

## Construction of O2O E-commerce Enterprise Supply Chain Risk Evaluation Index System Based on Semantic Clustering

Qiang Pan

School of Economics and Management, Chongqing University of Posts and Telecommunications,  
Chongqing 400065, China;

panq\_panda@163.com

### Abstract

**This paper uses the semantic clustering method to construct an O2O e-commerce enterprise supply chain risk evaluation index system consisting of 5 first-level indicators and 15 second-level indicators from the collected O2O e-commerce enterprise supply chain risk cases and news texts. The system comprehensively evaluates the supply chain risks of O2O e-commerce companies from five dimensions: market risk, cooperation risk, technical risk, operational risk, and environmental risk, and realizes the first step in the early warning and prevention of O2O e-commerce company supply chain risks.**

### Keywords

**Supply Chain Risk; Semantic Clustering; O2O E-commerce.**

### 1. Introduction

With the in-depth advancement of "Internet +", under the influence of new technologies, new applications and new models, the era of "dual-line" shopping experience has arrived, realizing the interconnection of online and offline resources, namely O2O (Online To Offline) e-commerce. In 2018, China's e-commerce transaction scale continued to expand and maintained a rapid growth trend. The annual e-commerce transaction volume achieved 31.63 trillion yuan, of which the online transaction scale of the life service market reached 156.207 billion yuan, and the online transaction scale of the Chinese in-store business market Reached 997.63 billion yuan. As a new driving force for economic development, O2O e-commerce has attracted great attention from the country and enterprises.

The development of the O2O market has made the competition among O2O e-commerce companies increasingly fierce. How to gain a foothold in the fierce market competition has become a problem that every O2O e-commerce company has to consider. For O2O e-commerce companies, whether it is product type competition, product price competition, or service level competition, in the final analysis, it is competition in the enterprise supply chain. In the actual operation of the supply chain, due to factors such as the complex and changeable organization structure of the supply chain, the unpredictable network environment, the excessive efficiency of the supply chain management and the neglect of risk management, etc., the risks of the enterprise supply chain are constantly emerging. In 2018, Dingdang Kuaiyao had a single profitable drug online, slow growth in user viscosity, and insufficient follow-up resistance risks, which showed a disadvantage in the competition. As an innovative Internet model, O2O e-commerce companies are advancing the deep integration of traditional industries and e-commerce, and reconfiguring and utilizing online and offline data. At the same time, it will inevitably lead to supply chain strategies, organization, process, and information. In terms of changes, the increasingly complex global economic environment also makes the supply chain face more risks [1]. This article uses semantic clustering to construct an O2O e-commerce enterprise supply chain risk evaluation index system, summarizes the O2O e-commerce enterprise supply chain risk evaluation index system, and realizes the first step in the early warning and prevention of O2O e-commerce enterprise supply chain risk.

## 2. Method introduction

### 2.1 Research method

This paper uses semantic clustering to refine supply chain risk evaluation indicators for O2O e-commerce companies. This method is a combination of grounded theory and cluster analysis [2]. Grounded theory is a process of inducing experience from the original text through continuous comparison, integration, and analysis of concept categories and data to establish a theory [3-4]. Cluster analysis is a process of classifying different research objects into different classes or clusters according to their similar characteristics. Objects in the same kind of cluster are very similar, but there are great differences between different clusters. Combining the two, it divides and categorizes and extracts concepts by repeatedly comparing the semantic similarity between text keywords, and then constructs an index system.

### 2.2 Research idea

This article analyzes the text from the search, using semantic clustering method to compare and analyze the O2O e-commerce enterprise supply chain risk description sentences in the search text, extract the independent index concepts from them, and then summarize and refine these concepts form an O2O e-commerce enterprise supply chain risk evaluation index system. The specific idea is divided into the following three steps: 1. Collection and processing of relevant description text; 2. Semantic clustering of keywords; 3. Construction of O2O e-commerce enterprise supply chain risk evaluation index system..

(1) Collection and processing of relevant description texts. Enter keywords such as "O2O supply chain risk", "O2O e-commerce business bottleneck", "O2O e-commerce problems" and other keywords on the China E-commerce Research Center, Ebang Dynamics, and Net Economics. The collected text includes: experts Scholars interpret the problems and challenges faced by the supply chain of O2O e-commerce companies in interview articles; related reports and typical case reports on the O2O e-commerce industry. Read and analyze the collected text, then merge and number sentences containing keywords in the text.

(2) Semantic clustering of keywords. Based on the grounded theory, keyword extraction and frequency statistics are performed on the numbered O2O e-commerce enterprise supply chain risk description sentences, so that the sentences that semantically express the same or similar concepts are classified into the same cluster.

