OmniaScience

Journal of Industrial Engineering and Management
JIEM, 2015 – 8(5): 1396-1408 – Online ISSN: 2013-0953 – Print ISSN: 2013-8423
http://dx.doi.org/10.3926/jiem.1483

Port Throughput Influence Factors Based on Neighborhood Rough Sets: An Exploratory Study

Weiping Cui, Lei Huang, Ying Wang

School of Economics and Management Beijing Jiaotong University (China)

13113151@bjtu.edu.cn, lhuang@bjtu.edu.cn, ywang1@bjtu.edu.cn

Received: April 2015
Accepted: October 2015

Abstract:

Purpose: The purpose of this paper is to devise an efficient method for the importance analysis on Port Throughput Influence Factors.

Design/methodology/approach: Neighborhood rough sets is applied to solve the problem of selection factors. First the throughput index system is established. Then, we build the attribute reduction model using the updated numerical attribute to reduction algorithm based on neighborhood rough sets. We optimized the algorithm in order to achieve high efficiency performance. Finally, the article do empirical validation using Guangzhou Port throughput and influencing factors' historical data of year 2000 to 2013.

Findings: Through the model and algorithm, port enterprises can identify the importance of port throughput factors. It can provide support for their decisions.

Research limitations: The empirical data are historical data of year 2000 to 2013. The amount of data is small.

Practical implications: The results provide support for port business investment, decisions and risk control, and also provide assistance for port enterprises' or other researchers' throughput forecasting.

Originality/value: In this paper, we establish a throughput index system, and optimize the algorithm for efficiency performance.

Keywords: rough sets, neighborhood rough sets, port throughout, throughout indicator system

1. Introduction

In recent years, China's national economy grows rapidly, total import and export trade keeps rising and the port throughput also grows steadily. Data of Ministry of Transportation Highway And Waterway Transportation Industry Statistical Bulletin shows that in 2012, the National port cargo throughput was 10.776 billion tons, made an increase of 7.3% over the previous year. From 2006 to the end of 2012, the national average growth rate of total port cargo throughput was about 13%, which in the year of 2009 due to the restructuring of the national economy, the growth rate slowed slightly, with an average increase of about 7%. The number of ports increases from previous year's 22 to 26 whose cargo throughput over one hundred million tons.

Various factors can affect the port cargo throughput, such as political, economic, cultural and natural environment. The factors' uncertainties and interactions affect the throughput growth trend.

Rough set theory is a powerful mathematical tool for dealing with inconsistent information in decision situations. Rough set methods can be applied as a component of hybrid solutions in machine learning and data mining. They have been found to be particularly useful for rule induction and feature selection (semantics-preserving dimensionality reduction). Rough set-based data analysis methods have been successfully applied in bioinformatics, economics and finance, medicine, multimedia, web and text mining, signal and image processing, software engineering, robotics, and engineering.

Based on the above discussion, this paper propose an efficient method based on rough sets to analysis the importance of Port Throughput Influence Factors. The rest of the paper is organized as follows. In Section 2, we provide some basic material about Rough set theory. In Section 3, we describe the Port Throughput Influence Index System (Hu, 2005), provide an attribute reduction algorithm and make some improvements. We test the performance of our model and present the empirical results in Section 4. Finally, Section 5 summarizes the conclusions of the paper. The experiment demonstrates that the proposed method can efficiently process attribute reduction. Also the results can provide assistance to management for making timely and accurate risk aversion, encourage enterprises to targeted adjust the business strategies.

2. Neighborhood Rough Sets

2.1. Deficiencies of the Classical Rough Set

The classical rough set theory (Pawlak, 1998; Pawlak & Skowron, 2007a) proposed by Prof. Pawlak simulates the human learning and reasoning process, and uses production rules to represent the learned knowledge (Jiang, Sui, 2015). It's easy to understand, accept and use. And it's particularly applicable when dealing with incomplete or inaccurate information. But as an effective granular computing model, it cannot directly process the numeric data exists in reality world.

The classical rough set theory is based on equivalence relations (Pawlak & Skowron, 2007b; Pawlak & Skowron, 2007c; Qin, Yang & Pei, 2008). However, equivalence relation is often hard to be satisfied because of its restrictions and limitations (Zhu & Wang, 2003). Classical rough set theory then has been extended from equivalence relation to some other relations, such as similarity relation, tolerance relation and arbitrary binary relation (Zhu, 2007a; Liu & Zhu, 2008). Notably, classical rough set theory has also been extended to covering-based rough sets (Zhu, 2007b; Kondo, 2006).

