

Application of Cluster Analysis for Business Processes in the Implementation of Integrated Economic and Management Systems

Korobova L.A.*

Department of Higher Mathematics and Information
Technology
Voronezh State University of Engineering Technologies
Voronezh, Russia
e-mail Lyudmila_korobova@mail.ru

Gladkikh T.V.

Department of Higher Mathematics and Information
Technology
Voronezh State University of Engineering Technologies
Voronezh, Russia
e-mail gtv1113@rambler.ru

Savvina E.A.

Department of Theory of Economics and Accounting
Policies
Voronezh State University of Engineering Technologies
Voronezh, Russia
e-mail katenok2207@yandex.ru

Lukina O.O.

Department of theory of Economics and Accounting
Policies
Voronezh State University of Engineering Technologies
Voronezh, Russia
e-mail oks.lukina@gmail.com

Kovaleva E.N.

Department of Higher Mathematics and Information
Technology
Voronezh State University of Engineering Technologies
Voronezh, Russia
e-mail kovaleva.lena@gmail.com

Tolstova I.S.

Department of Higher Mathematics and Information
Technology
Voronezh State University of Engineering Technologies
Voronezh, Russia
e-mail irin2102ka@mail.ru

Abstract — The use of logistic management methods at all levels of management arouses interest in world experience in providing economic recovery, solving crisis problems, and globalizing economic problems. Particular attention in logistics is paid to the principles of rational organization of the production process, i.e. organization of rhythmic, coordinated work of all parts of production in accordance with a single schedule and uniform release, ensuring maximum continuity of production processes, maximum reliability of planned calculations and the minimum complexity of planned works. To ensure the effective functioning of logistics companies, it is necessary to use modern management methods, improve coordination of processes in the logistics chains. It is also important to develop and implement advanced information technologies that support all key management processes of a logistics company and ensure its integration with partners in the process of goods distribution. The article considers the urgent task of developing a mathematical model of route planning that allows setting information about the movement of cargo with the definition of the shortest path. To improve the methods for solving the transport problem, we propose to use the cluster analysis method, which allows calculating the optimal route for the vehicle. Earlier, when finding a reference plan for production, one of the values was entered in the data table: tariff for the transportation of goods, transportation distance, cost, etc., which did not allow the calculation of the optimal route for delivering goods to a given point. The use of cluster analysis allows solving the problem of route planning in order to collect information about the movement of goods in online mode, make flight

schedules, and easily create reports and documents for enterprise logistics.

Keywords — *logistic management methods, cluster analysis, cluster, transport problem, enterprise logistics, linear programming, optimization problem, economic problems*

I. INTRODUCTION

Most enterprises objectively give priority to marketing and sales that are at the forefront of the business, but logistic losses often lead to financial losses [1]. Therefore, timely restoring of order in the logistics departments allows organizing the accuracy of accounting, thereby increasing the priority of the enterprise in the market of suppliers and buyers.

The main strategic objectives of logistics are: optimization of logistics costs; optimization of the route of the vehicle; reducing order processing time; reduction in the number of employees working in the logistics department; increasing the ability to control and monitor current indicators; improving the quality of services provided [2].

The majority of logistics specialists, when asked whether the company should implement a software system for automated control over the efficiency of using vehicles [3], will answer in the affirmative, since without software when planning the movement of even one machine, the number of

errors and miscalculations will reduce the benefit of the volumes of exported goods.

The enterprises with branches that are divided geographically need to clearly and competently build routes, both to customers and between their own warehouses. Today, suppliers and buyers are interested in information about the location of goods and the stages of its processing.

Quick and timely supply of raw materials (from both suppliers and own warehouses of the enterprise) provides the increase in the quality and speed of production of goods [4], as well as the attractiveness of the organization in the eyes of customers and, ultimately, ensures an increase of the profit. In this regard, industrial enterprises are faced with the need to strengthen the quality and speed of their own delivery services or logistics departments. The duties of logisticians are to calculate the optimal path, minimize transportation costs, conclude an agreement with suppliers and customers, which requires the greatest time costs [5]. Therefore, we have to solve the transport problem. At this moment, this is the main unsolved task, the solution of which is necessary for conducting a certain kind of business processes that arise during the implementation of integrated economic and management systems.

II. METHODOLOGY

The important problem is to develop methodological and instrumental support for the management of enterprise business processes based on the application of cluster analysis methods. The object of study is the subsystem "Logistics Management", which in real time monitors the process of delivery of goods and when a new order appears, analyzes the current location and workload of staff [6]. Based on this information, the system offers the most suitable contractor and makes changes to the route. To solve the transport problem, the clustering method is used.