(3) Construction of the supply chain risk evaluation index system for O2O e-commerce enterprises. Analyze the main keywords after clustering, summarize and condense the secondary indicators for the supply chain risk evaluation of O2O e-commerce enterprises, and then summarize the primary indicators based on the secondary indicators.

According to the first step above, the 108 texts of O2O e-commerce enterprise supply chain risk description texts collected were screened, and 96 papers were finally selected through in-depth reading based on whether they contained keywords describing the O2O e-commerce enterprise supply chain risk. The original text containing the supply chain risks of O2O e-commerce companies is serially numbered [1]-[96], and contains 353 independent sentences, which are sequentially numbered [A001]-[A353].

## 3. Index system construction

### 3.1 Semantic clustering based on keywords

#### 3.1.1 Keyword extraction of independent sentences

For the 353 independent sentences obtained above, keywords were extracted respectively. In order to ensure that keywords fully cover the supply chain risks of O2O e-commerce companies, in operation, as many original words in the sentence as possible are extracted as the keywords of the sentence. The example process of extracting some keywords is shown in Table 1.

Table 1. Example of extracting keywords from independent sentences

Independent statement number	Original sentence	Keyword extraction
[1] [A001]	Since O2O is also a kind of e-commerce, it will also face the problem of online traffic. Without online traffic, there will naturally be no sales. In order to gain traffic, many group-buying websites shop for advertisements, which has caused a boom in group-buying.	Online traffic
[7] [A024]	Consumers have uncertain concerns about the quality of products purchased by O2O e-commerce and the quality of offline services.	Product quality, service quality
[25] [A080]	In particular, third-party payment platforms will be attacked by hackers or criminals, and O2O e-commerce users' private information will be intercepted by criminals, and their private property will be lost.	Third-party payment, privacy information

### 3.1.2 Semantic clustering process

The keywords extracted from the independent sentences of O2O e-commerce enterprise supply chain risk in Table 1 are clustered and analyzed according to the method of semantic clustering [5]. For example, the keywords extracted in [A001] are "online traffic" and "Service quality" is temporarily classified into one category, and the keywords extracted in [A080] are "third-party payment" and "private information" are temporarily classified into another category. Next, perform keyword iteration to determine whether the remaining keywords can be classified into the same category as the previous category based on semantics. If they can, they are classified into the same category; if not, they are classified into a new category, and finally [A001]-[A353] The clustering is divided into seven categories, and the clustering results are shown in 2.

Table 2. Keyword semantic clustering results

Semantic category	Total sentences	Frequency of indicator appearance (%)	Clustering keywords (frequency of keywords)
First category	88	92.67	Online payment [67]; third-party payment [42]; information security [40]; mobile payment [36]; information platform [26]; equipment failure (including server failure, network failure) [14]; information leakage [12]; Information transmission (including data transmission) [6]; Privacy information [2];
Second category	53	55.20	Natural disasters [25]; laws (including issuance of documents and regulations) [32]; industry policies [18]; social environment [5]; economic policies [20];
Third category	62	64.58	Demand changes (including demand reduction, demand increase) [29]; customer preferences [15]; substitutes (including similar products, competing products) [20];
Fourth category	68	70.83	Product pricing (including price, price, price) [51]; funding issues [23]; customer satisfaction [18]; offline experience (offline service) [39]; current traffic [14]; online promotion (Including platform promotion) [20]; product quality [7]; online marketing [5]; physical store [8]; strategy [3]; customer stickiness [2];
Fifth category	49	51.04	Conflict of interest (including uneven and unfair distribution of benefits) [21]; cooperative trust [12]; morality [5]; cultural differences [2];
Sixth category	18	18.75	Logistics efficiency [12]; distribution [5]; self-pickup [2]
Seventh category	11	11.15	Merchant interaction [6]; coupons [3]; group purchase [3]

### 3.1.3 Semantic clustering results screening

It can be seen from Table 2 that the frequency of the sixth and seventh types of sentences in the 96 original texts is less than 20%, which is a small probability event, and it is discarded. The occurrence frequencies of the first to fifth types of sentences are all greater than 50%, which are high-probability events and should be retained.

## 3.2 O2O e-commerce enterprise supply chain risk secondary indicators

### 3.2.1 Refinement method of secondary index

Respectively summarize the condensed concepts from the reserved sentence keywords, and extract the secondary indicators of O2O e-commerce enterprise supply chain risk evaluation. The steps are as follows: (I) Select high-frequency keywords as the key words for investigation, analyze their definition and essence; (II) Whether it is possible Summarize most of the remaining keywords as the standard, and examine the meaning of the keywords to initially extract the secondary indicators ③ Determine the secondary indicators of the supply chain risk of O2O e-commerce enterprises.

