

Konzeption eines proaktiven Risikomanagements in Logistiknetzwerken

Concept of a proactive risk management in logistics networks

Jens Leveling
Arkadius Schier
Francesco Luciano
Prof. Dr. Boris Otto

Fraunhofer-Institut für Materialfluss und Logistik IML

Logistiknetzwerke von Unternehmen wachsen sehr schnell und werden immer komplexer. Unternehmen wissen oft nicht, von welchen anderen Unternehmen sie abhängig sind und welche geschäftskritischen Risiken sich daraus für sie ergeben. Aus diesem Grund wird in diesem Artikel ein Konzept eines proaktiven Risikomanagements in Logistiknetzwerken vorgestellt. Das Konzept basiert auf der Big Data Technologie und verwendet zur Identifikation von Risiken und zum Aufbau eines Logistiknetzwerkes neben internen Unternehmensdaten auch externe Daten, z. B. Social Media Plattformen oder andere Datenportale. Diese Daten werden ausgewertet und mit Risiken behaftete Beziehungen werden dem Bediener grafisch angezeigt. Zusätzlich dazu kann das System dem Benutzer mögliche Alternativen zur Vermeidung dieser Risiken aufzeigen und somit zur Entscheidungsunterstützung genutzt werden.

[Schlüsselwörter: Big Data, Supply Chain Visibility, Risikomanagement, Entscheidungsunterstützung, Supply Chain Management]

Logistics business networks are growing rapidly and becoming more and more complex. Companies often do not know of what other companies they depend and which business-critical risks are consequences of these dependencies. For this reason, a concept of a proactive risk management in logistics networks is presented in this article. The concept is based on the Big Data technology and used for the identification of risks and the development of a logistics network using in addition to internal company data, external data such as social media platforms or other data portals. These data are analyzed and risky relationships are graphically displayed to the operator. In addition, the system can identify possible alternatives to the user to avoid these risks and thus be used for decision support.

[Keywords: Big Data, Supply Chain Visibility, Risk Management, decision support, Supply Chain Management]

1 EINLEITUNG

Die Wettbewerbsfähigkeit deutscher Produktionsunternehmen hängt immer stärker von weltweit verzweigten, mehrstufigen Zuliefernetzwerken ab. Einerseits muss heutzutage auf die Bauteile von Technologieführern zugegriffen werden, andererseits sind die niedrigen Kosten in anderen Weltregionen am einen Ende des Liefernetzes Grundlage für den Erhalt von Arbeitsplätzen in Deutschland am anderen Ende der Produktionskette. Die Kehrseite dieser weltweit verteilten Produktion über mehrere Stufen, die durch lange Transportrelationen verbunden sind, liegt in der deutlich höheren Störanfälligkeit des gesamten Netzwerkes. Diese Störanfälligkeit wirkte sich in den letzten Jahren insbesondere auf deutsche Unternehmen aus. Beispiele hierfür sind Naturkatastrophen, politische Unruhen oder lokale sowie globale wirtschaftliche Krisen aber auch alltägliche Hindernisse wie beispielsweise Verkehrsstaus oder Zugverspätungen. Vor diesem Hintergrund rückt aktuell das Thema Risikomanagement von Logistiknetzwerken in den Fokus vieler Unternehmen.

1.1 MOTIVATION

Mithilfe der industriellen Nutzung und Auswertung großer Datenmengen (Stichwort: „Big Data Analytics“), eröffnen sich für das Risikomanagement von Logistiknetzwerken völlig neue Möglichkeiten. Denn sowohl im Unternehmen, als auch im Internet existieren unzählige Daten und Informationen, aus denen sich Risiken für das logistische Netzwerk ableiten lassen. Beispiele hierfür sind Informationen zur Sicherheitslage, Unwetter- und Katastrophenmeldungen und -historie, Finanzdaten von Ländern und Unternehmen aus dem Internet, Ergebnisse von Audits, Materialflussmassendaten oder Informationen zur Produkt- und Netzwerkstruktur der Unternehmen. Vor diesem Hintergrund ist die Grundidee, die Verknüpfung von logistischen Strukturinformationen mit Informationen zu Risiken für das Logistiknetzwerk zu einer unternehmensspezifischen, digitalen Risikoweltkarte als Frühwarnsystem zusammenzufassen. Hierbei können sowohl Struktur- als auch Risikoinformationen aus den Unter-

nehmenssystemen selbst oder aus dem Internet stammen. Das in diesem Beitrag vorgestellte Konzept schafft die Grundlage, um proaktive Maßnahmen zur Vorbeugung von Störungen und Unterbrechungen zu planen und durchzuführen. Durch die rechtzeitige Minimierung von Risiken werden die Folgekosten gesenkt und es kann eine Effizienzsteigerung erzielt werden.

1.2 AUSGANGSSITUATION

Derzeit werden die in den Unternehmen teilweise bereits vorliegenden Informationen gar nicht oder nicht flächendeckend zur Risikoanalyse verwendet. Einerseits existiert kein oder nur ein eingeschränkter Zugang zu relevanten Quellen und andererseits ist der Aufwand zur Analyse und Aufbereitung aufgrund des Datenvolumens und der Heterogenität der Daten sehr groß. Außerdem können aktuell größere Datenabfragen die Systemperformance negativ beeinflussen. Trotz des verstärkten Interesses ist Risikomanagement in der Realität der Unternehmen häufig gleichzusetzen mit einer reaktiven Störungsbehandlung. Die rechtzeitige Identifikation zukünftiger Störungen, die in den meisten Fällen erst die notwendige Reaktionszeit für die erforderlichen Maßnahmen zur Verfügung stellen würde, fehlt ebenso, wie die Transparenz über das logistische Netzwerk, auch und gerade über die erste Stufe des Netzwerkes hinaus.

