Ann Oper Res
DOI 10.1007/s10479-016-2342-x

Customer demand analysis of the electronic commerce supply chain using Big Data

Lei Li[1] · Ting Chi[1] · Tongtong Hao[1] · Tao Yu[2]

Abstract With the advent of the Internet and the flourishing of connected technology, electronic commerce has become a new business model that disrupts the traditional transactional model and is transforming the consumer's lifestyle. Electronic commerce leads to constantly changing customer needs, therefore quick action and collaboration between production and the market is essential. Meanwhile, the abundant transactional data generated by electronic commerce allows us to explore browsing behaviors, habits, preferences and even characteristics of customers, which can help companies to understand their customer's needs more clearly. Traditional supply chain management (SCM) simply cannot keep up with electronic commerce because demand forecasts are constantly changing. Customer demands create and affect the whole supply chain. The purpose of SCM is to satisfy the customers who support the company by paying for the products; so meeting changing customer needs should be incorporated into SCM by developing demand chain management (DCM). In this paper, we explore how DCM can perform better in the electronic commerce environment based on studying website behavior data and using data analytics tools. The results show that DCM performs much better when paired with the benefits of electronic commerce and Big Data than traditional SCM methods.

Keywords Supply chain · Demand chain · Electronic commerce · Big Data

1 Background

With the rapid development of information technology, electronic commerce has become an integral part of people's daily lives. The network has the capacity to provide food, cloth-

✉ Lei Li
lilei@tju.edu.cn

1 School of Management and Economics, Tianjin University, 92, Weijin Road, Nankai District, Tianjin, China

2 Mmdata Marketing Company, Building 1, Longxi Road, Changning District, Shanghai, China

Published online: 06 October 2016

ing, travel, entertainment, education and almost any other conceivable possibility. In a narrow sense, electronic commerce refers to global business trade activity using electronic tools such as the telegraph, telephone, radio, television, fax, computer or network, mobile communication, etc. (Turban et al. 2002). Electronic commerce represents a diverse portfolio of business arrangements based on computer networks, including goods and services exchange, advertising, consumer purchasing, intermediaries and other relevant models. The aforementioned body of possible transactions composes the general understanding of electronic commerce. In a broad sense, electronic commerce is a derivative of electronic business, and is a means for completing business affairs through an electronic medium. Companies can communicate internally or with suppliers, customers and partners and share business information or implement electronic business processes between enterprises. Connecting with technology in this way helps to improve the efficiency of production and distribution channels, inventory management, and the financial health of the enterprise. By December 2014, electronic commerce transactions in China exceeded more than 12.3 trillion Yuan with growth rate of 21.3 % (Taobao Analysis 2015). With the development of electronic commerce, traditional supply chain management models have been subject to innovative new changes, leading to a more intuitive and flexible user experience. The next generation of customer interactions with online business presents new challenges and opportunities for supply chain management.

In traditional supply chain modeling, upstream and downstream firms share benefits and information by connecting businesses to each other in order to realize mutual compensation and the advantages of common development. Classical supply chain interaction is characterized by the fact that retailers, suppliers and factories no longer act as isolated units, but rather as a collaborative body. It is important to focus on the sustainability of supply chain (Maria Vanalle and Blanco Santos 2014). This platform provides services or products to customers via different processes and activities, illustrated in Fig. 1 below:

Figure 1 draws clear distinctions between the parties that take part in the supply process: customers, vendors and manufacturers. These entities are linked by information, capital, materials service and the flow of the product (Lambert and Cooper 2000). Products and value are created with the flow of capital and information, and different flows travel in different directions. The consequence of this directionality is that the functions and management of each of the different flows are distinct. Suppliers organize manufacturing and assembly

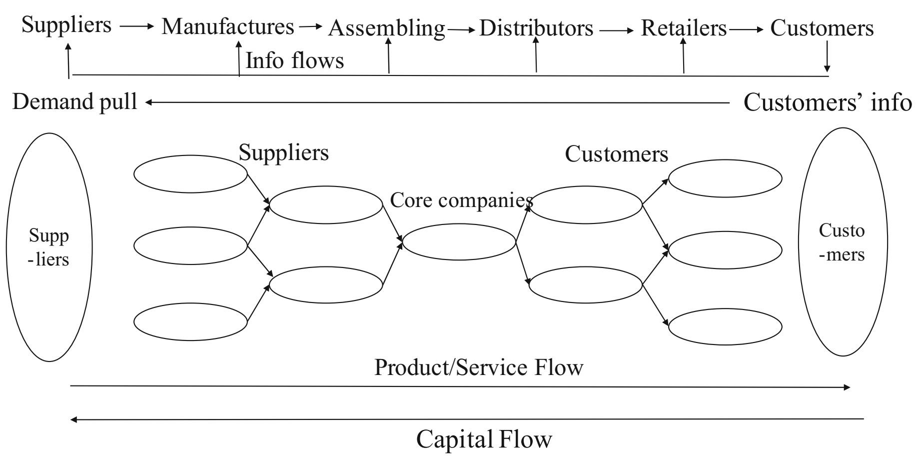

Fig. 1 The mechanisms and operation of the traditional supply chain

according to the customer needs at the end of the supply chain. Consequently, for efficient supply chain management, cooperation between all the companies in the chain is crucial to build a mutually beneficial relationship. The management of these complicated and sophisticated strategic relationships is the definition of SCM originally proposed by Gibson et al. in 2005. The Global Supply Chain Forum (GSCF) defined SCM as the integration of all flows that create value and satisfy the final customers (or stakeholders) (Drucker 1998). Apparently, the existence and adjustment of the definition of SCM are based on customer demands, which drive the flows to operate smoothly and dynamically. It's important to identify customer needs and design reasonable supply strategy using information flows. The process of crafting supply strategies around these information flows is called demand chain management (DCM). One of the causes of bullwhip effect is an inaccurate demand forecasts that results in inventory pressure and wasteful management of resources. The bullwhip effect refers to the distortion of demand information propagates upstream in an amplified form in supply chains. The causes of bullwhip effect include demand forecast, repertory management, batch order, delivery cycle, price wave, etc. The most direct factor is demand forecast.

