International Journal of Computer Science and Applications,
©Technomathematics Research Foundation
Vol. 13, No. 1, pp. 84 – 92, 2016

ORDER ALLOCATION FOR SERVICE SUPPLY CHAIN BASE ON THE CUSTOMER BEST DELIVERY TIME UNDER THE BACKGROUND OF BIG DATA

LIU Yan-Qiu

School of science, Shenyang University of Technology, Shenyang 110870, China

WANG Hao

School of science, Shenyang University of Technology, Shenyang 110870, China
E-mail: 757172232@qq.com

We mainly studied order allocation of logistics service supply chain under the background of big data. According to the characteristic of prediction of big data, designed the two-phase solution of "prediction of big data stage" and "model optimization stage" in logistics service supply chain . The first stage through analysis the massive amounts of customer's location data and clicks to predict the best delivery time and the number of customer's demand. The second stage is a three-echelon order allocation model with the objective of the logistics service cost is established. And select the suppliers to offer service which the customer was satisfied with base on the analysis of big data and the customer's requirements. The result shows that the prediction result of big data is more accurate and better provide decision-making basis for next stage in order allocation. So it can minimize the costs while the customer satisfaction was guaranteed.

Keywords: big data; logistics service supply chain; order allocation; association analysis; delivery time prediction; clicks ratio; location data.

1. Introduction

Logistics service supply chain optimization problem was first proposed by Ellram in the article "Journal of supply chain management" in 2004 and define the concept of logistics service supply chain. The main structure of the logistics service supply chain is logistics service integrator, functional logistics service providers and customers. The customers' demand orders are integrated by the Logistics service integrator after receiving the order from the customers. The orders are allocated to the functional logistics service providers which provide the corresponding logistics service to fulfill the customers. This process is called the order demand allocation. With the rapid development of information technology, such as the Internet, and people's living standard is higher and higher, people has higher requirement and more individual need to the service level of logistics. In the era of big data, in order to improve the reputation and competitiveness of enterprise, we should analyze customer interest according to the customer's data and make order allocation decisions based on the customer's interests. And then to optimize the cost of the whole logistics service supply chain. So it can bring long-term benefits for the enterprise.

In the study of logistics service supply chain optimization, we mainly focus on the problem of order allocation. References [1], [2] consider transaction costs between logistics service integrator and subcontractors in order allocation and solved with the genetic algorithm. Liu et al. [3-4] built three-echelon logistics service supply chain of order allocation , minimized the costs of the logistics service integrator and maximized satisfaction of the function logistics service providers as the goal. Fan et al. [5] proposed a dynamic integer-programming model with multiple tasks and targets in multiple periods. [6] introduced the model of Stackelberg Game into the logistics service supply chain. Fan et al [7] considered the characteristics of the bad into the supply chain and built a mixed integer programming model with multiple targets. They combined Simulated Annealing algorithm with a heuristics rule to solve this model. [8] proposed the value-oriented operation process model for service supply chain. Order Allocation Model of suppliers selected based on fuzzy linear programming was presented by Li et al [9]. Li Yang Zhen[10] introduced the time reliability into the logistics services supply chain and built the model of profit maximization based on the constraints of logistic capacity, which considered the quality level of logistics services and the reliability needs of our customers.

The present study on the optimization of the logistic service supply chain rarely combined with big data .For this reason; this paper presented two improvements in logistics service supply chain order allocation under the background of big data. On the one hand, instead of using probability distributions to anticipate customer's need, we predict the results according to the vast amounts of data of the customers. And the result is closer to the actual. Therefore this paper predicts the number of customer need base on analyzing the correlation among customer's Internet clicks ratio, browsing time and sales. However, previous studies usually shorten the delivery time to improve the quality of service in logistics service supply chain, but shorten the delivery time may not be the best delivery time for customers. Therefore, this article will predict best delivery time of customer base on analysis the behavior of customer which according to vast amounts of data of customer's location. On the other hand, it is based on the predictions of big data and the individual needs of customers to match the functional logistics service providers with customers. Thus we can select appropriate suppliers which customers are most satisfied with can improve customer service quality. So it can bring more profits for the long-term development of enterprises.