1.3 NEUHEITEN DES LÖSUNGSANSATZES

Der Hauptaspekt des in diesem Beitrag vorgestellten Konzeptes ist die Automatisierung der Ermittlung von Risiken aus Informationen sowohl im Unternehmen als auch aus dem Internet. Ein übergreifender Ansatz hierzu existiert derzeit nicht, was auch die Existenz manuell erstellter unternehmensunabhängiger Risikodatenbanken unterstreicht. Das Zusammenführen mehrerer kommerzieller und nicht-kommerzieller Datenquellen bildet hier eine grundlegende Idee. Neu ist vor allem das Kombinieren von Risikoereignissen für das Logistiknetzwerk mit Geolokationen sowie die Prüfung auf Aktualität und die anschließende Bewertung. Darüber hinaus ist für den Anwender die reine Kenntnis von bewerteten Risikoereignissen wegen der Möglichkeit von falsch positiven Meldungen nicht ausreichend. Ebenso wenig existiert ein Ansatz zum automatisierten Aufbau der Liefernetzwerkstruktur unter Einsatz von öffentlich zugänglichen Daten. Beide Aspekte werden von dem in diesem Beitrag beschriebenen Lösungsansatz berücksichtigt.

2 STAND DER WISSENSCHAFT UND TECHNIK

Das Risikomanagement von Liefernetzwerken wurde im Rahmen des vom BMWi geförderten Forschungsprojektes InKoRISK (inkorisk.research4.eu) vom Fraunhofer IML, Volkswagen, DB Schenker, Continental und der 4Flow AG detailliert untersucht. Vor diesem Hintergrund

hat sich als ein Hauptproblem des Risikomanagements die rechtzeitige Identifikation der Risiken im Netzwerk herausgestellt. Hierdurch konzentrieren sich die meisten Maßnahmen im Unternehmen auf die Reaktion bereits eingetretener Risiken, während das proaktive Management auf sehr aufwändige Lieferantenaudits der ersten Stufe des Netzwerkes beschränkt ist. Die Veröffentlichung [Emm12] beschreibt mit dem *Early Warning System* ebenfalls ein Risikomanagement für Logistiknetzwerke. Die Quelldaten dafür werden strukturiert aus ERP-, CRM-Systemen oder aus APIs von Internetdiensten (Google) ausgelesen. Big Data Technologien werden jedoch nicht berücksichtigt.

Zur Modellierung von Liefernetzwerken existiert mit dem am Fraunhofer IML entwickelten OTD-NET [Wag07] und dem darauf aufbauenden Framework OTD-Assist [Toth09] eine Modellierungsumgebung, die in zahlreichen Industrie- und FuE-Projekten eingesetzt wird, wie [Schw11] und [Cir12] aufzeigen. Die Modellierungsumgebung unterscheidet hierbei ein statisches Datenmodell, das Grundlage des OTD-Assist Frameworks ist und ein dynamisches Simulationsmodell für OTD-NET. Im bereits zuvor beschriebenen Projekt InKoRISK wurden dynamische Datenmodell-Konzepte zur Abbildung von Risiken in Liefernetzwerken entwickelt. Derzeit fehlt jedoch die Möglichkeit der Abbildung von Risiken im statischen Datenmodell des Liefernetzwerkes.

2.1 BIG DATA

Big Data beschreibt als übergeordneter Begriff Technologien für die Erfassung, Verwaltung und Analyse großer Datenmengen. In der Literatur wird bei der Definition von großen Datenmengen auf die drei großen V's Volume (Volumen), Velocity (Geschwindigkeit) und Variety (Vielfalt) verwiesen [Robak13, Wrobel12, Katal13]. Das *Volumen* charakterisiert die große Menge von Daten welche in der IT-Infrastruktur gespeichert sind. Alleine diese Speicherung birgt bereits Probleme für die Unternehmen und die in der Praxis eingesetzten Datenspeicher. Eine der größten Herausforderungen für IT-Infrastrukturen, die mit großen Daten konfrontiert sind, ist die Verfügbarkeit von Speicherplatz und einen effizienten Zugang zu gewährleisten. *Geschwindigkeit* beschreibt die große Menge von Daten, die in Echtzeit in den datenverarbeitenden Systemen eintrifft. Diese eintreffenden Daten müssen schnell gehandhabt, analysiert und zur weiteren Verwendung verdichtet und verarbeitet werden. Die *Vielfalt* bezieht sich auf die verschiedenen Strukturen von Informationen die in heterogenen IT-Landschaften von Unternehmen oder dem Internet wiederzufinden sind. Die gespeicherten Daten können entweder auf einer Struktur (Datentyp) beruhen oder gänzlich unstrukturierte Informationen sein (bspw. Fließtexte). Als viertes V wird ergänzend der Wert (Value) der Daten bezeichnet, den die zu verarbeitenden Daten besitzen und der zur weiteren Verwendung [Wrobel12] analysiert wird.

