

Data Science-basierte Identifizierung zusammenhängender operationeller Schadensereignisse

Übersicht Problemstellung und Lösungsansatz

Operationelle Risiken stellen sowohl für Finanzdienstleister als auch für Industrieunternehmen eine ebenso große wie schwer zu analysierende Bedrohung dar.

Zusammenhänge zwischen verschiedenen OpRisk-Ereignissen können hierbei trotz komplexer Modelle in der Praxis kaum identifiziert werden und tieferliegende Ursachen bleiben oft unerkannt. Andererseits haben sich **Data Science-Verfahren** für ähnliche Fragestellungen bereits etabliert und ermöglichen die **Untersuchung** großer Mengen unterschiedlicher Daten **nach Zusammenhängen**, etwa im Bereich der Analyse des Kaufverhaltens von Kunden im Onlinehandel.

Durch Anpassung auf die Gegebenheiten operationeller Risiken angepasst und können **bereits entwickelte Algorithmen** für die **Identifizierung von Interdependenzen** zwischen operationellen Schäden verwendet werden.

Unternehmen können damit **Kausalbeziehungen zwischen Schäden** aufspüren, wodurch die Suche nach gemeinsamen Ursachen erleichtert wird. Das gesamte angesammelte Wissen kann effizient genutzt werden, um künftige Schäden möglichst zu vermeiden bzw. frühzeitig zu antizipieren.

Operationelle Risiken

Arten

Operationelle Risiken lassen sich je nach Ursache den folgenden Kategorien zuteilen

- *Menschen*: z.B. Betrug, mangelnde Kenntnisse, Mitarbeiter-Fluktuation
- *Prozesse*: z.B. Transaktionsfehler, Projektrisiken, Reportingfehler, Bewertungsfehler
- *Systeme*: z.B. Programmierfehler, Abstürze
- *Externe Ereignisse*: z.B. Klagen, Diebstahl, Feuer, Überschwemmungen

Management

Für gewöhnlich werden operationelle Risiken nach Schadensausmaß und Wahrscheinlichkeit kategorisiert. Entsprechend kommen als Managementstrategien infrage:

- *Vermeidung*: Bei großen, unnötigen Risiken

- *Versicherung*: Bei großen, nötigen Risiken
- *Verminderung*: Insb. bei kleineren Risiken mit hoher Eintrittswahrscheinlichkeit
- *Inkaufnahme*: Bei Risiken, die zum Geschäftsmodell gehören

Desweiteren ist eine angemessene Erfassung operationeller Risiken auch auch bei der FMEA (Failure Mode and Effects Analysis / Fehlermöglichkeits- und -einflussanalyse) unabdingbar.

Methoden und Problematik

Die Handhabung operationeller Risiken ist insbesondere im Finanzdienstleistungssektor streng geregelt. So müssen **Banken** nach *Basel II/III* operationelle Risiken mit **Eigenkapital** unterlegen. Hierfür gibt es vorgeschriebene Rechenschemata wie den auf pauschalen Faktoren basierenden *Standardized Approach (SA)* und den *Advanced Measurement Approach (AMA)*. Letzterer basiert auf Verteilungsannahmen und wird künftig durch den *SA* ersetzt.

Methodisch unterscheidet man bei der Behandlung operationeller Risiken u.a. zwischen

- *Fragebögen und Self-Assessment*: Es werden eher qualitativ Wahrscheinlichkeiten und Ausmaße ermittelt
- *Aktuariellen Verfahren*: Diese basieren auf Verteilungsannahmen anhand vergangener Schäden
- *Key Risk Indicator-Verfahren*: Es werden einfach beobachtbare Maße identifiziert, die zur Frühwarnung dienen
- *Kausale Netzwerke*: Wirkzusammenhänge werden mittels Bayes-Statistik abgebildet

Zusammenhänge zwischen und Ursachen von operationeller Risiken können **entweder überhaupt nicht oder nur auf sehr komplexe und fehleranfällige Weise ermittelt** werden.