Compared with traditional supply chain dynamics, connected technology and electronic commerce provide a more flexible solution to avoid damaging bullwhip effects by forecasting demands much more precisely thanks to the advent of Big Data. In other words, Big Data can turn numbers into information, offering a competitive edge for forward-minded modern businesses. Using Big Data, online retailers could collect important information like what customers usually bought, how often they made purchases online, how they chose to navigate a retailer's website, and their interest in promotions and advertising. Going further, more detailed characteristics of individual shoppers are also available. Data like age and gender, or whether the design and layout of website was appealing to them all help create a clearer picture of the customer pool for retailers. The development of complicated algorithms to realize precision marketing is shown in Fig. 2.

The electronic commerce supply chain (ECSC) and the traditional supply chain (TCSC) models are characterized by the following differences (Grieger 2003; Wannenwetschl and Nicolai 2004).

First, each model draws from a different pool of customers. In the traditional model of supply chain management, the target customers served by certain corporations are well

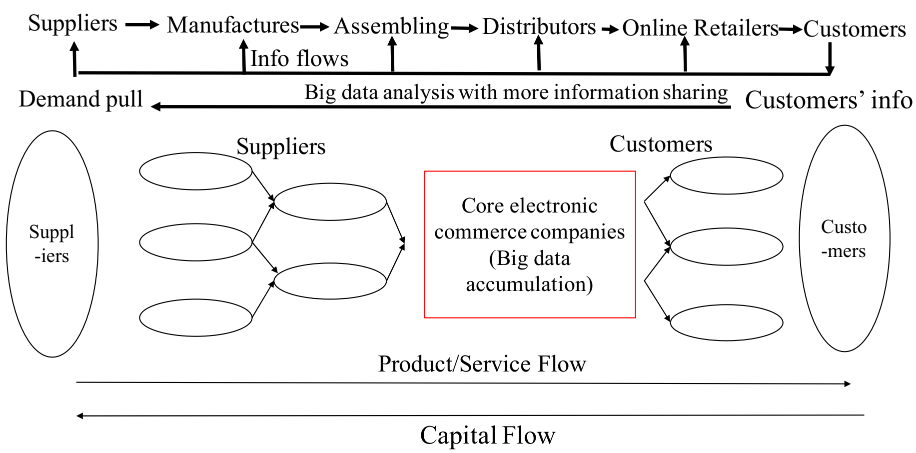

Fig. 2 The mechanisms and operation of the electronic commerce supply chain

defined. Thus, the providers of supply chain management services are well acquainted with the needs of their customers and are able to tailor services and products to adequately meet their personal appetites. With the advent of e-commerce and stronger demand for faster, more efficient, and more accurate logistics, the image of "the target customer" begins to become obscured. Customers can order the products or services online with following their own guidelines on personal consumption needs, price and convenience. The customer has become submerged under the network, and it's becoming increasingly difficult to extract key features of these target customers.

Next, electronic and traditional SCM call for a different flow of logistics. In traditional supply chain dynamics, there are no real-time updates on the status of logistic flow and the information of geographically-distributed customers. This model lacks the capability to personalize service to each customer due to the shortcomings of disconnected technology; it can only achieve consolidated transportations and undifferentiated services to centralized destinations. By contrast, the supply chain of e-commerce is fully customized according to the specific needs of each customer. Such logistics are distributive rather than centralized in order to achieve low freight-cost through consolidated transportation on the one hand, as well as superior services through differentiated distribution network and point-to-point delivery on the other hand.

Additionally, both models have different systems for managing all of the logistical information. Traditional supply chains typically manage the flow of logistical information through manual collection and transmission. This way of processing information is unilateral thereby making the information owned by the supplier and demander asymmetric. This information management system is normally based on a stand-alone machine, or at best a local-area-network system confined to an internal web. In e-commerce, the online logistics information is digitally collected and transmitted between the information supplier and demander, so the process is bilateral and symmetric. A software system manages the flow of information, which ensures the timeliness, accuracy and effectiveness of the whole process.

Electronic commerce and traditional models also have different focuses regarding supply chain management. Classical supply chain management always emphasizes the stability and consistency of the logistics flow, otherwise disorder or even chaos could erupt in the logistics. Any fluctuation or variation during the logistics process may bring huge losses to the upstream or downstream corporations. However, for e-commerce, the demands of customers are quite different and personalized, which creates a shifting seasonality, inconsistency and discontinuity in customer demands. Because data manipulation must be advanced and precise to handle this mercurial environment, we see the major value-add contributed by e-commerce is through its highly efficient information management schema. Therefore, the modern focus of supply chain management has become tailored to quick and flexible operations in response to fluctuating demand, providing high-quality services, and coming out with a low total cost.

Further, e-commerce and its predecessor entertain different operational models. The traditional supply typical employs "pushing" operations. By "pushing", we mean all the activities originate from the planning and production of the manufacturers. After this they must overcome the difficulty of transporting products through space and time to ultimately deliver them to the market or the end customers through careful logistics. The product flow is "pushed" forward from the manufacturing to the final customers. Alternatively, in the e-commerce supply chain, all activities originate from the specific demand of the end markets or customers. The production planning, distribution, warehousing and delivery are all based on the customer orders, and the flow of capital and goods are also revolving around the end markets. The difference between these two systems is akin to the discrepancy between a major t-shirt manufacturer and a custom t-shirt studio: one pushes a variety of standard products to the market, while the other generates content based on the needs of its customer base. Therefore, the supply chain of e-commerce is a typical "pulling" operation driven from the demands of end customers.

Finally, ESCM and TSCM have disparate flows of stocks and orders. In the traditional supply chain, the flows of stocks and orders are unidirectional, occurring without much communication between buyers and sellers. In e-commerce, the flows become bidirectional: each customer is able to customize, monitor or even revise their stocks and orders. At the same time, manufacturers and distributors can also adjust the stocks and orders according to customer needs with the intention to maximize the utility of proper SCM.