Transportation problem belongs to the class of linear programming problems and therefore has all the qualities of linear optimization problems, but at the same time it has several additional useful properties that made it possible to develop special methods for its solution.

A mathematical model of the transportation problem is as follows:

$Z(x) = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \rightarrow \min$, the mathematical

description of the full inventory of transportation providers:

$\sum_{j=1}^n x_{ij} = a_i, i = 1, \dots, m$, mathematical description meet the

demands of consumers is: $\sum_{i=1}^m x_{ij} = b_j, j = 1, \dots, n$, the

condition of non-negativity of all variables, tasks: $x_{ij} \geq 0, i = 1, \dots, m, j = 1, \dots, n$.

Let us add a mathematical description of the total inventory of providers and consumers that show type challenges. If equality occurs, then we have a problem with

the right balance, and the model is closed, otherwise we have a problem with in-correct balance, and the model

$$\text{is } \sum_{i=1}^m a_i = \sum_{j=1}^n b_j.$$

Now we will consider several mathematical methods that solve the transport problem. Such tasks are typical problems of linear programming. The minimum cost method allows building a support solution that is close enough to the optimal one, since it uses the cost matrix of the transport problem. Like the northwest corner method, it consists of a series of steps of the same type, on each of which only one cell of the table corresponding to the minimum cost is filled in, and only one row (supplier) or one column (consumer) is excluded from consideration. The next cell is filled according to the same rules as in the northwest corner method. A supplier is excluded from consideration if its inventory runs out. The consumer is excluded from consideration if his requests are fully satisfied.

The stages of solving transport problems by the method of potentials are shown in Figure 1.

In order to solve our transportation problem, we decided to move away from the classical methods for solving it due to the huge number of parameters that needed to be considered [7]. And they turned to the methods of cluster analysis, which allows identifying the most significant factors that determine the movement of vehicles and simplify the formulation of the transport problem. Cluster analysis is a multidimensional statistical procedure that collects information about a sample of objects and then organizes the objects into relatively homogeneous groups. The clustering task relates to statistical processing, as well as to a wide class of training tasks without a teacher. K-means clustering is a type of training without a teacher, which is used when there is a sample of unmarked data [8].

The k-means algorithm works iteratively to assign each data point to one of the k groups. Data points are clustered based on the similarity of objects. Results for clustering k-means algorithm are provided below:

- we obtain centroids from k clusters that can be used to mark new data;
- we obtain labels for the training data.

Before considering the data, clustering allows finding and analyzing groups that have formed correctly. Each centroid of a cluster is a set of object values that define the resulting groups. The study of centroid weights can be used to qualitatively interpret which group each cluster represents. The clustering algorithm k-means includes:

- general business cases where we use k-means;
- the steps associated with the execution of the algorithm;
- an example of a method of using the data for delivery.

The clustering algorithm is used to find groups that were not explicitly labeled in the data. This can be used to confirm

business assumptions about what types of groups exist, or to identify unknown groups in complex data sets. After starting the algorithm and defining groups, any new data can be easily

assigned to the desired group. It is a versatile algorithm that can be used to group data of any type, is shown in Fig. 2.

