Spieltheoretische Risikominimierung in IKT-Infrastrukturen

Stefan Schauer[1] · Sandra König[1]
Stefan Rass[2] · Martin Latzenhofer[1]

[1]AIT Austrian Institute of Technology GmbH
{stefan.schauer | sandra.koenig | martin.latzenhofer}@ait.ac.at

[2]Alpen-Adria-Universität Klagenfurt
stefan.rass@aau.at

Zusammenfassung

In diesem Artikel wird ein qualitativer Ansatz zur Risikobewertung vorgestellt, der auf kontinuierlich aktualisierten öffentlichen Informationen zu Schwachstellen von IKT-Systemen aufbaut. Über frei definierbare kategoriale Skalen werden diese Informationen aggregiert und dadurch Einschätzungen zur Verwundbarkeit und dem potentiellen Schaden im Falle eines Angriffs generiert. Durch den Einsatz von spieltheoretischen Methoden, welche mit diesen qualitativen Werten arbeiten können, werden potentielle Angriffsszenarien evaluiert und eine optimale Auswahl aus bestehenden Verteidigungsstrategien zur Minimierung des aktuellen Risikos identifiziert. Durch die Möglichkeit eines kontinuierlichen Updates der Informationen zu den Schwachstellen sowie zum betrachteten System kann eine momentane Risikoeinschätzung der IKT Infrastruktur stets schnell und effektiv aktualisiert werden.

1 Einleitung

Die Bewertung und vor allem die Minimierung von Risiken stellt ein zentrales Thema im Bereich der Informations- und Kommunikationstechnologie (IKT) dar. Insbesondere die stetigen Erneuerungen von (Software-)Komponenten innerhalb einer IKT-Infrastruktur und der damit verbundene kontinuierliche Wechsel von potentiellen Angriffsmöglichkeiten bringen aktuelle Ansätze aus dem Risikomanagement bisweilen an ihre Grenzen. Um eine adäquate Risikobewertung durchführen und entsprechende Maßnahmen zur Risikominimierung treffen zu können, sollten Informationen über die Komponenten (Assets) der IKT-Infrastruktur und ihren potentiellen Schwachstellen immer am neuesten Stand sein.

In der Praxis werden unterschiedliche Tools, wie etwa OpenVAS (http://www.openvas.org/index.de.html), Nessus (https://www.tenable.com/products/nessus-vulnerability-scanner) oder MetaSploit (https://www.metasploit.com), eingesetzt, um Schwachstellen in den Software-Komponenten einer IKT-Infrastruktur zu identifizieren. Hierfür bedienen sich diese Tools zumeist frei zugänglicher Informationen zu bestehenden Schwachstellen, welche z.B. in der National Vulnerability Database (NVD) der National Institute of Standards and Technology (NIST) gesammelt werden, und erstellen ein Ranking in Bezug auf den Schweregrads der einzelnen Schwachstellen. Eine Aussage über das Risiko einzelner Komponenten – sei es qualitativ oder quantitativ – können diese Tools aber nicht geben.

P. Schartner (Hrsg.) · D•A•CH Security 2016 · syssec (2016) 174-187.

Eine solche Risikobewertung ist jedoch zentraler Teil von etablierten Risikomanagement-Rahmenwerken, wie etwa der ISO 31000 [Inte09], der ONR 49000 [Aust14], COBIT for Risk [Info13] oder ähnlichen, wobei in diesen Rahmenwerken zumeist nur die Rahmenbedingungen für eine Risikobewertung beschrieben sind, jedoch wird keine konkrete Vorgehensweise vorgeschrieben. Hierfür müssen spezifische Methoden in den Organisationen entwickelt und umgesetzt werden. Dabei wird meist auf eine qualitative Bewertung von Komponenten und Schwachstellen durch Experten gesetzt, etwa anhand einer stufenweisen Skala „sehr niedrig", bis „sehr hoch". Solche Bewertungen sind in vielen Fällen partiell subjektiv, aber dennoch zuweilen unter Unsicherheit formuliert. Die Verwendung von frei zugänglichen Informationen zu Schwachstellen verringert hier sowohl die Unsicherheit als auch die Subjektivität.

Neben den etablierten Vorgehensmodellen werden in den letzten Jahren verstärkt Elemente der Spieltheorie im Bereich der IKT-Sicherheit und des IKT-Risikomanagements eingesetzt [AlBa10, RaSn11, Rass13, RSPG13, RRVG15]. Dabei wird die „natürliche" Konfliktsituation zwischen einem Angreifer (sei es ein menschlicher Akteur oder die Natur als zufälliger äußerer Einfluss) und einem Verteidiger direkt durch spieltheoretische Konzepte modelliert. Daraus ergibt sich der signifikante Vorteil, dass durch spieltheoretische Algorithmen für eine gegebene Situation optimale Angriffs- und Verteidigungsstrategien identifiziert werden können. Dabei gehen die meisten bestehenden Ansätze davon aus, dass eine präzise Bewertung des Risikos in Form des verursachten Schadens möglich ist. In der Praxis ist es jedoch oft schwierig den Schaden zu prognostizieren so dass die Modellierung durch eine Wahrscheinlichkeitsverteilung angebracht ist [ScKR15]. Hier knüpft dieser Artikel an und stellt einen Ansatz zur qualitativen Risikobewertung auf Basis spieltheoretischer Methoden und unter Verwendung öffentlich verfügbarer Informationen vor. Durch diese Informationen sollen Einschätzungen der Risikomanagementexperten einer Organisation automatisiert unterstützt werden. Darüber hinaus werden die kontinuierlichen Updates dieser Quellen genutzt, um damit das Modell auf den aktuellsten Informationen aufzubauen. Zusätzlich können diese Informationen aber auch frei mit weiteren Experteneinschätzungen kombiniert oder durch diese ersetzt werden. In diesem Kontext ermöglicht die Anwendung spieltheoretischer Modelle eine Minimierung des aktuellen Risikos in einer IKT-Infrastruktur. Diese Modelle bauen dabei auf den qualitativen Informationen aus den Expertenmeinungen auf und berücksichtigen ebenfalls potentielle Unsicherheiten in Bezug auf diese Informationen.