The port indicator system used in this article is all numeric data. So if we want to use the classical rough set theory to reduce the attributes (Hu, Yu & Xie, 2008). We must discretize the data, which will inevitably bring about irreversible loss of information, and directly influent the forecast results of port throughput. The neighborhood rough sets can avoid the influence on data accuracy of data discretization (Wang, 2006; Hu et al., 2008; Hu, Yu, & Liu, 2010). Numerical data set can be directly processed by neighborhood rough sets without the need for data discretization (Yu, Bai & Yun, 2013; Liu, Huang, Jiang & Zeng, 2014). So, this article uses reducing methods based on Neighborhood Rough Sets theory.

2.2. Reduction Theory Based on Neighborhood Rough Sets

Granulation and approximation are basic problems of rough set theory and granular computing. Pawlak's rough sets model is built on distinct equivalence relation of discrete space, domain space granulation generated by equivalence relation's dividing the domain. But to the real number space, the value of the object is not discrete, using equivalence relation will result in the value of individual properties' over fitting (Hu et al., 2008).

Neighborhood Rough Sets theory extends the classical rough sets theory. It translate the Symbolic data sets processing based on equivalence relation and indiscernibility relation into hybrid data processing based on the neighbor relationship between distance and neighborhood, and it can deal directly with continuous and hybrid data sets. Consequently, Neighborhood Rough Sets theory can avoid the important and potential loss of information caused by data preprocessing and discretization (Jin, Tung, Han & Wang, 2006; Wang, 2006).

The core concept of neighborhood rough set model is to extend the equivalent approximation of classical rough set model with neighborhood approximation, which enables it to support both numerical and discrete data types. This section will only introduce several necessary concepts on neighborhood rough set model and its reduct, some further details can be found in reference (Hu et al., 2008; Hu, Yu & Liu, 2010).

Definition 1 (Metric space): Given an N-dimensional real number space Ω, Δ: $R^N \times R^N \rightarrow R$, Δ is a measure of R^N. If Δ satisfies:

(1) $\Delta(x_1, x_2) \geq 0$ if and only if $x_1 = x_2$, $\forall x_1, x_2 \in R^N$, $\Delta(x_1, x_2) = 0$;

(2) $\Delta(x_1, x_2) = \Delta(x_2, x_1)$, $\forall x_1, x_2 \in R^N$;

(3) $\Delta(x_1, x_3) \leq \Delta(x_1, x_2) + \Delta(x_2, x_3)$, $\forall x_1, x_2, x_3 \in R^N$.

Then we call $<\Omega, \Delta>$ a Metric space.

Definition 2 (Neighborhood Particle): To a non-empty finite set of real numbers in the given real number space $U = \{x_1, x_2, \ldots, x_n\}$. For any object x_i on U, it's δ– is defined as a neighborhood:

$$\delta(x_i) = \{x_j | x_j \in U, \Delta(x_i, x_j) \leq \delta\} \tag{1}$$

In the formula below, $\delta \geq 0$, $\delta(x_i)$ represents the neighborhood information particles δ which are generated by x_i, it's short called neighborhood particle of x_i. For two-dimensional real numbers space, the norm neighborhood based on norm 1 is diamond, the norm neighborhood based on norm 2 is circle, and the norm neighborhood based on infinite norm is square. It's showed in Figure 1.