In order to maximum the efficiency of electronic commerce supply chain management, Big data is becoming a powerful tool (Mishra et al. 2016). From Figs. 1 and 2, we can also tell how Big Data benefits DCM and is distinguished from normal data in the four dimensions presented here (Waller and Fawcett 2013).

1.1 Variety

In the past, manual data-recording processes was time-consuming, tedious and prone to error prone. The data itself has always been simple and static, but it was the method of collection that prevented data science to reach its full capacity in the past. Commercial websites have enormous registers of detailed and accurate data, including log-on data, browsing data, and the click-through rate of advertisements, all of which can be used to make assumptions on the behavior and interests of target customers thereby promoting sales and generating more impactful customer forecasts.

1.2 Volume

The Internet generates an incredible volume of data every second. For example, taobao.com had 0.19 billion active users per day on 11st November 2015, which accounts for over 13 % of the total Chinese population (Mobile Electronic Commerce Industry 2015). Compared with manually created records, the quantity, dimension and depth of this information is quite impressive and gives companies a great opportunity to explore business intelligence.

1.3 Velocity

In many cases real-time information is extremely important, especially in the ever-changing climate of our market, which based on customer needs. For example, the environmentally driven haze in Beijing became so bad in September, 2015 that people were more likely to search for protective products on Internet (Index Analysis 2016). If a given supply chain was able to digest this information and increase production in a timely manner, they would have greatly increased their potential for profit.

1.4 Sustained and controllably-sourced data

Our collectively advancing prowess in information technology ensures the correctness and reliability of Big Data and our ability to manipulate it, both of which can help to avoid the errors of data collection more commonly found in antiquated collection techniques. It's vital that data sources are consistent and flexible to facilitate the normal operation of supply and sales.

Big Data has become a core competence in many industries by consolidating productivity, growth, and innovation. For this technology to truly sharpen the competitive edge, all parties in the supply chain must take Big Data seriously and tailor their products and services accordingly. In summary, Big Data and electronic commerce are favorable for both SCM and the demand forecast. This paper is organized as follows: first there is a brief literature review relating SCM to electronic commerce, customer demands and Big Data analysis. Next, we summarize the methodology that supports our empirical study. By unpacking different methods of customer need analysis via their collected browsing data, we ultimately offer our suggestions and conclusions outlined according to our research results.

2 Literature review

There are many mature theories and case studies of both SCM and DCM. We first reviewed theory development in the field of SCM and then focused on customer demand and the market of DCM. Then we will introduce how electronic commerce and Big Data improve DCM and provide a foundation for empirical study.

As defined earlier, SCM is the management and coordination of manufacturers, vendors, distributors, retailers and customers. SCM has attracted a lot of attention in the business community because of its cost-reducing properties and efficiency. Strictly, SCM is not just a single link but also a coordinated network of multiple business parties tied through tactical partnerships. The concept of SCM first appeared in the mid-1980's until eventually the framework that we know today was created by Cooper et al. The literature of SCM can be broken up into several categories, such as papers that specialize on specific purpose, multiple functions, operation mechanism or efficient coordination of the supply chain. The general purpose of SCM is to make full use of available resources, to satisfy customers, and to reduce cost. SCM has traditionally been divided according to three of its distinct aspects: the management parties, business processes and structure (Cooper et al. 1997). Figure 3 provides a visual depiction of this tri-fold relationship:

With the development of society and economics, the importance of synergies between customer demands and SCM has gained increasing attention. This dynamic style of SCM is referred as demand chain management (DCM) and has gained more recognition in recent years. DCM entails designing and adjusting the supply network by acutely analyzing the final customer's needs. Excellent DCM performance can help industries enhance competitiveness and has been shown to have a substantial positive impact on the overall supply chain performance. For example, Hilletoft came up with the idea that both demand and supply processes

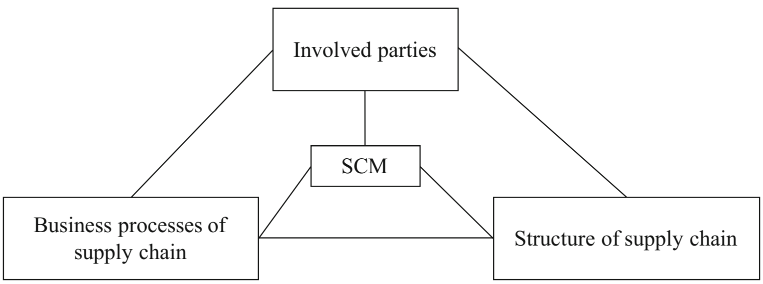

Fig. 3 Elements of supply chain management

are equally important in value creation, innovation and technology (Hilletofth 2011); while Jüttner et al. thought that there are three core factors of DCM: integrating the demand and supply processes, managing the relationship between marketing and SCM, and balancing the integrated processes and segments (Jttner et al. 2007); Laari et al. also stated that manufacturers can quickly answer their customer's demands by observing the environmental requirements upstream in the supply chain (Laari et al. 2016).

To address the irregularly changing nature of demand due to economic globalization, shifting government policies, climate changes and the progress of high-technology, researchers have proposed many theories and methodologies. Tsiakis et al. (2001) proposed a model based on a mathematical programming formulation and designed multi-echelon supply chain networks bound by demand uncertainty; separately, a stochastic programming method was developed to optimize the SCM in the face of demand uncertainty based on customer satisfaction and production expenditures (Gupta and Maranas 2003); meanwhile Fattahi et al. (2015)developed a mixed-integer linear programming (MILP) model to factor in price-sensitive demands; finally Cachon and Lariviere proposed that demand forecasts in sharing regulation could help the supply chain operate more smoothly and enable it to dynamically adjust capacity over time (Cachon and Lariviere 2001).