Der nächste Abschnitt beschreibt ausführlicher, wie Software-Schwachstellen anhand des Common Vulnerability Scoring Systems (CVSS) strukturiert bewertet werden. Abschnitt 3 demonstriert, wie die Informationen aus dem CVSS für eine qualitative Risikobewertung weiterverwendet werden können. Dabei wird vor allem auf die Einschätzung der Plausibilität eines Angriffs und des dabei zu erwartenden Schadens anhand von vorliegenden Informationen eingegangen. In Abschnitt 4 wird gezeigt, wie Methoden der Spieltheorie eingesetzt werden können, um aus den qualitativen Informationen Handlungsoptionen zur Risikominimierung auf Basis von bestehenden Verteidigungsstrategien zu entwickeln. Ein Beispiel illustriert in Abschnitt 5 die konkrete Anwendung des Ansatzes.

2 Exploitability und Impact Bewertung

Das *Common Vulnerability Scoring System* (CVSS) ist ein Modell zur Charakterisierung von Schwachstellen in Softwaresystemen [MeSR07a, MeSR07b], welches von dem „Forum of Incident Response and Security Teams" (FIRST – https://www.first.org) ins Leben gerufen wurde. Ziel des CVSS ist es, einzelne Schwachstellen anhand von unterschiedlichen Metriken

zu beschreiben und diese entsprechend eines vorgegebenen Scorings zu gewichten. Die Einschätzungen werden von Experten des FIRST und anderer Organisationen durchgeführt und öffentlich in diversen Datenbanken (wie etwa der National Vulnerability Database (NVD – https://nvd.nist.gov) des NIST) zur Verfügung gestellt. Somit kann anhand dieser Scorings eine IKT-Infrastruktur untersucht und eine Prioritätenliste der schwerwiegendsten Schwachstellen erstellt werden. Auf Basis dieser Liste können in weiterer Folge auch Gegenmaßnahmen identifiziert und koordiniert werden.

Für die Bewertung von Schwachstellen verwendet CVSS drei unterschiedliche Gruppen von Metriken [MeSR07a]: *Base Metric*, *Temporal Metric* und *Environmental Metric*. Hierbei stellt die Base Metric die zentrale Gruppe des CVSS dar und umfasst fundamentale Charakteristika einer Software-Schwachstelle, welche im Laufe der Zeit gleich bleiben und unabhängig vom Einsatzbereich sind. Im Gegensatz dazu beschreibt die Temporal Metric Eigenschaften, welche sich im Lauf der Zeit verändern können, aber innerhalb desselben Umfeldes gleich bleiben. Der Einfluss der Umgebung, in dem sich eine Schwachstelle befindet, wird in der Environmental Metric behandelt. Nachdem die Temporal und Environmental Metric je nach Einsatzgebiet stark variieren können, gibt es hierzu keine Experteneinschätzungen in den öffentlichen Datenbanken. Daher beschränkt sich dieser Artikel im verbleibenden Teil auf die Base Metric.

Die Base Metric ist nochmals in zwei Untergruppen *Exploitability* und *Impact* aufgeteilt. Die Untergruppe Exploitability umfasst drei Kategorien, welche eine Abschätzung darüber ermöglichen, wie einfach die Schwachstelle ausgenutzt werden kann. Zu diesen Kategorien gehören:

- der *Access Vector* (AV) als Indikator für die Sichtbarkeit der Schwachstelle und die Zugriffsmöglichkeit auf die Schwachstelle von außen
- die *Access Complexity* (AC) als Indikator für die Schwierigkeit, die Schwachstelle erfolgreich auszunützen
- die *Authentication* (Auth) als Indikator, wie oft eine Authentifizierung durchgeführt werden muss, um die Schwachstelle auszunutzen

Jeder dieser drei Kategorien hat wiederum drei Levels und einen spezifischen Score, der mit jedem einzelnen Level verbunden ist. Für den Access Vector beschreiben diese Levels, ob ein Angreifer einen physischen Zugriff oder einen lokalen Account am verwundbaren System hat (Level *Local*), ob sich der Angreifer in derselben Collision Domain befindet (Level *Adjacent Network*), oder ob er lediglich remote auf das verwundbare System zugreifen kann (Level *Network*).

Für die Kategorie der Access Complexity beschreiben die drei Levels *High*, *Medium* und *Low*, ob sehr spezielle Rahmenbedingungen, nur einige zusätzliche oder gar keine Anforderungen für die Ausnützung der Schwachstelle existieren. Je höher der Level der Access Complexity, desto höher auch die Schwierigkeit, diese Schwachstelle auszunutzen. Die dritte Kategorie Authentication ist ein Indikator dafür, ob für die Ausnutzung einer Schwachstelle mehrere Authentifizierungsschritte (Level *Multiple*), nur ein einziger Schritt (Level *Single*) oder gar keine Authentifizierung (Level *None*) notwendig ist.

Der Score für die Exploitability wird aus den entsprechenden Scores der einzelnen Levels für den Access Vector, die Access Complexity und die Authentication berechnet (siehe z.B. [MeSR07a] für Details zur Berechnung).

In der zweiten Untergruppe *Impact* werden die Auswirkungen einer Schwachstelle auf das System beschrieben [MeSR07a]. Hierbei unterscheidet CVSS zwischen den drei klassischen Sicherheitsaspekten

- *Confidentiality* (C), also der Vertraulichkeit von Informationen
- *Integrity* (I), also der Integrität von Daten oder eines gesamten Systems
- *Availability* (A), also der Verfügbarkeit eines System (oder einzelner Funktionen)

Auch bei der Untergruppe Impact ist jeder dieser drei Sicherheitsaspekte wiederum in drei Levels aufgeteilt, welche jeweils mit einem spezifischen Score verbunden sind. Diese drei Levels sind für alle drei Kategorien gleich benannt: *None*, *Partial* und *Complete*.

Der Level *None* bedeutet, dass ein Ausnutzen der Schwachstelle keine Auswirkungen auf die Vertraulichkeit, Integrität oder Verfügbarkeit von Informationen oder Daten des Systems hat. Besteht ein beträchtlicher, jedoch eingeschränkter Verlust, so ist dies in allen drei Kategorien durch den Level *Partial* beschrieben. Kann das Ausnutzen einer Schwachstelle einen totalen Verlust an Vertraulichkeit, Integrität oder Verfügbarkeit von Informationen oder Daten des Systems zur Folge haben, so ist der Level *Complete* erreicht.

Ebenso wie bei der Berechnung der Exploitability werden auch bei der Bestimmung des Impacts einer Schwachstelle die Scores des entsprechenden Levels heran gezogen. Der gesamte Impact wird dann ebenfalls mittels einer vordefinierten Formel berechnet und mit dem Ergebnis der Exploitability kombiniert (Details zur Berechnung sind in [MeSR07a] dargestellt).

