapitel ?? zu finden.

EZ: Warum nicht? Einige werben damit, Innentäter erkennen zu können, z.B. IBM QRadar

EZ: Warum? Welche?

Ein Ansatz, der diese Bedenken ausräumen oder zumindest lindern könnte, ist die Nutzung von Pseudonymen bei der Datenerfassung (siehe Abschnitt ??). Anstatt direkt identifizierende Merkmale eines Arbeitnehmers abzuspeichern, werden diese Merkmale durch ein Pseudonym ersetzt. Eine Liste dieser Ersetzungen wird verschlüsselt abgelegt. Im Fall eines Angriffs durch einen Innentäter kann die Liste entschlüsselt werden und relevante Ereignisse de-pseudonymisiert, also ihrem ursprünglichen Verursacher wieder zweifelsfrei zugeordnet, werden.

Um die Entschlüsselung nicht einzelnen (möglicherweise bösartig agierenden) Personen zu

1. Insbesondere die Angst vor Imageschäden, die auch in der *Spezialstudie Wirtschaftsschutz* erwähnt wird, könnte ein Grund für das Geheimhalten von Vorfällen sein.

ermöglichen, können sogenannte Schwellwertschemata eingesetzt werden (siehe Abschnitt ??). Durch sie wird die Entschlüsselung erst durch die Kooperation mehrerer Parteien möglich gemacht.

Bei diesem Ansatz muss jedoch auch beachtet werden, dass durch den Einsatz von Pseudonymen die Erkennung von Angriffen erschwert werden könnte. Beispielsweise könnte das Ändern von Pseudonymen in regelmäßigen Zeitintervallen und die dadurch entstehende Nicht-Verkettbarkeit von Ereignissen dafür sorgen, dass langfristig angelegte Angriffe nicht aufgedeckt werden.

Ziele der Arbeit

In dieser Arbeit soll es darum gehen, prototypisch ein solches Szenario auf Basis eines Open-Source-SIEM-Systems umzusetzen. Hierbei müssen einige Fragen betrachtet werden:

- An welcher Stelle des Systems kann eingegriffen werden, um die erfassten Daten zu verändern, und welche Auswirkungen hat dies?
- Wie erfolgt die angesprochene Pseudonymisierung technisch?
- Welche kryptographischen Schwellwertschemata können genutzt werden? Gibt es bereits quelloffene Implementierungen? Was muss selbst entwickelt werden? Wie erfolgt das Schlüsselmanagement?
- Können neben der Pseudonymisierung noch weitere Funktionen zur Veränderung von Daten sinnvoll sein und wie könnten diese umgesetzt werden?

EZ: Was genau ist das Szenario? Liegt dein Fokus nun auf den zusätzlichen Datenquellen und Erkennungslogiken oder auf den datenschutzrechtlichen Bedenken?

Gerade die letzte Frage sorgt dafür, dass zusätzliche Anforderungen an den zu entwickelnden Prototypen gestellt werden. Es sollte möglich sein, abhängig von den eingehenden Daten die entsprechend gewünschten Funktionen konfigurieren und den Prototypen in eventuell aufbauenden Arbeiten auch um zusätzliche Funktionen ergänzen zu können.

Zu ergänzen: Aufbau der Arbeit, Literatur

1.1 Related work

2 Grundlagen

Dieses Kapitel widmet sich den Grundlagen der in dieser Arbeit verwendeten Konzepte, Verfahren und Systeme. Zu Beginn werden für den Verlauf der Arbeit notwendige Definitionen gegeben. Anschließend werden die juristischen Hintergründe des Arbeitnehmerdatenschutzes erläutert, die den rechtlichen Rahmen für das Thema dieser Arbeit bilden.

Es folgen Erläuterungen zu SIEM-Systemen, in die - wie bereits in der Einleitung erläutert - die prototypische Umsetzung der datenschutzfreundlichen Speicherung erfolgen soll, sowie zu den zu verwendenden Datenschutztechniken Pseudonymisierung und kryptographische Schwellwertschemata.

2.1 Definitionen und Notationen

- Insider-Angriff
- Datenarten
- ...

2.2 Arbeitnehmerdatenschutz

(Mitarbeiterdatenschutz, Beschäftigtendatenschutz, Personaldatenschutz, Betriebsdatenschutz)

- Recht auf informationelle Selbstbestimmung als Ausprägung des Allgemeinen Persönlichkeitsrechts
- Besondere Rechtslage im Beschäftigungsverhältnis (aktuelle Rechtsprechung und Ausblick...)
- (Antibeispiele: Lidl, Bahn, Überwachungsaffäre der Deutschen Telekom)
- *Eine heimliche Überwachung von Mitarbeitern ist im Regelfall unzulässig, wie das Bundesarbeitsgericht jüngst entschieden hat (Urteil vom 27. Juli 2017, 2 AZR 681/16).*

Recht auf informationelle Selbstbestimmung

Das RaiS ist eine Ausprägung des allgemeinen Persönlichkeitsrechts (Schutz der Persönlichkeit einer Person vor Eingriffen in ihren Lebens- und Freiheitsbereich) nach Artikel 2, Absatz 1 in Verbindung mit Artikel 1, Absatz 1 des Grundgesetzes.