Erkennung von Zusammenhängen mittels Data Science-Verfahren

Assoziationsanalyse

Für die Analyse der Zusammenhänge mehrerer unterschiedlicher Ereignisse („*Items*“) bieten sich Methoden aus dem Gebiet der **Assoziationsanalyse** („*Association Analysis*“) an. Entsprechende Methoden aus dem Bereich der Warenkorbanalyse haben sich bereits seit einigen Jahren etabliert und finden insbesondere im Online-Handel (z.B. Buchvorschläge im Online-Handel), Suchmaschinenvorschlägen oder im Einzelhandel (Produkte in Regalen) Verwendung. Mittels der Assoziationsanalyse lässt sich das gemeinsame Auftreten unterschiedlicher Ereignissen **unmittelbar und ohne Verteilungsannahmen** identifizieren.

Die enorme Vielzahl möglicher Schlussfolgerungen lässt sich mittels **eigens hierfür entwickelter Maße** wie *Support*, *Confidence* und *Lift* effizient und sachgerecht einschränken. Die Analysen erfordern Programme auf der Basis entsprechender **Analysetools**, wie z.B. *Python*, *R* oder *RapidMiner*.

Analyse-Vorbereitung

Zunächst müssen die Schadensdaten in ein für die Analyse brauchbares **Format** gebracht werden. Je nach Schadensart sind zudem zeitliche **Aggregationen** (z.B. auf Tages-, Wochenbasis) vorzunehmen.

Zu häufig vorkommende oder bereits erklärte Schadensarten sind – anhand von **Experteneinschätzungen** – zu entfernen.

Analyse-Durchführung

Vor dem Analysebeginn sind die **Kriterien für die Schlussfolgerungsregeln** gemäß Support und Confidence festzulegen. Die Festlegung der Kriterien kann (z.B. bei R) durch Grafiken unterstützt werden.

Anschließend müssen die Schlussfolgerungen von Experten **plausibilisiert** werden. Die Schritte sind für alle relevanten zeitlichen Aggregationen zu wiederholen.

Anwendungsbeispiel: Analyse einer fiktiven Schadensdatenbank

Als Anwendungsbeispiel wurde eine fiktive Schadensdatenbank einer Bank für ein gesamtes Jahr konstruiert.

Es lagen insgesamt 23 mögliche Schadensarten vor, darunter z.B. eine Grippewelle, verspätete Reports, falsche Bewertungen, und Klagen wegen Falschberatung. Folgende Annahmen lagen dem Testbeispiel zugrunde:

- Fehlerhafte Transaktionen sind sehr häufig
- Mängel bei Outsourcer-Hotline werden bei Anfragen wg. PC-Headcrashes deutlich
- Reporting-Mitarbeiter fahren meist Auto und sind von Schneesturm betroffen
- Nach Abstürzen des Bewertungs-Systems kommt es zu Falschbewertungen
- Diebstähle während Arbeiten nach Brand im Besprechungsraum
- Personalengpässe bei Lieferanten führen zu gescheiterten Projekten
- Massive Kundenbeschwerden nach Abgang von Kundenbetreuern

Da die fehlerhaften Transaktionen sehr oft und unzusammenhängend auftraten, wurden sie als erstes entfernt:

Zunächst wurden – zur Ermittlung der relevanten Support- und Confidence-Maße sämtliche ermittelte Regeln grafisch dargestellt.

Ereignis	Anzahl in 2014
Fehlerhafte Transaktionen	21
Verspaetete Reports	13
Kuendigungswelle Kundenbetreuer	11
Systemabsturz Bewertungsprogramm	11
Falschbewertungen Produkte	7
Grippewelle	5
Fehlerhafte Eingaben	5
Headcrash PC-Festplatte	5
Outsourcer-Hotline nicht erreichbar	4
Kundenbeschwerden/Kuendigungen	4
Streik Busunternehmen	3
Unvollstaendige Reports	3
Diebstahl Kuechenequipment	3
Schneesturm	2
E-Mail-Server nicht verfuegbar	2
Personalengpass bei Lieferanten	2
Projekt gescheitert & abgebrochen	2
Diebstahl Bueroequipment	2
Brand Besprechungsraum	1
Klagen Falschberatung	1
Nachteiliges Presse-Interview	1
Shitstorm in Sozialen Medien	1
Verstoß Datenschutz	0

Tabelle 1: Häufigkeit der Schäden

Zunächst wurden – zur Ermittlung der relevanten Support- und Confidence-Maße sämtliche ermittelte Regeln grafisch dargestellt.

