Software selection based on analysis and forecasting methods, practised in 1C

A N Vazhdaev[1, a], T YChernysheva[1, b], E I Lisacheva[1, c]

[1, 2, 3] 26, Leningradskaya street, Yurga, Kemerovo region, 652055, Russia
Yurga Technological Institute (affiliated) National Research Tomsk Polytechnic University

e-mail: [a]wazdaev@ngs.ru, [b]chernyshevat@tpu.ru, [c]djlena.91@mail.ru

Abstract. The research focuses on the problem of a "1C: Enterprise 8" platform inboard mechanisms for data analysis and forecasting. It is important to evaluate and select proper software to develop effective strategies for customer relationship management in terms of sales, as well as implementation and further maintenance of software. Research data allows creating new forecast models to schedule further software distribution.

1. Introduction

One of the current requirements imposed to organizations is the utilization of information systems and technologies. Many organizations place confidence in businesses that specialize in software supply and maintenance to select proper software. In fact, the process of evaluating and selecting proper software is a complex and multistep process. Mathematical evaluation is performed by both statistic and expert methods [1]. When organizations use inboard mechanisms for data analysis and forecasting while working in authors information system of product programs evaluation and selection, they can develop an effective strategy for customer relationship management in terms of sales [2], as well as implementation and further maintenance of software [3].

2. "1C: Enterprise 8" platform inboard mechanisms

Mechanism for data analysis and forecasting is one of the "1C: Enterprise 8" platform options that help to develop commercial and analytical reports.
This mechanism allows [4]:
- to find consistent patterns of the information database input data;
- to control evaluation parameters by means of software-base method and interactively;
- to implement of programmic access to the evaluation results;
- to convert evaluated data into tabular style automatically;
- to create forecast models that automatically forecast subsequent events or some new objects values.

Initial data for evaluation can be obtained from information databases as well as from external sources. Recommendations, developed due to evaluation, make further operations more efficient.

Evaluated data is a sample of data behavior which can be displayed in a final document, or saved for future utilization. Further utilization of evaluated data allows creating a trend model to forecast new datum behavior in accordance with the existing model [5].

3. Different types of platform "1C: Enterprise 8"data analysis

Platform "1C: Enterprise 8" mechanism of data analysis and forecasting includes several types of data evaluation [4]:

1) General Statistics. This mechanism serves to gather information about data selected to be analyzed. This type of evaluation is intended for preliminary investigation of evaluated data source.

Evaluation shows a number of characteristics of discrete and continuous fields. When the report is converted into tabular style, pie charts are filled to display fields' structure. Figure 1 shows the use of this type of data evaluation in the information system developed by the authors of the paper.

Datum in a source can be of continuous or discrete types. Numbers and dates are continuous, while other types refer to discrete one. Each type contains its own specific information [4].

Discrete datum:

• Number of values. Number of values in a data source field;

• Number of unique values (with the exception of duplicate values);

• Mode (the most frequent value in a data source). If there is more than one value with equal frequency, the first one is to be taken as a mode;

• Frequency (depends on the number of times the value occurs in a data selection list);

• Relative frequency (is the ratio of a number of volume occurrence to the total number of values);

• Cumulative frequency (is a sum of values frequency and frequency amount of previous values in a data selection list);

• Cumulative relative frequency (is a sum of a cumulative frequency of a value and a sum of relative frequencies of previous values).

Continuous data:

• Number of values;

• Minimal value;

• Maximal value;

• Average;

• Swipe. Difference between maximal and minimal values;

• Standard deviation (meansquare deviation);

• Median value (value that lies in the middle of a data selection list).

It should be noted that if several fields of different types are analyzed simultaneously, evaluation is to be done independently so to avoid intercorrelation.