Jeder hat das Recht auf die freie Entfaltung seiner Persönlichkeit, soweit er nicht die Rechte anderer verletzt und nicht gegen die verfassungsmäßige Ordnung oder das Sittengesetz verstößt.

– Artikel 2, Absatz 1, GG

Die Würde des Menschen ist unantastbar. Sie zu achten und zu schützen ist Verpflichtung aller staatlichen Gewalt.

– Artikel 1, Absatz 1, GG

Es wurde im Volkszählungsurteil als Grundrecht anerkannt.

TBD siehe pdf markiert

– Volkszählungsurteil

Dieses Urteil enthält auch erforderliche Grundsätze bei der Datenverarbeitung wie Datensparsamkeit, Zweckgebundenheit, ... (richtet sich jedoch nur an staatliche Eingriffe)

Einschränkungen sind möglich, jedoch in Gesetzen festzuhalten (Abwägung zwischen Geheimhaltungsinteresse des Betroffenen und dem öffentlichen Informationsinteresse verarbeitender Stellen).

Das RiaS bildet die Grundlage für Gesetze wie das BDSG, die LDSG (diese jedoch nur für öffentlich-rechtliche Einrichtungen relevant) oder DSGVO der EU.

Art.8, EU-Grundrechtecharta

Schutz personenbezogener Daten

- (1) Jede Person hat das Recht auf Schutz der sie betreffenden personenbezogenen Daten.
- (2) Diese Daten dürfen nur nach Treu und Glauben für festgelegte Zwecke und mit Einwilligung der betroffenen Person oder auf einer sonstigen gesetzlich geregelten legitimen Grundlage verarbeitet werden. Jede Person hat das Recht, Auskunft über die sie betreffenden erhobenen Daten zu erhalten und die Berichtigung der Daten zu erwirken.
- (3) Die Einhaltung dieser Vorschriften wird von einer unabhängigen Stelle überwacht.

§32, Bundesdatenschutzgesetz

Ursprünglich:

<https://dejure.org/gesetze/BDSG/32.html>

Datenerhebung, -verarbeitung und -nutzung für Zwecke des Beschäftigungsverhältnisses

(1) Personenbezogene Daten eines Beschäftigten dürfen für Zwecke des Beschäftigungsverhältnisses erhoben, verarbeitet oder genutzt werden, wenn dies für die Entscheidung über die Begründung eines Beschäftigungsverhältnisses oder nach Begründung des Beschäftigungsverhältnisses für dessen Durchführung oder Beendigung erforderlich ist.

Zur Aufdeckung von Straftaten dürfen personenbezogene Daten eines Beschäftigten nur dann erhoben, verarbeitet oder genutzt werden, wenn zu dokumentierende tatsächliche Anhaltspunkte den Verdacht begründen, dass der Betroffene im Beschäftigungsverhältnis eine Straftat begangen hat, die Erhebung, Verarbeitung oder Nutzung zur Aufdeckung erforderlich ist und das schutzwürdige Interesse des Beschäftigten an dem Ausschluss der Erhebung, Verarbeitung oder Nutzung nicht überwiegt, insbesondere Art und Ausmaß im Hinblick auf den Anlass nicht unverhältnismäßig sind.

(2) Absatz 1 ist auch anzuwenden, wenn personenbezogene Daten erhoben, verarbeitet oder genutzt werden, ohne dass sie automatisiert verarbeitet oder in oder aus einer nicht automatisierten Datei verarbeitet, genutzt oder für die Verarbeitung oder Nutzung in einer solchen Datei erhoben werden.

(3) Die Beteiligungsrechte der Interessenvertretungen der Beschäftigten bleiben unberührt.

Mögliche Erweiterung durch Entwurf:

Deutscher Bundestag: Gesetzentwurf der Bundesregierung: Entwurf eines Gesetzes zur Regelung des Beschäftigtendatenschutzes, BT-Drs. 17/4230 vom 15. Dezember 2010

<http://dip21.bundestag.de/dip21/btd/17/042/1704230.pdf>

<http://www.arbeitnehmerdatenschutz.de/>

Bisher scheinbar keine Abstimmung, da auf die EU DSGVO gewartet wurde.

§26, BDSG(neu) - Datenschutz-Anpassungs- und Umsetzungsgesetz

Tritt Mai 2018 in Kraft.

Datenverarbeitung für Zwecke des Beschäftigungsverhältnisses

(1) Personenbezogene Daten von Beschäftigten dürfen für Zwecke des Beschäftigungsverhältnisses verarbeitet werden, wenn dies für die Entscheidung über die Begründung eines Beschäftigungsverhältnisses oder nach Begründung des Beschäftigungsverhältnisses für dessen Durchführung oder Beendigung oder zur Ausübung oder Erfüllung der sich aus einem Gesetz oder einem Tarifvertrag, einer Betriebs- oder Dienstvereinbarung (Kollektivvereinbarung) ergebenden Rechte und Pflichten der Interessenvertretung der Beschäftigten erforderlich ist. Zur Aufdeckung von

Straftaten dürfen personenbezogene Daten von Beschäftigten nur dann verarbeitet werden, wenn zu dokumentierende tatsächliche Anhaltspunkte den Verdacht begründen, dass die betroffene Person im Beschäftigungsverhältnis eine Straftat begangen hat, die Verarbeitung zur Aufdeckung erforderlich ist und das schutzwürdige Interesse der oder des Beschäftigten an dem Ausschluss der Verarbeitung nicht überwiegt, insbesondere Art und Ausmaß im Hinblick auf den Anlass nicht unverhältnismäßig sind.

