Proceedings of the 2014 Winter Simulation Conference
A. Tolk, S. Y. Diallo, I. O. Ryzhov, L. Yilmaz, S. Buckley, and J. A. Miller, eds.

BIG DATA FUELED PROCESS MANAGEMENT OF SUPPLY RISKS: SENSING, PREDICTION, EVALUATION AND MITIGATION

Miao He
Hao Ji
Qinhua Wang
Changrui Ren

IBM Research - China,
Building 19, Zhongguancun Software Park,
8 Dongbeiwang WestRoad,Haidian District,
Beijing, P. R. CHINA.

Robin Lougee

IBM T. J. Watson Research Center
1101 Kitchawan Road, Yorktown Heights,
NY, UNITED STATES

ABSTRACT

Supplier risks jeopardize on-time or complete delivery of supply in a supply chain. Traditionally, a company can merely do an ex-post evaluation of a supplier's performance, and handles emergencies in a reactive rather than a proactive way. We propose an agile process management framework to monitor and manage supply risks. The innovation is two fold - Firstly, a business process is established to make sure that the right data, the right insights, and the right decision-makers are in place at the right time. Secondly, we install a big data analytics component, a simulation component and an optimization component into the business process. The big data analytics component senses and predicts supply disruptions with internally (operational) and external (environmental) data. The simulation component supports risk evaluation to convert predicted risk severity to key performance indices (KPIs) such as cost and stockout percentage. The optimization component assists the risk-hedging decision-making.

1 Introduction

The fast-growing Internet and Mobile Internet empowers media and individuals to connect with the rest of the world anywhere at any time. The instantaneous access to a huge pool of data, which might be collected in real time or accumulated over time, enables companies to extract insights that were impossible before - we call it the *big data* era. Big data possesses four characteristics, i.e., volume, velocity, variety and veracity. Volume indicates the vast sea of available data. Velocity refers to the fact that data are generated at an unprecedented speed. Variety means the variety in format, including numeric, text, category data, etc. Veracity points out the difficulty of discerning useful data from useless data.

In this paper, we present an in-market research under development, which leverages big data to manage supply risk (interchangeably, sourcing risk). Supply risk is one of the four major risks (supply risk, production risk, distribution risk and demand risk) faced by supply chains. As sourcing is the initial operation of a supply chain, any risk events affecting the sourcing operations will have a cascaded effect down chain. There are many off-the-shelf software products for supplier management or contract management, most of which require suppliers to fill some questionnaires and evaluate a supplier by a weighted sum over pre-defined metrics. The scoring results usually stay unchanged until a contract expires or some unexpected (mostly negative) events happen.

Aware of the value of big data (a blend of internal and external data), we propose a novel three-phase supplier risk management process, which transforms the traditional supply management process into a responsive and dynamic one. Regularly, The first part of the process senses, predicts and alerts risk by fusing environmental and traditional data, where environmental data includes weather forecast, news, stock market, etc. Operational data are the historic transaction records with supplemental information such as delay, deficient rate. The second part models the major material flow of a supply chain, with a simulation engine available to evaluate the impact of supply fluctuations and disruptions. With the risk impact evaluated, the third part defines work flows to choose a good mitigation plan via what-if analysis by calling the simulation engine.

The rest of the paper is organized as follows. Section 2 reviews the literature of business process modeling and supply risk management. We note that the current supply risk management seldom incorporates environmental data such as weather or news. Section 3 describes the business process we build to manage the supplt risks and introduces the critical analytics and optimization technologies we are implementing. The paper is concluded in Section 4.

2 Literature Review

Supply chains are comprised of complex flows of goods, information and funds with tight links among companies and plants across the globe. Lengthened supply chains nowadays are extremely vulnerable to various risks. There are several ways to categorize supply chain risks based on purposes. Jüttner, Peck, and Christopher (2003) categorized risks by sources into environmental risks, organizational risks and network-related risks. Kleindorfer and Saad (2005) stated that there are two broad classes of supply chain risks - one that results from the mismatch of supply and demand, and one that rises from disruptions to normal activities. From the supply chain structure perspective, Chopra and Sodhi (2004) classified risks into three categories which are supplier-related, internal and customer-related. Our fous on the supply risk migitation fits into the Chopra and Sodhi (2004) framework.