2) Search for associations. This type of evaluation finds groups of objects that often go together or property value, and searches for association rules [4]. Search for associations can be used, for example, to determine items and services frequently purchased together. Figure 2 demonstrates search for associations in an information system. Evaluation helps to obtain:

• data information (number of objects and elements, average number of elements in an object, number of groups and association rules found);

• found groups elements. Group items are specified as well as an amount of cases and percentage of cases in which the group presents;

• found association rules. Indicates initial composition of elements, consequence (elements composition) and percentage of cases, credibility and relevance rules.
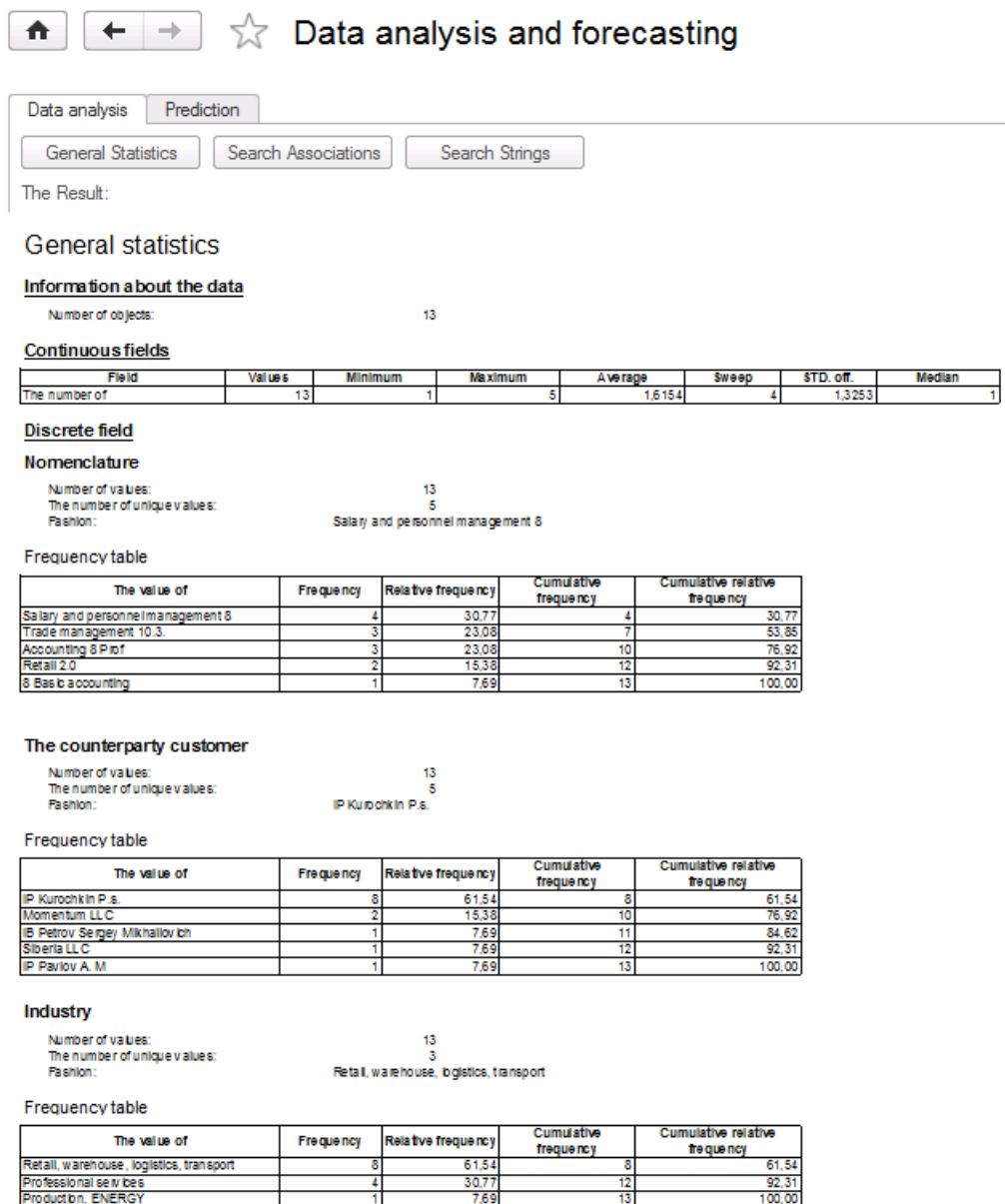
Figure 1. General Statistics type of evaluation.