(2) Erfolgt die Verarbeitung personenbezogener Daten von Beschäftigten auf der Grundlage einer Einwilligung, so sind für die Beurteilung der Freiwilligkeit der Einwilligung insbesondere die im Beschäftigungsverhältnis bestehende Abhängigkeit der beschäftigten Person sowie die Umstände, unter denen die Einwilligung erteilt worden ist, zu berücksichtigen. Freiwilligkeit kann insbesondere vorliegen, wenn für die beschäftigte Person ein rechtlicher oder wirtschaftlicher Vorteil erreicht wird oder Arbeitgeber und beschäftigte Person gleichgelagerte Interessen verfolgen. Die Einwilligung bedarf der Schriftform, soweit nicht wegen besonderer Umstände eine andere Form angemessen ist. Der Arbeitgeber hat die beschäftigte Person über den Zweck der Datenverarbeitung und über ihr Widerrufsrecht nach Artikel 7 Absatz 3 der Verordnung (EU) 2016/679 in Textform aufzuklären.

(3) Abweichend von Artikel 9 Absatz 1 der Verordnung (EU) 2016/679 ist die Verarbeitung besonderer Kategorien personenbezogener Daten im Sinne des Artikels 9 Absatz 1 der Verordnung (EU) 2016/679 für Zwecke des Beschäftigungsverhältnisses zulässig, wenn sie zur Ausübung von Rechten oder zur Erfüllung rechtlicher Pflichten aus dem Arbeitsrecht, dem Recht der sozialen Sicherheit und des Sozialschutzes erforderlich ist und kein Grund zu der Annahme besteht, dass das schutzwürdige Interesse der betroffenen Person an dem Ausschluss der Verarbeitung überwiegt. Absatz 2 gilt auch für die Einwilligung in die Verarbeitung besonderer Kategorien personenbezogener Daten; die Einwilligung muss sich dabei ausdrücklich auf diese Daten beziehen. § 22 Absatz 2 gilt entsprechend.

(4) Die Verarbeitung personenbezogener Daten, einschließlich besonderer Kategorien personenbezogener Daten von Beschäftigten für Zwecke des Beschäftigungsverhältnisses, ist auf der Grundlage von Kollektivvereinbarungen zulässig. Dabei haben die Verhandlungspartner Artikel 88 Absatz 2 der Verordnung (EU) 2016/679 zu beachten.

(5) Der Verantwortliche muss geeignete Maßnahmen ergreifen, um sicherzustellen, dass insbesondere die in Artikel 5 der Verordnung (EU) 2016/679 dargelegten Grundsätze für die Verarbeitung personenbezogener Daten eingehalten werden.

(6) Die Beteiligungsrechte der Interessenvertretungen der Beschäftigten bleiben unberührt.

(7) Die Absätze 1 bis 6 sind auch anzuwenden, wenn personenbezogene Daten, einschließlich besonderer Kategorien personenbezogener Daten, von Beschäftigten verarbeitet werden, ohne dass sie in einem Dateisystem gespeichert sind oder gespeichert werden sollen.

(8) Beschäftigte im Sinne dieses Gesetzes sind:

1. Arbeitnehmerinnen und Arbeitnehmer, einschließlich der Leiharbeiterinnen und Leiharbeiter im Verhältnis zum Entleiher,
2. zu ihrer Berufsbildung Beschäftigte,

3. Teilnehmerinnen und Teilnehmer an Leistungen zur Teilhabe am Arbeitsleben sowie an Abklärungen der beruflichen Eignung oder Arbeitserprobung (Rehabilitandinnen und Rehabilitanden),
4. in anerkannten Werkstätten für behinderte Menschen Beschäftigte,
5. Freiwillige, die einen Dienst nach dem Jugendfreiwilligendienstgesetz oder dem Bundesfreiwilligendienstgesetz leisten,
6. Personen, die wegen ihrer wirtschaftlichen Unselbständigkeit als arbeitnehmerähnliche Personen anzusehen sind; zu diesen gehören auch die in Heimarbeit Beschäftigten und die ihnen Gleichgestellten,
7. Beamtinnen und Beamte des Bundes, Richterinnen und Richter des Bundes, Soldatinnen und Soldaten sowie Zivildienstleistende.

Bewerberinnen und Bewerber für ein Beschäftigungsverhältnis sowie Personen, deren Beschäftigungsverhältnis beendet ist, gelten als Beschäftigte.