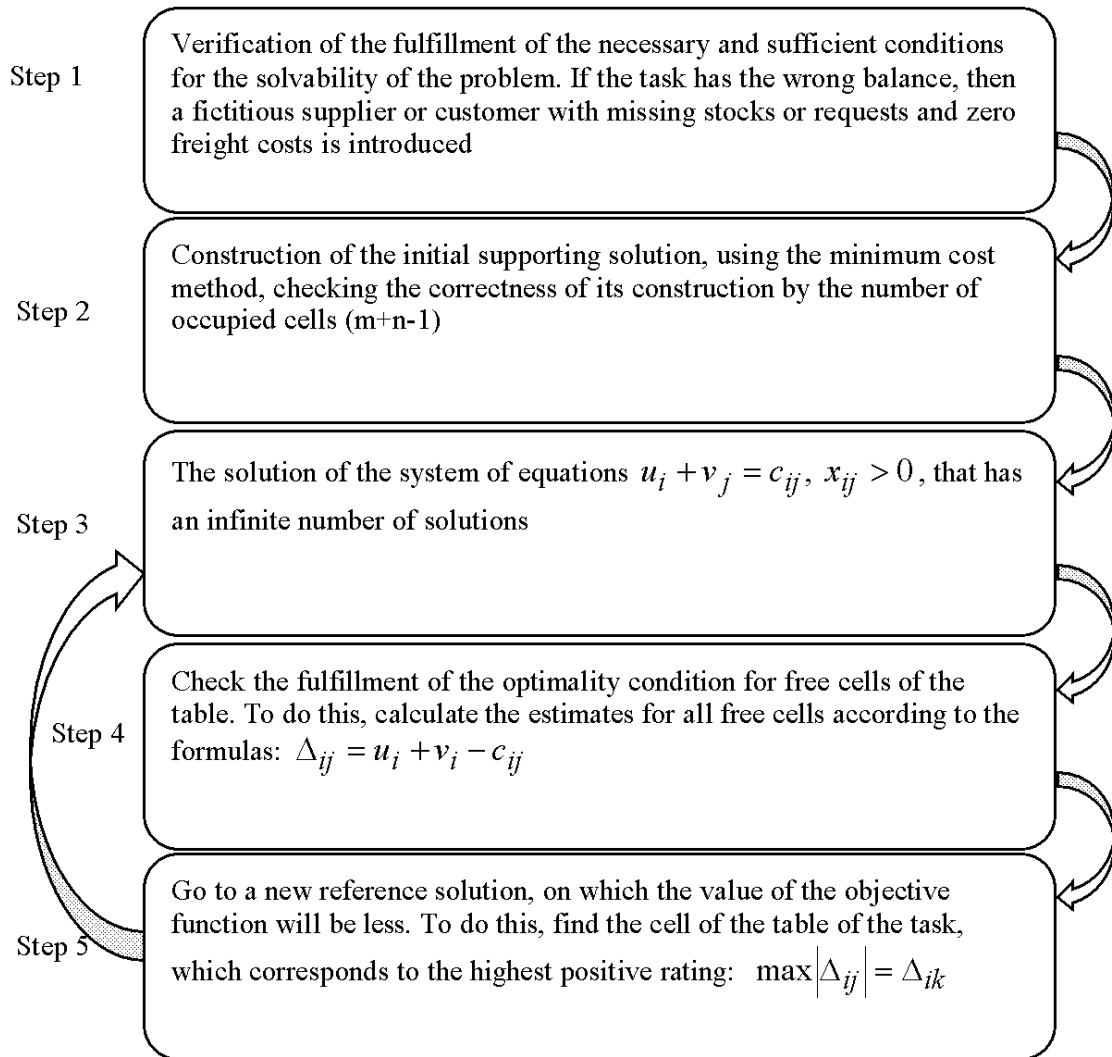


Fig. 1. The stages of solving transport problems by the method of potentials

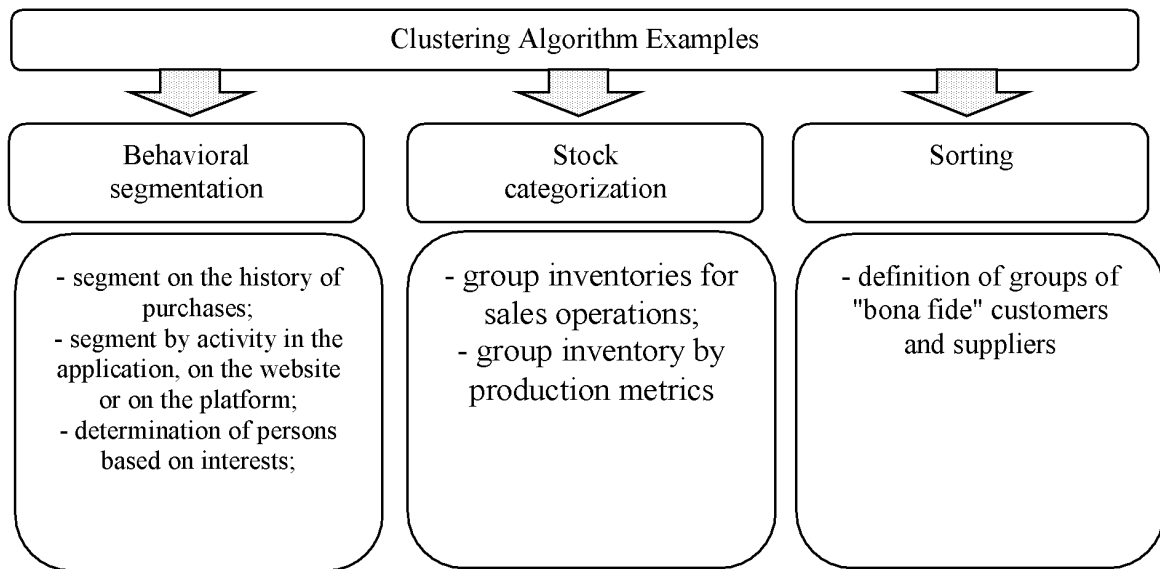


Fig. 2. Examples of using the clustering algorithm of k-averages

In addition, this algorithm supports the ability to monitor if the tracked data point switches between groups over time, and can also be used to detect significant changes in the data set.

In the clustering algorithm of k-means, we used iterative refinement to obtain the final result. For clustering the input

data were: number of clusters k and a dataset. The dataset consisted of a set of objects for each data point. The algorithm begins with initial estimates of k centroids, which are randomly chosen from the dataset. Then, the algorithm was divided into two steps, shown in Figure 3.

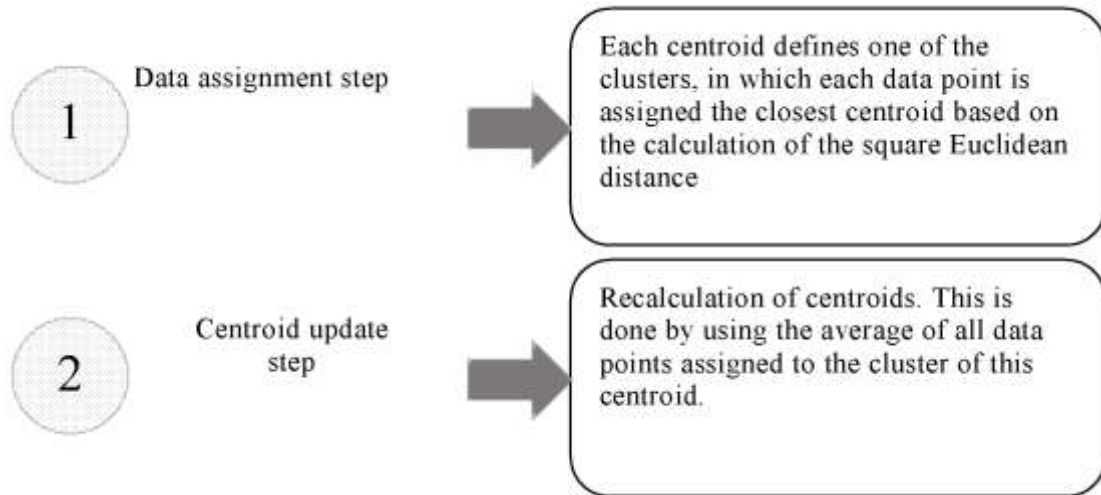


Fig. 3. The algorithm of cluster analysis method of k- means

III. RESULTS

The algorithm iterates between steps one and two until the stopping criteria are met (i.e., no data points change the clusters, the sum of the distances is minimized or a certain maximum number of iterations is reached).

This algorithm is guaranteed to converge to the result. The result may be a local optimum (i.e., not necessarily the best possible result), which means that evaluating more than one

run of the algorithm with randomized initial centroids can give a better result.

One of the indicators that we used to compare the results for different values of k - means, was the distance between the data points and their cluster centroids.

Let us consider the evolution of the solution to the transport problem and trace its movement from simple to complex. With the advent of two suppliers located in different regions of the country, the first difficulties in solving the

transportation problem immediately appear [9-10]. Here, calculation and comparison of several options is already required in order to prevent oncoming and excessively long transportation. The way out of this situation may be the application of economic and mathematical methods and electronic computer technology in this area.

Here is an example of using the cluster analysis method with a sample of data delivery of product information to the driver. Two driver functions are considered:

- average distance per day;
- average percentage of time during which the driver was at speed more than 60 km per hour.

The following diagram (Fig. 4) shows a data set for 4000 drivers with a spatial object the distance along the x-axis of the acceleration on the y axis.

The results of applying cluster analysis are presented in Fig. 5. The diagram (fig. 5) shows that the algorithm has identified four different groups. Drivers traveling at high speed have been separated from those who adhere to speed limits, in addition to rural and urban separation. The threshold for speeding is lower than for rural drivers, probably because city drivers spend more time at intersections and stops.