2. Problem Formulation and Modeling

In order to improve service levels of customer, this paper introduces customer personalized demand and predicted result which based on the analysis of big data into the logistics service supply chain, and establishes a three-echelon order allocation model with the objectives of the total service cost of logistics service integrator. According to the order allocation model we select functional logistics service providers to provide service which the customer was most satisfied with..

2.1. *Notations description and assumptions*

Assumptions are as follows:

(1)Assume that the quality and category of the product has no difference.

(2)A distributor can provide service formultiple clients, but a client can accept only one distributor's service.

(3)Assume that the data of location of customers is known.

Notations for the model as follows: n denotes the number of functional logistics service provider, m denotes the number of customers, A_{ij} denotes the attraction which the i th functional logistics service provider appeal to the j th customer , q_i denotes the service number of functional logistics service provider, c_{ij} denotes service costs of functional logistics service provider, p_{ij} denotes the service quote of the i th functional logistics service provider to the j th customer, Q_{ij} denotes the service level of the i th functional logistics service provider to the j th customer, p_j denotes the highest acceptable price of customer, Q_j denotes the expected level of service of customer, γ_j denotes the sensitivity coefficient of customer to the price change, θ_j denotes the sensitivity coefficient of customer to the level of service, H_i denotes the profit of the functional logistics service provider, S_i denotes the minimum profit of functional logistics service provider can participate in order allocation, x_i denotes the maximum service capacity of functional logistics service provider, c_i denotes the price of the product. Decision variables

$$y_{ij} = \begin{cases} 1 & \textit{under} \text{ the } U_{ij} \quad \text{i=1,}\cdots\text{n} \quad \text{j=1,}\cdots\text{m} \\ 0 & \text{else} \end{cases}$$

$$U_{ij} = maxA_{ij}$$

2.2. *Model building*

In this model, the total service cost of the whole logistics service integrator is composed of the quote of functional logistics service provider and the penalty cost of the delivery time.

$$\min Z = \sum_{i=1}^{n}\sum_{j=1}^{m} y_{ij} p_{ij} + c_{ij}^{'} - \sum_{i=1}^{n} q_i c_i \tag{1}$$

$s.t.$

$$y_{ij} = \begin{cases} 1 & \textit{under} \text{ the } U_{ij} \quad \text{i=1,}\cdots\text{n} \quad \text{j=1,}\cdots\text{m} \\ 0 & \text{else} \end{cases} \tag{2}$$

$$A_{ij} = \left[1+\left(p_j - p_{ij}\right)/\, p_j\right]^{\gamma_j} \left(Q_{ij}\,/\,Q_j\right)^{\theta_j} \tag{3}$$

$$c_{ij}^{'} = \begin{cases} 0 & \text{t} \le \text{T}_j \\ -\pi\left(t - T_j\right) p_{ij} & \text{t>T}_j \end{cases} \tag{4}$$

$$q_i = \sum_{j=1}^{m} y_{ij} \quad , \quad \text{m∓aT(A}\cap\text{ B)} \tag{5}$$

$$H_i = \sum_{j=1}^{m} y_{ij} \left(p_{ij} - c_{ij} \right) \geq S_i \tag{6}$$

$$c_{ij} = e^{\frac{1}{1-\frac{\eta Q_{ij}}{Q_{\max}}}} \tag{7}$$

$$\sum_{i=1}^{n} q_i = m \ , \quad q_i \leq x_i \ , \sum_{i=1}^{n} x_i \geq m, \ \ p_{ij} \leq p_j \tag{8}$$

$$\gamma_i \in [0,1] \ , Q_{\max} = 5 \ , Q_{ij} \in [1,5] \ , \quad \theta_i \in [0,1] \ , \ \ \eta \in [0,1] \tag{9}$$