Despite the large body of modern research revolving around SCM and DCM, these solutions are not enough to cope with the tricky demand forecast and collaboration problems. The emergence of electronic commerce also brought an explosion of information and a host of new purchasing models. Further, the Internet has created many challenges throughout the transaction processes as outlined in the differences we described in the background. For instance, the problem of building solid relationship between different parties and reducing conflicts of interest throughout the supply chain structure (McIvor et al. 2003); Lin thought electronic commerce supply chain models needed to deal with multi-enterprise data and processes cost-effectively while emphasizing the synchronization of a multiple-enterprise supply network (Lin et al. 2002); Boyson built a core business model by selecting a leader companies to solve this central information sharing and collaboration problem (Boyson et al. 2003).

In order to improve DCM performance, is it necessary for the collaborating companies to concentrate on improving the transformation of demand information at the end of the chain (Blocher and Chhajed 2008). Here Big Data and the nature of the electronic marketplace provide truthful, real-time and detailed data to help improve SCM and DCM through extracting value from the database (Ittmann 2015). For instance, Zage et al. demonstrated how to improve risk management in SCM using Big Data (Zage et al. 2013); Tan et al. (2015) explored how Big Data enhanced the innovation capabilities of the supply chain by using the deduction graph technique; different commercial websites need tailored SCM strategies in accordance with their own needs.

Furthermore, there are many articles about the value of information sharing in SCM that emphasize the importance of Big Data. Demand information-sharing is one of the most important business strategies in the current age of information (Cachon and Fisher 2000). Gavirneni et al. (1999) compared the results of three different levels of information sharing in the supply chain and concluded how capacity and inventory were influenced by the end-item. For example, Kovtun et al. (2014)assessed the value of demand-sharing among players in multistage supply chains under given conditions and found that demand sharing is not equivalent among upstream supply chain players. Ultimately Big Data and its analysis tools can help sharing information in supply chain more easily, clearly and efficiently.

So far, electronic commerce and the DCM conceptual foundation have been developed quickly, however, the empirical studies on these topics are scarce. This paper wants to empirically support the opinions above by focusing on a Chinese Real Estate website case study.

3 Methodology

Big Data analytics refers to the useful manipulation of huge data as characterized by the four points previously stated. As the hottest topic in the IT industry nowadays, the business value of Big Data is followed by data warehouse infrastructures, data security, data analysis, and data mining, and has gradually become a central focus of the industry. With the coming era of Big Data, we examine five basic aspects of data analysis:

3.1 Analytic visualizations

Both for data analysis experts and ordinary users, data visualization is one of the most basic requirements for data analysis tools. Visualization can project data intuitively, let the data speak for itself, and help the audience understand the presented findings more clearly.

3.2 Data mining algorithms

Visualization helps people understand Big Data, while data mining aids in the machine's digestion of the information. Clustering, segmentation, outlier analysis and other algorithms can help flush out in-depth internal data and dig out more value. The algorithms not only deal with large quantities of data but they do so in a timely manner, which is very important for the increasing volume of available material.

3.3 Predictive analytic capabilities

Data mining enables the analyst to apprehend data better. However, predictive analysis can give the expert advanced predictive judgments according to the visualization analysis and Data Mining results.

3.4 Semantic engines

Due to the new challenges of data analysis caused by the diversity of unstructured data, we need a series of tools to analyze and extract data. Semantic engines are structures in Software that need to be designed to extract information intelligently from the "document".

3.5 Data quality and master data management

Professional data quality and data management are some of the best management practices. These tools help guarantee quality analysis results through standardized processes and tools to deal with large volumes of data.

Currently, the most popular and effective scientific method is data mining. Data mining is also called knowledge discovery in database (KDD), and represents a deepening of the concept of "knowledge discovery" (Deogun et al. 1997; Grossman et al. 1999; Han et al. 2011). KDD first appeared in the 11th international joint conference on artificial intelligence in 1989. Data mining is an effort to extract implicit and useful information and knowledge from large, incomplete, noisy, fuzzy and random data. Fayyad et al. (1996) has put forward the generally accepted definition of data mining: to extract the knowledge people are interested in from a large database. The nature of this desired knowledge is hidden, unknown and potentially very useful, and it can be expressed as concepts, rule, regularities, patterns and many other forms. Data mining is an interdisciplinary subject involving many fields such as machine learning, model reorganization, induction and deduction, statistics, database theory and high-performance calculation methods. Data mining has entertained extensive attention from diverse groups recent years because it is an application oriented towards multiple-subject intersectional fields and is motivated by tangible applications. Data mining has proved useful in a variety of different areas, such as retail industries, communication techniques, banking, insurance, gene analysis, and stock-market analysis.

There are many types of data mining algorithms that specialize in identifying different forms of data. The general algorithms that apply to the e-commerce platform precision marketing are: clustering algorithms, classification and prediction algorithms, and associative rule algorithms.

Clustering analysis, also called group analysis, is a statistical analysis method that revolves around the classification of samples or indices. Based on similarities, clustering analysis makes use of the fact that related items share more key features within the same clustering patterns than those from other clusters. In business, clustering can help market analysts distinguish different consumer groups from the consumer database. Additionally, clustering can illuminate the spending patterns and habits of each class of consumer. As a module in data mining, clustering often sheds light on deeper information distributed in database, and outlines the characteristics of each category. For even more in depth analysis, clustering can focus on a particular class. Clustering analysis can also be used as a preprocessing step for other methods of analysis or algorithms integrated into a data mining platform.

Categorization or classification can map categorical data to a specific category using classification functions or modes, and can also be used to predict which class a piece of data is most likely to belong to base on a statistical justification. Classification structures automatically deduce the given data description from a historical data record, and then project the results onto future data previously unseen to the system. The chief application of classification algorithms for electronic commerce is to classify users according to their stated purpose or to label and analyze the characteristics of these different users. The structure of the classifier is based on a variety of forms, such as machine-learning methods, Bayesian methods, or parameter methods, etc. Predictive algorithms can estimate user access-behavior and product category requirements. Central classification algorithms include linear regression, multiple regression and nonlinear regressive forms.

