Future Generation Computer Systems 88 (2018) 151–155

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

Agricultural research recommendation algorithm based on consumer preference model of e-commerce

Qian Zhu [a,*], Yazhuo Li [b]

[a] Economics and Trade Department, Hebei Finance University, Baoding, China
[b] Department of Finance and Assets Management, Hebei University, Baoding, China

HIGHLIGHTS

- This paper introduces agricultural products with it's e-commerce and the accurate mining algorithms.
- Constructs the characteristic information of consumer preference model.
- Constructs the agricultural product feature model and recommendation algorithm.
- Based on the characteristics of agricultural products, the optimization of the traditional product recommendation.

ARTICLE INFO

Article history:
Received 30 March 2018
Received in revised form 25 April 2018
Accepted 15 May 2018
Available online 18 May 2018

Keywords:
Agricultural products
Electricity supplier
Agent algorithm

ABSTRACT

Computer technology in agricultural e-commerce has great potentiality of development, but agricultural e-commerce platform market is still inefficient, making the agricultural recommendation not ideal. This paper studies the application of computer algorithms in agricultural products electricity supplier, introduces agricultural products with its e-commerce and the accurate mining algorithms, constructs the characteristic information of consumer preference model, and constructs the agricultural product feature model and recommendation algorithm. In four aspects, an in-depth study of the exact recommendation algorithm was conducted. Based on the characteristics of agricultural products, the optimization of the traditional product recommendation process is a great help to improve the recommendation accuracy.

1. Introduction

Different from traditional agricultural product logistics, the target customers of fresh produce e-commerce logistics include many individual households. So its distribution nodes are numerous and wide. It also has many varieties of distribution and small batch (Hua Jiang et al. 2017) [1]. There are different logistics links between production and delivery of fresh produce from customers. In the whole process, we must ensure the quality and safety of products and the convergence of different links, which also increases the challenges for fresh produce e-commerce logistics (Kenji Taniguchi 2016) [2]. Therefore, ensuring the efficient and reliable logistics of fresh agricultural product e-commerce, shortening the delivery time of its logistics and reducing the damage of fresh agricultural products in the logistics process becomes an urgent problem to be solved. (Tama Shinnosuke 2017) [3]. Based on the above analysis, the reliability of postgraduate fresh agricultural product e-commerce logistics has a certain significance.

2. State of the art

Fresh e-commerce has been hailed as the "blue ocean" of the e-commerce industry and has gradually developed into an online lifestyle service, such as the United States, Local Harvest, Amazon, China's Jingdong Mall, SF Express and so on. Local Harvest provided consumers with a map retrieval system to facilitate the purchase of local agricultural products. Amazon also proposed two fresh delivery methods outside the door and at home. Under the influence of food consumption structure and network consumption mode, the development prospects of fresh agricultural products with the help of e-commerce platform are promising (Marija Dragicevi et al. 2017) [4]. Domestic and foreign scholars have carried out some research on the development of fresh electronic commerce, involving development status, development trends, and model research. Some scholars believe that industrial structure, product complexity and transaction threshold will affect the development of agricultural electricity suppliers (Ji Xing Zheng et al. 2015) [5]. Based on this, they put forward factors such as social and economic factors, consumer preferences and agricultural product complexity (Wei Wang 2016) [6]. Some scholars also summarized the

* Corresponding author.
E-mail address: elmart@126.com (Q. Zhu).

https://doi.org/10.1016/j.future.2018.05.036
0167-739X/

development trends of fresh agricultural products e-commerce in technology, information, and standards, analyzed the logistics distribution model of fresh agricultural products e-commerce, and put forward the Innovative model as "Third-party logistics + consumer self-raising + third-party distribution" (S.Brezuleanu et al. 2015) [7]. Since 2012 is the first year of the development of China's fresh produce e-commerce, the existing research focuses more on theoretical analysis. News reports lack data-based analysis and effective model verification account for the majority (L.Alsemgeest et al. 2016) [8]. The influence factors and optimization of the fresh e-commerce logistics are still in the research stage and the literature available for reference is limited. In terms of logistics system reliability, domestic and foreign scholars have applied the reliability theory to the logistics field and achieved certain results (Anonymous 2017) [9]. Some scholars have constructed a logistics risk indicator system to analyze and evaluate the reliability of logistics systems, and applied some heuristic algorithms to the reliability optimization of logistics systems, which provides the basis for the reliability study of fresh produce electricity supplier logistics for this article (D.Yu. Samygin et al. 2015) [10].