### 3.2.2 Example of refining secondary indicators

Taking the first type of sentences as an example, the keywords mainly include "online payment", "third-party payment", "information security", "mobile payment", "information platform", "equipment failure", "information leakage" and so on. The secondary index extraction process is as follows: (I) Select "online payment", "third-party payment", "information security", "mobile payment", and "information platform" as the key words to be investigated, and their frequency of occurrence is 67 and 42 respectively, 40, 36, 26, much higher than other keywords. Analyzing its definition, it can be found that online payment is a form of electronic payment, which is an instant payment method through a payment interface provided by a third party with the bank; third-party payment is a third-party payment that means An independent institution with strength and reputation guarantees, through the docking with the network to facilitate the transaction between the two parties to the transaction online payment mode; information security is to protect information and information systems from unauthorized access, use, disclosure, destruction, modification, and inspection , Recording and destruction; mobile payment refers to the mobile client using mobile phones and other electronic products to make electronic money payments; information platform is the digital and networked existence of information, one is that the information itself becomes the carrier of information, and the other is Information system based on digital network operation. (II) Preliminarily refined the secondary indicators "online payment risk" and "information security risk", and the analysis found that "online payment risk" basically summarizes the keywords of "third-party payment" and "mobile payment" in the first sentence, "information "Security risk" basically summarizes the keywords of "information security", "information leakage", "information transmission", and "private information", but the remaining keywords of "information platform" and "equipment failure (including server failure, network failure)" Unable to cover, after review, "equipment failure" is defined as equipment failure generally refers to an event or phenomenon in which the equipment loses or reduces its specified functions, which manifests as abnormal equipment production and operation, so it and the "information platform" are condensed as information infrastructure failures risk. (III) Determine online payment risk, information security risk, and information infrastructure failure risk as secondary indicators for supply chain risk evaluation of O2O e-commerce companies. Online payment risk is defined as the security problems encountered by users in electronic payments through Internet communication devices. Information security risks are defined as security risks such as information leakage, information loss, and information interruption caused by possible software and hardware defects and potential weaknesses in information security management due to the data and information processed in the process of information transmission. Information infrastructure failure is defined as the abnormal operation of information systems, platforms and servers due to man-made or natural reasons.

### 3.2.3 Refined results of secondary indicators

According to Table 2, the keywords of the first to fifth types of sentences are extracted in sequence, and a total of 15 secondary indicators such as online payment risk, natural risk, demand fluctuation risk, and online promotion risk are obtained. The results of the extraction are shown in Table 3.

Table 3. Refined results of secondary indicators

Semantic category	Secondary indicators	Indicator definition
First category	Online payment risks	Security problems encountered by users in electronic payment through Internet communication devices
	Information security risk	In the process of information transmission, the processed data and information may cause security risks such as information leakage, information loss, and information interruption due to possible software and hardware defects and potential weaknesses in information security management
	Information infrastructure failure risk	Abnormal operation of information systems, platforms, servers and other infrastructures due to man-made or natural reasons.
Second category	Natural risk	The possibility of floods, fires, earthquakes, etc. and other irresistible natural disasters and their possible unconventional comprehensive damage to the supply chain
	Legal policy risk	Risks faced by the supply chain due to uncertainty in policies and laws
Third category	Demand fluctuation risk	The fluctuating demand and the risk of failing to maintain a stable quantity
	Increased risk of alternatives	More products that can meet the same needs
	Product pricing risk	Risks of overpriced or underpriced products
	Capital chain risk	Risks such as breakage of the capital chain, tightening, and poor turnover in the use of enterprises
Fourth category	Customer satisfaction reduces risk	Risks from reduced customer satisfaction
	Offline experience risk	Risks caused by poor customer experience in offline physical stores
	Online promotion risks	The risk of insufficient online platform promotion and insufficient online traffic
Fifth category	Risk of conflict of interest	Conflict risk caused by unfair distribution of benefits among supply chain partners
	Trust risk	Mistrust between the main partners of the supply chain system
	Moral Hazard	The main partners of the supply chain system violate ethics and harm others and benefit themselves

### 3.3 O2O e-commerce enterprise supply chain risk first-level indicators

Compare and analyze the 15 secondary supply chain risk indicators of e-commerce companies extracted in Table 3. They respectively reflect the five aspects of technical risk, environmental risk, market risk, operational risk, and cooperation risk that O2O e-commerce companies face in the overall supply chain. Based on this, five first-level indicators of supply chain risk for O2O e-commerce companies are proposed.

