

EVALUATING THE QUALITY OF MODELING THE SCENARIO OF INFORMATION ANALYSIS ON A BRANCHED NETWORK

Summary: proposed solution of the problem of assessing the quality of scenario modeling based on the analysis of ontologies of the domain description using the structure representation in the form of a graph, which makes it possible to evaluate its properties using the corresponding metric. This assessment of possible reduction of cost to build a collection of streaming media on the branched network. Shown urgency of solving such problems for the analysis of the budget process on the branched network. A modified search algorithm wide construction of optimal scenarios based on analysis of ontologies describe the subject area.

Keywords: quality assessment, modeling, information gathering, ontology graph analysis, budget analysis, branched network.

Introduction

Today many methods to select the appropriate quality and efficiency of the plurality of possible scenarios or acceptable scenarios based on analysis of the nature and the factors that affect the planning scenario. In practice, the planning situation stands, different number of factors on which the decision and determine the fundamental differences between the procedures necessary formation scenarios. One of the very complex and pressing problems today that require building effective scenarios for its decision, the task of building and further optimization scenarios to analyze the budget process on an branched network. The problem of optimization scenarios collect information on streaming branched network is one of the most important tasks of handling large amounts of data [1].

The most important tool to address the budget process is the financial analysis of regional budgets. In the conditions of crisis of the economy and regions of the country, strengthening the role of regions and regional budgets in solving economic and social problems, problems of stability of regional budgets are very relevant.

Financial analysis allows the state budget to sum up over time, to determine the effectiveness of budget spending, predict the state budget for the next quarter or year and make decisions that will improve the financial condition of the state budget. One measure of financial condition is considered to be the level of financial stability budget. Resistance budget allows conclusions about the possibility of the normal functioning of public power and the strength of its financial base of operations.

The main result of the analysis of financial stability is to get the budget estimates of stability analysis for each subject over time and opinion on the state budget stability for this subject. Based on these results, you can make general conclusions about the financial condition of the state budget, to make predictions about the state forward and decide on actions to improve the economic situation to achieve stability and independence of the state budget.

Software development for financial analysis of the state budget and determine the stability of the budget can significantly reduce the time of analysis and facilitates the work of state institutions and authorities [2].

The overall structure of the network, determines the ontological analysis features for solving the problem of the budget process on an branched network rather complicated structure and a large number of elements for which it includes.

One of the most common solutions to such problems is characterized in that at build script based on a particular ontology that describes the space provided in determining benefits when evaluating the quality and effectiveness of scenarios considered only one factor. However, other factors are not taken into account. Another common approach is determined that the required scenario is determined taking into account two or more factors may be taken into account in varying degrees.

Determining the required quality and performance scenarios for situations where the planning is based on one factor may be using these strategies, using hierarchical graph describing the relevant ontology.

You can select the following basic quality assessment scenarios:

- on the vector of priorities, according to the hierarchy of the ontology,
- assessment of income and expenses from the sale of each scenario,
- subject to peer review,
- the results of the impact of implementation scenarios.

Determining the necessary quality and effectiveness of scenarios in accordance with the second situation where planning controls a number of factors, a number of features that require the selection of two different planning procedures in hierarchical systems of several factors.

An important element of consideration of the impact of a number of factors to determine the necessary quality and effectiveness of scenarios is a method taking into account the benefits of the effects of certain factors on the value assessment scenario. The choice can be built as one of the strengths of the factors and on the basis of the trade-offs between several or all factors

Automation solve such problems needs to build algorithms that can enable unambiguous choice scenario based on a hierarchical structure describing ontology data on which to construct scenarios.

The purpose is to develop models and algorithms determine on the basis of the constructed ontology defined some real hierarchical storage network and the formation of the current data. This problem concerns the choice of the most efficient scenarios by identifying and using the most appropriate scenarios to the terms of a specific request.

Constructing scenarios based on ontology network storage

There are many models for solving the problem of constructing scenarios based on ontologies describe the subject area, which are based on various mathematical methods. Among them are fairly common model for constructing scenarios based on the structure data storage on an branched network, which is typical analysis problems of regional budgets. Based on these models and their various modifications built modern information-analytical and information retrieval systems [3].

This notion of ontology is considered as follows and defined as:

$$O = \{T, A, R, D\}, \quad (1)$$

T - set of terms that denote objects and concepts of a domain,

A - set of attributes of concepts,

R - set of relationships (links) between terms

D - set of concepts and relationships.

In ontology graphically represented as a network, which peaks labeled domain concepts, and edges indicate the connection between them. The base is such hierarchical relationships of the "class - the class" and "part - whole" that define the structure of the branched structure of the network storage. Thus, ontology is a description of the structure of the subject area, providing insight together concepts and relationships between them.

To describe ontologies for tasks to select the appropriate quality and efficiency of the plurality of possible scenarios or acceptable scenarios widely used methods for describing and mapping are based on graph theory. The most widely descriptions in a hierarchical graphs tend to estimate both the weight and the number included in the scenario description edges of the graph.