The next useful technique is correlation analysis, also called association mining. The goal of this type of data manipulation is to find frequent patterns, relevance, or causal structure from transactional data, relational data or other information carriers. Alternatively, correlation analysis can be used to find the connection between different goods in a transactional database and ultimately generate individually catered suggestions based on past behaviors. For example, correlation analysis can make the connection between the different products customers choose to analyze the buying habits of customers. By knowing what goods and services customers frequently pair together, retailers can fine-tune their marketing strategies based on these associative relationships. Other applications include pricing table design, product promotion, the discharge of the goods and other classification based on the buying patterns of customers.

In this paper we use the decision-tree classification method, which can implement classification with the aid of branches that form a tree structure. The internal nodes of the tree represent a potential judgment, and the branch of the nodes correspond to the classification results, while the leaf nodes represent a class standard.

As a supervised learning method, decision tree algorithms must be trained with known data for each category in advance to create a reliable classifier. However, a decision tree may not always attain a well-trained state especially when dealing with huge amounts of data. Decision trees can be a reliable mode of prediction with formatting and preprocessing from machine-learning enabled systems. Decision trees represent a mapped relationship between the object properties and the object values. In a decision tree, each node signifies an object, each branch path represents a possible attribute values, and each leaf node represents the value of the object from the root node to the leaf nodes on that path. Decision trees have a single output, and an analyst can set up independent decision trees to handle different outputs if they desire more. The machine-learning technology associated with generating decision trees from data is called "decision tree learning". Plainly speaking, this concept simply represents the decision tree, a prediction-tree that can predict and classify future data based on classification derived from the known training data.

4 Model design and setup

Big Data produced by commercial websites is typically divided into two groups: behavior data and transactional data. A website's behavior data includes browsing information, user-visit statistics, click-page data, popular search keywords and so on. On the other hand, transactional data embraces business information, product details, payment information and quantities of exchanged goods and services. Online retailers can easily know their customers' consumption habits and characteristics to promote a better business relationship and discover new underlying markets. The core usefulness of data analysis is to understand customers' needs to provide feedback for business partners in the supply chain that can help them adjust and accordingly provide better products and services. Due the huge potential and commercial value of Big Data and analysis, specialized data companies are emerging in response to the needs of the market. With the resources to provide both website behavior data and transactional data analysis and measurements for individual online retailers, these companies have been very effective. Big Data not only changes the business environment but also the consumer's lifestyles. However, these companies only consider data from the perspective of sales promotion instead of the true needs of customers. Websites take advantage of data accumulation and analysis to construct customer information systems that classify them and extract valuable information to sell for profit.

Considering the availability of data, we chose to examine ten days of three mainstream commercial real estate agent websites as objects of research to try and find out how to use page-visit data to analyze different customers' needs for real estate. Then we discuss how our results can benefit SCM and DCM structures. The *MMData Company* provided us with data whose structure is depicted in Fig. 4.

4.1 Background

It is well known that China is undergoing a spectacular real estate boom, which has become a reliable source of both growth and of revenue. The property industry accounted for about 60 % of China's gross domestic product (GDP) in 2014 [34]. The numbers of property companies and their average prices have almost doubled when compared to data from 10 years ago. In developed cities like Beijing, Shanghai, Shenzhen, many people cannot afford the high

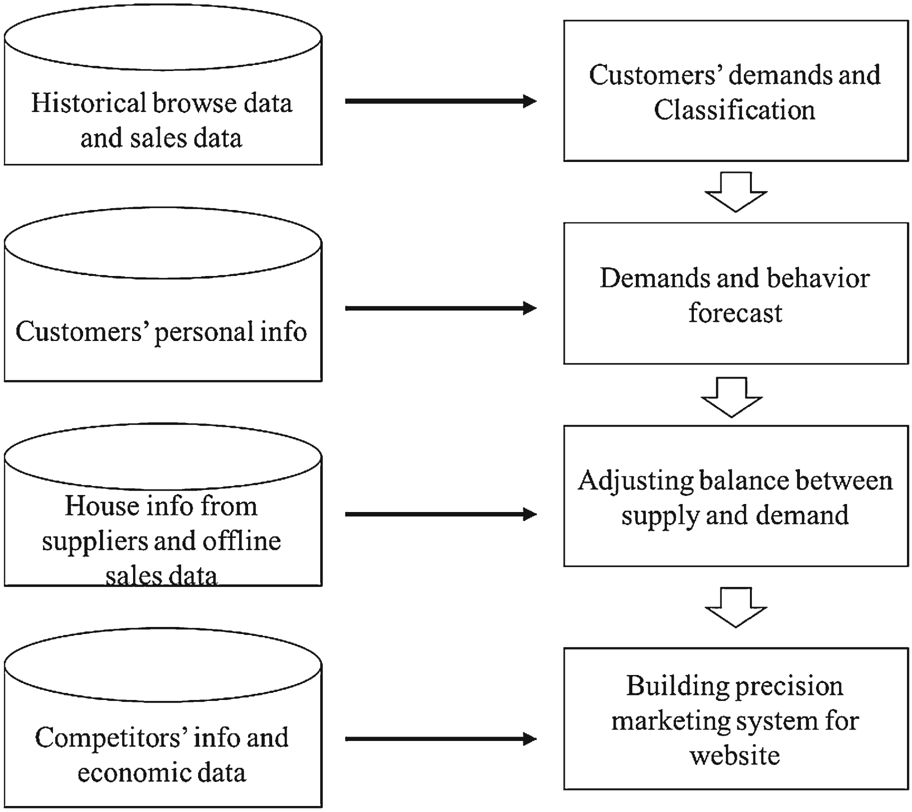

Fig. 4 Big Data applications in commercial property analysis

prices. There are plenty of reasons why Chinese people and investors are drawn to property, notably the traditional ideas, favorable government policies and low bank interest rates. There are also many more factors, like legislation, culture, taxes, physical geography and so on. The prosperous development of the property industry has hastened the development of many associated industries, like construction, decoration and the pipe equipment industry. Some experts are predicting the existence of a potential bubble and vulnerability of property industry. The mismatch between supply and demand in the traditional supply chain partly accounts for the possibility of this bubble. Property developers have overestimated the demand of the ordinary resident and focused more on the physical properties themselves rather than paying more attention to customer needs.