The objective function (1) represents that minimizes the total service cost of logistics service integrator, and the cost is composed as described above. The function (2) is the decision variables that select the i th functional logistics service provider which has higher attraction to the j th customer. (3)is the function of attraction which functional logistics service provider appeal to the customer. Without considering the difference in product quality, the higher level of logistics service that the providers offered, the higher attractiveness to customers and the same to higher logistics costs. But the higher logistics costs, the higher quote of the functional logistics service provider, which will reduce the attractiveness to customers. So we define attractiveness as function (3). The function (4) represents penalty cost of the delivery time and T_j is the best delivery time of the customer that is predicted by analysis of the big data. (5) represents the number of providers who assigned to offer service. (6) represents the profit constraint of functional logistics service provider which can participate in order allocation. (7) is the service costs of logistics, and η is the service cost coefficient.(8) represents respectively the delivery quantity that must meet the customer demand and the delivery quantity that is no more than the maximum service capacity of functional logistics service provider . And the quote of the provider who gets the order must no more than the highest acceptable price of the customer. Function (9) represents the scope of the variables.

3. Data preparation and processing

3.1. *Demand forecasting*

In the context of big data, it can predict the sales of goods according to the customer's hits, browsing time, shopping cart, reviews and all sales-related data. The product was known as an hot goods in near future, if the product hits soared recently. So there must be a correlation between clicks and sales. Then the browsing time guarantee the effectiveness of the hits which are not delayed clicks or malicious clicks. So this paper considers the correlation among the product hits, browsing time and sales to predict sales. Assume that the number of customer's demand is m, and product hits is A, the browsing time is B. We define one customer click on the web page of the product is 1, else is 0. The browsing time is more than 30 minutes is 1, else is 0. While the customer buy this product is 1,else 0. The data form is shown in table 1.

Table 1: network data of customer

Product hits A	Browsing time B	Sales m
1	1	1
1	0	1
1	0	1
0	0	0

Based on the confidence and support of the association rules to be judged:

$$S_{X \to Y} = \frac{\left|T\left(X \cap Y\right)\right|}{\left|T\right|} \tag{1}$$

$$C_{X \to Y} = \frac{\left|T\left(X \cap Y\right)\right|}{\left|T\left(X\right)\right|} \tag{2}$$

Function (1) indicates that the support of the rules. It represents the universality of simple association rules, if the support is low then the rules can't be use to judge. $\left|T\left(X \cap Y\right)\right|$ represents the number of event X and event Y which occur simultaneously. $\left|T\right|$ represents the total number of events. Function (2) indicates that the confidence of the rules that represents the accuracy of simple association rules. If the confidence is high, then the probability of event Y is high under the conditions of event X appearing. $\left|T\left(X\right)\right|$ indicates the number of event X. Only when the confidence and support rules achieve the set minimum support, the rules was set up. So this paper mines the correlation through using Apriori algorithm. Then the number of customer demand can get by the formula: m=a$T\left(A \cap B\right)$. a indicates the correlation among the product hits, browsing time and sales.

3.2. *Delivery time prediction*

Assuming that position coordinates of customers is $\left(x_n, y_n, t_n\right)$, (x_n, y_n) indicates horizontal and vertical coordinates of the location of customers , t_n is time data to the customer's location correspondingly. First of all, according to the customer's location data, the day of customer is divided into several sections by using segmentation techniques. And then all distances that are less than the minimum distance are clustered, changed that position coordinates the same as the position coordinates of cluster centers. At last the association rule is used to analyze user behavior.

3.3. *Data mining Algorithm*

Step 1. Enter the time series $\left\{P_{\mathrm{ij}}\right\}$ assume that i=1, deal with the first column of data. Define distance= $A\left(\mathrm{i}\right)$, distance(i,1)=$\sqrt{(A\left(1,1\right) - A\left(\mathrm{i},1\right))^2 + (A\left(1,2\right) - A\left(i,2\right))^2}$, index=distance $\leq$ 0.5, and select out all data which meet the condition $A(\text{index}1:2)=A(1,1:2)$, Output data and merge same data as Q_{11} = （$x_{11}, y_{11}, t_{11}, t_{1j}$）, circulation processing all the data and then get the segment matrix Q_{ij} .