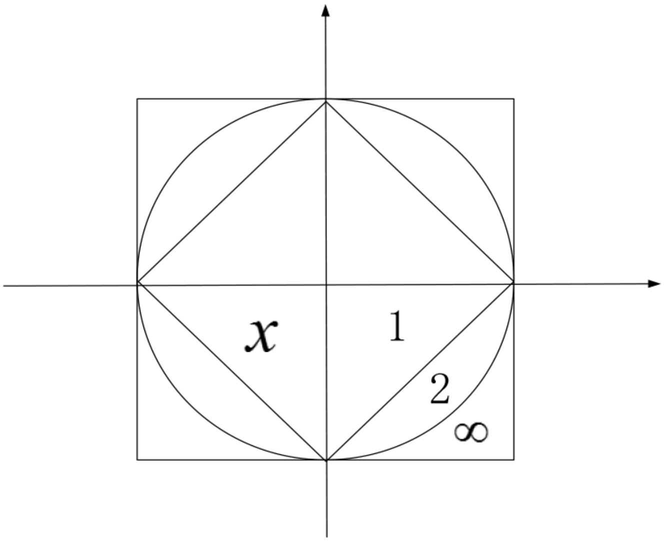

Figure 1. Two-dimensional spatial neighborhood particles

Definition 3 (Positive Region): To a Neighborhood Decision-making System $NDT = <U, A, D>$, $U = \{x_1, x_2, \ldots, x_n\}$, A is a descripition of a real type feature set, D is the decision attribute. U is divided into n equivalence classes by D: $X_1, X_2, \ldots, X_n$, $\forall B \subseteq A$, the lower and upper approximations of D with respect to a partition B are defined as:

$$\underline{N_B}D = \bigcup_{i=1}^{N} \underline{N_B}X_i \tag{2}$$

$$\overline{N_B}D = \bigcup_{i=1}^{N} \overline{N_B}X_i \tag{3}$$

Decision boundary is defined as:

$$BN(D) = \overline{N_B}D - \underline{N_B}D \tag{4}$$

the lower approximations of D is also called the decision positive region, Denoted as $pos_B(D)$.

Definition 4 (Attribute Importance): To a Neighborhood Decision-making System $NDT = <U, A, D>$, $B \subseteq A$, $\forall a \in A - B$. Define the relative importance of A to B as:

$$SIG(a, B, D) = \gamma_{B \cup a}(D) - \gamma_B(D) \tag{5}$$

3. Attribute Reduction of Port Throughput's Influencing Factors

3.1. Analysis of Throughput's Influencing Factors

Factors that have influence on the port throughput are very complicated, in general, can be divided into two categories: macro and micro influencing factors (Zhang, Yan & Xu, 2006). Macro factors mainly refers to the objective regional factors, such as the size of the hinterland area, the social product develop level, Export-oriented economic development level and the number of import and export commodities. Micro factors refers to the port's self-construction conditions, including natural conditions and social economic factors, such as topography, waterways, hydrological and meteorological conditions, vehicle type, ship type, handling and loading ability and technology level, labor organization and management level, type of cargo through the port handling services. All of the above factors are likely to become the important factors which influent the port throughput capacity.

In this section, we build the Port Throughput Influence Index System on the basis of personal practical research and analysis and the research of scholars at home and abroad. The index system including seven aspects: natural environment, political environment, hinterland economy, hardware, service level, collection and distribution capabilities and port logistics. Specific information shown in Table 1.

A level index	B level index	A level index	B level index
Natural Environment	Location Climatic conditions Channel depth Shoreline length	Political Environment	Port Development Policy Regional port planning Port Environment
		Port Logistics	Port transport level Port fixed investment
Hinterland Economy	GDP The added value of primary industry The added value of secondary industry The added value of tertiary industry Foreign trade turnover	Hardware	Number of berths Anchorage capacity Handling equipment capability Cargo carrying capacity
Service Level	Consolidated expenses The efficiency of port operations Port's ships average stopping time Modern management level	Collection and distribution Capabilities	Road freight Sea freight Rail freight

Table 1. Port Throughput Influence Index System

The Port Throughput Influence Index System showed in Figure 1 contents some un-quantifiable indicators, such as natural environment, political environment. In this article, we mainly use the quantifiable indicators to analyze the port throughput forecast. The used indicators are as follows: GDP, the added value of primary industry, the added value of secondary industry, the added value of tertiary industry, foreign trade turnover, port fixed investment, area added value (transportation, warehousing, postal), road freight, sea freight, rail freight and throughput of peripheral port.

3.2. Attribute Reduction

In this article, we use the Forward-Greedy Numerical Attribute Reduction Algorithm based on the neighborhood model. The algorithm is a heuristic algorithm, which uses dependent function to build the Forward-Greedy Numerical Attribute Reduction Algorithm. Its basic idea is: the reduction collection starts with an empty set, each time calculates all remaining properties' attribute importance, choose the attribute whose attribute importance value is the max to add into the reduction set, until all remaining properties' attribute importance value are 0, which means that add any new attribute, the dependent function value in the decision-making system will make no change (Kryszkiewicz, 1998).