EU Verordnung 2016/679 (Datenschutzgrundverordnung)

Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur **Aufhebung der Richtlinie 95/46/EG** (Datenschutz-Grundverordnung)

Ersetzt: *Richtlinie 95/46/EG des Europäischen Parlaments und des Rates vom 24. Oktober 1995 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr*

<https://www.datenschutzbeauftragter-info.de/eu-datenschutz-grundverordnung-und-beschaefigtendatenschutz/>

Art. 88, Datenverarbeitung im Beschäftigungskontext

(1) Die Mitgliedstaaten können durch Rechtsvorschriften oder durch Kollektivvereinbarungen spezifischere Vorschriften zur Gewährleistung des Schutzes der Rechte und Freiheiten hinsichtlich der Verarbeitung personenbezogener Beschäftigtendaten im Beschäftigungskontext, insbesondere für Zwecke der Einstellung, der Erfüllung des Arbeitsvertrags einschließlich der Erfüllung von durch Rechtsvorschriften oder durch Kollektivvereinbarungen festgelegten Pflichten, des Managements, der Planung und der Organisation der Arbeit, der Gleichheit und Diversität am Arbeitsplatz, der Gesundheit und Sicherheit am Arbeitsplatz, des Schutzes des Eigentums der Arbeitgeber oder der Kunden sowie für Zwecke der Inanspruchnahme der mit der Beschäftigung zusammenhängenden individuellen oder kollektiven Rechte und Leistungen und für Zwecke der Beendigung des Beschäftigungsverhältnisses vorsehen.

(2) Diese Vorschriften umfassen angemessene und besondere Maßnahmen zur Wahrung der menschlichen Würde, der berechtigten Interessen und der Grundrechte der betroffenen Person, insbesondere im Hinblick auf die Transparenz der Verarbeitung, die Übermittlung personenbezogener Daten innerhalb einer Unternehmensgruppe oder einer Gruppe von Unternehmen, die eine gemeinsame Wirtschaftstätigkeit ausüben, und die Überwachungssysteme am Arbeitsplatz.

(3) Jeder Mitgliedstaat teilt der Kommission bis zum 25. Mai 2018 die Rechtsvorschriften, die er aufgrund von Absatz 1 erlässt, sowie unverzüglich alle späteren Änderungen dieser Vorschriften mit.

§ 75 Abs. 2 Satz 1, Betriebsverfassungsgesetz

weitere?

2.3 SIEM-Systeme

SIEM-Systeme dienen dazu Daten in Netzwerken zu sammeln, um so einen zentralisierten Überblick über das Netzwerk zu erhalten und Bedrohungen erkennen und verhindern zu können.

Der Begriff *Security Information and Event Management* (SIEM) wurde von zwei Analysten des IT-Marktforschungsunternehmens Gartner geprägt, das auch jährlich einen Bericht über aktuelle Trends im Bereich der SIEM-Systeme veröffentlicht. Er setzt sich zusammen aus *Security Event Management* (SEM), das sich mit Echtzeitüberwachung und Ereigniskorrelation befasst, sowie *Security Information Management* (SIM), in dessen Fokus Langzeiterfassung und Analyse von Log-Daten steht [NK11].

Ein SIEM-System sollte nach [Det+15] die folgenden Aufgaben erfüllen können:

- **Event Correlation:** Da diese Aufgabe zentral für den Inhalt der Arbeit ist, wird sie unten stehend näher erläutert.
- **Network Behaviour Anomaly Detection:** Beschreibt die Erkennung von Anomalien auf Netzwerkebene durch die Erkennung von vom Normalzustand abweichenden Kommunikationsverhalten.
- **Identity Mapping:** Abbildung von Netzwerkadressen auf Nutzeridentitäten.
- **Key Performance Indication:** Zentrale Analyse sicherheitsrelevanter Informationen und Netzwerkdetails.
- **Compliance Reporting:** Überprüfung der Einhaltung von durch Regelungen vorgeschriebenen Anforderungen wie Integrität, Risiko und Effektivität.
- **API:** Bereitstellung von Schnittstellen zur Integration heterogener Systeme im Netzwerk.
- **Role based access control:** Zuständigkeitsabhängige Sichten auf sicherheitsrelevante Ereignisse.

Eine besondere Bedeutung im Kontext dieser Arbeit kommt hier der Behandlung von sicherheitsrelevanten Ereignissen (Events) zu, die beispielsweise von Intrusion-Detectionen-Systemen oder aus den Log-Daten von Firewalls, Switches oder anderen Netzwerkgeräten stammen können.

Um diese Ereignisse zu erhalten, muss ein SIEM-System nach [DRS14] vor ihrer Speicherung insbesondere drei Aufgaben wahrnehmen. Zu Beginn werden die Daten aus Logeinträgen oder empfangenen Systemmeldungen herausgelesen (Extraktion).

Anschließend müssen die extrahierten Daten in ein SIEM-spezifisches Format übersetzt werden, um eine sinnvolle Weiterverarbeitung zu gewährleisten (Homogenisierung). Hierbei werden relevante Felder eines SIEM-Events wie Datumsangaben, Adressen oder Aktionen aus den empfangenen Daten befüllt. Dieser Schritt wird in anderen Quellen auch als Normalisierung oder Mapping bezeichnet.

Optional können darauf folgend gleichartige Events in bestimmten Fällen anschließend zusammengefasst werden, um aussagekräftigere Informationen zu erhalten (Aggregation).

Liegen die Events nun in einem vorgegebenen Format im System vor, so können sie weiterhin mit dem System bekannten Umgebungsdaten über Benutzer, Geräte oder Bedrohungen verknüpft werden, um ihre Relevanz besser einschätzen zu können.

Anschließend lassen sich vorgegebene Regeln anwenden, um aus der Korrelation von Ereignissen aus verschiedenen Datenquellen auf eine Bedrohung schließen zu können, die in den einzelnen Events nicht erkennbar wäre (Event Correlation).