2.2 DATENQUELLEN

Um Datenquellen zu kategorisieren, bietet sich als eine Möglichkeit die Unterteilung in interne und externe Datenquellen an. Als *interne Daten* (Unternehmensdaten) werden die Informationen bezeichnet, welche innerhalb der operativen Geschäftstätigkeit des Unternehmens anfallen. Innerhalb der Business-IT-Systemlandschaft eines Unternehmens werden diese erzeugt, zusammengetragen, kombiniert und in Datenbanken vorgehalten. Ein Beispiel für solche Daten sind strukturierte Datensätze in der Datenbank eines ERP-Systems. Ebenfalls zu dieser Kategorie von Daten gehören Informationsätze welche in modernen Produktionssystemen erzeugt werden und bspw. Auskunft über den Zustand der eingesetzten Industrieanlagen, den Ausschuss der Produktion oder Bewegungskordinaten von intelligenten Lagereinheiten geben, die mit Funkfrequenzidentifikation (RFID) ausgestattet sind.

Externe Daten liefern beispielsweise Social Media Plattformen wie u. a. Facebook oder Twitter. Weitere Vertreter dieser Kategorie sind Datensätze von Datenportalen. Social Media Plattformen stellen Daten in Form von Datenströmen (aktuelle Posts, Nachrichten und Twitter Tweets) zur Verfügung. Diese sind unstrukturiert und die Semantik ist stark unterschiedlich und wechselt sehr schnell. Weiterhin ist der Zugang zu diesen Datenquellen nur beschränkt öffentlich zugänglich. Die Plattform Twitter beispielsweise limitiert Anfragen der eigenen API auf eine bestimmte Anzahl und bietet darüber hinaus nur maximal ein Prozent aller aktuellen Tweets zur Verarbeitung an [Baron]. Diesen Umstand machen sich spezialisierte Unternehmen wie DataSift zu Nutze. Ihr Geschäftskonzept ist die Archivierung von Social Media-Daten [Ghemawat04]. Diese Dienste ermöglichen einen Zugriff auf umfangreichere Social Media Datenbestände, allerdings nicht kostenfrei [Baron]. Des Weiteren können neben Twitter auch Quellen, wie die Google Search API, genutzt werden.

Außerdem können offenen Daten, welche von staatlichen und nicht-staatlichen Einrichtung zur Verfügung gestellt werden, verwendet werden. Diese Quellen sind sowohl für den kommerziellen als auch für den nicht-kommerziellen Gebrauch frei einsetzbar. Die Europäische Union bietet mit dem Portal open-data.europa.eu einen Katalog für viele in der EU zur Verfügung gestellten offenen Datenquellen an. Eine häufig verwendete offene Datenquelle ist Eurostat (epp.eurostat.ec.europa.eu), welche von dem Statistischen Amt der Europäischen Union bereitgestellt wird. Eurostat stellt zum Beispiel (geo)politische Informationen zu Regionen und Ländern zur Verfügung [SCC10]. Entsprechende offene Quellen können im Hinblick auf die Informationsgewinnung über Verkehrswege, Risiken aufgrund von politischen Unruhen oder Unterbrechungen durch Baustellen auf Autobahnen für die Logistikkette eingesetzt werden. Ein Großteil dieser Daten kann unter Zuhilfenahme von Web-

Technologien wie Sematic Web (SW) [Bern01, Raj08, BernLee, Pfisterer09] oder durch Verwendung von Linked Open Data (LOD) [Hofman11] durchsucht werden.

2.3 DATENMANAGEMENT UND -ANALYSE

Bei der Verarbeitung großer und strukturell heterogener Daten sind innovative Ansätze der Datenverarbeitung erforderlich. Herkömmliche Technologiekomponenten, wie SQL-Datenbanken sind nicht in der Lage, diese Datenmengen zu verarbeiten, insbesondere in Szenarien, in denen eine Echtzeitverarbeitung erforderlich ist. Herkömmliche Business-IT-Systeme fokussieren primär auf die Unternehmensprozesse innerhalb der Organisation. Die benötigten Datenstrukturen werden individuell entworfen und innerhalb vorgegebener Schemata innerhalb relationaler Datenbanken gespeichert. Moderne Enterprise-Architekturen hingegen werden auf Basis der Datenarchitektur eines Unternehmens geplant, um beispielsweise Big Data Analysen zu ermöglichen [Robak12]. Dabei sind innovative Datenbankkonzepte entstanden, um dynamische und unstrukturierte Datensätze zu verarbeiten.

Diese neuartigen Datenbanktechnologien werden als NoSQL (englisch für Not only SQL) Datenbanken bezeichnet. Diese Datenbanktypen verwenden keine festgelegten Tabellenschemata. Hierdurch können Datenstrukturen angepasst und verändert werden, ohne das existierenden Tabellen oder Einträge in den Datenbanken verändert werden müssen [Bakshi12]. Beim Entwurf dieser Technologie wurde insbesondere auf Anforderungen wie Skalierbarkeit und die Unterstützung verteilter Zugriffe geachtet [End11]. Für die Modellierung von Lieferketten sind Graphen Datenbanken, welche zu den NoSQL-Datenbanken gezählt werden, aufgrund ihrer engen Beziehung zu Anwendungen wie der Transportplan-Optimierung, Navigationssystemen oder sozialen Netzwerken, interessant [Soori12]. Graphen Datenbanken organisieren die Daten in einer Graphen Struktur. Die Informationen werden innerhalb der Knoten gespeichert. Zusatzinformationen werden als Objekte definiert [Soori12]. Schließlich beschreibt eine Kante die Verbindung zwischen verschiedenen Einträgen.