3) Search for sequences. This type of evaluation helps to find successive chains of events in a da-ta source [4]. This type of evaluation allows to search through the hierarchy, which allows monitoring not only a specific sequence of events, but also a sequence of parent groups. Figure 3 shows the use of search sequences in an information system.

The main result of this type of evaluation is templates sequences. These templates contain following information [8]:

• composition of template sequence;
• number of cases with this sequence;
• maximum intervals between events (two events - interval a);
• minimum intervals between events (two events - interval a);
• percentage of cases in which the sequence was executed;
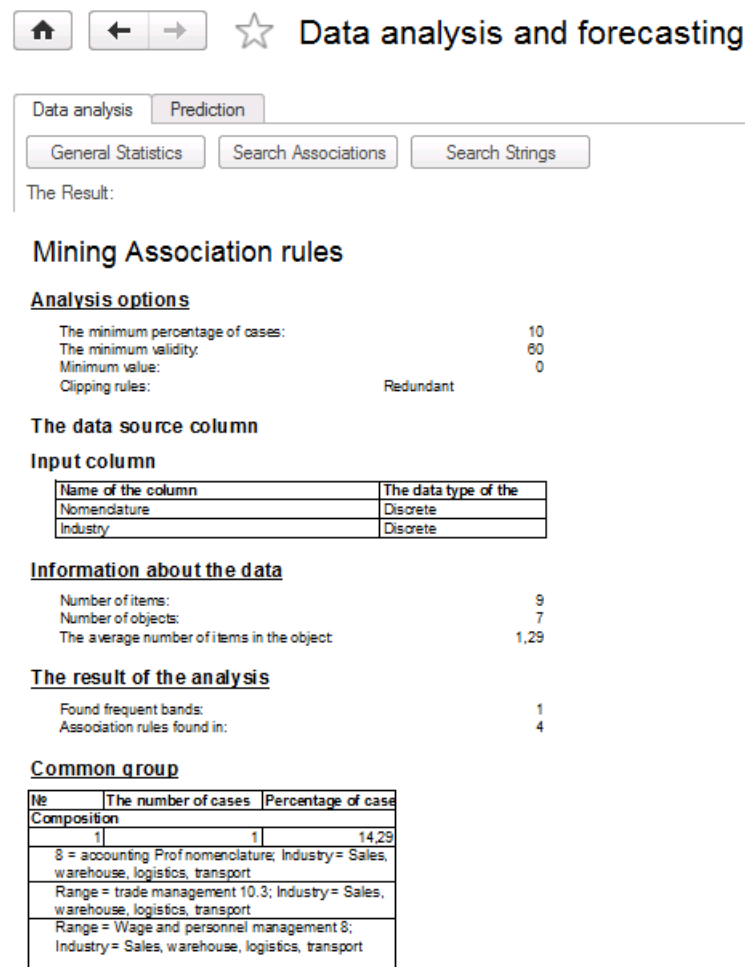• average intervals between events (two events - interval a).

Figure 2. Search for associations type of evaluation.

4) Cluster analysis. Clustering helps to select a number of relatively homogeneous groups, which are segments or clusters, among objects of the same nature [5]. The analysis is based on calculating distance between groups, which are clusters. It is possible to calculate distance with different methods (based on different metrics). Platform "1C: Enterprise 8" supports following metrics:

• Euclidean metric. Distance between two objects is calculated by the equation:

$$S = \sqrt{\sum (W_i * (X_i - Y_i)^2)} \qquad , \qquad (1)$$

where

X_i, Y_i –two objects attribute values (distance between these objects is calculated);