In order to create a route sheet, it is necessary to build a route for the vehicle. In turn, to build a route, it is necessary to

check the possibility of the capacity of the delivered goods in existing vehicles. If the product does not fit in the vehicle, it is necessary to divide the existing delivery points into groups or clusters. Each element of the cluster has a common feature by which grouping occurs – it is their proximity to each other. For clustering, the k-means clustering mechanism is used, which is built into the technology platform “1C: Enterprise 8.3”.

After clustering, two clusters are formed at the output, from which one is selected with the minimum total weight of the delivered goods. After choosing a cluster, the mass is formed, as well as the overall dimensions of the delivered goods and the available vehicles are sorted in case the capacity of all available goods is reached. Next, delivery points from the selected cluster must be built in the sequence in which the goods will be delivered [11–13]. For this, the mathematical method of the nearest neighbor is used. From the warehouse location point, the closest one from the cluster to the end of the points is searched. If the mass or dimensions of the delivered goods exceeds the capacity or carrying capacity of the vehicle, the points are added to the list until these product indicators do not exceed the parameters of the machine. When adding points the sum of the masses of the delivered goods and their dimensions are recorded in variables. After the formation of the delivery list, a route sheet is formed, in which detailed information on the delivered product, its quantity, and also the delivery address is entered.

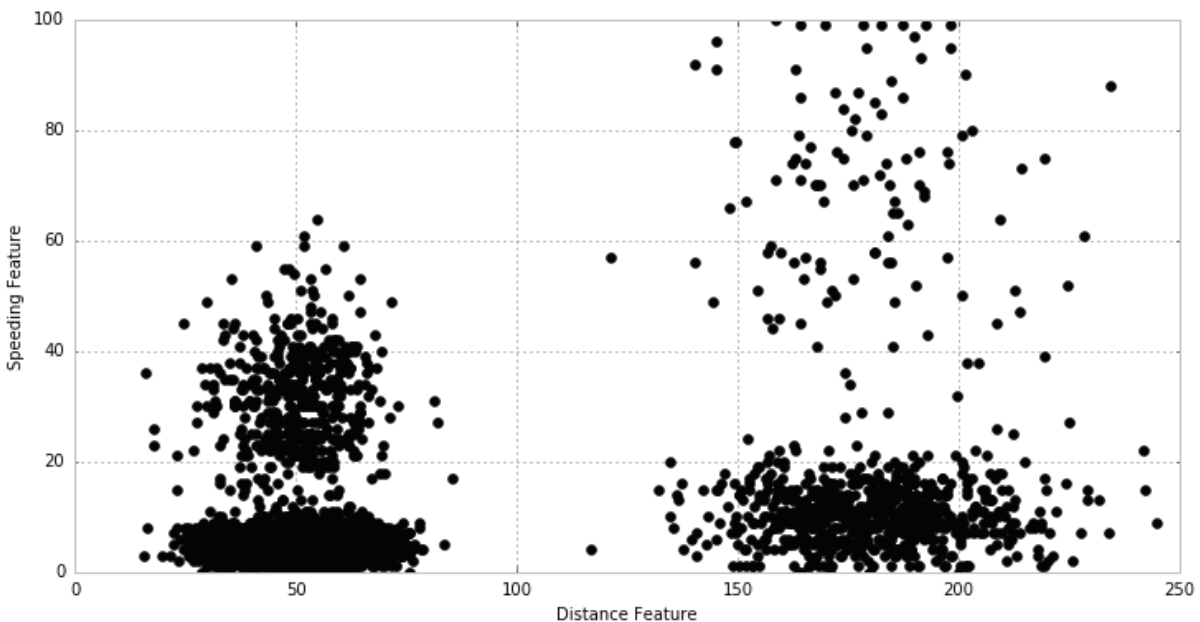


Fig. 4. Dataset diagram

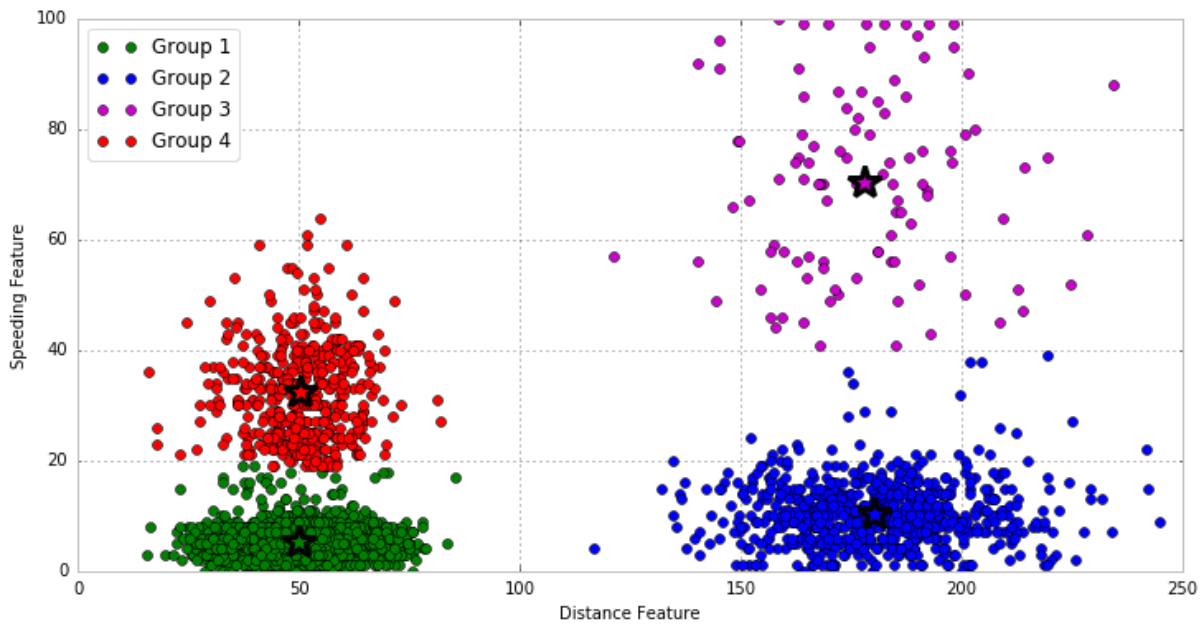


Fig. 5. Group cluster centroids

Thus, the clustering method can be used to solve the transport problem. Constructing actual routes and comparing them with planned routes closes a whole series of issues [14]. This gives the manager a guarantee of staff integrity and makes it possible to adjust the coefficients and initial conditions specified by the program for calculation. Perhaps the average loading and unloading time of the goods is longer than expected or some roads included in the route are closed. Then they must be excluded from the routes. As a result, we have convenient monitoring of work in the transport logistics of the enterprise due to the ability to compare the planned route and actual data [15]. The manager sees the actual route and its mileage. Information on lateness and deviations from

the route is displayed on the map. The introduction of stable and understandable business processes of the company, the provision of up-to-date data on the status of orders and resource loading, restoration of order in documents, as well as optimization of interaction between departments is a significant increase in work efficiency of the logistic company.

Thanks to the automatic accounting of these and other parameters, the constructed routes will be optimal. According to the experience of our clients, this saves up to 10% in transportation costs and allows reducing the cost of paying for driver processing. The algorithm for using the clustering method is presented in Figure 6.