Wäre hier ein Beispiel für Event Correlation nötig/hilfreich?

2.4 Pseudonymisierung

Ein Pseudonym¹ bezeichnet nach [PH10] einen Identifikator eines Subjekts ungleich seinem echten Bezeichner. In §3, Absatz (6a) des BDSG wird zusätzlich noch der „Zweck [eines Pseudonyms], die Bestimmung des Betroffenen auszuschließen oder wesentlich zu erschweren“[**TODO**] ergänzt. Beispiele für Pseudonyme lassen sich in verschiedensten Bereichen finden: E-Mail-Adressen, Sozialversicherungsnummern oder auch Autoren, die unter einem Pseudonym ihre Schriften veröffentlichen.

Evtl. auch Telemediengesetz erwähnen

Es lassen sich nach [PWP90] unterschiedliche Arten von Pseudonymen unterscheiden. Eine Eigenschaft, die zur Unterscheidung herangezogen werden kann, ist die (Un-)Kenntnis des Zusammenhangs zwischen dem Pseudonym und dem zugehörigen Subjekt zu Beginn seiner Verwendung. Hier kann zwischen öffentlichen, nicht-öffentlichen und anonymen Pseudonymen unterschieden werden. Ein öffentliches Pseudonym ist beispielsweise die Telefonnummer einer Person, die von Beginn an im Telefonbuch mit der Identität der Person verknüpft ist. Eine Kontonummer zu einem Bankkonto, dessen Inhaber bei der Kontoeröffnung nur der Bank bekannt ist, bildet ein nicht-öffentliches Pseudonym. Ein anonymes Pseudonym bildet beispielsweise die DNA-Sequenz eines Menschen, die oftmals nicht einmal ihm selbst bekannt ist. Diese Eigenschaft eines Pseudonyms kann sich im Laufe der Zeit ändern, wenn Informationen über den Zusammenhang zwischen Pseudonym und zugehörigem Subjekt veröffentlicht werden.

Eng in Verbindung mit Pseudonymen steht auch der Begriff der Verkettbarkeit. Verkettbarkeit bezeichnet dabei die Eigenschaft, dass ein Außenstehender mit hoher Wahrscheinlichkeit entscheiden kann, ob zwei Objekte in einem System unabhängig sind[PH10].

Auf Basis dieser Eigenschaft lassen sich nun ebenfalls verschiedene Arten von Pseudonymen ausmachen: Diese Arten unterscheiden sich durch Nutzung des Pseudonyms in verschiedenen Kontexten. Personen-, Rollen-, Beziehungs-, Rollen-Beziehungs- und Transaktionspseudonym sind mögliche Ausprägungen dieser Eigenschaft. Eine Übersicht über diese Pseudonymarten und deren Zusammenhang zur Verkettbarkeit ist Abbildung 2.1 zu entnehmen.

Personenpseudonyme beschreiben Pseudonyme wie beispielsweise die Sozialversicherungsnummer, die stellvertretend für die eindeutige Identität des Subjekts in der Gesellschaft stehen. Ereignisse, die mit solchen Pseudonymen verbunden sind, lassen sich über die gesamte Gültigkeitsspanne des Pseudonyms verketteten.

Rollen-, Beziehungs- und Rollen-Beziehungspseudonyme stellen Pseudonyme da, die das Subjekt nur in einer besonderen Funktion oder mit einem bestimmten Kommunikationspartner bzw. in der Kombination beider Möglichkeiten nutzt. Ein Beispiel hierfür wäre eine E-Mail-Adresse wie *administrator@unternehmen.de* in einem Unternehmen, die nicht das Subjekt

1. ursprünglich aus dem Griechischen stammend: *pseudonumon* - falsch benannt

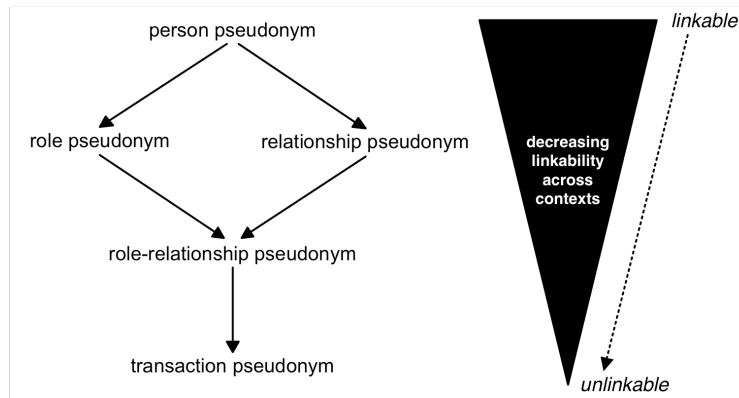


Abbildung 2.1: Pseudonym-Verband entsprechend ihrer Nutzung in verschiedenen Kontexten. Entnommen aus [PH10].

als solches sondern nur in seiner Rolle in dem Unternehmen beschreibt. Ereignisse verknüpft mit diesen Pseudonymarten lassen sich nur in bestimmten Kontexten, jedoch nicht über diese hinweg, verknüpfen. Beispielsweise könnten zwei Kommunikationspartner mit derselben Person kommunizieren ohne dies feststellen zu können, wenn die Person unter unterschiedlichen Beziehungspseudonymen auftritt.