In enger Verbindung zu diesen Daten-Management-Konzepten für den Umgang mit großen Datenmengen stehen Technologien der Datenanalyse. Große verteilte Datenspeicher erfordern Batch-Analysen. In der Regel werden Batch-Analysen mit verteilten Aufgaben auf großen Datenspeichern durchgeführt, um Informationen zu suchen und extrahieren [Bakshi12]. Das Apache Hadoop Framework wird häufig für diese Batch-Analysen eingesetzt. Es besteht aus zwei Hauptkomponenten: Dem Hadoop File System (HDFS) für die Verwaltung von großen Datenmengen und einem MapReduce-basierten Framework für die Ausführung der Datenanalysen [Bakshi12, Katal13].

Eine Herausforderung bei der Nutzung der Datenanalyse ist neben der Analyse vorhandener großer Datenmengen die parallele Verarbeitung und Analyse der in Echtzeit und unregelmäßig eintreffenden Datenströme. Für diesen Anwendungsfall ist die Lambda-Architektur entwickelt worden, um Datenstrom-Analysen (englisch Stream Analytics) zusätzlich zu den Batch-Analysen [Robak13, Marz12] durchzuführen. Die Lambda Architektur erlaubt es jederzeit, Analyseergebnisse der aktuellen Datenbestände zu nutzen. Eine Batchanalyse auf den im verteilten Datensystem abgelegten Daten benötigt eine bestimmte Zeit n . Alle Daten, welche nun innerhalb des Zeitraums n erstellt werden bzw. als Datenstrom anliegen, werden im Rahmen der Datenstrom-Analyse direkt analysiert. Datenströme können auch durch Technologien wie das Complex Event Processing (CEP) verarbeitet werden [Robak13, Xingyi08]. Ein CEP kann beispielsweise für die Integration und Verarbeitung von Ereignissen von RFID- und Barcode-Lesegeräten sowie Sensoren eingesetzt werden.

2.4 BIG DATA-LÖSUNGEN FÜR LOGISTIKNETZWERKE

Heutzutage ist eine der Kernentwicklungen innerhalb von Logistiknetzwerken die Optimierung der Transparenz und Sichtbarkeit. Dies ist ein komplexes Problem aufgrund der Interaktion zwischen den beteiligten Personen, Prozessen, Technologien und Informationsflüssen [Heaney13]. Das Ziel der Verbesserung von Sichtbarkeiten und Transparenz in Logistiknetzwerken ist es, aktuelle Aktivitäten und potentielle Risiken entlang einer Lieferkette aufzuzeigen. Diese Informationen können Entscheider verwenden, um Unterbrechungen zu detektieren und auf diese zu reagieren [Mong13]. Verfügbare Datensätze sind oft veraltet, aber die Informationen sind für die Entscheidungsfindung [Barratt07, Mong13, Griffiths07 und Hofman11] erforderlich. Dieses Problem muss durch neue IT-Lösungen für das Lieferkettenmanagement und insbesondere für die Risikoidentifikation adressiert werden. Zwei auf Big Data basierende Lösungen für das Lieferkettenmanagement werden nachfolgend vorgestellt.

In [Hofman11] werden LOD (Linked Open Data) und SW (Semantic Web) Konzepte verwendet, um die SCV (Supply Chain Visibility) zu erhöhen. Es werden Geschäftsvorfälle ausgewertet und analysiert, um Umweltdaten zu sammeln und daraus die Supply-Chain-Netzwerke zu beschreiben. Diese werden mit internen Daten kombiniert, verdichtet und analysiert. Die Ergebnisse werden verwendet, um Informationen zu potentiellen Risiken und Unterbrechungen für alle Standorte und Transportwege innerhalb der Lieferkette abzuleiten. Für dieses Vorgehen wird eine Beschreibung in Form einer Ontologie für die ausgetauschten Daten entlang der Lieferketten benötigt [Haase22].

Ein weiterer Ansatz ist das RiskVis Framework, welches verfügbare Lösungen zur Visualisierung von Wert-

schöpfungsketten auf Basis von Google Maps und Baidu Maps mit Analyse-Ansätzen für interne und externe Daten vereint. Die Analyse wird anhand von Echtzeit-Daten durchgeführt. Das Ergebnis ist eine Echtzeit-Visualisierung der Versorgungskette für die Risikoüberwachung. Diese liefert die erforderlichen Informationen für Entscheider und kann zur Reduzierung der Komplexität der Lieferkette eingesetzt werden [Mong13].

3 KONZEPT DES PROAKTIVEN RISIKOMANAGEMENT IN LOGISTIKNETZWERKEN

Das Konzept für ein proaktives Risikomanagement in Logistiknetzwerken umfasst fünf Komponenten, welche in der Abbildung 1 dargestellt sind. Nachfolgend werden die Komponenten genauer betrachtet.