Step2.Inpu Q_{ij} t, find out Q_{ij} which meet the conditions according to the formula $\left|x_{ij}-\delta\right| \leq x_{11}, \left|y_{ij}-\delta\right| \leq y_{11}$, and named it as Q_{11}. Cycle to find coordinates of all the same position and clustering.

Step3. Calculate the number of each Q_{ij} , calculate correlation between all adjacent Q_{ij} according the function $S_{Q_k \to Q_{k'}} = \left|T\left(Q_k \cap Q_{k'}\right)\right| / |T| \geq \alpha, C_{Q_k \to Q_{k'}} = \left|T\left(Q_k \cap Q_{k'}\right)\right| / \left|T\left(Q_k\right)\right| \geq \beta$, select out all Q_{ij} which meet the conditions $k \neq k'$.

4. Data Simulation

4.1. *Demand Forecasting*

The selected data must be authoritative and objectivity to ensure the quality of data. The experiments selected data is from the network of easy car in January 2012. The number of customer hits is 9882039, the number of browsing time is 1232178, the sales of the car is 76663, then these data are organized into the form as Table 1. The number of meeting the condition in click and browsing time is 84245. So the correlation among the click, browsing time and sales is 91%, the click and the browsing time has strong ability to explain to sales. Thus, it can predict the future sales based on the click and browsing time in the next period.

4.2. *The best delivery time prediction of customers*

Randomly selecting location data of 10 users within a week to analyze, we only consider the time that from 10:00 am to 19:00 pm because of the actual delivery time. Data as table 2

Table2: user’s location data in a week

Monday	**Tuesday**	**Wednesday**	**Thursday**
(10,9.5,10:00)	(10.1,9.55,10:00)	(10,9.6,10:00)	(10.01,9.68,10:00)
(10.01,9.5,10:59)	(10.2,9.58,10:30)	(10.1,9.6,10:41)	(10.08,9.67,11:00)
(10.09,9.82,11:38)	(10.09,9.80,11:40)	(10.54,10,12:00)	(10.5,10.2,12:37)
(10.1,9.95,12:35)	(10.06,9.8,12:50)	(10.36,10.45,13:30)	(10.2,9.8,14:35)
(10.3,10,13:30)	(10.2,10.1,13:50)	(10,9.55,14:17)	(9.8,9.9,15:30)
(10,9.8,14:30)	(10,9.85,14:45)	(10,9.86,15.37)	(9.9,10,17:00)
(9.9,10,16:30)	(9.9,10,16:16)	(9.96,10,16:59)	(4.31,4,18:00)
(5,6.3,17:29)	(5.01,6.21,17:40)	(5,6.3,17:20)	(4.3,4,19:00)
(5.1,6.2,18:30)	(4.29,4,18:45)	(4.35,4.2,18:30)	

Friday	Saturday	Sunday
(10,9.6,10:01)	(4.32,4.01,10:00)	(4.3,4.1,10:00)
(10.1,9.61,10:50)	(4.3,4.02,10:50)	(4.35,4.05,11:30)
(10.32,10.1,11:50)	(4.29,4,11:30)	(4.3, 4.1,12:59)
(11,12,12:30)	(15,15.5,12:10)	(11,10,14:45)
(9.9,9.9,14:20)	(15.5,16,13:40)	(11.02,10.10,17:19)
(10,10,17:10)	(15.2,15.7,14:45)	(11.1,10.15,18:00)
(4.31,4.1,18:20)	(15.6,16,17:30)	(4.3,4.1,19:00)
(4.31,4,19:00)	(15.3,16.1,18:35)	