There are many academic papers investigating how to hedge against supply risks. Christopher and Lee (2004) pointed out that the "confidence," resulting from en-to-end visibility, will increase the robustness of a supply chain and therefore a robust supply. Faisal, Banwet, and Shankar (2006) proposed to use Interpretive Structural Modeling (ISM) to model the enablers which can be adopted to mitigate supply chain risks. The ISM method models the dependence among enablers to identify the most effective action to take. Manuj and Mentzer (2008) developed a risk management strategy by employing a $2X2$ matrix of supply chain environment based on supply and demand risks. Yang, Tang, and Yan (2012) quantitatively analyzed the supply risk by defining the risk as the multiplication of probability of the risk and the consequence when the risk occurs. Fang, Zhao, Fransoo, and Van Woensel (2013) used approximate dynamic programming to solve for the optimal sourcing strategy (single, dual, multiple and contingent sourcing). Establishing their work on visibility into risks, none of the above research provided an execution plan on how to obtain the "visibility" into risk factors. Meng and Shi (2010) constructed the supply risk prediction model by using support vector machine (SVM). However, only internal risk factors are considered in their prediction model.

Stepping into the interconnection era enabled by Internet, Mobil Internet and numerous sensors installed in cities, we are collecting unprecedented amount of data. Companies are realizing the potential to extract actionable insights from big data, structured or unstructured and internal or external. In supply chain management, this type of mind-changing research work is emerging as well. According to IPCC (2013), awareness of climate risks is growing among leading companies and industry stakeholders due to the increased frequency and magnitude of some types of extreme weather and natural disasters (such as floods, tornadoes, hurricanes, snowstorms). A five-step framework was introduced in Environment Agency (2013) to increase the supply chain resilience to changing climate and extreme weather. Laderach, Lundy, Jarvis, Ramirez, Portilla, Schepp, and Eitzinger (2011) used global climate models (GCMs) to quantify the impact of climate change on coffee supply chains. Mitigating terror attack impact on supply chain has become a hot research topic after the 9/11 attack. Supply chain management and certain operational approaches under terrorist attack were well studied by Sheffi (2001). A so-called just-in-case strategy was proposed in Joseph and Subbakrishna (2002) to cope with terrorist attacks. Other supply risks (like bankruptcies,