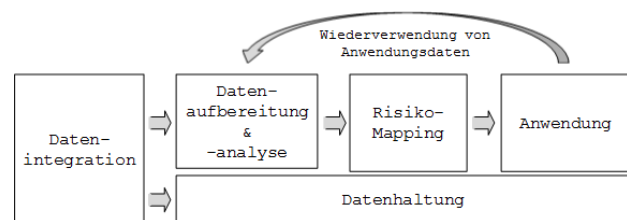


Abbildung 1. Konzept-Übersicht für ein proaktives Risikomanagement in Logistiknetzwerken

3.1 DATENINTEGRATION

Bei der Datenintegration müssen drei Punkte berücksichtigt werden, die in den nachfolgenden Abschnitten näher erläutert werden: 1. Die Verwendung von konstanten Datenquellen. 2. Die Integration von Datenströmen. 3. Die Terminologie für die Datenanalysen.

3.1.1 KONSTANTE DATENQUELLEN

Konstante Datenquellen bezeichnen interne IT-Systeme und Datenbanken aber auch Web-Dienste, beispielsweise Wetter-Dienste oder offene Datenquellen. Erstens ändern diese Quellen und die Datensemantik sich nicht oder nur selten. Außerdem bestehen entsprechende Quellen, IT-Systeme und Datenbanken in der Regel über Jahre. Zweitens umfassen konstante Quellen häufig ein großes Datenvolumen. Die Daten müssen aus diesem Grund erst auf ein verteiltes Datensystem überführt werden. Dies muss vor dem Hintergrund erfolgen, da die bisherigen Datenspeicherlösungen keine performante Datenanalyse und -Aufbereitung ermöglichen.

3.1.2 DATENSTRÖME

Im Gegensatz zu konstanten Datenquellen verfügen Datenströme über eine stark wechselnde Semantik. Das Datenvolumen ist ebenfalls stark schwankend. Es reicht von wenigen Datensätzen bis zu hin zu mehreren Gigabyte innerhalb einer Sekunde. Datenströme werden nicht nur von externen Diensten, wie den sozialen Medien, erzeugt.

Sie können auch unternehmensintern generiert werden. Webseiten können zum Beispiel durch Web-Crawler und Technologien wie LOD oder SW gesucht und analysiert werden. Die Analyseergebnisse werden zur Integration als Datenstrom bereitgestellt. Außerdem kann die Kommunikation zwischen internen IT-Systemen, zum Beispiel die Übermittlung von neuen Datensätzen ebenfalls als Datenstrom wahrgenommen und integriert werden. Entsprechende Daten können durch ein CEP erfasst und vorverarbeitet werden.

3.1.3 TERMINOLOGIE

Die Entwicklung einer Terminologie adressiert den Bedarf, das vorhandene Fachwissen mit den Datenanalysealgorithmen explorieren zu können. Ein Softwareentwickler kennt in den meisten Fällen die Fachdomäne nicht detailliert. Aus diesem Grund wird bei der Entwicklung und Ausführung von Analysediensten auf eine einheitliche Terminologie zurückgegriffen. Diese enthält für die Ermittlung von Informationen über das Logistiknetzwerk und entsprechende Risiken Begriffe in verschiedenen Sprachen.

3.2 AUFBEREITUNG, ANALYSE UND DATENHALTUNG

Die beiden Komponenten *Datenaufbereitung* und *analyse* sowie *Datenhaltung* sind eng miteinander verbunden, da die Datenanalysen zum Beispiel mit dem Hadoop Framework direkt auf dem verteilten Dateisystem HDFS ausgeführt werden. Analyseergebnisse werden außerdem wieder in das Dateisystem abgelegt. Nicht nur für die Analysen, sondern auch ganz allgemein bieten verteilte Ansätze in der Datenhaltung einen schnellen Datenzugriff, insbesondere auch auf größere Datenmengen. Die Analyseergebnisse können wiederum für weitere Analysen und für die Erstellung einer Statistik als Grundlage dienen. Die Analyse der Daten erfolgt zwei geteilt. Zuerst müssen Informationen über das Logistiknetzwerk ermittelt werden. Hierzu dienen beispielsweise Bestell- und Lieferaufträge sowie bisherige Transportrouten als Grundlage. Diese ermöglichen zusammen mit externen Daten wie Straßenkarten und Schiffsrouten die Zusammenstellung von Logistiknetzwerken. Im zweiten Schritt werden Risiken zu einzelnen Knoten und ganzen Routen ermittelt. Dies ist besonders wichtig, damit die ermittelten Informationen bei der weiteren Nutzung auch direkt einem oder mehreren Knoten eines Logistiknetzwerkes zugeordnet werden können. Für die Ermittlung von Risiken kann auf weitere Datenquellen, beispielsweise Verkehrsmeldungen oder Wetterdienste, zurückgegriffen werden. Die Verwendung von sozialen Daten ist für die Ermittlung und Früherkennung von Risiken, gerade aus einer globalen Perspektive, interessant.