Wie bereits angemerkt, werden die CVSS-Scores für bestehende Schwachstellen in Softwaresystemen im Internet auf mehreren Seiten, wie etwa der NVD, zur Verfügung gestellt. Hierbei orientiert man sich an den *Common Vulnerabilities und Exposures* (CVE), einer strukturierten Auflistung von aktuell bekannten Schwachstellen von Cybersecurity-Systemen, um eine einheitliche Namensgebung der Schwachstellen und somit eine Vergleichbarkeit in verschiedenen Systemen zu gewährleisten. Zudem ist es durch die Sammlung der Scores von Schwachstellen in einer zentralen Datenbank möglich, diese Informationen tagesaktuell abzurufen und für den Einsatz von Analyse-Software (Vulnerability Scanner, Topological Vulnerability Analysetools, etc.) zu verwenden.

Nachdem CVSS lediglich ein Ranking über die Schweregrade von Schwachstellen erlaubt, kann der Einsatz von CVSS-Scores nur im Rahmen qualitativer Risikoeinschätzungen erfolgen, zumal quantitative Aussagen (etwa Wahrscheinlichkeiten) hieraus nicht formal ableitbar sind. Die Scores für die einzelnen Levels, wie in [MeSR07a, MeSR07b] beschrieben, sind speziell auf die Berechnung des CVSS-Scores zugeschnitten und können somit nur bedingt weiterverwendet werden. Zudem kann nicht festgestellt werden, auf welcher Basis die Bewertung einzelner Schwachstellen ausgewählt wurde – es ist lediglich bekannt, dass hierfür Expertenmeinungen herangezogen werden. In diesem Fall könnte nur die qualitative Information (welcher Level ausgewählt wurde) anstatt der spezifischen Scores für eine weitere Verwendung verwendet werden. Wie diese qualitativen Informationen genutzt werden können, wird im nächsten Abschnitt genauer beschrieben.

3 Qualitative Bewertung von Schwachstellen

Die beiden in Abschnitt 0 beschriebenen Gruppen der Base Metric aus dem CVSS – Exploitability und Impact – geben einen Aufschluss über die Schwierigkeit, eine Schwachstelle auszunutzen, und die Auswirkungen, die eine erfolgreiche Ausnutzung auf die Sicherheitsaspekte

Confidentiality, Integrity und Availability hat. Um auf dieser Grundlage spieltheoretisches Risikomanagement zu betreiben, müssen in einem ersten Schritt die einzelnen Kategorien auf eine einheitliche, qualitative Skala gebracht werden. Für die weitere Betrachtung wird hier eine fünfteilige Skala gewählt, welche durch die Stufen „Very Low“ (VL), „Low“ (L), „Moderate“ (M), „High“ (H) und „Very High“ (VH) beschrieben wird. Alternativ kann auch eine feinere Granularität oder eine andere Benennung der Stufen verwendet werden.

3.1 Bewertung der Verwundbarkeit

Im Bereich der CVSS Exploitability wird nun eine Abbildung von den drei Kategorien *Access Vector* (AV), *Access Complexity* (AC) und *Authentication* (Auth) auf diese fünfteilige Skala definiert (siehe Tabelle 1 für ein Beispiel einer solchen Abbildung).

Tab. 1: Mapping der CVSS Exploitability auf eine Verwundbarkeitsskala

AV	Local			Adjacent			Network		
AC / Auth	High	Medium	Low	High	Medium	Low	High	Medium	Low
Multiple	VL	VL	L	L	L	M	M	M	H
Single	VL	L	M	L	M	H	M	H	VH
None	L	M	M	M	H	H	H	VH	VH

Somit hätte eine Schwachstelle, die remote von überall erreichbar ist („Access Vector“ ist „Network“), keine speziellen Einschränkungen beim Zugang hat („Access Complexity“ ist „Low“) aber zumindest eine korrekte User/Passwort-Kombination zur Authentifizierung benötigt, eine Verwundbarkeit vom Level „Very High“ (VH). Ist die gleiche Schwachstelle nur mit einem lokalen Account ausnutzbar („Access Vector“ ist „Local“), so verringert sich die Verwundbarkeit auf das Level „Moderate“ (M).

3.2 Bewertung des Auswirkungen

Auf ähnliche Weise kann auch der CVSS Impact auf rein qualitative, fünfteilige Skala abgebildet werden. Dafür werden die Kategorien „None“, „Partial“ und „Complete“ der drei Sicherheitsziele *Confidentiality* (C), *Integrity* (I) und *Availability* (A) verwendet (siehe Tabelle 2 für ein Beispiel einer solchen Abbildung). Daraus leitet sich ab, dass eine Schwachstelle, die einen eingeschränkten Verlust an Vertraulichkeit („Confidentiality“ ist „Partial“) und Integrität („Integrity“ ist „Partial“) der Daten und keinen Verlust der Verfügbarkeit („Availability“ ist „None“) zur Folge hat, eine gesamte Auswirkung von „Low“ (L) aufweist. Bei einem totalen Verlust der Vertraulichkeit oder der Integrität der Informationen steigt der Level der gesamten Auswirkung auf „Moderate“ (M).

Tab. 2: Mapping des CVSS Impact auf eine Auswirkungsskala

C	None			Partial			Complete		
I / A	None	Partial	Complete	None	Partial	Complete	None	Partial	Complete
None	VL	VL	L	L	L	M	M	M	H
Partial	VL	L	M	L	M	H	M	H	VH
Complete	L	M	M	M	H	H	H	VH	VH

Für beide Skalen gilt, dass bei dieser Betrachtung alle Kategorien dieselbe Bedeutung haben, also die gleiche Gewichtung besitzen. Wenn einzelne Kategorien bei der Bewertung ein höheres Gewicht haben sollen als andere, z.B. weil die Zugriffsmöglichkeit von außen für die generelle Verwundbarkeit eine höhere Bedeutung hat oder die Vertraulichkeit von Informationen für die gesamte Auswirkung wichtiger ist, wird die Zuordnung entsprechend anders aussehen. Zusätzlich bleibt anzumerken, dass die Beispiele in Tabelle 1 und Tabelle 2 gegebenenfalls auf die Anforderungen eines konkreten Anwendungsfalls angepasst werden müssen. Somit können spezifische Rahmenbedingungen einer Organisation in diese Abbildung auf die gemeinsame Skala einfließen, z.B. wenn der Verlust der Verfügbarkeit höhere Auswirkungen hat als der Verlust von Vertraulichkeit oder Integrität.