The result of data segmentation as table 3

Table3: customer's segmentation data

d_1	d_2	d_3	d_4
Q_{11}=(10,9.5,10:00,16:30)	Q_{11}=(10.1,9.55,10:00,16:16)	Q_{11}=(10,9.6,10:00,16:59)	Q_{11} = (10.01,9.68,10:00,11:30)
Q_{12}=(5,6.3,17:29,18:30)	Q_{12}=(5.01,6.21,17:40)	Q_{12}=(5.6,3,17:20)	Q_{13}=(4.31,4.18,18:00,19:00)
Q_{13}=(4.3,4.19,19:00)	Q_{13}=(4.29,4.18,18:45,19:00)	Q_{13}=(4.35,4.2,18:30,19:00)	

d_5	d_6	d_7
Q_{11}=(10,9.6,10:01,17:10)	Q_{13}=(4.32,4.01,10:00,11:30)	Q_{13}=(4.3,4.1,10:00,19:00)
Q_{13}=(4.31,4.1,18:20,19:00)	Q_{62}=(15,15.5,12:10)	Q_{72} = (11,10,14:45,18:00)
	Q_{63}=(15.5,16,13:40,18:35)	
	Q_{64}=(16,17,19:00)	

We can get the correlation between Q_{11} and Q_{12}, Q_{12} and Q_{13} higher than other data base on the result . So the user's convenience can be improved ,if we chose these 2-time-periods to delivery. Because of the order delivery address of customer is known, so we chose the delivery time which closer to the order delivery address. Q_{11} indicates the address of customer's company, Q_{12} indicates the address of supermarket which near to the customer's home, Q_{13} indicates the address of home. If the customer fill the home address into the order form address, then output the time Q_{12}, so we get the delivery time of the customer is 17:20pm. Compare the predicted results with expectation of customer; it can meet customer acceptable time.

4.3. *Logistics Service Supply Chain Simulation*

There are 8 functional logistics service providers in this example. a is 90% , $T(A \cap B)$ is 12, Assume that the sensitivity coefficient of all customers to the price is 1, The sensitivity coefficient of customer to the level of service is 0.9，0.9，0.7，0.75，1，0.7，0.75，0.85，0.8，0.7, service cost coefficient η is 0.8, S_i is 106、100、101、100、102、100、99、103, x_i is 50、50、55、60、60、60、65、50, penalty coefficient π is 0.002, the price of the product c_i is all 100, Different functional logistics service provider for different tasks has different levels of service, and the data as table 4 to table 6

Table 4: logistics service level of functional logistics service provider

		user									
		1	2	3	4	5	6	7	8	9	10
provider	1	3.85	3.8	3.87	3.99	3.79	3.81	3.90	3.93	3.75	3.76
	2	3.74	3.70	3.85	3.97	3.76	3.76	3.69	3.81	3.72	3.73
	3	4.32	4.31	4.33	4.29	4.33	4.36	4.01	4.02	4.09	4.23
	4	3.99	3.90	3.89	3.95	3.92	3.93	3.94	3.89	3.88	3.93
	5	3.50	3.51	3.59	3.49	3.43	3.56	3.61	3.58	3.51	3.51
	6	3.28	3.30	3.31	3.31	3.28	3.31	3.34	3.42	3.39	3.40
	7	3.55	3.56	3.55	3.50	3.49	3.50	3.51	3.60	3.58	3.61
	8	4.32	4.40	4.34	4.50	4.39	4.41	4.32	4.48	4.39	4.41

Table 5: e quote of functional logistics service provider

		user									
		1	2	3	4	5	6	7	8	9	10
provider	1	167	167	168	169	160	165	169	171	168	165
	2	170	171	172	169	165	168	161	165	168	163
	3	180	181	179	175	182	183	180	182	188	183
	4	179	178	180	177	177	180	180	179	179	180
	5	159	160	155	155	158	160	160	160	155	159
	6	140	141	142	142	145	145	143	147	145	150
	7	160	161	160	159	159	160	160	161	162	160
	8	181	181	182	185	185	185	183	187	183	190

