Intelligent Automation And Soft Computing, 2020
Vol. 26, no. 3, 617–624
DOI: 10.32604/iasc.2020.013940

Application of the DRGs and the Fuzzy Demand in the Medical Service Resource Allocation Based on the Data Mining Algorithm

Fanxiu Dong

Xinjiang Medical University, Xinjiang 830000, China
Address: Medical Record Management Office, Xinjiang Uygur Municipality People's Hospital, 91 Tianchi Road, Tianshan District, Urumqi, Xinjiang, 830001

ABSTRACT
At present, the allocation of the medical service resources is directed at a single service resource, and there are many unreasonable problems, which causes medical cost to be high. Based on this, the application of the DRGs and the fuzzy demand in the medical service resource allocation based on the data mining algorithm is proposed. The application research of the DRGs and the data mining algorithm is simply analyzed, then the uncertain demand estimation is applied to the fuzzy demand processing based on the fuzzy demand theory and the medical service resources are configured under the established demand satisfaction rate. Through the example analysis, it is proven that the proposed algorithm has higher reliability, and the optimal allocation of resources can greatly reduce the total medical cost and confirm the feasibility of the resource optimization.

KEY WORDS: DRGs; data mining; fuzzy demand; resource allocation.

1 INTRODUCTION

THE allocation of the medical service resources means allocating the resources of the hospital to all parts to meet the needs of the patients and ensure the quality of medical care. In recent years, medical services often appear in the register difficult, and bed utilization rates of underground problems not only increased the management costs, but also increased cost of medical services, resulting in patient missed treatment time (Liu Y et al, 2016). The low efficiency of medical institutions and unreasonable allocation of medical resources are important causes of the rise in cost and many scholars begin to analyze and solve the problem of service efficiency of the medical resources. The allocation of the medical service resources needs to be set up in the reasonable unit cost, by the product line management. The medical system put forward the DRGs mechanism, the application is more successful, and successfully solved the problem of the doctor's expense (Tseng S F et al, 2015). The Medical Service resources cover the resources of technology, manpower and material resources. At present, the research of the medical resource allocation focuses on the management of operating room scheduling, medical resource dispatching, medical equipment utilization and so on (Mihea C et al, 2013). Most of the research focuses on the allocation of the single core medical service resources, which has great limitations and cannot reflect the overall allocation of resources (Liu M et al, 2015). It is proposed that the combination of the DRGs and the data mining fuzzy set algorithm to realize the optimal allocation of the medical service resources has some practical significance in promoting the allocation of the medical service resources.

Specific contributions of this paper include: The application of the disease-based payment and the fuzzy demand in the allocation of medical service resources is briefly analyzed. The application research of the disease-based payment and the data mining algorithm is analyzed. A latest demand estimate to the fuzzy demand processing based on the fuzzy demand theory is applied. The optimal configuration of resources can greatly reduce the total medical cost and the medical service resources are configured with the established satisfaction to prove that the algorithm has high reliability and feasibility.

The rest of this paper is organized as follows: Section 2 discusses related work, followed by the DRGs information management data algorithm. The

CONTACT Fanxiu Dong 422637858@qq.com

uncertainty theory, fuzzy prediction of the resource allocation, and the medical service resource allocation are analyzed in Section 3. Data processing and the result analysis are discussed in Section 4 and Section 5 concludes the paper with summary and future research directions.

2 RELATED WORK

THE DRGs information management has been used for some time in hospitals and a lot of data resources have been accumulated. The data can be used to extract individual drug use, diagnostic patterns, and personal information that can provide important information support for medical diagnostics. The data quantity is huge, the work search is difficult, and it is not easy to match to the optimized grouping scheme. Therefore, the data mining algorithm can be used in order to find out the valuable information from the massive data, and then improve the working level (A M B et al, 2012). At present, the algorithm, which is combined with the DRGs is mainly the support vector machine, neural network, k means, fuzzy set algorithm and so on. The BP neural network algorithm can divide the DRGs information into several levels in operation. Each level is computed, and then the mapping is done for different mappings, which exist in a two-dimensional plane, which completes the weight set under the constraint (Yan Y H et al, 2017). The Support Vector Machine (SVM) solves the problem of the data analysis with less data quantity and nonlinearity and introduces the adaptive resonance theory to further analyze the valuable information in the DRGs packet (Huang Y H et al, 2013), then the classification is dug. The K-means algorithm is a distance-based algorithm, which classifies the data in the DRGs and then analyzes the similarity between the data.