W_i– weight number of an attribute;

i – attribute number from 1 to n;
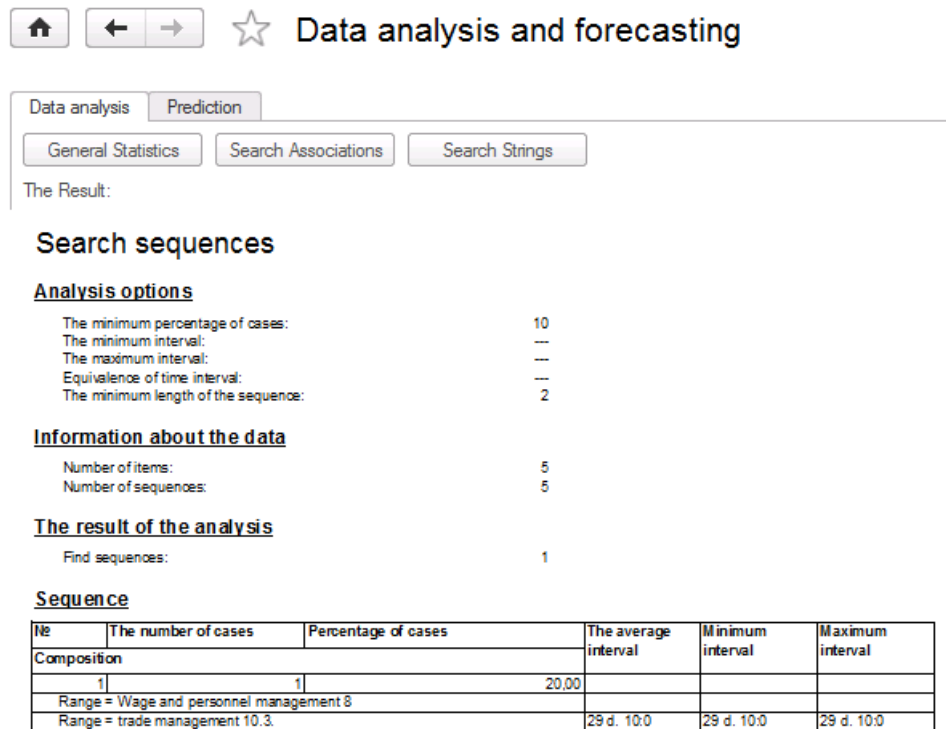
n – number of attributes.

Figure 3. Presentation of a Search for sequences evaluation method results.

• The Euclidean metric squared. Distance between two objects is calculated by equations:

$$S = \sum (W_i * (X_i - Y_i)^2)$$
(2)

where

X_i, Y_i – two objects attribute values (distance between these objects is calculated);

W_i – weight number of an attribute;

i – attribute number from 1 to n;
n – number of attributes.
•   City metric. Distance between two objects is calculated by equations:

$$S = \sum (W_i * |X_i - Y_i|)$$
(3)

where

X_i, Y_i – two objects attribute values (distance between these objects is calculated);

W_i – weight number of an attribute;

i – attribute number from 1 to n;
n – number of attributes.
•   Dominance metric. This distance metric is calculated between two objects according to equations:

$$S = MAX(W_i * |X_i - Y_i|)$$
(4)

where

X_i, Y_i – two objects attribute values (distance between these objects is calculated);

W_i – weight number of an attribute;

i – attribute number from 1 to n;

n – number of attributes.

When distance between objects is calculated, several algorithms that allow ranging objects into clusters can be opted.Clusterization method is chosen on the base of principals through which an object refers to this or that group and algorithms of clusterization [6]. Any clusterization algorithm aims at minimizing variability within clusters, and at maximizing variability between clusters. Following clusterization methods are supported:

• Short-range communication – object: we tap to a group which is at a minimal distance to the nearest object;

• Long distance – object: we tap to a group which is at a minimal distance to the furtherest object;

• Middle-end: we opt for objects which are at the top of a data selection list. They are considered as cluster centers. Then we opt for another object in accordance with a distance to the cluster centers. Cluster center, to which an object was added, is to be recalculated. The procedure is repeated until the complete enumeration of all objects. Then we start a new object selection (starting with the first one). The procedure is repeated as long as cluster centers vary;

• Center of gravity: we tap objects to a group with a minimal distance from the center of gravity.

Currently, the authors of cluster analysis implemented in its information system [7].

5) Decision tree. This method of evaluation allows developing a hierarchical structure of classification rules in a form of a tree. To make a decision tree, it is necessary to choose a target attribute, according to which it will be possible to make a classifier, and a number of input attributes that will be used to develop rules.