Thanks to the Internet and information technology, property's agencies and electronic commerce have gradually developed. electronic commerce is mainly used in two kinds of property management: portal sites and agency sites. Property developers use portal sites to extend their publicity, enhance their product reputation, create famous products, and compete in the international real estate market. Electronic commerce makes up for the disadvantage of traditional media advertisement.

For agency sites, property developers provide a platform for the exchange of property information, views, and professional knowledge, to serve as a bridge for customers and developers. The advent of these sites has greatly accelerated the circulation of property market information and precision marketing. Consequently, the market has turned from a buyer's

Table 1 A summary of the analyzed real estate agencies

Name	Brief introduction	Date established	Number of daily active users
W1	An electronic information platform and property supply chain service center	1995	600,000
W2	Made up of four sectors: new houses, second-hand houses, rental properties and commercial real estate	2007	200,000
W3	Focused more on property, financing and transaction management as an integrated company and O2O platform	2001	100,000

market into a seller's market where customers have bigger control of the price and quantity, which further increases the competiveness of real estate agencies in China. In this paper, we choose three main famous agencies of China to focus on: $W2$, $W3$ and $W1$ as summarized in Table 1. Empirical data have been collected mainly from a database containing Shanghai residents' real estate browsing activity both on APP and the computer.

4.2 Decision tree

We observed the browsing behaviors decision-making processes of customers on these three websites over the course (20 G of data in total). The average number of daily page-visits was 3,344,000 for $W2$, 211,00 for $W3$ and 2,395,000 for $W1$. It turns out that different websites have different focuses and design but also share common characteristics. As economic globalization continues to mature, and the Chinese people continue their affinity for real estate, agencies are all beginning to add an "Overseas Property" module. Also, "Financing" modules have become a welcomed addition to agency websites because of the value associated with real estate investment. Apparently $W1$.com has entered the decor market and even provides loans for buyers since it aims to provide a one-stop service experience and to be a professional and comprehensive website. Differences are also embodied in the navigation and property classification systems, which are summarized through Big Data and customer preference. Our decision tree is pictured below in Fig. 5.

In Fig. 5, all the indexes on decision tree mean the proportion of website visits. For example, the New houses and the renting represent the distribution of supply between the rental and owner-occupied markets. The indexes can help us sense, serve, and satisfy the needs of customers in a well-defined market.

4.3 Results

In 2015, newly built residential purchases generated over 7.6 trillion Yuan in transactions (China's National Bureau of Statistics 2015). We can tell from the website browsing data that new houses are still in an uptrend but with slow growth. Over 10 % of page-visits to all three websites involved second-hand residences, making them far more popular than new houses. The Chinese property market is shifting form a new-house oriented market to second-hand house oriented environment. Government incentive policies play an important role in the second-hand house market. We found that the scarcity, higher prices and often remote locations of the newer developments are driving demand for these units down. Conversely,

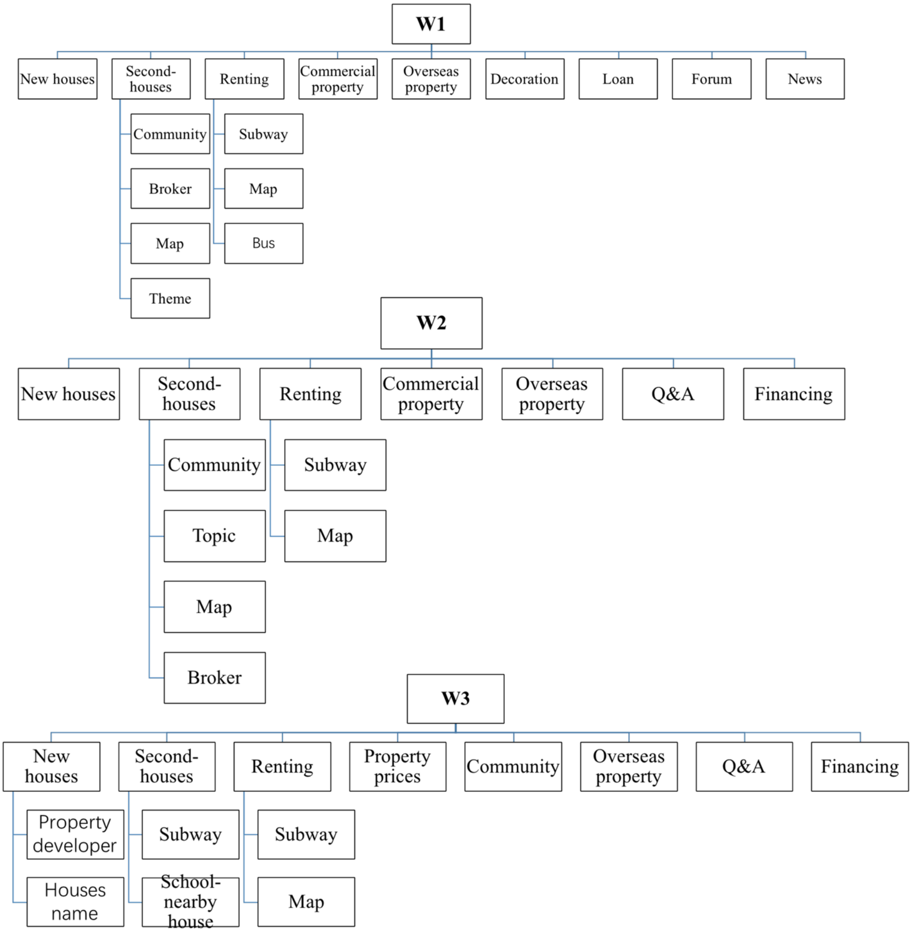

Fig. 5 Customer decision trees for $W2$, $W3$, and $W1$

the prime locations of used residences (like being located in a preferable school district, for example) and lower prices have driven demand for these units up. Property suppliers should find more second-hand residential resources and establish the perfect service system to meet their customers' needs. This data also reminds upstream property developers to pay close attention to the demand for new houses in order to avoid dead stock and wasting their resources.