Table 6: highest acceptable price and the desired service level of customer

	user									
	1	2	3	4	5	6	7	8	9	10
Q_j	4	4	3.5	3.5	4	3	3.5	4	4	3
p_j	180	185	170	170	185	160	175	180	183	161

The result of the model that is the 3th provider offer service to customer1,2, the 6th functional logistics service provider offer service to customer 3、4、6、7、8、10, the 8th functional logistics service provider offer service to customer 5,9. The delivery time of customer 3,4,6 which the 6th functional logistics service provider to offer is 20min late, 10min late, 30min late and the others are not late. The result of the logistics service costs is 580.78 Yuan.

5. Conclusion

This paper anticipates customers’ demand on the basis of analysis data of customer's hits and browsing time, then change uncertain demand into certain demand , which is convenient to solve. The behavior of clients is analyzed according to the customer's location data, then the best delivery time of customers can be gotten out. On the basis of the result as above a three-echelon order allocation model is built, this model considers the attraction, service levels, logistics cost, delivery time penalty cost. The results showed that analyzing demand and behavior of customers in the context of big data can better help companies understanding customer’s interest, so that it can provide better service for customer. Meanwhile , it can reduce the cost of logistics services under the premise of improving the level of customer service, thereby, enhancing the credibility of the enterprise, which is conducive to long-term development of enterprises to get more benefits.

There are still many limitations in this paper. For example , we define the attraction only from two respect of service level and quote. And we will consider more factors in future research, so that it can improve the accuracy of customer's individual needs. Because that the analysis of big data has just started, this paper consider less factors in association of data analysis. In order to improve the accuracy of the data mining, we will consider more factors in future research. At the same time, we not only introduce big data into logistics service supply chain, but also introduce it into other areas of the logistics.

References

Shan-shan L.(2014):Genetic Algorithm to Logistics Service Supply Chain Order Allocation Problem, Operations Research And Management Science, **23**(05),pp.35-41

Zhi-jun G, Wei L, Zhi-qiang F, Wei-ping Z,(2013): Multi-objective Optimization Model of Order Allocation of LSSC Considering Transaction Costs, Systems Engineering, **30**(7), pp. 35-39.

Wei-hua L, Si-Yuan Q, Shi-quan Z.(2012):Order allocation in three-echelon logistics service supply chain under stochastic environments, Computer Internet Manufacturing Systems, **20**(7),pp. 35-39.

Weihua L Meiying G, Wenchen X, (2014): An order allocation model in logistics service supply chain based on the pre-estimate behavior and competitive-bidding strategy. International Journal of Production Research **52**(8):2327-2344.

Chen F, Xiaoli W, Jin C, (2014):Service Order Allocation Based on Multi-objective Dynamic Programming, JOURNAL OF TONGJI UNIVERSITY NATURAL SCIENCE, **42**(9),pp. 1452-1458.

Jian-feng L, Shi-ping C, Rong-hua Y.(2013): Research on Pricing Decisions and Efficiency in a Two-Level Logistics Service Supply Chain, Chinese Journal of Management Science, ,**21**(2),pp.84-90.

Zhiqiang F (2012): Simulated annealing algorithm to supply chain order allocation problem. Computer Engineering and Applications, **48**(25),pp. 28-33.

Tian-yang L, Ting H, Han-chuan X,(2015): Value-oriented operation process model for service supply chain, Computer Integrated Manufacturing Systems, **20**(1),pp. 235-244.

Amin,S.H.,Razmi J, Zhang G.(2011): Supplier Selection and Order Allocation Based on Fuzzy SWOT Analysis and Fuzzy Linear Programming. Expert Systems with Application **38**, pp. 334-342.

Yang-zhen L. (2014):Reliability and optimization of logistics service supply chain . Statistics and Decision ,**05**, pp. 0186-0188.