

# Data Mining for Supply Chain Management in Complex Networks

**Mahesh S. Raisinghani**

*University of Dallas, USA*

**Manoj K. Singh**

*i2 Technologies, USA*

## INTRODUCTION

Supply chain comprises the flow of products, information, and money. In traditional supply chain management, business processes are disconnected from stock control and, as a result, inventory is the direct output of incomplete information. The focus of contemporary supply chain management is to organize, plan, and implement these flows. First, at the organizational level, products are manufactured, transported, and stored based on the customers' needs. Second, planning and control of component production, storage, and transport are managed using central supply management and replenished through centralized procurement. Third, the implementation of the supply chain involves the entire cycle from the order-entry process to order fulfillment and delivery. Data mining can create a better match between supply and demand, reducing or sometimes even eliminating the stocks.

Data mining thus has become an indispensable tool in understanding needs, preferences, and behaviors of customers. It is also used in pricing, promotion, and product development. Conventionally, data mining techniques have been used in banking, insurance, and retail business. This is largely because of the fact that the implementation of these techniques showed quick returns. Data mining is being used for customer profiling where characteristics of good customers are identified with the goals of predicting new customers and helping marketing departments target new prospects. The effectiveness of sales promotions/ product positioning can be analyzed using market-basket analysis to determine which products are purchased together or by an individual over time, which products to stock in a particular store, and where to place products in each store (Groth 2000; Kopanakis & Theodoulidis, 2003; Weir, 1998). In addition, data mining is used in a variety of other industries such as the financial, healthcare, and telecommunications industry, among others.

There are a lot of opportunities and applications of data mining even beyond the obvious. One of the potential areas is "Supply Chain Management." One of the realities of the demand and supply in the manufacturing

industry is that no matter how well balanced a system is, there is an element of uncertainty that creates a mismatch between demand and supply. The objective of this article is to identify those areas in the supply chain where most of the uncertainty exists and to determine suitable data-mining methods to accurately predict uncertainty. The underlying assumption of this paper is that a data warehouse has been implemented before the data-mining techniques can be applied.

## A KEY BUSINESS PROBLEM: UNCERTAINTY

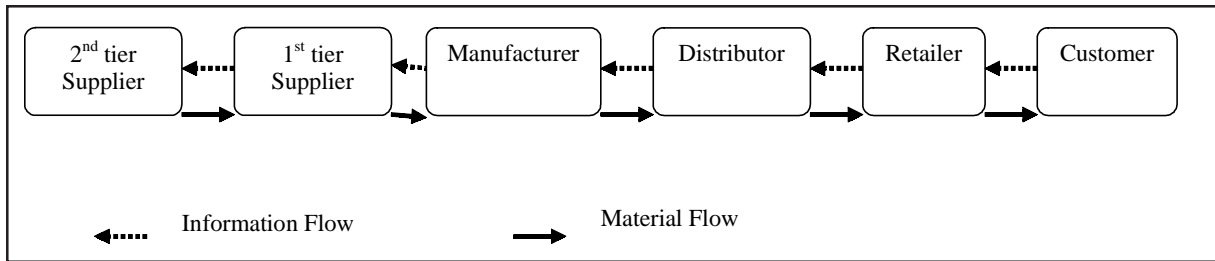
There are two issues that plague supply chain management—variation in the demand and supply and variation in the speed and extent of communication within the supply chain. Variation in demand and supply is due to the inherent uncertainty also present in the processes. Accurately predicting the uncertainties in demand, supply, and processes and then formulating action plans around the prediction is the essence of supply chain management (SCM).

Before we can address the problem of uncertainty in supply chain and explain the use of data-mining techniques, we need to understand the basic process of SCM and where uncertainty exists. In its most simplified form, a supply chain can be depicted as the flow of information from a customer's customer to a supplier's supplier and then the flow of material in the reverse direction, as shown in Figure 1.

The whole supply chain can be conceptually broken down as Supply – Process – Demand. Traditional forecasting planners of supply chains use demand and supply forecasting as a means of controlling uncertainty. However, there are three major drawbacks in those methods, namely:

1. Incorrect forecasting model
2. Incorrect number of parameters
3. Incorrect coefficients values of these parameters.

Figure 1. Information and material flows in the supply chain



Each of these three problems can be solved using data mining. The models are chosen from a finite set of pre-defined models in data mining. The model can be recreated as many times as needed in order to extract previously unknown patterns and relationships in data. When forecasting using data-mining techniques, the program can detect even minor effects of some parameters.