Der zentrale Vorteil einer gemeinsamen Risiko-Bemessungsskala stellt die Vergleichbarkeit mit anderen (etwa bereits im Unternehmen bestehenden) Ansätzen dar. Durch die rein qualitativen Einschätzungen können die Informationen über Schwachstellen, welche z.B. automatisiert aus der NVD oder über ein entsprechendes Tool erhoben werden, mit Expertenmeinungen, die auf einer subjektiven Einschätzung zwischen „Very Low“ und „Very High“ basieren, verglichen werden. Auch die Einbindung von anderen Quellen, die Schwachstellen nicht anhand des CVSS-Schemas bewerten, ist dadurch möglich.

Auch im Bereich der Auswirkungen bietet eine qualitative Einschätzung den Vorteil, dass die zugrunde liegenden Kategorien auf die speziellen Anforderungen eines Unternehmens zugeschnitten werden können. So kann die Skala bei einer Organisation rein textuell beschrieben werden, wobei ein anderes Unternehmen konkrete monetäre Beschreibungen (mit unteren und oberen Grenzen) für die einzelnen Risikostufen definieren kann. Entsprechend kann auch ein Impact der Kategorie „High“ für ein KMU etwas anderes bedeuten, als für einen international agierenden Konzern.

3.3 Einschätzung der Plausibilität eines Angriffs

Die Abbildungen in Tabelle 1 und Tabelle 2 geben zwar eine Einschätzung über die Verwundbarkeit bzw. die potentiellen Auswirkungen von Schwachstellen, allerdings muss für eine Risikobewertung auch das Potential des Angreifers berücksichtigt werden. Je nach Angreifer ist die Wahrscheinlichkeit, eine bestimmte Schwachstelle erfolgreich auszunutzen, unterschiedlich groß. Einem versierten Angreifer mit entsprechenden Ressourcen wird es mit höherer Wahrscheinlichkeit gelingen, auch komplexere Schwachstellen auszunutzen, als etwa ein „Script Kiddie“ mit relativ eingeschränkten Ressourcen. Die NIST spezifiziert in der Special Publication SP 800-30 [Nati12] eine Menge von Charakteristika, die einen Angreifer beschreiben und auf seine Fähigkeiten, seine Absichten und seine Hintergründe eingehen. Für die Betrachtung in diesem Artikel wird vorerst nur eine Einschätzung zu den generellen Fähigkeiten eines Angreifers herangezogen (siehe Tabelle 3). Weitere Informationen können natürlich ebenso in das Modell mit aufgenommen werden.

Um nun eine Aussage darüber treffen zu können, wie plausibel eine erfolgreicher Angriff eines Gegners auf eine bestimmte Schwachstelle ist, müssen die Fähigkeiten des Angreifers und die Verwundbarkeit der Schwachstelle in Relation zueinander gebracht werden (siehe Tabelle 4). Hierbei wird aber außer Acht gelassen, mit welcher Wahrscheinlichkeit eine bestimmte Schwachstelle von einem Angreifer ausgewählt wird. Das Ziel ist es, auch hier wieder zu einem qualitativen Ergebnis zu kommen, also die Plausibilität eines erfolgreichen Angriffs durch qualitative Kategorien zu beschreiben.

Tab. 3: Beschreibung der Fähigkeiten eines Angreifers

Level	Beschreibung
Very High (VH)	Der Angreifer verfügt über eine sehr ausgeprägte Expertise und sehr ausgiebige Ressourcen. Zudem ist es ihm möglich, mehrfache, wiederholte und koordinierte Angriffe durchzuführen.
High (H)	Der Angreifer verfügt über eine ausgeprägte Expertise und ausgiebige Ressourcen. Zudem ist es ihm möglich, mehrfache, wiederholte und koordinierte Angriffe durchzuführen.
Moderate (M)	Der Angreifer besitzt eine fundierte Expertise und substanzielle Ressourcen. Zudem verfügt er über einige Möglichkeiten, um mehrfache Angriffe durchzuführen.
Low (L)	Der Angreifer besitzt eine geringe Expertise und eingeschränkte Ressourcen. Zudem verfügt er nur über wenige Möglichkeiten, um einen erfolgreichen Angriff durchzuführen.
Very Low (VL)	Der Angreifer besitzt sehr geringe Expertise und sehr eingeschränkte Ressourcen. Zudem verfügt er kaum über Möglichkeiten, um einen erfolgreichen Angriff durchzuführen.

Diese Einschätzung bezieht sich jedoch lediglich auf den erfolgreichen Angriff unter Ausnutzung einer einzelnen Schwachstelle. Bei einem Großteil der Angriffe auf eine IT-Infrastruktur wird aber zumeist eine Reihe von Schwachstellen auf unterschiedlichen Systemen ausgenutzt, bevor das Zielsystem erreicht wird. Für diese unterschiedlichen Angriffspfade muss eine entsprechende qualitative Bewertung gefunden werden, um eine Aussage über das Risiko treffen zu können.

Tab. 4: Plausibilität eines erfolgreichen Angriffs abhängig von Fähigkeiten (F) & Verwundbarkeit (V)

V \ F	Very Low	Low	Moderate	High	Very High
Very Low	VL	VL	L	L	M
Low	VL	L	L	M	H
Moderate	L	L	M	H	H
High	L	M	H	H	VH
Very High	M	H	H	VH	VH

Die Berechnung der Bewertung einzelner Angriffspfade kann auf unterschiedliche Weise erfolgen. Die Grundlage für diese Einschätzung stellt die Plausibilität eines erfolgreichen Angriffs auf die einzelnen Schwachstellen dar. Basierend darauf ergibt sich eine Worst-Case Abschätzung als einfachste Methode: für jeden möglichen Pfad werden alle Schwachstellen auf diesem Pfad zusammen mit der Plausibilität eines erfolgreichen Angriffes betrachtet. Die höchste dieser Kategorien über alle Schwachstellen repräsentiert die Plausibilität eines erfolgreichen Angriffs über den gesamten Pfad (Maximum-Prinzip).

Ein weiterer Ansatz ist, für jede der fünf qualitativen Kategorien einen quantitativen Repräsentanten (also eine Zahl, im statistischen Kontext einen Rang) zuzuordnen, welcher in weiterer Folge in eine Formel eingesetzt werden kann. Dieser Ansatz wird etwa von der NIST verfolgt [SiOu11], wobei jeder Kategorie eine Zahl zwischen 0 und 1 zugewiesen wird und diese Zahlen entsprechend multipliziert werden. Das Ergebnis wird als „Wahrscheinlichkeit“ für einen erfolgreichen Angriff interpretiert. Alternativ kann es aber wiederum in eine der fünf Kategorien umgewandelt werden, welche dann die Plausibilität für einen erfolgreichen Angriff über diesen Pfad beschreibt.