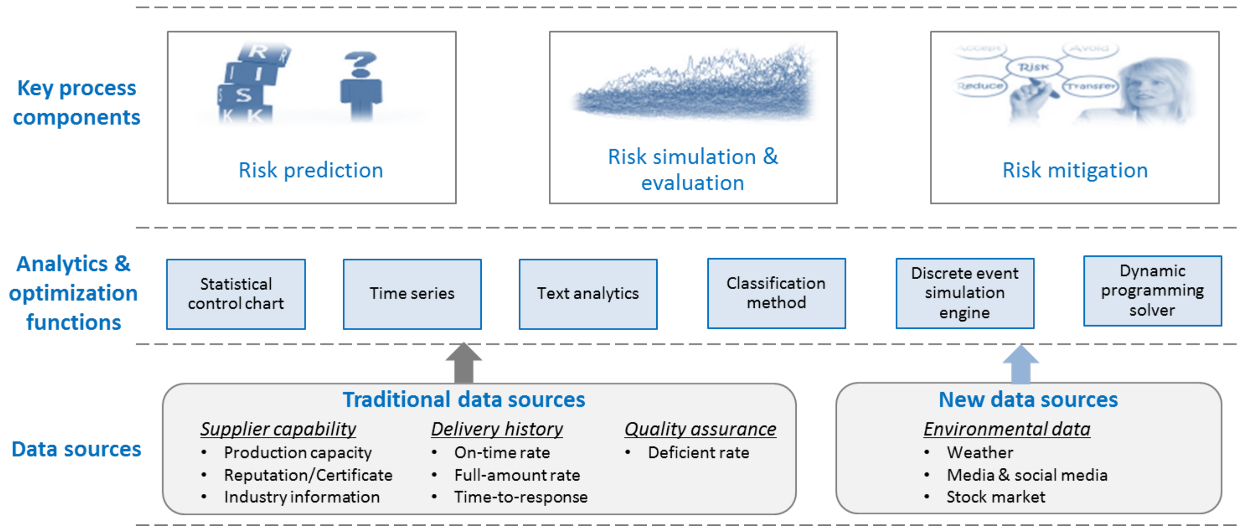

Figure 1: The architecture to support the proposed supply risk management process.

political changes and national unrest) will also definitely bring enormous influence on supply chain. The above work laid foundation for incorporating a single source of external data. Our work creates a full picture across silos.

A very relevant research was published by Schmitt and Singh (2009), where the authors collected a large spectrum of risk data and used monte carlo simulation to build risk profiling for each node in the supply chain. A discrete event simulation model was used to assess the impact of risks and effectiveness of mitigation plans as we propose. Our framework differentiate itself by the capability to sense and predict risks.

We implement the conceptual framework in the form of business processes modelling because it is a mature industrial application to "*enact*, *control* and analyze the operational processes involving humans, organizations, applications, documents, and other sources of information." (Van Der Aalst, Ter Hofstede, and Weske 2003). Besides, the business process modelling helps clarify the set of activities with their logical order and dependence considered (Aguilar-Saven 2004). Therefore, it is easier to enact a real process with a business modelling in place.

3 A Process View of Supply Risk Management

The architecture that supports the proposed business process is depicted in Figure 1. We introduce the architecture in a bottom-up manner.

3.1 The architecture

3.1.1 Data sources

Before contracting a supplier, a company usually evaluates its suppliers with relatively stationary information such as corporate size, production capability, reputation, and certificates. If the company has done business with the suppliers before, the historic performance in delivery and quality is also worthwhile to consider. For contracted suppliers, the company usually monitors their delivery and quality. Alerts are filed if some suppliers fail to meet the specified criteria.

We compliment the traditional data sources with new ones, which are made accessible by parties outside the supply chain. We start from weather, media & social media, and the stock market, and the number of data sources is increasing. In this paper we choose news to demonstrate how we extract supply risk from environmental data.

3.1.2 Analytics & optimization functions

The analytics and optimization functions support three key components in the business process, which are risk prediction, simulation and mitigation. Note that multiple functions are assembled to support a key component. Take the risk prediction component for an example, the text analytics extracts the key message from news; with key messages fed into some classifier, the risks are automatically classified based on its nature (e.g., political turmoil, nature disaster, economical downturn). The risk natures, in turn, affect the severity of impact that will be imposed on the supply.

3.1.3 Key process components

The key process components will be invoked according to the business process definition. They could be triggered either periodically or by events. Roughly speaking, risk prediction, risk simulation and risk mitigation are executed sequentially in the process.

3.1.4 Process management of supplier risk

Figure 2 shows the blueprint of the supply risk management process that we propose. The horizontal swimlanes divide the organizational responsibility. From up to down, the business units are data team, analytics team, supply chain team and emergence team, respectively. The verticals are the four major sections of the process, i.e., sensing, prediction, evaluation and mitigation.

There is a time trigger at the first block, which invokes automatic retrieval of external data based on schedule such as hourly. The data are parsed into structured ones before being stored into databases. The retrieved data (weather, news, stock market, etc.) are analyzed to predict if any significant risks are approaching (the fourth block). If yes, the first event trigger calls the risk magnitude prediction. The risk magnitude embodies the percent of supply in short, and the delayed supply arrival. If the system detects high-magnitude emerging risks, it alerts the supply chain team and feeds predicted risks into the simulation engine. Simulation converts the supply shortage or delay into the reduced service level, increased cost, and rising stockout percentage. If the simulated impact (dollars at risk) deserves an emergency response, the second event trigger sends message to the emergency handling team. The team will evaluate multiple mitigation alternatives via what-if analysis and select the best one.