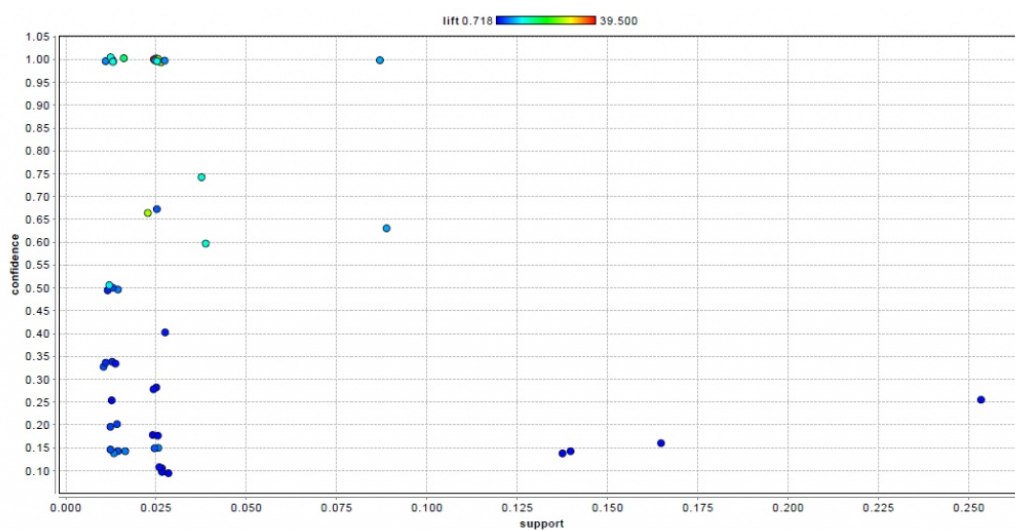


Abbildung 1: Auftragung der Schluss-Regeln auf Tagesbasis

Nach Einschränkung der Confidence auf ein Mindestmaß von 0,6 ergibt die u.g. Liste.

Nr	Premises	Conclusion	Support	Confidence	Lift
1	Headcrash PC-Festplatte	Outsourcer-Hotline nicht erreichbar	0,01	0,6	38,1
2	Systemabsturz Bewertungsprogramm	Falschbewertungen Produkte	0,03	0,64	23,09
3	Unvollstaendige Reports	Verspaetete Reports	0,01	0,67	13,03
4	Unvollstaendige Reports	Schneesturm	0,01	0,67	84,67
5	Unvollstaendige Reports	Verspaetete Reports, Schneesturm	0,01	0,67	84,67
6	Outsourcer-Hotline nicht erreichbar	Headcrash PC-Festplatte	0,01	0,75	38,1
7	Schneesturm	Verspaetete Reports	0,01	1	19,54
8	Brand Besprechungsraum	Verspaetete Reports	0	1	19,54
9	Falschbewertungen Produkte	Systemabsturz Bewertungsprogramm	0,03	1	23,09
10	Schneesturm	Unvollstaendige Reports	0,01	1	84,67
11	Verspaetete Reports, Unvollstaendige Reports	Schneesturm	0,01	1	127
12	Schneesturm	Verspaetete Reports, Unvollstaendige Reports	0,01	1	127
13	Verspaetete Reports, Schneesturm	Unvollstaendige Reports	0,01	1	84,67
14	Unvollstaendige Reports, Schneesturm	Verspaetete Reports	0,01	1	19,54
15	Systemabsturz Bewertungsprogramm, Streik Busunternehmen	Falschbewertungen Produkte	0	1	36,29
16	Falschbewertungen Produkte, Streik Busunternehmen	Systemabsturz Bewertungsprogramm	0	1	23,09
17	Systemabsturz Bewertungsprogramm, E-Mail-Server nicht verfuegbar	Falschbewertungen Produkte	0	1	36,29
18	Falschbewertungen Produkte, E-Mail-Server nicht verfuegbar	Systemabsturz Bewertungsprogramm	0	1	23,09

Tabelle 2: Identifizierte Zusammenhänge auf Tagesbasis

Von den gefundenen Koinzidenzen erweisen sich – nach Plausibilisierung – die grün markierten als valide.