(1) Technical risk. Including "online payment risk", "information security risk", "information infrastructure failure risk" three secondary indicators. Among them, online payment risk refers to the risks encountered by users in electronic payment through Internet communication equipment; information security risks include information loss, information leakage, information transmission errors, information interruption, etc.; information infrastructure failure risks refer to information

systems, platforms, and servers The operation of the infrastructure is abnormal or stopped due to man-made or natural reasons.

(2) Environmental risks. Including "natural risk" and "policy and legal risk" two secondary indicators. For example, natural risks and policy and legal risks are the two major environmental risks that O2O e-commerce companies face during the normal operation of their supply chains, and they are also the basic guarantee for the development of O2O e-commerce companies' documents.

(3) Market risk. Including "demand fluctuation risk" and "increased substitutes risk" two secondary indicators. The risk of demand fluctuations is the risk of fluctuations in demand and failure to maintain a stable quantity; the risk of increasing substitutes is the increase of products that can meet the same demand; the instability of both will lead to market turmoil for O2O e-commerce companies.

(4) Operational risk. Including five secondary indicators: "product pricing risk", "funding chain risk", "customer satisfaction reduction risk", "offline experience risk", and "online promotion risk", which reflect the supply chain operation process of O2O e-commerce companies, Product pricing needs to be reasonable, funds run smoothly, to give customers a good offline physical store experience, online promotion is effective, and customer satisfaction is improved.

(5) Cooperation risks. Including "conflict of interest risk", "trust risk", and "moral hazard" three secondary indicators. Among them, the conflict of interest risk refers to the conflict risk caused by the unfair distribution of benefits among supply chain partners; trust risk refers to the mistrust between the main partners of the supply chain system; moral hazard refers to the main partners of the supply chain system that violate morality and harm others and themselves.

The final construction of the O2O e-commerce enterprise supply chain risk evaluation index system is shown in Table 4.

Table 4. O2O e-commerce enterprise supply chain risk evaluation index system

First-level indicators	The secondary indicators
Technology risk	Online payment risks
	Information security risk
	Information infrastructure failure risk
Environmental risk	Natural risk
	Legal policy risk
Market risk	Demand fluctuation risk
	Increased risk of alternatives
Operational risk	Product pricing risk
	Capital chain risk
	Customer satisfaction reduces risk
	Offline experience risk
Cooperation risk	Online promotion risks
	Risk of conflict of interest
	Trust risk
	Moral Hazard

#### 4. Conclusion

The construction of the O2O e-commerce enterprise supply chain risk evaluation index system involves a wide range and complex content, and its research is slowly advancing. However, the current vigorous development of O2O e-commerce urgently needs to improve the evaluation system as soon as possible to play the role of guidance and prevention. This paper constructs an O2O e-commerce company supply chain risk evaluation index system through semantic clustering, summarizes the O2O e-commerce company supply chain risk evaluation index system, and provides O2O e-commerce companies with a scientific basis for evaluating their supply chain risks, and

effectively uses risk factors. The information provides decision-making reference for relevant O2O e-commerce companies.

## References

- [1] Liu Jianguo, Kong Yudan, Zhou Huan, Li Jun. Methodology of supply chain risk management based on WSR theory [J]. Journal of Systems Engineering, 2018,33(03):298-307.
- [2] Gao Xirong, Yang Jian, He Jie. Establishment of the Index System for Idea Value Assessment by Semantic Clustering [J]. Science & Technology Progress and Policy, 2017(05): 121-126. China National Standardization Management Committee. Specifications of Crane Design (China Standardization Press, China 2008), p. 16-19.
- [3] GLASER B G, STRAUSS A L, PAUL A T, et al. Grounded Theory. Strategien qualitativer Forschung [J]. PFLEGE, 2006, 19(4): 0260-0260. Q. D. Zeng, Q. E. Li: Progress in Civil Engineering, Vol. 32 (2012) No. 9, p. 3077-3080.
- [4] Gao Xirong, Yang Na. The Construction of Thesis Evaluation Index System Based on Social Network Analysis Method [J]. Science & Technology Progress and Policy, 2017, 35(04):97-102+144.
- [5] HE Xijun, ZHANG Tingting, WU Yuying, JIANG Guorui. Technology demand recognition model based on semantic similarity under the supply-demand matching perspective [J]. Systems Engineering — Theory & Practice, 2019,39(02):476-485.