3.3 RISIKO-MAPPING

Die Ergebnisse der Datenanalyse aus 3.2 werden im nächsten Schritt an die Risiko-Mapping-Komponente

weitergeleitet. Diese bildet die identifizierten Risiken auf die ermittelten Informationen zum Logistiknetzwerk ab. Dies erfolgt durch die Verwendung von Graphen Datenbanken und ist angelehnt an den Aufbau und Verwaltung von sozialen Netzwerken wie Facebook. Logistiknetzwerke verfügen über eine Vielzahl von Kunden und Zulieferbeziehungen für verschiedene Produkte und Dienstleistungen. Diese Punkte verfügen alle über Standorte, Adresse, Personen, Bewertungen und viele weitere Informationen. Graphen Datenbanken bilden entsprechende Informationen in Form von Knoten ab. Straßenkarten werden häufig auch in Graphen Datenbanken vorgehalten. Diese Softwaretechnologie bildet außerdem Daten und Verweise auf weitere Datensätze vergleichbar ab [Bern01, Raj08]. Wie in Abbildung 2 ersichtlich, sind die Beziehungen der Knoten durch Kanten abgebildet. Neue Knoten und Kanten lassen sich sehr einfach hinzufügen und verwalten. Ein Knoten kann beliebig viele Informationen (Meta-Informationen) enthalten, welche den Knoten genauer beschreiben. Ein Graph kann zum Beispiel mit einer Tiefen- oder Breitensuchen durchsucht werden.

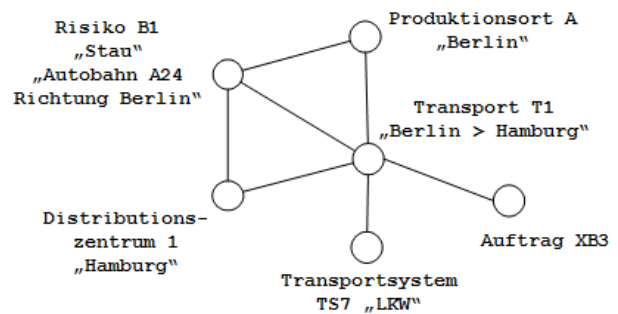


Abbildung 2. Graph eines Logistiknetzwerkes und Risiko-Mapping

Die Aufgabe der Komponente Risiko-Mapping besteht in der Erweiterung und Aktualisierung des vorgehaltenen Logistiknetzwerkes durch jedes neue Ergebnis der Datenanalyse. Im zweiten Schritt werden die identifizierten Risiken hinzugefügt und bestehende Risiken aktualisiert.

3.4 ANWENDUNG

Anwendungen müssen die Analyseergebnisse effizient nutzen können. Dafür können weit verbreitete Technologien wie RESTful Dienste eingesetzt werden. Über Parameter der Dienste können Abfragen und Tiefen- sowie Breitensuchen gezielt auf der Graphen Datenbank ausgeführt werden. Die Daten werden von den Diensten in einem allgemeinen Format, beispielsweise XML oder JSON, bereitgestellt und müssen von der aufrufenden Anwendung in deren internem System – sofern notwendig – übertragen werden. Die Anwendungsdaten können zur Verbesserung der Datenanalysen wieder in die Datenhaltungskomponenten überführt und somit als weitere Datenquelle eingesetzt werden. Entsprechende Daten können unter Verwendung von Analyseergebnissen entweder neue erzeugte Daten oder Protokollausgaben der

Anwendungen sein. Die Wiederverwendung dieser Daten kann auch für eine Optimierung der Analysealgorithmen eingesetzt werden.

4 REALISIERUNG

Die Realisierung des beschriebenen Konzeptes erfordert eine Nutzung und Verfügbarkeit von internen Daten aus unterschiedlichen IT-Systemen. Diese müssen entweder in einen neuen Datenspeicher in Form eines verteilten Dateisystems überführt werden oder stehen in Form von Datenströmen zur Verfügung. In vielen Unternehmen sind hierfür Anpassungen von Unternehmensrichtlinien erforderlich. Der Schutz und damit die Vermeidung der Nutzung von Privacy-bezogenen Daten muss außerdem sichergestellt sein. Lösungen für diese beiden Themen sind nicht Bestandteil dieses Beitrags, werden aber unter anderem in dem vom Bildungsministerium für Bildung und Forschung geförderten Forschungsprojekt PREsTiGE (Privacy-erhaltende Methoden und Werkzeuge für cloud-basierte Geschäftsprozesse) betrachtet (siehe <http://prestige.wifa.uni-leipzig.de/startseite.html>).

Für die Umsetzung des in diesem Beitrag vorgestellten allgemeinen Konzeptes ist eine schrittweise Einführung erforderlich. Im ersten Schritt werden die einzelnen IT-Komponenten installiert, konfiguriert und in Betrieb genommen. Anschließend können erste Daten aus internen konstanten Datenquellen überführt werden. Im dritten Schritt werden erste Datenanalysen entworfen und in Betrieb genommen. Anschließend werden weitere Datenquellen integriert und neue Datenanalysen entwickelt, bisherige aktualisiert oder ersetzt. Für die Integration von externen Datenquellen und Datenströmen müssen weitere IT-Komponenten und System installiert und konfiguriert werden. Die Verwendung und Wartung entsprechender Lösungen erfordert die Bereitschaft, fortwährend das System zu erweitern, zu pflegen und bestehende Lösungen zu ersetzen und grenzt sich damit eindeutig von bisherigen Vorgehensweisen ab, welche in einem festen zeitlichen Rahmen ein neues IT-System eingeführt haben.