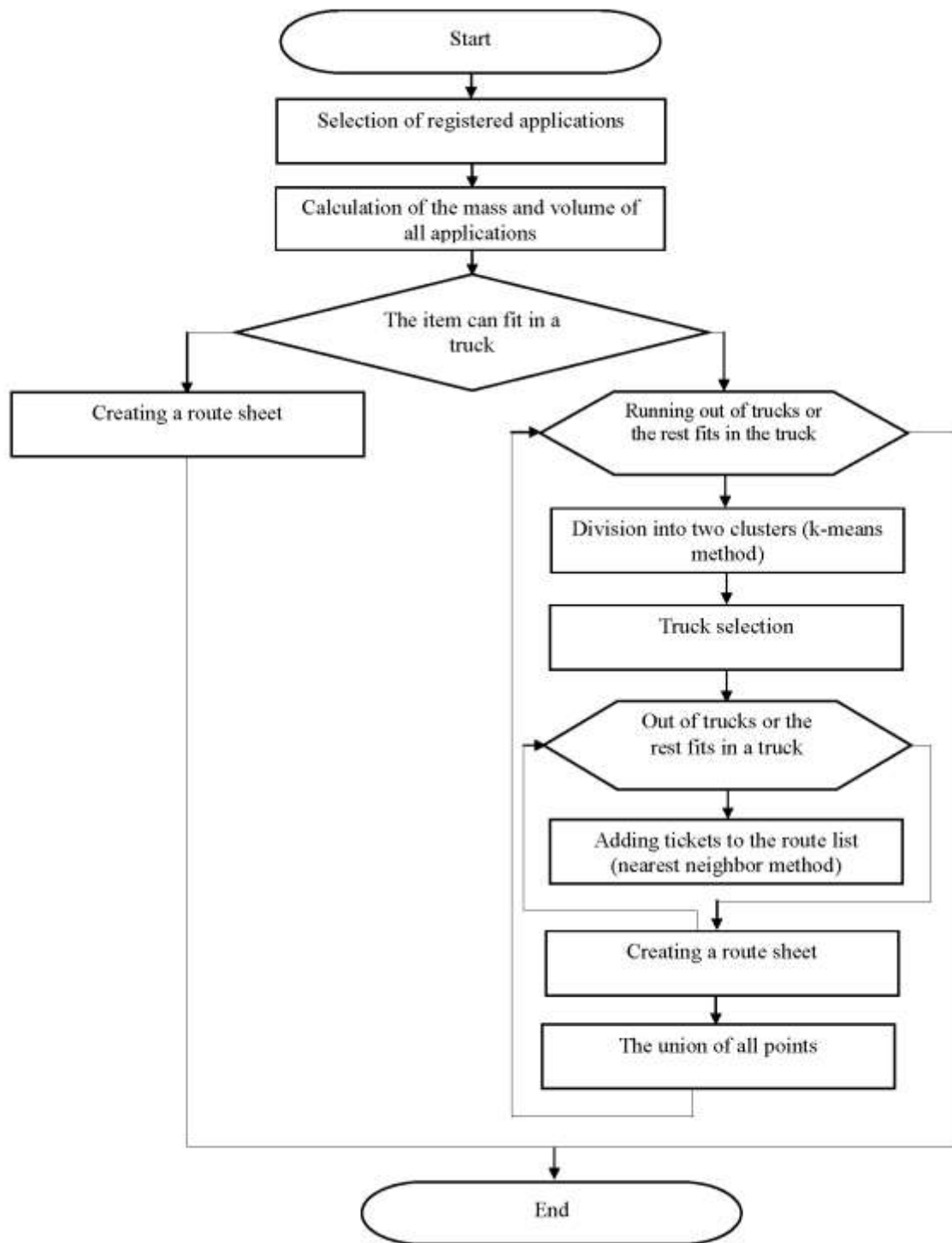


Fig. 6. Algorithm for the process of filling out an application for delivery using cluster analysis

IV. CONCLUSION

With the introduction of an automated control system for the transportation of goods, it is possible at any time to get the current picture of order fulfillment to customers, so the development of the system allows logisticians to take into account and optimally manage the delivery of goods to the

production enterprise. The application of the cluster analysis allows planning transportation and detailed analytics inherent to systems of this level. The identification of deviations from the established standards will help eliminate the problem areas of cargo transportation. The information flow directly controls both the financial flow and the physical movement of goods. Information is an integral and decisive factor in the

effectiveness of “Logistics Management” and the tasks of increasing the transparency of logistics activities, its flexibility and efficiency are becoming a priority for business [16].

Business opportunities in the implementation of integrated information and management systems are presented in Fig. 7.

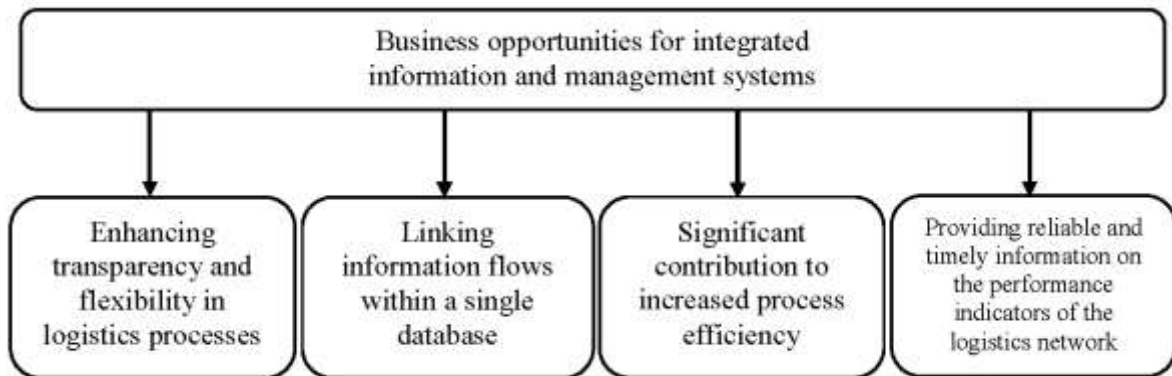


Fig. 7. Business opportunities in the implementation of integrated information and management systems

The “Logistics Management” subsystem distributes the work so that the performers arrive at their destinations in delivery windows convenient for the recipient. The system distributes the work and calculates the time of arrival, taking into account many factors: the requirements for the vehicle and its type, information about historical traffic jams, features and operating conditions of each driver.