Transaktionspseudonyme stellen die stärkste Form der Unverkettbarkeit da. Bei dieser Art von Pseudonymen wird für jedes Ereignis ein neues Pseudonym verwendet, das daher nur ein einziges Mal auftritt und Verkettbarkeit verhindert.

Abgrenzung zur Anonymität sinnvoll?

Ist der „Schnitt“ zwischen Kapitel 2 und 3 hier sinnvoll?

2.5 Schwellwertschemata

Mit der Verbreitung technischer Systeme, die kryptographische Verfahren nutzen, musste auch das Problem der sicheren Aufbewahrung und Verteilung kryptographischer Schlüssel betrachtet werden. Die Sicherheit dieser Schlüssel ist essentiell für den Betrieb solcher Systemen. Das einfache Speichern eines Schlüssels an einem Ort besitzt eine hohe Verlustwahrscheinlichkeit, da ein einzelner Fehler wie unbeabsichtigtes Löschen oder Speichermedienausfall den Schlüssel unwiederbringlich zerstören kann. Das mehrfache Speichern eines Schlüssels an verschiedenen Orten erhöht hingegen die Gefahr eines Schlüsseldiebstahls oder -missbrauchs, da auch die Angriffsfläche vergrößert wird. Bei möglichen Lösungen dieses Problems müssen also immer die Integrität und die Vertraulichkeit eines Schlüssels gegeneinander abgewogen werden. [Gem97]

Ausgehend von diesen Überlegungen entwickelte Shamir das erste (k, n) -Schwellwert-Schema²: Ein Geheimnis D wird so in n Teile D_1, \dots, D_n zerlegt, dass durch Kenntnis von mindestens k Teilen das Geheimnis wieder aufgedeckt werden kann, aber jede Kombination aus höchstens

2. Im selben Jahr veröffentlichte auch Blakley eine Lösung dieses Problems, die auf den Schnittpunkten von Hyperebenen über endlichen Feldern beruht [Bla79].

$k - 1$ Teilen keine Informationen über D liefert. Shamirs Lösung bediente sich der Polynominterpolation auf der Basis modularer Arithmetik [Sha79]:

Die Menge aller Ganzzahlen modulo einer Primzahl p bilden einen Körper, in welchem Polynominterpolation möglich ist. Soll nun das Geheimnis D (das o.B.d.A. als Ganzzahl angenommen wird) aufgeteilt werden, so wird eine Primzahl p mit $p > D$ und $p > n$ gewählt. Weiterhin wird ein Polynom

$$q(x) = a_0 + a_1x + \dots + a_{k-1}x^{k-1} \text{ mit } a_0 = D$$

derart gewählt, dass a_1, \dots, a_{k-1} gleichverteilt aus der Menge $\{0, \dots, p\}$ stammen. Die einzelnen Teile des Geheimnisses werden nun als

$$D_1 = q(1), \dots, D_i = q(i), \dots, D_n = q(n)$$

jeweils modulo p berechnet.

Um nun aus diesen einzelnen Teilen wieder das ursprüngliche Geheimnis zu erhalten, wird das Verfahren der Polynomdivision verwendet.

Erweiterte Beschreibung

Auf Basis dieses Verfahrens kann also die Integrität eines Schlüssels erhöht werden, da nun selbst bei Verlust von $n - k$ Teilen der Schlüssel noch wiederhergestellt werden kann. Auf der anderen Seite ist die Vertraulichkeit jedoch höher als bei der mehrfachen Speicherung des Schlüssels im Original, da mindestens k Teile des Schlüssels zur Wiederherstellung vorliegen müssen.

Das Problem dieser Lösungen bezogen auf den in dieser Arbeit behandelten Anwendungsfall ist jedoch, dass das Geheimnis nach erstmaligem Aufdecken bekannt ist. Wünschenswert wäre ein Verfahren, bei dem nur ein entsprechend verschlüsseltes Datum (bspw. der gesuchte Eintrag in einer Pseudonym-Tabelle) aufgedeckt werden kann, ohne dass der kombinierte Schlüssel selbst bekannt wird.

In [Des87] wird dieses Problem das erste Mal im Kontext von verschlüsselten Nachrichten an Gruppen betrachtet: Ein Sender möchte eine Nachricht an eine Gruppe von Empfängern senden, die nur in Zusammenarbeit die Nachricht entschlüsseln können sollen. Hier wird auch die zentrale Forderung aufgestellt, den mehrfachen Nachrichtenaustausch zwischen Sender und Empfänger(n) bei der Entschlüsselung (sogenannte Ping-Pong-Protokolle) zu vermeiden.

In [Des93] spricht der Autor bei dieser Klasse von Verfahren von *threshold decryption* und fordert weiterhin, dass praktisch einsatzbare Systeme auch *non-interactive* sein sollten, also bei der Entschlüsselung keinen aufwendigen Datenaustausch zwischen den Teilnehmern notwendig machen.