The fuzzy set theory has been widely used to characterize uncertainty. When there is a lack of certain data or available historical data, the fuzzy set theory provides an alternative and convenient framework for dealing with uncertain medical service needs. The most likely values in the range and the range of possible changes in the medical service needs requires expert estimates and designations. According to the membership function definition of the triangular fuzzy variable, the hospital administrator only needs to estimate the variable value range and estimate the most likely value.

The unreasonable allocation of the medical service resources has always been a major problem in the medical and health service industry. The unreasonable allocation of the medical resources, and the high health care investment does not necessarily lead to a high health output but will increase medical costs. To fundamentally solve the problem of poor medical care for our residents and expensive medical treatment, we need to realize the rational allocation of the medical resources as much as possible to meet the needs of patients. In view of the current situation of the resource allocation of the medical service providers combined with product line management in order to simplify the medical service process, the use of the DRGs and fuzzy demand, is to establish a goal to meet the level of demand for certain medical services, is the model of the medical service resource expenditure minimization.

3 METHODOLOGY

3.1 The Uncertainty theory

THE uncertain problem of the social economic production and service is the theory of uncertainty, which mainly covers the fuzzy set theory, rough set theory, set pair theory, inclusion degree theory and so on. Based on the database, the rough set theory looks for the decision rule from the database and takes the relation of the object combination in the different classes as the research objective (Murty M S N, 2013). The rough set uses the classification method to describe the object, which cannot be precisely described as the rough set. In the set of objects, R represents the hierarchical relationship of U, and the approximate space equivalence class is expressed as:

$$U / R = \{[x_i]_R \mid x_i \in U\} \quad (1)$$

For X that is considered to belong to U, exists:

$$\underline{R}(x) = \{x_i \mid [x_i]_R \in X\} \quad (2)$$

The rough set theory has been applied for a long time, and two kinds of generalization methods have been formed at present, which are the algebraic method and the constructive method respectively. The constructive forensic approximate space is given as the cut-in point, and the rough set is studied and popularized. In the data mining application, the relational table in the relational database is treated as the decision information by the rough set theory, then the uncertain support knowledge is realized, and the data mining is achieved.

The theory of the fuzzy sets is a theory appearing in the late 20th century. In recent years, the fuzzy clustering analysis, fuzzy decision prediction, fuzzy programming and fuzzy control have begun to appear (Lappa E et al, 2013). Assume that X contains only two domains, the fuzzy set A of X is represented by Zadeh, and the formula is:

$$\text{A}=u_A(x_1)/x_1+u_A(x_2)/x_2+\ldots+u_A(x_n)/x_n \quad (3)$$

The expression + denotes a continuous symbol. The fuzzy set theory is widely used and is used to deal with the problem of uncertainty in this paper (Geissler A et al, 2013).

3.2 The Resource Allocation Fuzzy Prediction

The demand factor is to determine the key factors of production, and fully grasp the market demand information, which can make a reasonable production plan, avoid all kinds of risks, realize the minimum cost investment to get the maximum benefit. The traditional processing method of the uncertain demand is to transform the uncertain problem into the deterministic problem and then analyze it, and the uncertain demand processing methods are mainly the multistage analysis, probability analysis, demand fuzzy analysis, interval analysis and so on. The multi-stage analysis means to divide each stage to open, then adopt different demand processing methods, and then solve the problem of uncertainty. The interval analysis means using interval programming and an algorithm to solve the problem of uncertainty and using the interval analysis to deal with uncertainties. The probabilistic analysis refers to the use of the probabilistic LUN theory to deal with the problem of uncertainty, including a variety of algorithms, such as random simulation, scenario analysis and so on. The lupin optimization is the optimization of the system in the internal and external uncertainty environment, using discrete scenarios to describe the uncertain parameters.