In the below, we elaborate the key technologies implemented in the risk sensing, prediction, evaluation and simulation.

3.2 Sensing of risks

We share an intermediate result of risk sensing from CNN News. We filtered out articles containing predefined keywords. Part of the keywords are shown in Table 1. Figure 3 shows that the number of news with regard to risks varies much across time, resulting from the unpredictable nature of risks. We notice that the May 8 peak was caused by the Nigeria abduction of girls conducted by the terrorists.

Table 1: Risk keywords (partial).

Natural Disaster	**Terrorist Attack**	**Political Unrest**
earthquake	aircraft hijacking	rebel
landslide	massacre	rebel army
flood	human bomb	unrest
hurricane	terror attack	anti-government
tsunami	suicide attack	riot

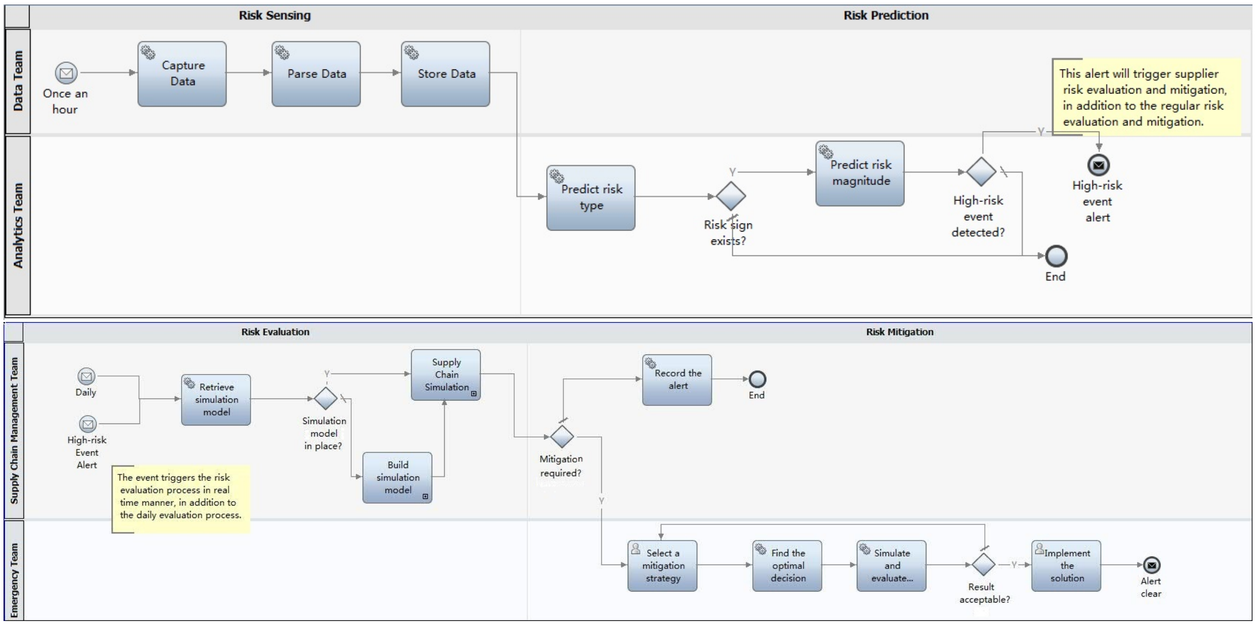

Figure 2: The supply risk management process.