Although the new housing market is becoming worse, commercial real estate has been developing steadily because storefronts have higher investment values than residential properties. However, the percentage of commercial real estate page visits is much lower when compared with the popularity of residential inquiries. On one hand, the demand for residential units in society is lower and casual. On the other hand, people are used to looking for commercial properties through offline market channels and with more prudence and caution. Despite this we can still see the potential of the commercial real estate market, as well as the tourist and hotel property market. Suppliers should explore their market prospects using Big Data before starting to develop property in order to avoid making any financial mistakes.

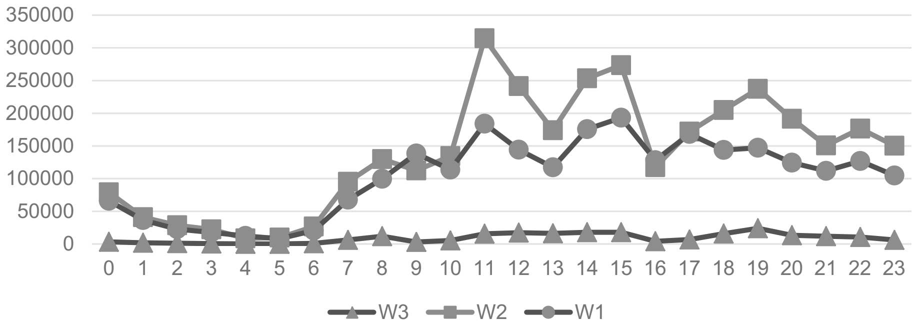

Fig. 6 Number of page-visits over the course of a day

"Overseas Property" and "Financing" services are helping to maintain strong momentum in property development. Because of the virtues and convenience of the stable market due to the nature of permanent property, investment in overseas real estate has become more and more popular among Chinese buyers in recent years. Since many property prices were too high to afford, there are many people who need to borrow money from bank. In the past, property agencies and bank agencies worked together to provide services to buyers, which made for a long and complicated process. Property websites have seized the opportunities and started to provide a one-stop service, making the process much easier. They can also bolster their profits by charging an agency fee in exchange for this useful service. Another contribution of Big Data is that property suppliers can easily collect their customers' information and preferences from these financing requests.

From Fig. 6, We also found that customers like to visit property websites during peak hours, like from 11:00 AM to 12:00 PM, from 1:00 PM to 3:00 PM, and from 6:00 PM to 8:00 PM. This information can help property agencies to update property sales information, push sales reminders and improve the quality of products.

5 Conclusion

Electronic commerce and Big Data have not only changed business patterns but also modes of production. The traditional SCM and DCM can't apply to this new economic environment because each party in the supply chain only considers its own business without real-time information sharing and an accurate demand forecast. Now, suppliers and online retailers can know customers' demands more clearly, but they also must remain flexible in order to keep up with the changing market. In today's economic climate, understanding the customer's situation and developing effective responses to differing needs through the coordination of marketing and SCM can be a source of superior customer value creation (Liu et al. 2015). Those companies who cannot make full use of Big Data are in jeopardy as time goes on. Big Data is an excellent tool to know the customers and the market very quickly and exactly, helping to solve the unclear demands problem. The variety and configurability of Bid Data's tools also help companies deal with all kinds of problems and seek for measures from any angle like market, policy and so on.

This paper has introduced DCM as a model that combines the strengths of marketing and SCM by shifting the focus to the customer thereby designing customer-centric supply chains. Next we introduced precision marketing geared at how to apply Big Data and electronic commerce specifically to the property supply chain. Precision marketing and Big Data help online retailers make light of their customers' behaviors, habits and preferences, all of which have helped break up the "mysterious customers" phenomenon. After that we used the websites of $W3$, $W2$ and $W1$ as examples and to learn about the most important functions of precision marketing: demand analysis. Apparently, the second-hand house market is flourishing while the new houses market is not. Developers should be well aware of the unfavorable situation and aim to strike a balance. With increasing amounts of financing and investment in overseas real estate, property agencies keep up with the trends and provide more personalized services. Internet technology also facilitates the diversification of property marketing. For example, we found that commercial, tourist, and retirement properties are becoming more and more popular. Additionally, the housing market is not limited to only renting or purchasing. Creative methods like house sharing will increase the efficiency of land and reduce financial pressure. There are also many other Big Data applications in the property market. For example, Haowu.com bought personal information from property demanders and built a big database aimed at matching projects between the supply and demand channels to bring in more profit.

In summary, SCM should make the best of electronic commerce and Big Data and enjoy the benefits they bring. In the meantime, it is crucial to develop the sustainability of supply chain. This paper introduced only the tip of the iceberg due to unavailability of specific data. But most of all, with more data analysis and further empirical study, we have found that SCM will be able to outperform its previous generations as we enter the era of Big Data.

Acknowledgments The authors gratefully acknowledge the support provided by the National Natural Science Foundation of China (Grant No. 71203153); Humanity and Social Science Youth foundation of Ministry of Education of China (Grant No. 16YJC630051); Philosophy and Social Sciences Planning Foundation of Tianjin (Grant No: TJGL16-016) Project of Science and Technology Development Strategy Research and Planning of Tianjin (Grant No. 2014ZLZLZF00002).

Authors' contribution Lei Li and Tao Yu conceived and designed the structure and provided the data; Ting Chi and Tongtong Hao performed the experiments and analyzed the data; Ting Chi wrote the paper.

References

Blocher, J. D., & Chhajed, D. (2008). Minimizing customer order lead-time in a two-stage assembly supply chain. *Annals of Operations Research*, *161*(1), 25–52.