For the uncertainty of demand, the methods used in the domestic research include; stochastic linear programming, deterministic linear programming, fuzzy linear programming and so on, in which the fuzzy set theory is widely used, and the results can be obtained more accurately on the basis of the lack of uncertain data. The use of fuzzy sets in the medical service needs is processed and evaluated according to the possible range of requirements. The fuzzy analysis method is used to describe the market demand by fuzzy variables, then the fuzzy method is used to deal with the uncertain demand. The probability distribution is obtained, and the result is fit for the actual production. Fuzzy demand processing has been applied to the problems of production and vehicle routing, and the uncertain demand is mainly the triangular fuzzy number. Assume that I represents a regular fuzzy set of real numbers, for any λ, suppose I is a closed interval. The I is considered a fuzzy number when the following conditions are met:

$$u_A(\lambda x_1 + (1-\lambda)x_2) > \min(u_A(x_1), u_A(x_2)) \quad (4)$$

If function F(x) satisfies the condition:

$$F(x)=F(1-x) \quad (5)$$

It is assumed that in a range greater than 0, the function is a non-increment function, and F (x) becomes the reference parameter of the fuzzy number. Assume that L and R are both reference functions of F(x), if the membership function of Fuzzy I satisfies the following formula, the formula is the L-R fuzzy number:

$$u_I(x) = \begin{cases} L(\theta - x)/a, x < \theta \\ R(x-\theta)/b, \theta < x \end{cases} \quad (6)$$

a is the left and right extension value, b is the datum, and θ is the mean value of the I or the most probable value.

When the membership function of L-R type fuzzy number I satisfies the formula (7), it is considered that I is a triangular fuzzy number, l represents a lower bound, and u represents an upper bound, and L is a triangular fuzzy number when the upper and lower bounds are the same. The fuzzy set theory is widely used in the fields of industry, agriculture and information.

$$u_I(x) = \begin{cases} (x-l)/(\theta - l), x < \theta \\ (x-u)/(\theta - u), \theta < x < u \\ 0, \text{else} \end{cases} \quad (7)$$

The demand for the medical services is a special commodity, but consumers are not health products and consumers buy health services. There are many factors affecting the demand of general products and services, such as price, income, consumer preference and so on. As the demands of health, medical services are different from other commodities, and it will be influenced by moral risk factors and inducement factors. The medical service needs are divided into two aspects of prevention and treatment, combined with the existing analysis. There are three types of medical service demands and there is uncertainty and inducement in the market of medical services. These factors use the influence of product demand to decide the uncertainty together. The fuzzy algorithm is introduced to analyze the demand of service by using the triangular fuzzy algorithm. The membership function of the triangular fuzzy demand is shown in Figure 1.

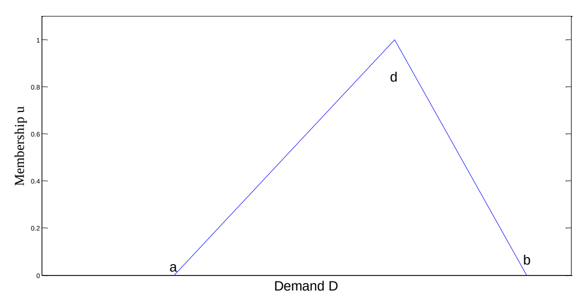

Figure 1. The Membership Function of the Triangular Fuzzy Demand.

Triangular fuzzy variables are used to describe the service needs of a certain periods of time, and the function is defined according to the membership degree of triangular fuzzy variables. Hospital managers are only required to assess the range of variables and the most likely value, and to be able to obtain the most likely value of medical services. The reliability distribution function of the triangular fuzzy number is used to measure the fuzzy demand and ensure the processing dependability.

3.3 The Medical Service Resource allocation

The medical service organizations can be regarded as a kind of special service enterprise. The DRGs are a description of the Medical service product system, accounting for the pathological combination cost. The DRGs are the first theory proposed by Yale University, the Chinese name for the disease diagnosis Related group, describes the medical service products widely used in domestic and foreign hospitals. The domestic DRGs research is generally associated with a disease, prompting the DRGs to combine the real situation and the DRGs to cover all kinds of diseases. The DRGs prepaid system is used to improve the resource utilization and reduce medical support. The application and research of foreign DRGs are very early, and there are not many domestic studies. In the traditional medical organization, the Department does not assume the corresponding responsibility, through the individual physician to coordinate the diversity of the Medical Department, and then provide services to patients. The cost accounting system based on the product line provides and the integrated framework for patients. The total cost of the medical service resources covers the surplus cost and the shortage cost, considering that the medical service resources are more special. The unreasonable allocation of resources will cause the waste of resources. The resource Process allocation section is introduced and expressed as:

$$\zeta_{\mathrm{i}} = (d_i^l, d_i, d_i^n) \qquad (8)$$

This formula expresses the fuzzy demand of the disease-type-related groups. E is used to indicate the desired needs of the disease-related group, expressed as:

$$E(\zeta_i) = (d_i^l + 2d_i + d_i^n)/4 \qquad (9)$$

E reflects the model of the disease demand forecast, using E to represent the resource unit vector, U represents the basic consumer unit, C represents the unit vector corresponding to the resource unit expenditure vector, and C represents the unit expenditure, w means the average consumption. The resource consumption matrix of a class L disease diagnosis Related Group is expressed as:

$$\mathrm{R_L} = \begin{Bmatrix} w_{11}, \dots, w_{1\mathrm{n}} \\ \dots \\ w_{\mathrm{n}1}, \dots, w_{\mathrm{nn}} \end{Bmatrix} \qquad (10)$$

The line vector represents the disease-related group, consumption, x represents the number of units, R represents the total budget of resources, and TC represents the overall expenditure.

To determine the cost of treatment, first clinical patient data needs to be obtained. Second, the detailed list of each patient's diagnosis and treatment is described, and records are integrated to obtain each patient's service framework. Finally, the detailed cost is obtained by definition in the service unit, and the standard cost is analyzed to obtain the cost of the patient in a DRG group. The unreasonable allocation of medical resources is one of the important causes of high medical cost. Questioning a healthy output, but also the doctor in question, the need to realize the service resources reasonable disposition as soon as possible and satisfy the populace demand. The product line management simplifies the process of the resource allocation by combining the fuzzy demand and the DRGs to realize the medical service resources and minimizing the cost of resources. To establish the objective function, the formula is expressed as:

$$\min TC = \sum_{j=1}^{N} c_j \sum_{i=1}^{L} x_{ij} \qquad (11)$$

There is $x_{ij} > E(\zeta_i) w_{ij} \cdot \vartheta_i^o$ in the equation, which responds to the diagnostic group, based on the expected demand of the confidence measure. E indicates the patient's predictive number and meets the actual situation when the quality of the medical service is met. Assume that the objective function is minimized, it should be able to meet the constraint conditions:

$$\mathrm{st.}\min\{\frac{x_{iN}}{E(\zeta_i) w_{iN}}\} > \theta_i^0 \qquad (12)$$

In the equation, i=1, 2, 3,..l., this constraint can be further changed, and the formula is expressed as:

$$\min TC = s.t. \begin{cases} \dfrac{x_{i1}}{E(\zeta_i) w_{i1}} > \theta_i^0 \\ \sum_{j=1}^{N} c_j \sum_{i=1}^{N} x_{ij} < B \\ 0 < x_{ij} < B / c_j \end{cases} \qquad (13)$$

In the model, the L Group Disease Diagnosis Related group reacted to a certain disease type. In order to ensure the core treatment advantage, the key medical resources will be allocated to these disease medical treatments. Assume that the diagnostic level of each group is at the same level, for a disease diagnosis, the following conditions are met:

$$\frac{x_{i1}}{E(\zeta_i)w_{i1}} = \frac{x_{iN}}{E(\zeta_i)w_{iN}} \quad (14)$$

The optimal allocation of resources can be realized and the cost of the resource input is minimized.

In the case of the fixed capital investment, in order to obtain the maximum benefit, the above analysis is combined to meet the needs of the hospital, and the resource allocation is carried out in order to simplify the calculation, existing the assumption is:

$$\theta_1^0 = \ldots = \theta_L^0 = \theta^0 \quad (15)$$

According to this formula, all kinds of the consumption information and forecast information of hospital is used. The medical service resources are the input, and the simulation with MATLAB compares and analyzes the resource configuration.