5 FAZIT / AUSBLICK

Das Supply Chain Management erfordert neue fachliche und technische Lösungsansätze sowie eine Veränderung des Denkens über den Umgang mit großen Datenmengen und der zunehmenden Komplexität der Logistiknetzwerke. Die in diesem Artikel aufgeführten bereits existierenden Lösungsansätze sind nicht in der Lage dies vollständig zu leisten. Um dem oben genannten Ziel näher zu kommen, wurden bei dem hier vorgestellten Konzept die Architektur und ihre Komponenten für ein proaktives Risikomanagement in Logistiknetzwerken entwickelt. Das Konzept wird zukünftig mit interessierten Partnern (z.B.

aus der Automobil- oder Chemiebranche) schrittweise umgesetzt.

LITERATUR

- [Bakshi12] K. Bakshi, "Considerations for big data: architecture and Approach", Aerospace Conference, 2012 IEEE, ISBN 978-1-4577-0556-4, pp 1 – 7, 2012
- [Baron] P. Baron, "Big Data für IT-Entscheider", Carl Hanser Verlag München, ISBN 978-3-446-43339-7, 2013
- [Barratt07] M. Barratt, and A. Oke, "Antecedents of supply chain visibility in retail supply chains: A resource-based theory perspective," *Journal of Operations management*, vol. 25, pp. 1217-1233, 2007
- [Bern01] T. Berners-Lee, et al., "W3C Semantic Web Activity" [Online]. Available: <http://www.w3.org/2001/sw>, 2001
- [BernLee] T. Berners-Lee, et al., "Scientific American: The Semantic Web", [Online] http://www.sciam.com/print_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF212001
- [Cir12] Cirullies, J.; Schwede, C.; Toth, M. Intra-Simulative Ecological Assessment of Logistics Networks: Benefits, Concepts, and Tool Enhancement. In: Laroque, C.; Himmelspach, J.; Pasupathy, R.; Rose, O.; Uhrmacher, A. M. (Hrsg.): *Proceedings of the 2012 Winter Simulation Conference*, 2012.
- [Emm12] Emmenberger S.; Laurenzini E.; Thönssen B.: "Improving supply-chain-management based on semantically enriched risk descriptions". *Proceedings of 4th Conference on Knowledge Management and Information Sharing (KMIS2012)*, Spain, 2012.
- [End11] S. Edlich, A. Friedland, J. Hampe, B. Brauer, M. Brückner, „NoSQL Einstig in die Welt nicht relationaler Web 2.0 Datenbanken“, 2nd Edition, Hanser Verlag, ISBN 978-3-446-42753-2, 2011
- [Ghemawat04] S. Ghemawat, H. Gobiuff, S.T. Leung, "The Google File System", <http://static.googleusercontent.com/med>

- ia/research. google.com/de/archive/, 2004
- [Griffiths07] J. L. Griffiths, A. Phelan, K. A. Osman, and A. Furness, "Using item attendant information and communications technologies to improve supply chain visibility," *Agile Manufacturing*, 2007. ICAM 2007, pp. 172 – 180, 2007
- [Haase22] P. Haase, D1.1.5 –Updated version of the networked ontology model, EU F6 NeOn: Lifecycle support for networked ontologies, 2009
- [Heaney13] B. Heaney, "supply chain visibility – A Critical Strategy to Optimize Cost and Service", [Online] <http://aberdeen.com/Aberdeen-Library/8509/RA-supply-chain-visibility.aspx>, 2013
- [Hofman11] W. Hofman, "supply chain visibility with Linked Open Data for supply chain Risk Analysis", *Workshop on IT Innovations Enabling Seamless and Secure supply chains*, pp. 20-31, 08.2011
- [Katal13] A. Katal, M. Wazid, R. H. Goudar, "big data: Issues, Challenges, Tools and Good Practices", *IEEE Sixth International Conference on Contemporary Computing (IC3)*, pp. 404-409, 08.2013
- [Marz12] N. Marz and J. Warren, "Big data. Principles and best practices of scalable realtime data systems". Manning Publications, MEAP Edition, Manning Early Access Program big data version 7, 2012.
- [Mong13] R. Siow Mong, Z. Wang, et.al., "RiskVis: supply chain Visualization with Risk management and Real - time Monitoring", *IEEE International Conference on Automation Science (CASE)*, 207-212, 2013
- [Pfisterer09] D. Pfisterer, et al., "SPITFIRE: Toward a Semantic Web of Things", *Communications Magazine, IEEE (Volume: 49, Issue: 11)*, pp. 40-48, 2009
- [Raj08] S. K. Rajapaksha, et al., "Internal Structure and Semantic Web Link Structure Based Ontology Ranking", *ICIAFS 2008. 4th International Conference on Information and Automation for Sustainability*, pp. 86-90, 2008
- [Robak13] S. Robak, B. Franczyk, M. Robak, "Applying big data and linked data concepts in supply chains management," *2013 Federated Conference on Computer Science and Information Systems (FedCSIS)*, pp. 1215 – 1221, 2013
- [SCC10] supply chain Council (SCC), *supply chain Operations Reference (SCOR) Overview Version 10.0*, 2010, <https://supply-chain.org/f/Web-Scor-Overview.pdf>
- [Schw11] Schwede, C.; Toth, M.; Wagenitz, A.: Funktionsübergreifende Zusammenarbeit in Unternehmen durch Logistische Assistenzsysteme. In: Dieter Spath (Hrsg.): *Wissensarbeit - Zwischen strengen Prozessen und kreativem Spielraum*. GITO-Verlag, Berlin [Schriftenreihe der Hochschulgruppe für Arbeits- und Betriebsorganisation], S. 435-447, 2011.
- [Soori12] C. Sooriaarachchi, T. Gunawardena, B. Kulasuriya, T. Dayaratne, "A study into the capabilities of NoSQL databases in handling a highly heterogeneous tree", *IEEE 6th International Conference on Information and Automation for Sustainability (ICIAFS)*, ISBN 978-1-4673-1976-8 , pp. 106-111 ,2012
- [Toth09] Toth, M.; Wagenitz, A. Neue Wege für die effektive Planung logistischer Netzwerke - Dynamische Verfügbarkeitsplanung mit Hilfe von Assistenzsystemen. In: *Industrie Management*, H. 2, S. 55-58, 2009:.
- [Wag07] Wagenitz, A. : *Modellierungsmethode zur Auftragsabwicklung in der Automobilindustrie*, Univ., Diss., Dortmund, 2007.
- [Wrobel12] S. Wrobel, „big data – Vorsprung durch Wissen“, *Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS*. Presentation, www.iais.fraunhofer.de, 2012.
- [Xingyi08] J. Xingyi, "Efficient Complex Event Processing over RFID Data Stream", *Seventh IEEE/ACIS International Conference on Computer and Information Science*, ISBN 978-0-7695-3131-1, pp. 75-81, 2008