Boyson, S., Corsi, T., & Verbraeck, A. (2003). The e-supply chain portal: A core business model. *Transportation Research Part E: Logistics and Transportation Review*, *39*(2), 175–192.

Cachon, G. P., & Fisher, M. (2000). Supply chain inventory management and the value of shared information. *Management Science*, *46*(8), 1032–1048.

Cachon, G. P., & Lariviere, M. A. (2001). Contracting to assure supply: How to share demand forecasts in a supply chain. *Management Science*, *47*(5), 629–646.

China statistical yearbook of 2015. China's National Bureau of Statistics. http://www.stats.gov.cn/tjsj/ndsj/2015/indexch.htm. Accessed 10 July 2016.

Cooper, M. C., Lambert, D. M., & Pagh, J. D. (1997). Supply chain management: More than a new name for logistics. *The International Journal of Logistics Management*, *8*(1), 1–14.

Deogun, J. S., Raghavan, V. V., Sarkar, A., & Sever, H. (1997). Data mining: Research trends, challenges, and applications. *Roughs Sets and Data Mining: Analysis of Imprecise Data*, 9–45.

Drucker, P. F. (1998). Managements new paradigms. *Forbes Magazine*, *10*, 98.

Fattahi, M., Mahootchi, M., & Husseini, S. M. (2015). Integrated strategic and tactical supply chain planning with price-sensitive demands. *Annals of Operations Research*. doi:10.1007/s10479-015-1924-3

Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., & Uthurusamy, R. (1996). *Advances in knowledge discovery and data mining*. Menlo Park: AAAI press.

Gavirneni, S., Kapuscinski, R., & Tayur, S. (1999). Value of information in capacitated supply chains. *Management Science*, *45*(1), 16–24.

Grieger, M. (2003). Electronic marketplaces: A literature review and a call for supply chain management research. *European Journal of Operational Research*, *144*(2), 280–294.

Grossman, R., Kasif, S., & Moore, R., et al. (1999). A report of three NSF workshops on mining large, massive, and distributed data. *Data Mining Research: Opportunities and Challenges, National Science Foundation from the Information and Data Management Program, the Algebra and Number Theory Program, from the Statistics and Probability Program*.

Gupta, A., & Maranas, C. D. (2003). Managing demand uncertainty in supply chain planning. *Computers and Chemical Engineering*, *27*(8), 1219–1227.

Han, J., Kamber, M., & Pei, J. (2011). *Data mining: Concepts and techniques*. Amsterdam: Elsevier.

Hilletofth, P. (2011). Demand-supply chain management: Industrial survival recipe for new decade. *Industrial Management and Data Systems*, *111*(2), 184–211.

Index Analysis. (2016). Baidu Index. http://index.baidu.com/?tpl=trend&word=%CE%ED%F6%B2. Accessed 10 July 2016.

Ittmann, H. W. (2015). The impact of big data and business analytics on supply chain management. *Journal of Transport and Supply Chain Management*, *9*(1), 9.

Jttner, U., Christopher, M., & Baker, S. (2007). Demand chain management-integrating marketing and supply chain management. *Industrial Marketing Management*, *36*(3), 377–392.

Kovtun, V., Giloni, A., & Hurvich, C. (2014). Assessing the value of demand sharing in supply chains. *Naval Research Logistics (NRL)*, *61*(7), 515–531.

Laari, S., Töyli, J., Solakivi, T., & Ojala, L. (2016). Firm performance and customer-driven green supply chain management. *Journal of Cleaner Production*, *112*, 1960–1970.

Lambert, D. M., & Cooper, M. C. (2000). Issues in supply chain management. *Industrial marketing management*, *29*(1), 65–83.

Lin, F-r, Huang, S-h, & Lin, S-c. (2002). Effects of information sharing on supply chain per- formance in electronic commerce. *Engineering Management, IEEE Transactions on*, *49*(3), 258–268.

Liu, Y., Li, H., Peng, G., et al. (2015). Online purchaser segmentation and promotion strategy selection: Evidence from Chinese E-commerce market. *Annals of Operations Research*, *233*(1), 263–279.

Maria Vanalle, R., & Blanco Santos, L. (2014). Green supply chain management in Brazilian automotive sector. *Management of Environmental Quality: An International Journal*, *25*(5), 523–541.

McIvor, R., Humphreys, P., & McCurry, L. (2003). Electronic commerce: Supporting collaborationin the supply chain? *Journal of Materials Processing Technology*, *139*(1), 147–152.

Mishra, D., Gunasekaran, A., Papadopoulos, T., et al. (2016). Big data and supply chain management: A review and bibliometric analysis. *Annals of Operations Research*, 1–24. doi:10.1007/s10479-016-2236-y.

Mobile Electronic Commerce Industry Development on Singles' Day. (2015). TrustData. http://www.itrustdata.cn/. Accessed 11 November 2015.

Tan, K. H., Zhan, Y., Ji, G., Ye, F., & Chang, C. (2015). Harvesting big data to enhance supply chain innovation capabilities: An analytic infrastructure based on deduction graph. *International Journal of Production Economics*, *165*, 223–233.

Taobao Analysis. (2015). TrustData. http://www.iresearch.com.cn/data/246308.html. Accessed 25 December 2015.

Tsiakis, P., Shah, N., & Pantelides, C. C. (2001). Design of multi-echelon supply chain networks under demand uncertainty. *Industrial and Engineering Chemistry Research*, *40*(16), 3585–3604.

Turban, E., King, D., Lee, J., et al. (2002). *Electronic commerce: A managerial perspective 2002*. Upper Saddle River: Prentice Hall. ISBN 0, 13(975285), 4.

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics*, *34*(2), 77–84.

Wannenwetschl, H. H., & Nicolai, S. (2004). *E-supply-chain-management.* Wiesbaden: Gabler.

Zage, D., Glass, K., & Colbaugh, R. (2013). Improving supply chain security using big data. *Intelligence and Security Informatics (ISI), 2013 IEEE International Conference on 2013* (pp. 254–259). IEEE.