Li Sanle has put forward an optimization method for this algorithm (Li, 2012). In this paper, the algorithm is further optimized by adjusting the calculation sequence. First, calculate the relative *D* nucleus of condition property *B*, and assign relative *D* nucleus to *RED*. Then calculate all the condition property equivalence relations in the decision table except relative *D* nucleus. Finally, calculate each equivalence relation using the boundary inspired factor.

$$E_B = Card\left(\overline{N_B}D\right) - \frac{Card\left(\underline{N_B}D\right)}{Card\left(U\right)} \tag{6}$$

Descend E_B, obtain the minimum attribute value of E_B, and then assign the attribute value to *RED*, use $pos_{RED}(D) = pos_A(D)$ as the ending condition of the improved recursive algorithm. The $pos_{RED}(D)$ represents a set of objects, which can be subsumed to decision class *D* in domain *U* according to condition attributes in *RED*.

Repeat the above property selection process, and retain the most suitable property. Use the new generated *RED* and the rest attribute to generate new equivalence relations. The inspired factor EB acts as the measure standard for selecting attributes. When the recursive operations satisfy the constraints, concentrate all the attributes and delete the attributes whose E_B values are big and with low classification capability when combined with relative *D* nucleus. Ultimately, get the best attribute reduction.

The improved algorithm is described as follows:

Input: $NOT = <U, A, D,>$

Output: Reduce *RED*

Step 1: $\forall a_i \in A$, calculate the neighborhood relation Na_i;

Step 2: Calculate $Core_D(A)$, $RED \Leftarrow Core_D(A)$, if $pos_{RED}(D) = pos_A(D)$, turn to Step 7; else, turn to Step 3;

Step 3: $\forall a_i \in A - RED$, calculate $E_{a_i \cup RED}$;

Step 4: Choose the mix attribute a_k in $E_{a_i \cup RED}$, $a_k = \arg\min_{a_i \in A-RED} E_{a_i \cup RED}$, when there are two or more minimum, calculate with the attribute that has the least attribute values.

Step 5: $RED \Leftarrow RED \cup a_k$, if $pos_{RED}(D) \neq pos_A(D)$, turn to Step 3; Else, turn to Step 6;

Step 6: Assume the result from Step 5 as $P = \{a_1, a_2, \ldots, a_k\}$ $(k \leq m)$; $\forall a_i \in \frac{p}{Core_D(A)}$, get $a_{\max} = \max_{\forall a_i \in RED \backslash B} E_{a_i \cup RED}$, set $RED = RED \backslash a_{\max}$, if $pos_{RED}(D) = pos_A(D)$, delete $a_{\max}$; Else, remain $a_{\max}$, and traversal all the a_i;

Step 7: Export RED.

4. Experimental Evaluation

On the basis of job in section 3.1, considering availability of data, we search the Statistical Yearbook and choose nine influential factors to finish attribute reduction, such as GDP, foreign trade turnover, the added value of primary industry, the added value of secondary industry. This article selects Guangzhou Port and its hinterland economy data as the Empirical Analysis sample. The arranged data is showed in Table 2.

No	1	2	3	4	5	6	7	8	9	10
2000	3819	233.8	488.49	3664.66	3370.81	4.09	26141	282.4	2.39	1.112
2001	4341	230.37	509.3	4173.55	3811.56	2.87	26982	320.2	1.95	1.282
2002	4842	279.3	532.69	4701.45	4301.12	1.88	30839	358.6	2.77	1.532
2003	5795	349.4	547.84	5962.1	4943.16	2.76	37081	397.3	3.16	1.718
2004	7523	448	596.84	7233.15	5742.25	23.7	39693	424.4	3.54	2.152
2005	9719	534.9	596	9138.6	8382.02	28.4	44258	676	3.8	2.503
2006	11729	637.7	604.43	11045.56	9820.21	33.2	47077	668.9	4.11	3.028
2007	13879	734.9	659.09	13569.21	13213.68	23.1	50301	732.11	4.38	3.413
2008	16322	819.5	711.48	14964.54	14069.59	10.7	53737	754.8	4.57	3.695
2009	17201	767.4	725.14	15344.41	16040.62	12.8	56329	784.9	4.04	3.75
2010	20501	1037.8	800.65	18394.09	17893.45	14.1	62071	746.3	4.58	4.096
2011	23618	1161.7	910.36	21291.04	21669.79	15.6	65462	446.9	4.67	4.29
2012	25220	1171.3	994.2	22198.36	24704.77	17.2	69730	502.2	4.59	4.33
2013	27364	1210.3	980.5	24210.22	26537.02	18.7	73246	530.1	4.75	4.73