In [BBH06] werden diese Systeme formalisiert. Ein *Threshold-Public-Key-Encryption-System* besteht aus fünf Schritten:

EZ zu Schritt 3:
 $u_i = (i, \dots) ???$

1. $Setup(n, k, \Lambda)$ liefert ein Tripel (PK, VK, SK) , bestehend aus dem öffentlichen Schlüssel PK , einem Verifikationsschlüssel VK und einer n -elementigen Liste aus *Private Key Shares*, von denen jeder Teilnehmer einen *Share* erhält. Λ wird als initialer Sicherheitsparameter bezeichnet.
2. $Encrypt(PK, M)$ liefert die verschlüsselte Nachricht C .
3. $ShareDecrypt(PK, i, SK_i, C)$ liefert ein *Decryption Share* $\mu = (i, \mu^i)$ des i -ten Teilnehmers, das im 5. Schritt zusammen mit weiteren *Shares* zur Entschlüsselung der Nachricht genutzt wird.

4. $\text{ShareVerify}(PK, VK, C, \mu)$ überprüft ein *Decryption Share* auf Validität.
5. $\text{Combine}(PK, VK, C, \mu_1, \dots, \mu_k, \dots)$ verknüpft die *Decryption Shares* von mindestens k Teilnehmern und liefert die Nachricht M zurück.

Anforderungen an diese Schritte sind [BBH06] zu entnehmen.

Hier ergänzen.

3 Stand der Wissenschaft und Auswahl von Verfahren

3.1 SIEM-Systeme

- Was gibt es evtl. gerade im Bereich der Insider-Angriffe schon?
- Was ist OSSIM und wie sieht die Architektur aus?

3.1.1 OSSIM

Eine quelloffene SIEM-Lösung, die im Rahmen dieser Arbeit genutzt werden wird, ist OSSIM, ein SIEM-System der Firma AlienVault, das auf Basis weiterer quelloffener Lösungen aus dem Netzwerksicherheits-Bereich unter anderem die oben genannten Funktionen bereitstellt.¹

3.2 Pseudonymisierung

- A name or another bit string. Identifiers, which are generated using random data only, i.e., fully independent of the subject and related attribute values, do not contain side information on the subject they are attached to, whereas non-random identifiers may do. E.g., nicknames chosen by a user may contain information on heroes he admires; a sequence number may contain information on the time the pseudonym was issued; an e-mail address or phone number contains information how to reach the user. [PH10]

- evtl. Common Criteria

- VANET

- TIMSI

Der Begriff der Pseudonymisierung beschreibt die Benutzung von Pseudonymen zur Identifizierung von Subjekten. Pseudonymisierung sagt dabei erst einmal lediglich etwas über die Verwendung eines Verfahrens aus, jedoch nichts über die daraus entstehenden Auswirkungen auf die Identifizierbarkeit eines Subjekts oder die Zurechenbarkeit bestimmter Aktionen. Hierfür spielen weitere Eigenschaften von Pseudonymen wie die folgenden eine Rolle:

- garantierte Eindeutigkeit von Pseudonymen
- Möglichkeit von Pseudonymänderungen
- begrenzt häufige Verwendung von Pseudonymen
- zeitlich begrenzte Verwendung von Pseudonymen

1. AlienVault OSSIM: The World's Most Widely Used Open Source SIEM
<https://www.alienvault.com/products/ossim>

- Art der Pseudonymserstellung

Die Ausprägungen dieser Eigenschaften werden auch im Rahmen dieser Arbeit für das umzusetzende System zu bewerten sein.

3.3 Schwellwertschemata

- **Grundlagen** Von Shamirs Secret Sharing bis heute
- **Konkretes Verfahren** Im Detail erläutern - vmtl. Desmedt auf Basis von ElGamal
- Zusätzlich Erweiterungen dieses Verfahrens wie verteilte Schlüsselgenerierung.

Ein solches System, das auf dem ElGamal-Algorithmus und damit dem Diskreten-Logarithmus-Problem basiert, veröffentlichten die Autoren in [DF90]. Dieser Ansatz setzt in der *Setup*-Phase auf eine zentrale vertrauenswürdige Stelle zur Erzeugung der Schlüssel und *Shares*. In [Ped91] stellt der Autor basierend auf diesen Ergebnissen ein Verfahren vor, das bei der Schlüsselgenerierung ohne eine vertrauenswürdige Instanz auskommt. Dieses Verfahren wird in [Gen+99] noch einmal verbessert.

Basierend auf dem jetzigen Recherchestand würde sich diese Kombination von Verfahren gut für den angestrebten Anwendungszweck eignen. Konkrete offene Implementierungen wurden jedoch bisher nicht gefunden, so dass möglicherweise eine eigene Implementierung umgesetzt werden muss.

Neben diesem Verfahren gibt es noch weitere Ansätze basierend auf RSA [Des93; Ngu05] oder dem Paillier-Kryptosystem [Pai99; DJ01], die jedoch deutlich komplexer zu sein scheinen.

Details

TBW

4 Implementierung

- **Anforderungen** Welche Anforderungen muss der Prototyp erfüllen?
- **Umsetzung** Architektur und Implementierung des Prototypen.
- **Einbindung in OSSIM** An welcher Stelle wird er in das SIEM-System integriert? Abwägung der verschiedenen Möglichkeiten.
- **Umsetzung des Schwellwertschemas** Implementierung des Schwellwertschemas und Integration in den Prototypen. Hier auch Key Management.