Auf Wochen- und Monatsbasis wurde analog vorgegangen:

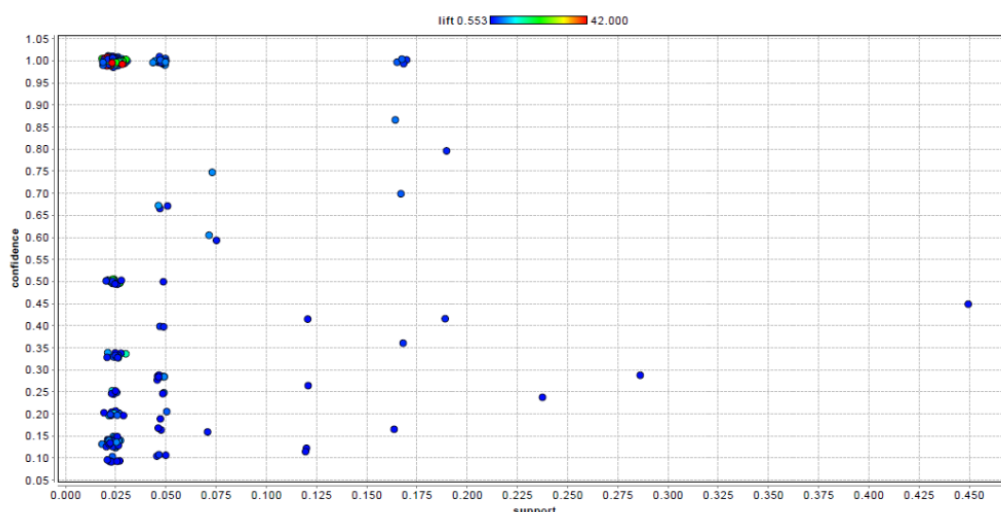


Abbildung 2: Auftragung der Schluss-Regeln auf Wochenbasis

Nr	Premises	Conclusion	Support	Confidence	Lift
1	Shitstorm in Sozialen Medien	Kundenbeschwerden/Kuendigungen	0,02	1	12,5
2	Diebstahl Bueroequipment	Diebstahl Kuechenequipment	0,02	1	16,67
3	Diebstahl Bueroequipment	Projekt gescheitert & abgebrochen	0,02	1	25
4	Brand Besprechungsraum	Personalengpass bei Lieferanten	0,02	1	25
5	Diebstahl Kuechenequipment, Projekt gescheitert & abgebrochen	Diebstahl Bueroequipment	0,02	1	50
6	Diebstahl Bueroequipment	Diebstahl Kuechenequipment, Projekt gescheitert & abgebrochen	0,02	1	50
7	Diebstahl Kuechenequipment, Diebstahl Bueroequipment	Projekt gescheitert & abgebrochen	0,02	1	25
8	Projekt gescheitert & abgebrochen, Diebstahl Bueroequipment	Diebstahl Kuechenequipment	0,02	1	16,67

Tabelle 3: Identifizierte Zusammenhänge auf Wochenbasis

Ereignis 1	Ereignis 2
Kundenbeschwerden/Kuendigungen	Projekt gescheitert & abgebrochen
Kundenbeschwerden/Kuendigungen	Personalengpass bei Lieferanten
Kundenbeschwerden/Kuendigungen	Kuendigungswelle Kundenbetreuer
Projekt gescheitert & abgebrochen	Personalengpass bei Lieferanten
Projekt gescheitert & abgebrochen	Diebstahl Kuechenequipment
Projekt gescheitert & abgebrochen	Diebstahl Bueroequipment
Projekt gescheitert & abgebrochen	Brand Besprechungsraum
Personalengpass bei Lieferanten	Diebstahl Kuechenequipment
Personalengpass bei Lieferanten	Diebstahl Bueroequipment
Personalengpass bei Lieferanten	Brand Besprechungsraum
Diebstahl Kuechenequipment	Diebstahl Bueroequipment
Diebstahl Kuechenequipment	Brand Besprechungsraum
Diebstahl Bueroequipment	Brand Besprechungsraum

Tabelle 4: Mögliche Zusammenhänge auf Monatsbasis

Nach entsprechender Plausibilisierung möglicher kausaler Beziehungen konnten **sämtliche bei der Erstellung gebrauchten Annahmen in den Daten identifiziert** werden.

Kontakt

Dr. Dimitrios Geromichalos
 Founder / CEO
 RiskDataScience UG (haftungsbeschränkt)
 Theresienhöhe 28, 80339 München
 E-Mail: riskdatascience@web.de
 Telefon: +4989244407277, Fax: +4989244407001
 Internet: www.riskdatascience.net

Web-App zu Assoziationsanalysen: <https://association-analysis.riskdatascience.net/>