The “Logistics Management” subsystem monitors the delivery process in real time and when a new order appears, analyzes the current location and workload of the staff. Based on this information, the system offers the most suitable contractor and makes changes to the route of movement.

References

[1] V.Ya. Vilisov, “Analysis of the transport model with approximation of the preferences of decision makers”, *Comput.: other Appl. inform. Sci. articles*, vol. 2, pp. 43–56, 2012.

[2] E. Kolesnichenko, A. Khorev, O. Rudakova, N.N. Kharchenko, “Conflict management: practical experience of Russian enterprises”, pp. 3454–3461, 2019 [Proc. of the 33rd Int. Business Information Management Association Conference, IBIMA 2019: Education Excellence and Innovation Management through Vision 2020].

[3] A.I. Khorev, K.V. Chekudaev, “Possibilities of an information logistic concept for enterprise management”, *FES: Finance. Econ.*, vol. 2, pp. 20–21, 2009.

[4] S.V. Kosnikov, *Mathematical methods in economics*. Moscow: Science, 2017.

[5] V.N. Kraev, “Methods of making managerial decisions”, *Manag., staff recruitment*, vol. 2, pp. 43–56, 2014.

[6] E.S. Bestembek, A.F. Sekerbaeva, “Tasks of research and development of a multifactor model of transport and storage solutions”, *Actual probl. of transport and energy: ways of their innovat. solut.*, pp. 156–158, 2016.

[7] S.V. Bukharin, A.V. Melnikov, S.N. Chernyaeva, L.A. Korobova, “The method of immersion the problem of comparing technical objects in an expert shell in the class of artificial intelligence algorithms”, vol. 225, p. 012208, July 2017 [IOP Conf. Ser. Materials Science and Engineering International Conference on Materials, 2017].

[8] R.M. Grant, *Contemporary strategy analysis*. Oxford: Blackwell Publ., 2002.

[9] R. Ireland, M. Hitt, “Achieving and maintaining strategic competitiveness in the 21st century: The role of strategic leadership”, *Acad. of Manag. Executive*, vol. 4, no. 19, pp. 63–77, 2005.

[10] B. Snowdon, G. Stonehouse, “Competitiveness in globalized world: Michael Porter on the Microeconomic Foundations of the Competitiveness of Nations, Regions and Firms”, *J. of Int. Busin. Studies*, vol. 2, no. 37, pp. 163–175, 2006.

[11] P. Dey, N. Petridis, K. Petridis, C. Malesios, “Environmental management and corporate social responsibility practices of small and medium-sized enterprises”, *J. of cleaner product.*, vol. 195, pp. 687–702, 2018.

[12] D.G. Collings, K. Mellahi, W.F. Cascio, “Global talent management and performance in multinational enterprises: A multilevel perspective”, *J. of Manag.*, vol. 45, no. 2, pp. 540–566, 2019.

[13] T. Hilorme, M. Chorna, L. Karpenko, M. Milyavskiy, “Innovative model of enterprises personnel incentives”, *Evaluat. Acad. of Strategic Manag. J.*, vol. 17, no. 3, pp. 1–6, 2018.

[14] S.G. Tikhomirov, I.A. Avtsinov, Y.A. Turovsky, A.S. Surovtsev, A.A. Adamenko, E.N. Kovaleva, “Hardware-software complex for managing biotechnological systems using intelligent information technologies”, *Proc. of Voronezh State Univer. Ser. System Anal. and Inform. Technol.*, vol. 3, pp. 158–165, 2019.

[15] I.V. Platonova, E.V. Gorkovenko, K.V. Chekudaev, “The system of indicators for assessing the innovative mechanism of an enterprise”, *Econ. and Entrepreneurship*, vol. 1, p. 360, 2013.

[16] G.V. Belyaeva, R.V. Nuzhdin, T.D. Artemenko, “Information-instrumental support of business analysis procedures”, *Econ. Right. Innovat.*, vol. 4, pp. 55–60, 2019.