3. Methodology

3.1. The construction of reliability distribution model

The reliability and costs are relatively abstract concept. Costs include human, material, and financial resources that increase unit reliability, so it is difficult to obtain statistical data between cost and reliability. In order to overcome this problem, Dale proposed a generalized cost function, which is built on the basis of the feasibility f_i, the unit minimum reliability $R_{i,\min}$ and the unit maximum reliability $R_{i,\min}$.

$$C_i(R_i, f_i, R_{i,\min}, R_{i,\max}) = e(1-f_i)(\frac{R_i - R_{i,\min}}{R_{i,\max} - R_i}). \tag{1}$$

F is in the range of 0 to 1, where the larger the value, the more feasible it is to increase the unit. The generalized cost function is a non-linear growth function of cost with respect to a unit. Theoretically speaking, it takes a considerable cost to achieve maximum reliability. On the contrary, the cost to achieve low reliability is very low, which is consistent with the fact that the cost increases as the reliability increases. The characteristics of the generalized cost function provide the foundation for the integration of the various factors in the distribution of system reliability. Therefore, based on the generalized cost function, this paper synthetically considers the importance and complexity of the system to build a logistics system reliability distribution model of fresh agricultural electricity supplier.

$$\min z = \sum_{i=1}^{n} \frac{1}{\omega_i} e^{u_i} (\frac{R_i - R_{i,\min}}{R_{i,\max} - R_i}) \tag{2}$$

$$s.t. \prod_{i=1}^{n} R_i \geq R^* \tag{3}$$

$$R_{i,\min} < R_i < R_{i,\max}, i = 1, 2, \ldots, n. \tag{4}$$

In the above formula, R_i is the reliability of the ith cell; $R_{i,\min}$ is the minimum reliability of R_i; $R_{i,\max}$ is the maximum of R_i; R^* is the expected reliability for the system; ω_i is the unit importance factor; u_i is the unit complexity improvement factor; z is the cost objective function. The objective function of the model 2 is based on the generalized cost function, and increases importance improvement factor and complexity improvement factor. First, the function introduces the importance degree coefficient, and constrains the cost of different units by enlarging and reducing the generalized cost function. The importance degree coefficient is obtained based on the importance of each influencing factor in the previous chapter, and is taken as the reciprocal of the importance degree. The value range is between 0 and 1. The higher the degree of importance is, the smaller the reciprocal of it is, uniting that ensure high importance achieve higher reliability. Secondly, instead of the feasibility parameter, the complexity of the unit by the proportion of the number of subsystem units to the total number of system units. The higher the complexity of the unit, the less feasible the improvement is, and the cost of actual investment increases. The constraint condition of the model 3 indicates that the reliability after optimization distribution is equal to or greater than the reliability of the system. Formula (4) indicates that the reliability of each cell allocation needs to be greater than its respective minimum reliability and less than its maximum reliability.

3.2. Model parameter estimation and verification

Assume that the data related to the logistics system of a fresh produce e-commerce company is shown in the table. Since the problems or failures of each influencing factor will affect the reliability of the entire logistics system, all the influencing factors are connected in series. The reliability of the external environment cannot be subjectively allocated, so only three factors affecting information technology, facilities, equipment, and personnel operations are considered. Set the minimum reliability $R_{i,\min}$ of each unit to 0.85, the maximum reliability $R_{i,\max}$ to 0.99, as Table 1.

Assume that the entire system is reliable (RM = 0.83), using fmincon from Matlab. Reliability optimization is got by calculation under different allocation methods, The equal distribution method uses formulas to calculate, where the reliability of each factor obtained is 0.9398. The allocation method Considering the complexity uses the formula to calculate the reliability of the factors obtained, which is as shown in the table. The final improved cost coefficient method is based on the reliability distribution model constructed in this paper. Although the final system reliability of the three methods has reached the required value of 0.83, it is known that the improved cost coefficient method adopted in this paper has the lowest cost under the same reliability requirements, which proves the feasibility of this model. The results of the optimization show that reliability of distribution of three factors of information technology, facility equipment and personnel operation increased by 0.0269, 0.1213, 0.1245 compared with the initial value 0.85. Facilities are of the highest degree of importance, but the complexity is also the highest. Therefore, considering the complexity and degree of importance, the value of facility and equipment reliability can be slightly lower than that of personnel. In addition, the added value of reliability allocation for facilities, equipment and personnel operations is significantly higher than that of information technology.

3.3. Agent algorithm

Cluster analysis is to use mathematics to analyze the analysis of a given object. It divides a set of non-existent classmark samples according to rules into multiple subsets, ensuring that the data in the same cluster has a great degree of similarity, yet data in different clusters do not have similarity. The partitioning method gives a data set that covers n objects or data in them, which can divide the data set into k subsets, where any subset here represents a cluster ($k \leq n$). That is, dividing the data into k groups. They all meet the following requirements: Each group contains at least one object and any object must belong to a group. The requirement afterwards is to be able to relax in the partial fuzzy division method. Given the number k that must be divided, a partitioning method needs to create an initial partition at the beginning. Afterwards adopting a cyclic repositioning technique, that is, by changing the objects