4 RESULT ANALYSIS AND DISCUSSION

4.1 Data Sources and Processing

THE types of hospital patients and the number of cases is large. The article research and analysis of a simple disease were used as an example, the hospital, and a single patient information resource was investigated. The data pre-processing, rough set theory software Rosseta is used to analyze the decision table, and redundant attributes are eliminated. The genetic algorithm is used for the attribute reduction in the decision table, and no redundant data is found. The condition attributes of the study included gender, hospital level, surgery, diagnosis, nursing, etc. The hospitalization cost is the decision attribute. The diagnosis, treatment service consumption and combined treatment were determined by using the well-defined case type, combined with the medical conditions related to the surgical treatment of the disease. Table 1 is a single case data collection.

A case of a single disease was used to investigate the hospital patient information. The calculated data was collated and the demand rate for the quarter was met at 43.53%. The Delphi method was used to predict the most probable value and range of patients, and to obtain a predictive triangular fuzzy demand of (240,260,320). The comparison between the unit cost and total cost is shown in Figure 2.

Table 1. **The Single Case Data Set.**

$\vec{RU}$	u_j	**Consumption**	**Unit cost**	**total cost**
u_1	First-degree nursing	2	65.37	131.4
u_2	Two level nursing	2	89.4	178.6
u_3	Operation time	1	148.5	148.8
u_4	An aesthesia	1	42.6	42.7
u_5	test	1	64.7	64.7
u_6	affine	1	26.5	26.5
u_7	Other	1	189.2	189.2
Total				781.9

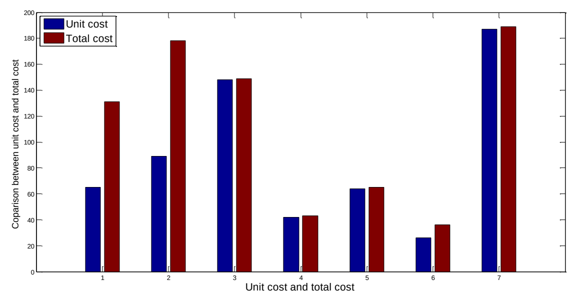

Figure 2. **The Comparison between the Unit Cost and Total Cost.**

4.2 Result Analysis

Based on the basic situation of the database variables, the relevant research is combined. The diagnosis, nursing, surgery, time, etc. are used as input variables, with the cost as the output variable. The learning set and test set are combined with the rough set data mining algorithm to establish the regression analysis model. After the calculation, a total of 7 basic variables were included, including nursing, operation time, anesthesia, examination, and radiation, which divided the nursing into one level nursing and two-level nursing. The test set is diagnosed by using the obtained rules, and Table 2 is the calculated result, and the total compliance rate is 84.2%.

Table 2. **Diagnostic Results of the Test Sets using the Existing Rules.**

Actual cost	**Rule**	
	< 700	> 700
< 700	1507	147
> 700	200	339

The ROC curve is plotted according to the calculation results, and Figure 3 is the result of the calculation. It can be seen from the graph that the area under the curve and the 95% ci value are 0.77 and the prediction ability is high. Based on the above results, the test set number is used to evaluate the rough set after the rule is established. The standard error of the ROC curve is 0.016, according to α=0.05 Test standard, and the fuzzy set algorithm is better for the hospitalization cost classification.