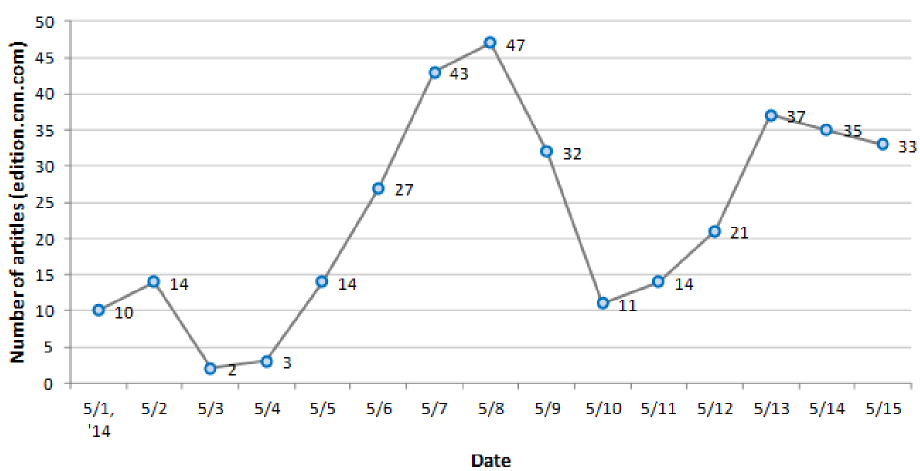

Figure 3: The number of articles involving risk keywords (published at editions.cnn.com, by day).

3.3 Big data fused risk prediction

The risk prediction packages two key technologies - *risk classification* and *risk severity prediction*. Risk classification identifies the risk types such as terror attacks, natural disasters, and anti-government rebels. Of course, there are some data sources which manifesto the types explicitly. Weather data, for instance, explicitly provide the disaster types (tornado, tsunami, hurricane, etc.). The risk type determines the model or parameter choice in risk severity prediction.

Figure 4: Natural disaster mapping of delay risk prediction.

3.3.1 Risk classification

We adopt a generative learning called multinomial event model to classify risk types. As the first step, we build a dictionary contains common words and index the words by l. The model views a piece of news, or more generally a document, as generated in the following steps. First, we choose the news type $p(y)$ from a multinomial distribution ϕ. The length of a document is n which varies across documents. We choose the j^{th} $(j = 1,\ldots,n)$ word according to some other multinomial distribution $p(x_j|y;\phi_y)$ conditioned on the news type. Therefore, the overall probability of a document is expressed as $p(y)\prod_{j=1}^{n} p(x_j|y)$. With document index i incorporated as the superscript, we have the below log-likelihood to maximize

$$\mathscr{L} = \prod_{i=1}^{m}[p(y^{(i)}|\phi)\prod_{j=1}^{n_i} p(x_j^{(i)}|y^{(i)};\phi_y)]. \tag{1}$$

The best estimate of the parameters is

$$\phi_{y=k} = \frac{\sum_{i=1}^{m} I(y^i = k)}{m}, \tag{2}$$

$$\phi_{l|y=k} = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i} I(x_j^{(i)} = l \wedge y^{(i)} = k)}{\sum_{i=1}^{m} I(y^{(i)} = k)n_i}, \tag{3}$$

where $I(\bullet)$ is the indicator function.

3.3.2 Risk severity prediction

For the risk severity prediction, we face the challenges of limited historic data as a result of the rarity of risk events. However, expert domain knowledge is integrated in our prediction module to quantize the risk impact effectively. We brief the risk impact prediction module as in the following.

For a special risk event E_t which is expected to occur at time t, the quantitative risk of this event can be defined as a combination of two components:

$$Risk(E_t) = P(E_t) \times L(E_t); \tag{4}$$

where $P(E_t)$ is the likelihood of the event occurring, $L(E_t)$ is loss resulting from the event occurring.

Based on the quantitative risk results, a final evaluation of risk severity can be obtained: insignificant, minor, serious or catastrophic. A further risk prediction of days to delay can be obtained via mapping rules provided by experts. Figure 4 shows an example.

3.4 Supply chain risk evaluation via simulation

The simulation engine is essential to transform the shortage or delay in supply to KPIs such as stockout percentage, service level, sourcing cost, etc. The simulation captures the joint impact of different suppliers according to the bills of materials (BOM) which specify how the raw materials are combined to produce final products. There might be thousands of parts or raw materials required by a manufacturer, which makes it hard to analytically evaluate the impact except for building a simulation model.