The research product is a decision tree each joint of which contains a condition or classification error that reveal conditions at which developed rules contradict to reality. To classify a new object, it is necessary to answer questions in each joint and pass all the chain from the root up to the leaves of the tree, going to daughter joints if the answer is positive or to neighbor joints if the answer is negative. An evaluation criteria set helps to regulate accuracy of the tree [4,5].

It is possible to make the tree less complicated. Some rules allow converting tree joints into leaves, excluding excessive branching [9].

During the process of converting a joint into a leaf it is necessary to take into account:

$$O = \sqrt{\frac{\left(b + \frac{k}{2}\right) * \left(s - \left(b + \frac{k}{2}\right)\right)}{s}} \tag{5}$$

where

a - number of errors in a joint, b - number of errors in daughter joints, k - number of leaves in a joint, s - number of cases.

We decide to convert a joint into a leaf if following conditions are fulfilled:

$$a + 0{,}5 \prec= b + \frac{k}{2} + O \tag{6}$$

4. Results and Discussion
The task of forecasting on a "1C: Enterprise 8" platform is generally divided into two basic procedures [10]:

1. Training the model based on a definite data selection list.

2. Utilization of a trained model in the process of actual data forecast.

Results of a S*earch for Associations* analysis help to develop a model of association forecast automatically. It means that a buying organization, based on other cases analogy, will more likely purchase a recommended item alongside with those software products it bought before. Sequence forecast model is developed through the evaluation of a Search for associations analysis.

5. Conclusion

Datum obtained from a database and external sources with their further evaluation by means of different types of analysis available on a "1C: Enterprise 8" platform allows developing specific recommendation for further actions. Evaluation helps to develop a trend model, which allows businesses to forecast new data behavior in accordance with a model which is already in operation.

Using the described 1C: "Enterprise 8" platform mechanisms of analysis and forecasting data during the process of choosing proper software for franchisee firms′ clients, allows businesses to find patterns in sale programs and develop forecast models able to plan further sales of software automatically.

References
[1]   Vollmann T, Berry W, Whybark D C, Jacobs F R 2004 *Manufacturing planning and control systems for supply chain management*: *The Definitive Guide for Professionals (5th ed.)* (McGraw-Hill)
[2]   Saloner G, Shepard A, Podolny J 2001 *Strategic Managemen* (Wiley)
[3]   Toporkov V V, Yemelyanov D M 2014 *Economic model of scheduling and fair resource sharing in distributed computations* Programming and computer software (USA: Pleiades Publishing) pp 35-42
[4]   Kataev M Y, Emeljyanenko V A, Ponomareva M G 2014 *The model of strategic planning at the enterprise with process management approach* Applied Mechanics and Materials (Switzerland: Elsevier) pp 572-578
[5]   Kataev M Y, Bulysheva L A, Emeljyanenko A A, Emeljyanenko V A 2013 *Enterprise systems in Russia: 1992–2012* Enterprise Information Systems (Great Britain: Elsevier) pp 169-186
[6]   Telipenko E V, Zakharova A A 2014 *Bankruptcy risk management of a machine builder* Applied Mechanics and Materials (Switzerland: Elsevier) pp 617-622
[7]   Chernysheva T Y, Gnedash E V, Zorina T Y, Lenskaya N V 2014 *Information systems project risk assessment: expert approach* Applied Mechanics and Materials (Switzerland: Elsevier) pp 539-543
[8]   Chernysheva T Y 2013 *Preliminary risk assessment in IT projects* Applied Mechanics and Materials (Switzerland: Elsevier) pp 220-223
[9]   Chernysheva T Y, Zhukov A G 2013 *Hierarchical model evaluation and selection of software prod-ucts for the enterprise* Polzunovsky Gazette (Krasnoyarsk: AltSTU) pp 25-28
[10]   Lisacheva E I, Vazhdaev A N 2013 *Decision Support System franchisee organization for the selection of software for buyers* Polzunovsky Gazette (Krasnoyarsk: AltSTU) pp 224-228