Weitere Methoden für die Bewertung einzelner Angriffspfade sind ebenfalls denkbar: z.B. kann eine Relation auf den fünf Kategorien definiert werden, die eine Verkettung der einzelnen Kategorien wieder auf eine der Kategorien abbildet. Eine andere Möglichkeit ist es, die Anzahl der Ausprägungen einzelner Kategorien festzuhalten, etwa wie oft die Kategorie „Very High“

entlang eines Pfads auftritt, und diese Information in einem Histogramm zusammenzufassen. Aufgrund dieser Fülle an Möglichkeiten wird für den in diesem Artikel beschriebenen Ansatz die Auswahl einer Methode für die Bewertung eines Angriffspfades dem Anwender überlassen. Dadurch kann der Anwender eine für seine bestehenden Methoden und seine Infrastruktur passende Methode wählen, wodurch ein angemessenes Gleichgewicht zwischen Genauigkeit und Einfachheit der Umsetzung gewährleistet wird. Für die weiteren Berechnungen ist es lediglich notwendig, dass das Ergebnis wiederum qualitativ, also durch eine Kategorie zwischen „Very Low“ und „Very High“, beschrieben wird.

3.4 Einschätzung des erwarteten Schadens

Neben der Plausibilität eines erfolgreichen Angriffs über einen bestimmten Pfad, wie im vorherigen Abschnitt 3.3 beschrieben, kann auch eine qualitative Einschätzung über den zu erwartenden Schaden gegeben werden. Diese bezieht sich entsprechend auf die gesamten Auswirkungen einer Schwachstelle in Bezug auf die Bereiche *Confidentiality*, *Integrity* und *Availability*, sowie auf die Bewertung einzelner Angriffspfade. Diese beiden Größen müssen zueinander in Relation gebracht werden, um den zu erwartenden Schaden bestimmen zu können (siehe Tebelle 5).

Tab. 5: Plausibilität eines erfolgreichen Angriffs abhängig von Fähigkeiten (F) & Verwundbarkeit (V)

P \ A	Very Low	Low	Moderate	High	Very High
Very Low	VL	VL	L	L	M
Low	VL	L	L	M	H
Moderate	L	L	M	H	H
High	L	M	H	H	VH
Very High	M	H	H	VH	VH

Somit erhält man zu jedem Angriffspfad eine Einschätzung über den zu erwartenden Schaden. Wie bereits im vorherigen Abschnitt 3.3 diskutiert, müssen auch hier diese Ergebnisse wieder gesammelt und konsolidiert werden. Zu diesem Zweck werden ebenfalls die Ergebnisse aller möglichen Pfade in einem Histogramm zusammengefasst. Durch diese Darstellungsweise gehen keinerlei Informationen verloren (wie etwa beim Einsatz des Maximum-Prinzips) und die Histogramme können direkt in den spieltheoretischen Ansatz zur Risikominimierung integriert werden (siehe Abschnitt 4.2).

4 Risikominimierung mittels Spieltheorie

Im Allgemeinen stehen zur Verminderung eines Risikos unterschiedliche Maßnahmen zur Verfügung. Hierbei stellt sich die Frage, wie diese Maßnahmen optimal eingesetzt werden können. Diese Entscheidung ist von besonderer Relevanz, wenn nur begrenztes Budget zur Verfügung steht. Lässt sich für eine Anzahl von Angriffen und Gegenmaßnahmen der erwartete Schaden bemessen, so kann Spieltheorie bei der optimalen Umsetzung möglicher Maßnahmen helfen.

4.1 Beschreiben der Risikosituation

Für die spieltheoretische Analyse ist eine möglichst vollständige Beschreibung der Risikosituation notwendig. Diese Beschreibung umfasst die folgenden drei Komponenten:

1. Liste aller möglichen Bedrohungsszenarien
2. Liste aller möglichen Maßnahmen zur Verminderung des Schadens

3. Für jede Kombination aus Bedrohungsszenario und Gegenmaßnahme eine Beschreibung des zu erwartenden Schadens

Während die Bedrohungsszenarien und die Maßnahmen zur Schadensminderung recht klar definiert werden können, ist eine Beschreibung des Schadens nur selten in exakten Werten (z.B. in Euro) möglich. Wie bereits weiter oben angesprochen, bringt eine Einschätzung dieses Schadens in Form von qualitativen Kategorien (wie in Abschnitt 3.2 beschrieben) einen deutlichen Vorteil mit sich. Darüber hinaus gilt auch hier das Erfordernis, alle relevanten Sicherheitsziele in einer einheitlichen (ggf. kategorialen) Skala zu bewerten.

Im Kontext der Spieltheorie ist folgende *Notation* üblich:

1. Bedrohungsszenarien werden als Strategien eins Angreifers aufgefasst (dieser Angreifer kann auch die Natur sein, z.B. wenn ein Naturereignis einen Schaden verursacht)
2. Gegenmaßnahmen werden als Strategien des Verteidigers aufgefasst
3. Der erwartete Schaden wird als Auszahlungsfunktion (im Englischen üblicherweise als (negativer) Payoff oder (negativ-wertige) Utility) bezeichnet

Während traditionelle Spieltheorie den Payoff zu maximieren sucht, ist der hier vorgestellte Ansatz darauf ausgerichtet, den Schaden zu minimieren. Das Vorgehen ist völlig analog, unterscheidet sich somit nur im Vorzeichen der Auszahlungsfunktion (welches im Fall von Schadensminimierung negativ ist).

4.2 Spieltheoretische Analyse

Basierend auf der Beschreibung der Risikosituation kann mittels Spieltheorie eine optimale Entscheidung getroffen werden, die den erwarteten Schaden minimiert. Während bestehende Ansätze [AlBa10, Rass13] davon ausgehen, dass das Risiko bzw. der Schaden im Eintrittsfall exakt bekannt sind, umfasst der vorliegende Ansatz auch Situationen, in denen die Konsequenzen unsicher sind [Rass15, ScKR15]. Im Allgemeinen beruht der spieltheoretische Ansatz auf folgender *Annahme*:

> Einem Angreifer wird keine spezielle Intention oder ein dezidiertes Ziel unterstellt, sondern es wird die Betrachtung auf den eigenen Schaden gelegt: man geht davon aus, dass es Ziel eines Angriffs ist, so hohen Schaden wie möglich zu verursachen (*worst case*).