## DATA MINING: METHODS AND PROCESS

Data mining is the process of extracting ideas in data. It can also be defined as “a decision support process that tries to discover patterns and relationships that are beyond the realm of human experience and imagination in the large database and present them to a knowledgeable user for review and examination” or “as the process of extracting previously **unknown, valid, and actionable** information from large databases and then using the information to make crucial business decisions” (Groth, 2000).

Data mining not only uses a discovery-based approach in which pattern matching and other algorithms are employed to determine the key relationships in the data but also describes the steps that must be taken to ensure meaningful results.

## DATA-MINING METHODS

Data mining is used to build six types of models aimed at solving business problems: classification, regression, time series, clustering, association analysis, and sequence discovery.

### Classification

A predictive model is generated based on the historical data. These models are used to assign instances to a group or class by calculating the value of a categorical

variable. The value of this categorical variable is generally binary in nature. It can include multiple but discrete values.

### Regression

Regression is used to predict values for categorical variables. The values are continuous, real numbers, i.e., it has decimal values, and it has no fixed range in which the values of the variables are fitted.

### Time-series forecasting

This method uses a series of existing values and their attributes to forecast future values, except that the values of the categorical variables are dependent on time. Using various data-mining tools, the distinctive features of time can be exploited.

### Clustering

Clustering is used to segment a database into clusters, with the members of each cluster sharing a number of interesting properties. These clusters are not predefined and have two basic uses: 1) summarizing the contents of the target databases; and 2) as inputs to the other methods like supervised learning.

### Association

Association is used to describe behavior that is captured in the database. This method relates the occurrences of various events by identifying patterns or groups of items.

### Sequencing

Sequencing defines items that are likely to occur together on a sequence basis. This could help marketers in timing their promotions to correlate with the sequential buying order exhibited by their customers.

4 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/data-mining-supply-chain-management/14323](http://www.igi-global.com/chapter/data-mining-supply-chain-management/14323)

## Related Content

---

### E-Mail as a Strategic Tool in Organizations

Denise Johnson McManus and Houston Hume Carr (2005). *Encyclopedia of Information Science and Technology, First Edition* (pp. 1030-1033).

[www.irma-international.org/chapter/mail-strategic-tool-organizations/14381](http://www.irma-international.org/chapter/mail-strategic-tool-organizations/14381)

### An Integrated Strategic Framework for E-Government Initiatives

Fatemeh Saghafi, Behrouz Zarei, Akbar Kary Dolat Abadi and Kolsoom Abbasi Shahkooh (2013). *Managing Information Resources and Technology: Emerging Applications and Theories* (pp. 67-79).

[www.irma-international.org/chapter/integrated-strategic-framework-government-initiatives/74500](http://www.irma-international.org/chapter/integrated-strategic-framework-government-initiatives/74500)

### Developing a Web Service Security Framework

Yangil Park and Jeng-Chung Chen (2009). *Encyclopedia of Information Science and Technology, Second Edition* (pp. 1072-1078).

[www.irma-international.org/chapter/developing-web-service-security-framework/13708](http://www.irma-international.org/chapter/developing-web-service-security-framework/13708)

### Cluster-Based Vehicle Routing on Road Segments in Dematerialised Traffic Infrastructures

Christophe Feltus (2022). *Journal of Information Technology Research* (pp. 1-14).

[www.irma-international.org/article/cluster-based-vehicle-routing-on-road-segments-in-dematerialised-traffic-infrastructures/299373](http://www.irma-international.org/article/cluster-based-vehicle-routing-on-road-segments-in-dematerialised-traffic-infrastructures/299373)

### Implementation of Electronic Records in a Medical Practice Setting

Jami M. Clark (2014). *Cases on Electronic Records and Resource Management Implementation in Diverse Environments* (pp. 211-225).

[www.irma-international.org/chapter/implementation-electronic-records-medical-practice/82650](http://www.irma-international.org/chapter/implementation-electronic-records-medical-practice/82650)