4.1 Anforderungen

4.2 Architektur und Umsetzung

4.3 Einbindung in OSSIM

4.4 Umsetzung der Pseudonymisierung

4.5 Implementierung und Integration des Schwellwertschemas

4.6 Evaluation

Inklusive Problemen, ...

5 Alternativen

- **Alternativen** Welche alternativen oder ergänzenden Vorgehensweisen zu Pseudonymisierung + kryptographisches Schwellwertschema gibt es und welche Eigenschaften, Vor- und Nachteile besitzen sie?
- **Umsetzung** Wie könnten diese Alternativen im Prototypen umgesetzt werden?

5.1 Alternative1

5.2 ...

5.3 Vorgehen zur Integration

6 Fazit

TBW

Literatur

- [BBH06] Dan Boneh, Xavier Boyen und Shai Halevi. *Chosen ciphertext secure public key threshold encryption without random oracles*. In: *Topics in Cryptology – CT-RSA 2006*. Hrsg. von David Pointcheval. Springer, 2006, S. 226–243.
- [bit16] bitkom. *Spezialstudie Wirtschaftsschutz*. bitkom. 2016. URL: <https://www.bitkom.org/Bitkom/Publikationen/Spezialstudie-Wirtschaftsschutz.html> (besucht am 24.06.2017).
- [Bla79] George Robert Blakley. *Safeguarding cryptographic keys*. In: *Proceedings of the AFIPS 1979 National Computer Conference*. Hrsg. von Richar E. Merwin. AFIPS Press, 1979, S. 313–317.
- [Bra+15] Nicholas Bradley u. a. *IBM 2015 Cyber Security Intelligence Index*. IBM. 2015. URL: <http://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=SEW03073USEN> (besucht am 24.06.2017).
- [Des87] Yvo Desmedt. *Society and Group Oriented Cryptography: a New Concept*. In: *Conference on the Theory and Application of Cryptographic Techniques 1987*. Hrsg. von Carl Pomerance. Springer, 1987, S. 120–127.
- [Des93] Yvo Desmedt. *Threshold cryptosystems*. In: *Advances in Cryptology — AUS-CRYPT '92*. Hrsg. von Jennifer Seberry und Yuliang Zheng. Springer, 1993, S. 1–14.
- [Det+15] Kai-Oliver Detken u. a. *SIEM approach for a higher level of IT security in enterprise networks*. In: *2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*. IEEE. 2015, S. 322–327.
- [DF90] Yvo Desmedt und Yair Frankel. *Threshold cryptosystems*. In: *Advances in Cryptology - CRYPTO' 89 Proceedings*. Hrsg. von Gilles Brassard. Springer, 1990, S. 307–315.
- [DJ01] Ivan Damgard und Mads Jurik. *A generalisation, a simplification and some applications of Paillier's probabilistic public-key system*. In: *Public Key Cryptography – Proceedings of the 4th International Workshop on Practice and Theory in Public Key Cryptosystems*. Hrsg. von Kwangjo Kim. Bd. 1992. Springer, 2001, S. 119–136.
- [DRS14] Kai-Oliver Detken, Thomas Rossow und Ralf Steuerwald. *SIEM-Ansätze zur Erhöhung der IT-Sicherheit auf Basis von IF-MAP*. In: *Proceedings of the DACH Security 2014: Bestandsaufnahme, Konzepte, Anwendungen und Perspektiven*. Hrsg. von Peter Schartner und Peter Lipp. syssec, 2014.
- [Gem97] Peter Gemmell. *An introduction to threshold cryptography*. In: *CryptoBytes, a technical newsletter of RSA Laboratories* 2.7 (1997), S. 295–310.

- [Gen+99] Rosario Gennaro u. a. *Secure distributed key generation for discrete-log based cryptosystems*. In: *Advances in Cryptology — EUROCRYPT'99*. Hrsg. von Jaques Stern. Springer, 1999, S. 295–310.
- [Ngu05] H.L. Nguyen. *RSA Threshold Cryptography*. 2005. URL: <https://www.cs.ox.ac.uk/files/269/Thesis.pdf> (besucht am 02. 07. 2017).
- [NK11] Mark Nicolett und Kelly M Kavanagh. *Magic quadrant for security information and event management*. In: *Gartner Research Note G00212454* (2011).
- [Pai99] Pascal Paillier. *Public-key cryptosystems based on composite degree residuosity classes*. In: *Advances in Cryptology — EUROCRYPT'99*. Hrsg. von Jaques Stern. Springer, 1999, S. 223–238.
- [Ped91] Torben Pedersen. *A threshold cryptosystem without a trusted party*. In: *Advances in Cryptology — EUROCRYPT'91*. Hrsg. von Donald W. Davies. Springer, 1991, S. 522–526.
- [PH10] Andreas Pfitzmann und Marit Hansen. *A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management*. 2010. URL: http://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.34.pdf (besucht am 27. 06. 2017).
- [PWP90] Birgit Pfitzmann, Michael Waidner und Andreas Pfitzmann. *Rechtssicherheit trotz Anonymität in offenen digitalen Systemen*. In: *Datenschutz und Datensicherung DuD 14.5-6* (1990), S. 243–253.
- [Sha79] Adi Shamir. *How to share a secret*. In: *Communications of the ACM 22.11* (1979), S. 612–613.