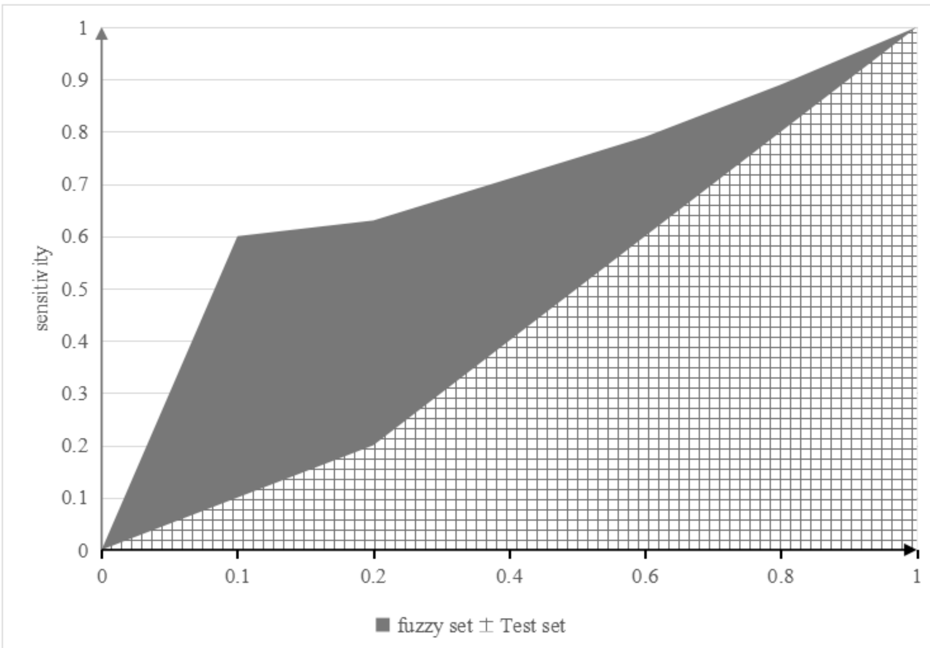

Figure 3. **The ROC Curve for the Diagnosis of a Test Sample.**

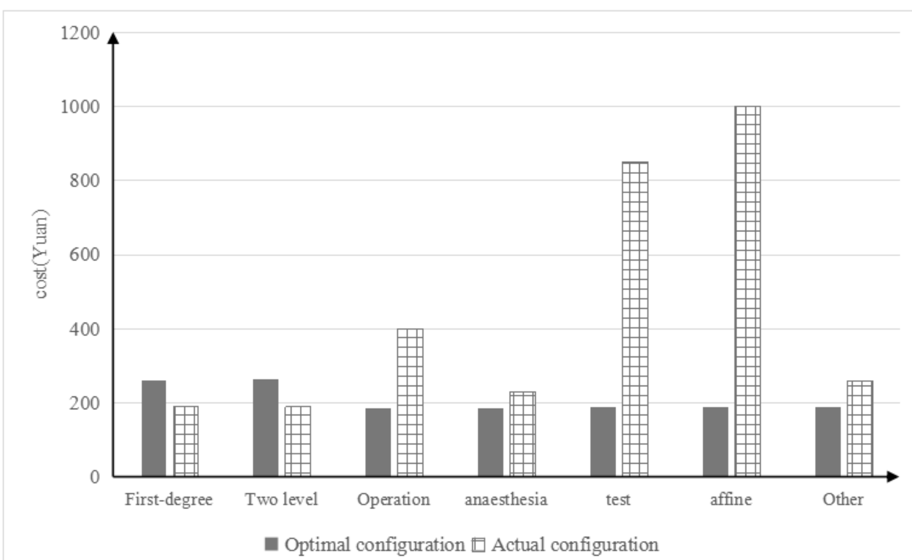

Figure 4. **The Comparison of the Two Schemes.**

It is assumed that the actual demand satisfies the constraint condition and satisfies the actual resource total budget constraint condition. According to the formula to optimize the configuration scheme and the actual resource allocation scheme, Figure 4 is a comparison of the two scenarios. Under the actual budget level, the MATLAB software is used to solve. The best scheme of resource allocation is, u_1 = 683.5, u_2 = 722.2, u_3 = 368.5, u_4 = 505.6, u_5 = 352.0, u_6 = 716.8, u_7 = 332.2, under which the resource input can satisfy the fuzzy demand of the estimating credibility.

After the end of the quarter, the survey received 300 actual cases. According to the actual data and the reliability analysis of the triangular fuzzy number, the expert can predict the reliability of the fuzzy demand is 0.833. The reliability and rationality of the demand estimation can be seen from this data, which can ensure the scientific fuzzy demand optimization model. It can be seen from the graph that the optimized resource allocation scheme based on the proposed algorithm is slightly higher in the resource allocation of the nursing level, the other resources allocation is obviously lower than the actual resource allocation, which indicates that the unreasonable allocation of hospital resources, especially in the resource monitoring. According to the ideal allocation ratio, a more reasonable resource consumption standard can be specified, and the total expenditure of the optimized resource allocation decreases obviously. That is, in the case of satisfying the demand, the optimized scheme only accounts for 33% of the actual expenditure. It can also be seen from the data that the optimized allocation resource can reach the ideal satisfaction rate, and it also indicates that the allocation of the hospital resources is unreasonable. The cost trend is shown in Figure 5.