Remark: the source of Table 2 is China statistical yearbook(1990-2013).
Index Number 1: GDP (Total GDP of Guangzhou, Foshan and Dongguan, Unit hundred million Yuan).
Index Number 2: Foreign trade turnover of Guangzhou (Unit hundred million Dollars).
Index Number 3: The added value of primary industry (Unit hundred million Dollars).
Index Number 4: The added value of secondary industry (Unit hundred million Yuan).
Index Number 5: The added value of tertiary industry (Unit hundred million Yuan).
Index Number 6: Construction Investment of Guangzhou Port (Unit hundred million Yuan).
Index Number 7: Collection and distribution capabilities of Guangzhou Port (Unit Ten thousand Tons).
Index Number 8: Transportation, warehousing and postal added value of Guangzhou (Unit hundred million Yuan).
Index Number 9: Throughput of peripheral port cargo (Unit Hundred million tons).
Index Number 10: Decision attribute, real throughput of Guangzhou Port (Unit Hundred million tons).

Table 2. Original Data of the Throughput of Guangzhou Port

When calculate sample's neighborhood, standardize the numeric attributes to interval [0,1] (Jian, Liu, Fang, Dang, Zhu, Wu, et al., 2007), in order to reduce the impact of results because of dimension inconsistent of attributes. Data is showed in Table 3.

No	1	2	3	4	5	6	7	8	9	10
2000	0.03819	0.02338	0.48849	0.0366466	0.0337081	0.0409	0.26141	0.2824	0.239	0.1112
2001	0.04341	0.023037	0.5093	0.0417355	0.0381156	0.0287	0.26982	0.3202	0.195	0.1282
2002	0.04842	0.02793	0.53269	0.0470145	0.0430112	0.0188	0.30839	0.3586	0.277	0.1532
2003	0.05795	0.03494	0.54784	0.059621	0.0494316	0.0276	0.37081	0.3973	0.316	0.1718
2004	0.07523	0.0448	0.59684	0.0723315	0.0574225	0.237	0.39693	0.4244	0.354	0.2152
2005	0.09719	0.05349	0.596	0.091386	0.0838202	0.284	0.44258	0.676	0.38	0.2503
2006	0.11729	0.06377	0.60443	0.1104556	0.0982021	0.332	0.47077	0.6689	0.411	0.3028
2007	0.13879	0.07349	0.65909	0.1356921	0.1321368	0.231	0.50301	0.73211	0.438	0.3413
2008	0.16322	0.08195	0.71148	0.1496454	0.1406959	0.107	0.53737	0.7548	0.457	0.3695
2009	0.17201	0.07674	0.72514	0.1534441	0.1604062	0.128	0.56329	0.7849	0.404	0.375
2010	0.20501	0.10378	0.80065	0.1839409	0.1789345	0.141	0.62071	0.7463	0.458	0.4096
2011	0.23618	0.11617	0.91036	0.2129104	0.2166979	0.156	0.65462	0.4469	0.467	0.429
2012	0.25220	0.11713	0.9942	0.2219836	0.2470477	0.172	0.69730	0.5022	0.459	0.433
2013	0.27364	0.12103	0.9805	0.2421022	0.2653702	0.187	0.73246	0.5301	0.475	0.473

Table 3. Granular Data of Guangzhou Port Throughput Impact Indicator

On the basis of comprehensive reference on literature information, we set the neighborhood radius in 0.1 to 0.4. And in order to avoid data overlap in the boundary neighborhood, using left open right closed interval for neighborhood of δ. So the neighborhood relations constitute a full coverage of domain.

In this problem, we with a no duplicate completely coverage of domain, the principle is priority select the max coverage of U of neighborhood. This article set the step length of δ as 0.05, through adjusting relations between δ and neighborhood to reach the best attribute reduction. In this article, we use MATLAB to reduce attribute. Results are as follows:

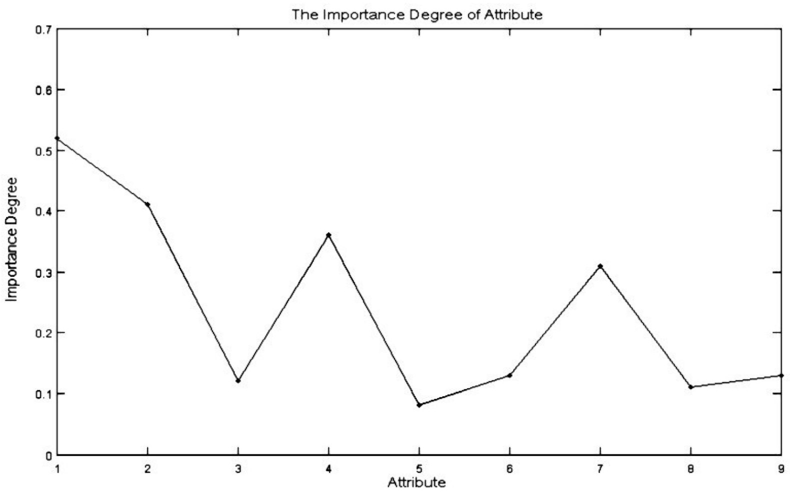

Figure 2. The Importance Degree of Attribute

The improved Forward-Greedy Numerical Attribute Reduction Algorithm based on the Neighborhood Rough Sets directly starts from relative nucleus, reducing the amount of computation. And this algorithm concentrates the attributes reduction to the best reduction by using deleting method. When there are many condition attributes, the algorithm can priority calculates the relative nucleus, which will save lots calculating time. At the same, this algorithm uses the new inspired factor, solve the common problem from different aspects.

According to the attribute importance in Figure 2, the key factors of Guangzhou Port's throughput are GDP, foreign trade turnover of Guangzhou, and the added value of secondary industry, collection and distribution capabilities of Guangzhou Port. Though different port may have different key influential factors of throughput, using this method can also get different ports' key influential factors. And key influential factors of Guangzhou Port may have certain reference significance for other ports. From the year of 1978, GDP and waterway transportation turnover grows synchronously, provides the market and opportunity for water transport. Foreign trade turnover of Guangzhou directly reflects Guangzhou's total cargo transport volume, which is closely related to the development of manufacturing and foreign trade markets of Guangzhou and surrounding areas.

Guangzhou as southern China's largest mineral trading center such as coal, secondary industry is still the main economic development of Guangzhou and the surrounding areas. A highly developed collection and distribution system can improve cargo turnover efficiency, and increase port product capability. Through the above analysis, the experimental results have been given a reasonable explanation, which proves that this method is accurate and reliable.

5. Conclusion

This article aims at reducing attributes for factors of port throughput using the theory of Neighborhood Rough Sets, finding out the key influential factors for port throughput. In order to provide support for port business investment, decisions and risk control, and also provide assistance for port enterprises' or other researchers' throughput forecasting. On the basis of practical research and analyze of factors for port throughput, this article selects Guangzhou Port and it's hinterland economy data as the Empirical Analysis sample. Uses the improved Forward-Greedy Numerical Attribute Reduction Algorithm based on the Neighborhood Model to get Guangzhou Port Throughput's key influential factors.

This article mainly has following results:

1. Port Throughout Indicator System. Firstly, analyzed the entire port industry's throughput influential factors from various aspects, limit the entire port industry's throughput influential factors system. Then combined with Guangzhou Port's practical data, establish the Port Throughout Indicator System considering internal and external factors on the basis of analyzing Guangzhou Port's formation mechanism.

2. Attributes Reduction of Neighborhood Rough Sets. There are many factors, which can influent port throughput, so how to select the most important factor as the forecast factor is the key point for a successful forecast model. This article used the improved Forward-Greedy Numerical Attribute Reduction Algorithm based on the neighborhood model to reduce the Port Throughout Indicator System, then chose the most important factor according to the reduce results. This method avoids the interference of unrelated factors to the model results, while improves models' operating efficiency.

Acknowledgement

This research was supported by the Fundamental Research Funds for the Central Universities under grant 2014JBM055.

References

Hu, Q.H., Yu, D.R., & Liu, J.F. (2010). Neighborhood Rough Set Based Heterogeneous Feature Subset Selection. *Information Sciences,* 178(18), 3577-3594. http://dx.doi.org/10.1016/j.ins.2008.05.024

Hu, Q., Yu, D., & Xie, Z. (2008). Neighborhood Classifiers. *Expert Systems with Applications,* 34(2), 866-876. http://dx.doi.org/10.1016/j.eswa.2006.10.043

Hu, K.C. (2005). The present situation and development strategy of China port logistics industry. *Shipping Management,* 27(1), 32-35. http://dx.doi.org/10.3969/j.issn.1000-8799.2005.01.012