Formal bedeutet diese Annahme, dass ein Nullsummenspiel betrachtet wird und damit nur der Payoff für den Verteidiger angegeben werden muss. Sind die Payoffs definiert, wählt der Verteidiger seine Strategie so, dass der Schaden möglichst klein wird, während der Angreifer versucht ihn zu maximieren. Es sei angemerkt, dass diese Annahme nicht in Konflikt mit der Einschätzung der Fähigkeiten des Angreifers (wie in Tabelle 3 beschrieben) steht. Die Einschätzungen in Tabelle 3 ermöglichen eine bessere Aussage über die Erfolgswahrscheinlichkeit eines Angriffs, mindern den maximal möglichen Schaden jedoch nicht.

4.2.1 Wahl der Strategien

Die Angriffsstrategien umfassen all jene Möglichkeiten eines Angreifers, Schaden am System zu verursachen. Dabei ist es wichtig, eine möglichst vollständige Liste der Angriffsstrategien zu bekommen, da die Risikoabschätzung nur für die betrachteten Strategien gültig ist. Neben beobachteten Vorfällen aus der Vergangenheit können hier Verwundbarkeitsanalysen (z.B. mit-

tels Nessus oder ähnlichen Tools) helfen. Parallel dazu beinhalten die Strategien für den Verteidiger all jene Mittel, die dem Verteidiger zur Verfügung stehen, um den potentiellen Schaden zu vermindern.

4.2.2 Beschreibung des erwarteten Schadens

Für jede Kombination aus Angriffs- und Verteidigungsstrategie muss der erwartete Schaden angegeben werden, da dieser erst einen Vergleich der verschiedenen Situationen ermöglicht. Der Begriff „erwartet“ ist hierbei nicht als Erwartungswert i.S.d. Statistik, sondern als (subjektive) Experteneinschätzung zu interpretieren. Dieser erwartete Schaden kann nur selten exakt beziffert werden. Eine adäquate Beschreibung bietet sich durch Verteilungsfunktionen (Dichten) an [RaKS15, RaKS16]. Während das Be-stimmen dieser Verteilungen in der Praxis meist schwierig ist, steht häufig zumindest eine kategoriale Einstufung zur Verfügung. Es wird daher im Folgenden davon ausgegangen, dass für jedes Paar von Angriffs- und Verteidigungsstrategie der erwartete Schaden durch eine der Kategorien von „Very Low“ bis „Very High“ dargestellt werden kann (siehe Abschnitt 3.4).

Ein Vergleich von so beschriebenen Schäden erfolgt lexikographisch:

- Es wird jenes Szenario bevorzugt, welches weniger Ausprägungen (Beobachtungen, Expertenmeinungen, etc.) eines Schadens vom Level VH aufweist.
- Sind die Ausprägungen eines Schadens vom Level VH gleich, so werden die Ausprägungen des Schadens vom Level H betrachtet.
- Dies wird bis zum Erreichen des niedrigsten Levels (VL) wiederholt.
- Stimmen die Ausprägungen aller Levels überein (also wenn die Histogramme identisch sind), werden die jeweiligen Payoffs als identisch erachtet (in diesem Fall kann eines der Szenarien ggf. als redundant aus der Betrachtung ausgenommen werden).

Sobald alle Strategien und die entsprechenden Payoffs untersucht wurden, beschreiben die Methoden der Spieltheorie die bestmöglichen Aktionen sowohl für den Angreifer als auch für den Verteidiger. In anderen Worten liefert das Gleichgewicht aus der Spieltheorie Antworten auf die folgenden beiden Fragestellungen:

1. Wie müssen die Assets der IKT-Infrastruktur geschützt werden, damit der erwartete Schaden minimal ist?
2. Wie wahrscheinlich ist es, dass ein Angreifer eine bestimmte Angriffsstrategie auswählt?

Aus der Konstruktion (*Nash-Gleichgewicht*) [NeMo44] folgt unmittelbar, dass ein Angreifer, sollte er von der berechneten Auswahl an Angriffsstrategien abweichen, ein schlechteres Ergebnis erzielt, also einen geringeren Schaden am System verursacht.

5 Anwendungsbeispiel

Zur Illustration der in Abschnitt 3 und 4 vorgestellten Methoden beschreibt dieses Kapitel ein detailliertes Anwendungsbeispiel in einem vereinfachten Umfeld. Hierzu werden drei Assets betrachtet: eine Web-Applikation (A_1), die auf einem Apache Web Server (A_2) gehostet wird, welcher wiederum auf einem Microsoft Windows Server 2012 (A_3) läuft. Durch Einsatz eines Vulnerability Scanners wird festgestellt, dass auf diesen drei Assets insgesamt acht Schwachstellen existieren (V_1 bis V_8), welche durch die CVSS-Metriken (siehe Tabelle 6) bemessen werden. Für eine bessere Nachvollziehbarkeit wurden für den Apache Web Server und den

Windows Server die Daten existierender Schwachstellen herangezogen (siehe CVE-Nr. in Tabelle 6). Lediglich für die Web-Applikation wurden beliebige Werte für die CVSS-Metriken gewählt.

Tab. 6: Liste der Schwachstellen inkl. der CVSS-Metriken

Asset	Schwach-stelle	CVE Nr.	CVSS Exploitability			CVSS Impact		
			AV	AC	Auth	C	I	A
A_1	V_1	N/A	N	L	S	C	C	C
A_1	V_2	N/A	A	M	S	P	P	N
A_1	V_3	N/A	N	M	N	C	C	P
A_2	V_4	CVE-2013-6111	N	M	N	N	P	N
A_2	V_5	CVE-2014-4721	N	H	N	P	N	N
A_3	V_6	CVE-2016-0099	L	L	N	C	C	C
A_3	V_7	CVE-2016-0037	N	L	N	N	N	P
A_3	V_8	CVE-2016-0016	L	M	N	C	C	C

Anhand der Abbildungen in Tabelle 1 und Tabelle 2 werden diese Informationen in Tabelle 7 entsprechend auf die fünfstufige Verwundbarkeitsskala und Schadensskala heruntergebrochen. Zusätzlich können auf Basis dieser Schwachstellen Verbindungen zwischen den einzelnen Assets dargestellt werden, welche in weiterer Folge als Angriffspfade dienen können. Hierzu werden die zusätzlichen Informationen und Charakteristika der Schwachstellen, wie sie etwa in der NVD zu finden sind, herangezogen. Darin ist z.B. festgehalten, ob in Angreifer durch das Ausnutzen einer Schwachstelle zusätzliche Rechte auf einem System erlangen kann (*escalate privileges*) und sich dadurch Zugang zu weiteren Applikationen am selben System oder anderen Systemen im Netzwerk verschaffen kann.