Jens Leveling, M.Sc. ist seit 2010 wissenschaftlicher Mitarbeiter im Bereich Anwendungsentwicklung für Cloud- und Datenbasierte Anwendungen am Fraunhofer-Institut für Materialfluss und Logistik IML. Er studierte zwischen 2005 und 2010 Informatik an der Fachhochschule Dortmund und ist seit Anfang 2014 Teamleiter für das Themengebiet Big Data Services. Seit Juli 2014 im Rahmen des Fraunhofer Innovationszentrum Logistik und IT (FILIT). Seine Forschungstätigkeiten umfassen die Entwicklung von Konzepten und Anwendungen im Bereich Cloud Computing, Big Data und Industrie 4.0 mit Fokus auf die Logistik.

Telefon: +49 231 9743-221, Fax: +49 231 9743-77-221,
E-Mail: jens.leveling@iml.fraunhofer.de
Fraunhofer-Institut für Materialfluss und Logistik IML,
Joseph-von-Fraunhofer-Straße 2-4, 44227

Dipl.-Inform. Arkadius Schier ist seit 2006 wissenschaftlicher Mitarbeiter am Fraunhofer-Institut für Materialfluss und Logistik IML. Zuerst in der Abteilung Leit- und Steuerungstechnik und seit 2010 in der Abteilung Software Engineering. Der Schwerpunkt seiner Forschungstätigkeit liegt im Bereich, verteilte Systeme, Cloud Computing und Datenschutz für Cloud-basierte Systeme. Er studierte Informatik zwischen 2002 und 2006 an der Fachhochschule Dortmund und ist seit 2008 Lehrbeauftragter an der Fachhochschule Dortmund für den Bereich „Drahtlose Sensornetzwerke“.

Telefon: +49 231 9743-481, Fax: +49 231 9743-77-481,
E-Mail: arkadius.schier@iml.fraunhofer.de
Fraunhofer-Institut für Materialfluss und Logistik IML,
Joseph-von-Fraunhofer-Straße 2-4, 44227

Francesco Luciano, B.Sc. ist Wissenschaftliche Hilfskraft in der Abteilung Software Engineering am Fraunhofer-Institut für Materialfluss und Logistik IML, Dortmund. Schwerpunkt seiner Arbeiten sind Industrie 4.0, Internet der Dinge und Integration externer Daten durch semantische Ermittlungsverfahren. Francesco Luciano studiert Wirtschaftsinformatik an der Universität Duisburg-Essen.

Telefon: +49 231 9743-529, Fax: +49 231 9743-77-529,
E-Mail: francesco.luciano@iml.fraunhofer.de
Fraunhofer-Institut für Materialfluss und Logistik IML,
Joseph-von-Fraunhofer-Straße 2-4, 44227

Prof. Dr.-Ing. Boris Otto ist Inhaber des Audi-Stiftungslehrstuhls Supply Net Order Management und Hauptabteilungsleiter Information Management & Engineering am Fraunhofer-Institut für Materialfluss und Logistik IML, Dortmund. Schwerpunkte seiner Forschung und Lehre sind Geschäfts- und Logistiknetzwerke, Unternehmensdatenmanagement sowie Enterprise Systems und Electronic Business. Boris Otto studierte in Hamburg Wirtschaftsingenieurwesen, promovierte bei Prof. Hans-Jörg Bullinger an der Universität Stuttgart zum Dr.-Ing. und habilitierte sich an der Universität St. Gallen bei Prof. Hubert Österle.

Telefon: +49 231 9743-655, Fax: +49 231 9743-77-655,
E-Mail: boris.otto@iml.fraunhofer.de
Fraunhofer-Institut für Materialfluss und Logistik IML,
Joseph-von-Fraunhofer-Straße 2-4, 44227