Jin, W., Tung, A.K.H., Han, J., & Wang, W. (2006). Ranking Outliers Using Symmetric Neighborhood Relationship. *Lecture Notes in Computer Science,* 3918, 577-593. http://dx.doi.org/10.1007/11731139_68

Jiang, F., & Sui, Y. (2015). A novel approach for discretization of continuous attributes in rough set theory. *Knowledge-Based Systems,* 73, 324-334. http://dx.doi.org/10.1016/j.knosys.2014.10.014

Jian, L.R., Liu S.F., Fang, Z.G., Dang, Y.G., Zhu, J.J., Wu, H.C. et al. (2007). The Performance Evaluation on Discipline Construction of Teaching and Research University Based on Dominance Rough Set Theory. *Journal of Industrial Engineering and Engineering Management,* 21(3), 132-136. http://dx.doi.org/10.3969/j.issn.1004-6062.2007.03.025

Kondo, M. (2006). On the structure of generalized rough sets. *Information Sciences,* 176(5), 586-600. http://dx.doi.org/10.1016/j.ins.2005.01.001

Kryszkiewicz, M. (1998). Rough set approach to incomplete information systems. *Information Sciences,* 112(4), 39-49. http://dx.doi.org/10.1016/S0020-0255(98)10019-1

Liu, Y., Huang, W.L., Jiang, Y.L., & Zeng, Z.Y. (2014). Quick attribute reduct algorithm for neighborhood rough set model. *Information Sciences,* 271(7), 65-81. http://dx.doi.org/10.1016/j.ins.2014.02.093

Liu, G., & Zhu, W. (2008). The algebraic structures of generalized rough set theory. *Information Sciences,* 178(21), 4105-4113. http://dx.doi.org/10.1016/j.ins.2008.06.021

Li, S.L. (2012). Improving on an Attribute Reduction Algorithm Based on Neighborhood-Based Rough Set. *Microcomputer Information,* 26(36), 268-269. http://dx.doi.org/10.3969/j.issn.2095-6835.2010.36.106

Pawlak, Z. (1998). Rough sets theory and its applications to data analysis. *Cybernetics and Systems,* 29(1), 661-688. http://dx.doi.org/10.1080/019697298125470

Pawlak, Z., & Skowron, A. (2007a). Rudiments of rough sets. *Information Sciences,* 177(1), 3-27. http://dx.doi.org/10.1016/j.ins.2006.06.003

Pawlak, Z., & Skowron, A. (2007b). Rough sets: some extensions. *Information Sciences,* 177(1), 28-40. http://dx.doi.org/10.1016/j.ins.2006.06.006

Pawlak, Z., & Skowron, A. (2007c). Rough sets and Boolean reasoning. *Information Sciences,* 177(1), 41-73. http://dx.doi.org/10.1016/j.ins.2006.06.007

Qin, K., Yang, J., & Pei, Z. (2008). Generalized rough sets based on reflexive and transitive relations. *Information Sciences,* 178(21), 4138-4144. http://dx.doi.org/10.1016/j.ins.2008.07.002

Wang, H. (2006). Nearest Neighbors by Neighborhood Counting. *IEEE Transactions on Pattern Analysis & Machine Intelligence,* 28(6), 942-953. http://doi.ieeecomputersociety.org/10.1109/TPAMI.2006.126

Yu, Z.M., Bai, X.L., & Yun, Z.Q. (2013). A study of rough sets based on 1-neighborhood systems. *Information Sciences,* 248(6), 103-113. http://dx.doi.org/10.1016/j.ins.2013.06.031

Zhu, W., & Wang, F.Y. (2003). Reduction and axiomatization of covering generalized rough sets. *Information Sciences,* 152(8), 217-230. http://dx.doi.org/10.1016/S0020-0255(03)00056-2

Zhu, W. (2007a). Topological approaches to covering rough sets. *Information Sciences,* 177(6) 1499-1508. http://dx.doi.org/10.1016/j.ins.2006.06.009

Zhu, W. (2007b). Generalized rough sets based on relations. *Information Sciences,* 177(22), 4997-5011. http://dx.doi.org/10.1016/j.ins.2007.05.037

Zhang P., Yan Y.X., & Xu C.X. (2006). Prediction Model of Port Throughput Based on System Dynamics. *Port and Waterway Engineering,* 32(4), 26-37. http://en.cnki.com.cn/Article_en/CJFDTOTAL-SYGC200605007.htm