The discrete event simulation engine takes uncertain supply amount and delivery time as input, and output KPIs. It is composed with the basic simulation modules as shown in Table 2, and the design follows the principles in (Law 2007).

Table 2: Main module description of the simulation engine.

Module	**Description**
Main program	0. Invoke the initialization function (one-time) 1. Invoke the timing function (repeatedly) 2. Invoke the event function (repeatedly) 3. Invoke reporting function upon simulation termination
Initialization function	1. Set simulation clock = 0 2. Initialize system state and statistical counters 3. Initialize the event list
Timing function	1. Determine the next event type 2. Advance the simulation clock to the next event time
Event function	1. Update the system state 2. Update statistical counters 3. Call random variates generator to generate random events or generate deterministic events according to system evolving rules, and add events into event list
Reporting function	1. Compute the KPIs of interest 2. Display the results in tables or diagrams
Graphical Modeling module	1. Enable the drag-and-drop modeling 2. Enable the parameter input from graphical interface

Note that the graphical modeling module predefines common elements in supply chain such as the supplier, the factory, the warehouse, and the outlet. We also build in common state variables including inventory level, inventory in the pipeline, backorder queue and so on to facilitate modeling. The statistical counters however requires more script coding in order to generate KPIs of interest.

3.5 Supply chain risk mitigation

The supply chain mitigation is all about decision making. The emergency team needs to choose the best one from alternative mitigation plans. Here again, the simulation engine is applied to support the what-if analysis. We could simulate several typical supply risk mitigation policies -

- to wait for the suppliers to recover from disruption or fluctuation; or
- to contract some emergent supplier(s) with extra cost; or
- to change the product design/ingradients (and BOM is changed in this case); or
- to build inventory ahead of time; or
- to contract to redundant supplier(s) from the moment.

By modifying the supply chain according to the mitigation plan, setting new parameters and running the simulation, the business user could preview KPIs associated with different mitigation plans and select a reasonable one which strikes a balance between cost and downstream customer satisfaction.

4 Conclusion

In this paper, we present a conceptual framework, and an ongoing project at IBM Research too, to manage the supply risks in an agile way. The big data fueled supply risk management capitalizes on the blend of internal and external data, structured or unstructured, to sense the risk. The impact of risk is quantified by simulation. Further more, the solution supports what-if analysis to identify the best mitigation plan out of several.

The solution assembles a couple of analytics and optimization technologies to deliver an end-to-end solution in supply risk management. It makes the supply chain become agiler in risk identification and more responsive in risk mitigation.

From a solution provisioning view, we intend to develop a platform which could plug in most of data sources with minimal configuration. Some of the risk sensing algorithms could be reused to parse new data sources, but it is not always the case. Another shining point of our platform is that it is going to be consumable via API. We design to allow the data uploading to the cloud, and view the results from browsers (Internet Explorer/Firefox/Chrome).

REFERENCES

Aguilar-Saven, R. S. 2004. "Business process modelling: Review and framework". *International Journal of production economics* 90 (2): 129–149.

Chopra, S., and M. S. Sodhi. 2004. "Managing risk to avoid supply-chain breakdown". *MIT Sloan management review* 46 (1): 52–62.

Christopher, M., and H. Lee. 2004. "Mitigating supply chain risk through improved confidence". *International Journal of Physical Distribution & Logistics Management* 34 (5): 388–396.

Environment Agency 2013. *Assessing and managing climate change risks in supply chains*. Environment Agency.

Faisal, M. N., D. Banwet, and R. Shankar. 2006. "Supply chain risk mitigation: Modeling the enablers". *Business Process Management Journal* 12 (4): 535–552.

Fang, J., L. Zhao, J. C. Fransoo, and T. Van Woensel. 2013. "Sourcing strategies in supply risk management: An approximate dynamic programming approach". *Computers & Operations Research* 40 (5): 1371–1382.

IPCC 2013. "Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaption (SREX): Special Report of the Intergovernmental Panel on Climate Change".

Joseph, M., and S. Subbakrishna. 2002. "Targeting a just-in-case supply chain for the inevitable next disaster". *Supply Chain Management Review* 6 (5): 18 – 23.