Tab. 7: Liste der Verwundbarkeit und des Schadens für die Schwachstellen

Asset	Schwach-stelle	Verwundbarkeit	Schaden
A_1	V_1	VH	VH
A_1	V_2	M	L
A_1	V_3	VH	VH
A_2	V_4	VH	VL
A_2	V_5	H	L
A_3	V_6	M	VH
A_3	V_7	VH	VL
A_3	V_8	M	VH

Für diese Infrastruktur kann nun ein spezifischer Angreifer betrachtet werden. Hierfür nimmt man an, dass es sich um einen Angreifer mit Fähigkeiten der Kategorie „Very High“ handelt, der das Asset A_3 (unter Ausnutzung der Schwachstellen V_6 bis V_8) als Ziel hat. Wie in Abbildung 1 dargestellt, können mehrere Pfade zu einer der Schwachstellen führen – der Einfachheit halber werden im Folgenden lediglich zwei Pfade (die Pfade „rot“ und „blau“, siehe Abbildung 1) betrachtet. Wie in Abschnitt 3.3 spezifiziert, bestimmen die Fähigkeiten eines Angreifers seine Möglichkeit, einen erfolgreichen Angriff durchzuführen. Analog ändern sich auch die individuellen Bewertungen der Schwachstellen (in Abbildung 1 sind die Werte bereits an den Angreifer der Kategorie „Very High“ entsprechend dem Mapping in Diese Einschätzung bezieht sich jedoch lediglich auf den erfolgreichen Angriff unter Ausnutzung einer einzelnen Schwachstelle. Bei einem Großteil der Angriffe auf eine IT-Infrastruktur wird aber zumeist eine

Reihe von Schwachstellen auf unterschiedlichen Systemen ausgenutzt, bevor das Zielsystem erreicht wird. Für diese unterschiedlichen Angriffspfade muss eine entsprechende qualitative Bewertung gefunden werden, um eine Aussage über das Risiko treffen zu können.

Tabelle 4 angepasst). Für jeden der Pfade wird abgeschätzt, wie plausibel ein Angriff über diesen Pfad ist, woraus sich eine Bewertung aller möglichen Angriffspfade ergibt.

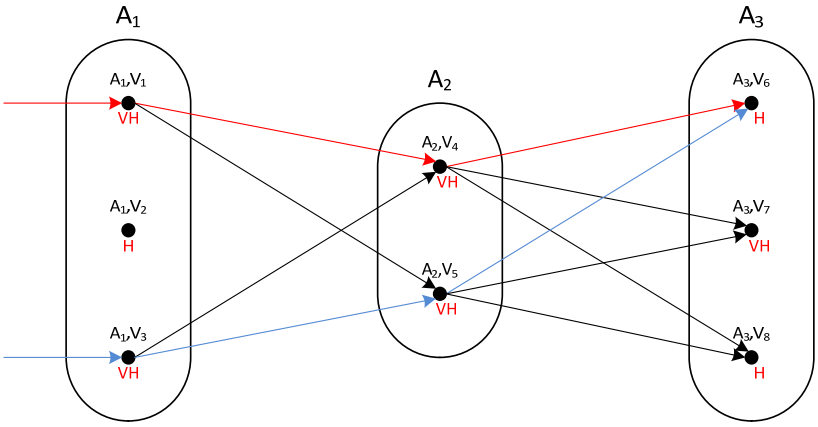


Abb. 1: Darstellung unterschiedlicher Angriffspfade im Anwendungsbeispiel

In Bezug auf den spieltheoretischen Ansatz beschreiben nun die drei Schwachstellen V_6 bis V_8 des Ziel-Assets des Angreifers auch seine möglichen Angriffsstrategien:

- Strategie a_1: Angriff von A_3 unter Ausnutzung der Schwachstelle V_6
- Strategie a_2: Angriff von A_3 unter Ausnutzung der Schwachstelle V_7
- Strategie a_3: Angriff von A_3 unter Ausnutzung der Schwachstelle V_8

Wie bereits angesprochen, kann jede dieser Strategien über mehrere Angriffspfade im Netzwerk umgesetzt werden. Analog dazu ergeben sich direkt die folgenden drei Strategien für den Verteidiger:

- Strategie d_1: Verteidigung von V_6 durch Einspielen eines Patches
- Strategie d_2: Verteidigung von V_7 durch Einspielen eines Patches
- Strategie d_3: Verteidigung von V_8 durch Einspielen eines Patches

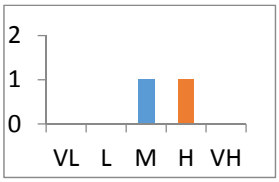
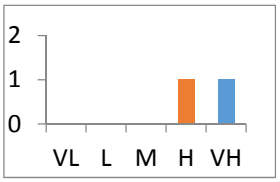
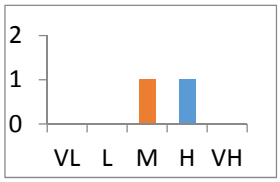
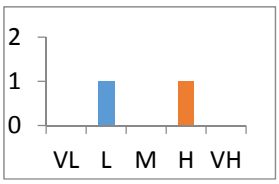
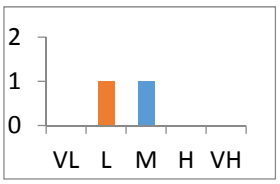
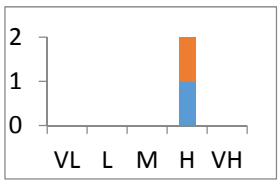
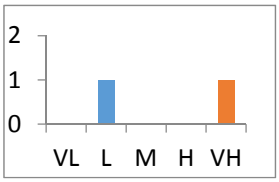
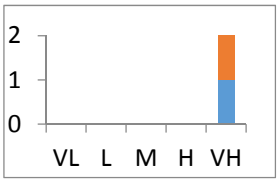
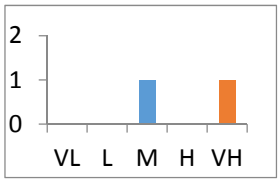
Alternativ sind natürlich auch andere Strategien für den Verteidiger denkbar, etwa die Installation einer zusätzlichen oder verbesserten Security-Software (z.B. Firewall, Intrusion Detection oder Intrusion Prevention System, etc.) oder die Durchführung von Awareness-Schulungen.