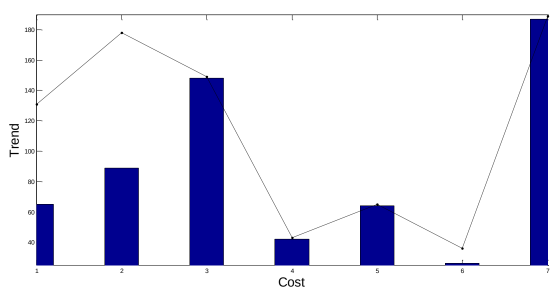

Figure 5. The Cost Trend.

5 CONCLUSION

THE use of the DRGs paid technology can improve the quality of the medical service, but it is aimed more at the single disease and unreasonable in the resource allocation. Based on the application of the data mining algorithm, the DRGs, the fuzzy demand in the medical service resource allocation, the simple analysis of the DRGs application background, and the fuzzy demand of the credibility theory in the medical resource allocation is introduced, combining the credibility theory and the triangular fuzzy algorithm to predict the uncertain demand. The forecast demand is closer to the actual demand. The fuzzy medical service demand is handled, and the data resource configuration is realized. The case analysis shows that the DRGs and the fuzzy demand in the resource allocation can optimize the allocation of the medical resources, improve the allocation of the resources in nursing, reduce the resource allocation in diagnosis and operation, and make the resource allocation more reasonable. It is necessary to point out that the classic DRGs is used in the article to optimize the allocation of the medical resources, not to consider the cost of the standard resource consumption unit. Although the uncertainty of the resource allocation is forecasted, the research data is obviously insufficient, which remains to be further studied.

6 REFERENCES

A M B, B H O. DRGs in Transfusion Medicine and Hemotherapy in Germany. [J]. Transfusion Medicine & Hemotherapy Offizielles Organ Der Deutschen Gesellschaft Fur Transfusions medizin Und Immunhamatologie, 2012, 39(2):60.

Geissler A, Quentin W, Busse R. [Do German DRGs appropriately explain the resource consumption of hospitals? An empirical patient-level cost-data analysis across ten episodes of care]. [J]. Gesundheitswesen, 2014, 76(5):284.

Huang Y H, Chang C W. Constructing a Data Mining Sensor to Classify Tw-DRGs Medical Specialties[J]. Applied Mechanics & Materials, 2013, 263-266:969-972.

Lappa E, Giannakopoulos G. E-health Information Management according Types of DRGs and ICD Classification Systems: Greek Perspectives and Initiatives [J]. Procedia - Social and Behavioral Sciences, 2013, 73(73):246-250.

Liu M. Optimal scheduling of logistical support for medical resources order and shipment in community health service centers[J]. Journal of Industrial Engineering & Management, 2015, 8(5):1362-1379.

Liu Y, Li Z, Liu J. A double standard model for allocating limited emergency medical service vehicle resources ensuring service reliability[J]. Transportation Research Part C, 2016, 69:120-133.

Mihea C, Kidong C. The Perceptions of Medical Staffs in terms of Human Resources, Training Programs, and Service Quality of Medical Tourism[J]. Journal of Tourism Sciences, 2011, 35:341-371.

Murty M S N, Kumar G S, Rao B V A. On Controllability of Fuzzy Dynamical Matrix

Lyapunov Systems[J]. Annals of West University of Timisoara - Mathematics, 2013, 51(2):73-87.

Tseng S F, Lee T S, Deng C Y. Cluster analysis of medical service resources at district hospitals in Taiwan, 2007-2011. [J]. Journal of the Chinese Medical Association, 2015, 78(12):732-45.

Yan Y H, Kung C M, Chen Y. The exploration of medical resources utilization among inguinal hernia repair in Taiwan diagnosis-related groups: [J]. Bmc Health Services Research, 2017, 17(1):708.

7 DISCLOSURE STATEMENT

NO potential conflict of interest was reported by the authors.

8 NOTES ON CONTRIBUTORS

Fanxiu Dong Received Bachelor's degree. Interest, medical record management.