Jüttner, U., H. Peck, and M. Christopher. 2003. "Supply chain risk management: outlining an agenda for future research". *International Journal of Logistics: Research and Applications* 6 (4): 197–210.

Kleindorfer, P. R., and G. H. Saad. 2005. "Managing disruption risks in supply chains". *Production and operations management* 14 (1): 53–68.

Laderach, P., M. Lundy, A. Jarvis, J. Ramirez, E. P. Portilla, K. Schepp, and A. Eitzinger. 2011. *Predicted impact of climate change on coffee supply chains*, Chapter 42, 703–723. Springer Berlin Heidelberg.

Law, A. M. 2007. *Simulation modeling & analysis*. McGraw-Hill.

Manuj, I., and J. T. Mentzer. 2008. "Global supply chain risk management strategies". *International Journal of Physical Distribution & Logistics Management* 38 (3): 192 – 223.

Meng, D., and C. Shi. 2010. "Research on the prediction model of aupply risk based on SVM". *The Proceedings of 2010 International Conference on Web Information Systems and Mining* 1:67 – 71.

Schmitt, A. J., and M. Singh. 2009. "Quantifying supply chain disruption risk using Monte Carlo and discrete-event simulation". In *Proceedings of the 2009 Winter Simulation Conference*, edited by M. D. Rossetti, R. R. Hill, B. Johansson, A. Dunkin, and R. G. Ingalls, 1237–1248. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Sheffi, Y. 2001. "Supply chain management under the threat of international terrorism". *The International Journal of Logistics Management* 12 (2): 1–11.

Van Der Aalst, W. M., A. H. Ter Hofstede, and M. Weske. 2003. "Business process management: A survey". In *Business process management*, 1–12. Springer.

Yang, P., J. Tang, and C. Yan. 2012. "A case study of global supply chain risk management". *The Proceedings of the 24th Chinese Control and Decision Conference (CCDC)*:1996–2000.

AUTHOR BIOGRAPHIES

MIAO HE is a Research Scientist in the Business Analytics and Optimization Department at IBM Research - China. Since joining IBM Research in 2009, she has focused on smarter commerce research, in particular demand-driven supply chain optimization, customer analytics and recommender systems. Her research is driven by solving applied industry problems for IBM and its clients in the growth markets. She holds a Master of Science degree in Management Science and Engineering from Tsinghua University, China. Her email address is hmhem@cn.ibm.com.

HAO JI is a Researcher at IBM China Research Laboratory. He joined IBM Research in 2012 after receiving his Ph.D. degree in Communication Engineering from Beijing University of Posts and Telecommunications in Beijing, P.R. China. His research interests include data mining, natural language process and supply chain management. His e-mail address is jihao@cn.ibm.com.

QIN HUA WANG is a researcher at IBM Research - China. She joined IBM Research in 2007 after receiving her M.S degree in Industrial Engineering from Tsinghua University in Beijing, P. R. China. Her research interests include supply chain management, business process management and enterprise resource planning system. Her email address is wangqinh@cn.ibm.com.

CHANG RUI REN is a Senior Manager and Senior Technical Staff Member in the Business Analytics and Optimization Department at IBM Research - China. He received his Ph.D. degree in Control Science and Engineering in 2005 from Tsinghua University, Beijing, China. He subsequently joined IBM Research - China, where he is dedicated to developing innovative industry solutions based on analytics and optimization techniques. His research interests include supply chain optimization, smarter commerce, business process management and predictive asset maintenance. He is a member of the IBM Academy of Technology. His email address is rencr@cn.ibm.com.

ROBIN LOUGEE is the Global Research Lead for the Consumer Products Industry at the IBM T. J. Watson Research Center in Yorktown Heights, NY. She received a Ph.D. in Mathematical Sciences from Clemson University. Her research deals with mathematical modeling and computational optimization. Robin is a founder of the open-source initiative for operations research, COIN-OR (www.coin-or.org). Her email address is rlougee@us.ibm.com.