Der Payoff des Spiels wird durch den verursachten Schaden am Ziel-Asset (A_3) dargestellt. Dieser Schaden ist sowohl von der ausgenutzten Schwachstelle (V_6 bis V_8 haben unterschiedliche Auswirkungen) als auch von dem gewählten Angriffspfad (unterschiedliche Pfade versprechen unterschiedliche Aussichten auf Erfolg) abhängig. Wie bereits in Abschnitt 3.4 und Abschnitt 4.2 erläutert, wird der Schaden durch Histogramme auf Basis gesammelter Schadensbemessungsdaten/-einschätzungen dargestellt. Diese zeigen den verursachten Schaden für jeden gewählten Pfad (siehe Tabelle 8, hier sind zur besseren Übersicht nur der rote und blaue Pfad aus Abbildung 1 betrachtet worden).

Dieses Spiel lässt sich basierend auf dem in Abschnitt 4.2.2 beschriebenen Vergleich von Payoffs analysieren. Dabei fällt auf, dass die Gegenmaßnahme d_3 nie verwendet werden sollte,

da es immer eine bessere Alternative gibt (d_3 ist eine dominierte Strategie). Aus den verbleibenden Strategien kann eine optimale Lösung (ein Nash-Gleichgewicht) bestimmt werden.

Tab. 8: Darstellung der Spiel-Matrix mit Histogrammen als Payoffs

	a_1	a_2	a_3
d_1			
d_2			
d_3			

6 Zusammenfassung und Ausblick

Die in diesem Artikel vorgestellte Methode zur Minimierung von Risiken in IKT-Infrastrukturen beschreibt einen neuen Ansatz, der auf Elementen der Spieltheorie aufbaut und dadurch eine spezifische Aussage darüber erlaubt, welche Aktionen zur Risikominimierung gesetzt werden müssen, um eine optimale Auswahl an möglichen Gegenmaßnahmen zu ergreifen. Die Anwendung von aktuellen Erkenntnissen aus der Spieltheorie ermöglicht es dabei nicht nur, auf rein qualitativen Risikoeinschätzungen aufzubauen, sondern alle verfügbaren Informationen „gleichberechtigt“ in Form empirischer Verteilungen (Histogrammen) in die Analyse einfließen zu lassen. Hierdurch entfällt die sonst übliche Aufgabe, Meinungen und Einschätzungen zu konsolidieren und zu harmonisieren (was etwa im Falle widersprüchlicher Daten schwierig werden kann).

Darüber hinaus wurde gezeigt, wie eine Einbindung von Informationen über Schwachstellen aus öffentlichen Datenbanken sowie von Bewertungen dieser Schwachstellen mittels CVSS in die Methode erreicht werden kann. Dadurch ist gewährleistet, dass der Ansatz und damit auch die Auswahl an Handlungsoptionen immer auf aktuellen Informationen über die Komponenten einer IKT-Infrastruktur fußt. Zudem bleibt das System offen für subjektive Einschätzungen von Security-Experten einer Organisation oder kann sogar aufgrund des qualitativen Ansatzes ausschließlich auf diesen Experteneinschätzungen aufbauen.

Danksagung

Diese Arbeit wurde teilweise durch das EU Projekt MITIGATE (Grant-Nr. 653212) im Rahmen des H2020 Rahmenprogramms finanziert.

Literatur

[AlBa10] T. Alpcan, T. Basar: Network Security: A Decision and Game Theoretic Approach. Cambridge (2010).

[Aust14] Austrian Standards Institute (Hrsg.): ONR 49000 – Risikomanagement für Organisationen und Systeme – Begriffe und Grundlagen – Umsetzung von ISO 31000 in die Praxis.(2014)

[Info13] Information Systems Audit and Control Association (ISACA) (Hrsg.): COBIT 5 for Risk. Rolling Meadows, IL 60008 USA (2013)

[Inte09] International Organization for Standardization (ISO) (Hrsg.): ISO 31000 Risk management – Principles and guidelines, Geneva, Switzerland (2009)

[MeSR07a] P. Mell, K. Scarfone, S. Romanosky: A Complete Guide to the Common Vulnerability Scoring System Version 2.0, https://www.first.org/cvss/v2/guide (2007).

[MeSR07b] P. Mell, K. Scarfone, S. Romanosky: The Common Vulnerability Scoring System (CVSS) and Its Application to Federal Agency Systems. NIST Interagency Report 7435 (2007).

[Nati12] National Institute of Standards and Technology (NIST): Guide for Conducting Risk Assessments. NIST Special Publication 800-30 Revision 1, Gaithersburg, MD 20899-8930, USA (2012)

[NeMo44] J. von Neumann und O. Morgenstern, *Theory of games and economic behavior*, Princeton University Press, 1944.

[Rass13] S. Rass: On game-theoretic network security provisioning. Journal of Network and Systems Management, Springer (2013) 47-64.

[Rass15] S. Rass: On Game-Theoretic Risk Management (Part One) – Towards a Theory of Games with Payoffs that are Probability-Distributions. ArXiv e-prints (2015).

[RaKS15] S. Rass, S. König, S. Schauer: Uncertainty in Games: Using Probability-Distributions as Payoffs. In: Decision and Game Theory for Security, Lecture Notes in Computer Science, Springer (2015) 346-357.

[RaKS16] S. Rass, S. König, S. Schauer: Decision with Uncertain Consequences – A Total Ordering on Loss-Distributions, to be published.

[RaSn11] L. Rajbhandari, E. Snekkenes.: Mapping between classical risk management and game theoretical approaches, in Proc. Communications and Multimedia Security (CMS), LNCS, Springer, 2011

[RSPG13] S. Rass, S. Schauer, A. Peer, J. Göllner: „Sicherheit auf Basis Multikriterieller Spieltheorie“. DACH Security 2013, Nürnberg; 17.09.2013 - 18.09.2013; in: P. Schartner, J. Trommler: „DACH Security 2013“, S. 289 - 301, (2013)

[RRVG15] S. Rass, B. Rainer, M. Vavti, J. Göllner, A. Peer, S. Schauer: Secure Communication over Software-Defined Networks. Mobile Networks and Applications, Vol. 20, No. 1, 2015, pp. 105-110

[ScKR15] S. Schauer, S. König, S. Rass: Mathematical Aspects of Risk Management in Interconnected Utility Networks. In: Advances in Mathematics and Computer Science and their Applications (MAMECTIS 2016), WSEAS (2016).

[SiOu11] A. Singhal, X. Ou: Security Risk Analysis of Enterprise Networks Using Probabilistic Attack Graphs. NIST Interagency Report 7788 (2011).