There are many methods to assess the number of ontology, but there is no comprehensive approach to this issue. These methods examine some specific aspects, but most do so incompletely. An important factor for assessing ontology is perspective from which it is viewed. It depends on which domain it is applied and that is a priority for the user. It is therefore necessary to determine which of the methods is most appropriate for a particular case.

Today has models which were to quantitative and qualitative advantages of one over the other. Therefore, the development of new algorithms and methods based on the combination of the advantages of different approaches in building models is very relevant and needs to develop new solutions.

Evaluation of scenario models based on network storage ontology

The structural approach to the ontology is most appropriate to the task, because the presentation as a graph structure makes it possible to measure its properties using the metrics by which you can determine its quality and form recommendations for its improvement.

An important aspect of the structural approach is the need to assess the effectiveness of building ontological model-based estimation cost reduction finder graph describing ontology, compared with non-directional search path, which corresponds to a script that is based on ontology described as a graph.

To estimate the reduction of search costs by the graph describing the ontology, one can conduct an assessment of the effectiveness based on assumptions, which is quite simple and can accurately assess the value selected for comparison characteristics. These estimates can be built on the difference between full and partial passage through the column, which is a graphical representation of ontology analyzed in describing the scenario.

The basis for such procedures can be obtained before passing on information about the count in the implementation of other requests that overlap with current, that the information obtained in the previous passage through the column. This information is stored in the appropriate knowledge base for each ontology, and hence for a the corresponding graph. In addition, for the measurement of the efficiency of the approach as a whole, reducing the degree of generality, we can consider some typical types of graphs for describing ontologies. This makes it possible to determine the boundary for efficiency assessment approach in general.

In general, this situation can be seen that search query data for ontology that consists of n levels already known based on the analysis of previous requests stored in the knowledge base, the results of the passage (ie, partial search) for m levels where

$$m < n. \quad (2)$$

Then the effect of using previously obtained information for this case will be

$$E_{min} = 2^n / (2^n - 2^{n-m}) \quad (3)$$

Thus the number of inspections during the passage of the count be reduced to E again.

As shown in the Miller ontology evaluation suggest that the number of ties in one concept in fully connected graph describing ontology should not exceed 9[4]. Thus, we can assume that in most cases the actual total number of incoming and outgoing edges do not exceed 9.

Based on this assumption the maximum benefit from the use of previously received information possible way to construct a scenario for this case can reach

$$E_{max} = 9^n / (9^n - 9^{n-m}) \quad (4)$$

To assess the effectiveness of building ontology model can be used methods are based on finding the shortest path in the graph that describes it.

In general, we can write that implementation choices to count the ways to achieve the goal T according to the request:

$$T = \{x / A(x)\}, \quad (5)$$

x - where all possible avenues on the graph.

$A(x)$ - the characteristic feature that captures the essence of a specific request and will lead to results.

Then

$$\forall a \exists x \forall c (c \in x \Leftrightarrow c \leq a) , \quad (6)$$

c - all paths in the graph leading to the goal,

a - the minimum distance on the column corresponding to the target.

Thus, there is always a minimum distance on the column corresponding to the goals that demand, then there is always

$$T_{min} = \min (a). \quad (7)$$

Problem finding the shortest path in the graph (Single Source Shortest Path - SSSP) [5] are formed based on common characteristics graph model describing ontology.

Let a given graph with weights arcs and dedicated source of top – u . The sequence of arcs called by that goes from the top u to the top v .

It is necessary for each vertex v , accessible from the top source u , specify the path that has the smallest possible total weight: $P^*(u, v)$.

$$f^*(v) = f(P^*(u, v)) = \min f(P(u, v)). \quad (8)$$

Objectives can be set for both targeted and for an undirected graph. This kind of problem can apply in both cases. The problem satisfies the principle of optimality if the path is part of the shortest path, that is the shortest path from the source to the top u w. The principle of optimality means that for each vertex v shortest path instead of just enough to keep it past the arc.

Consider the weighted graph with n vertices and m arcs with weights. For it is necessary to construct the shortest path from vertex u (initial) to top v (final).

The shortest path from u to v can be restored if the time starting from the top v , place ribs in the opposite direction until, until the top has u . Ribs can be restored bust curves for each vertex:

$$e_v^* \in \{(u, v) \in E \mid f^*(v) = f^*(u) + f(u, v)\} \quad (9)$$

To find the shortest distance (v) on a set of curves, it is necessary to build a tree on which to search with the ability to parallelize. The most common processing tasks to streaming information on the network is branched breadth-first search algorithm [6,7].

This algorithm can be used to solve the problem of finding the shortest path for all pairs of vertices. For this algorithm should run for each vertex.

For optimization problem scenario analysis of the budget process on an branched network based ontologies describe the structure of storing data in a hierarchical graph (Fig. 1) search algorithm in width allows us to calculate the shortest distance from the selected vertex directed graph all the other nodes. In addition, the breadth-first search can solve the problem of checking range, ie, whether there are ways between the top of the source and the other vertices of the graph. The algorithm is based on the vertices traversing the levels of hierarchy. The idea of breadth-first search algorithm allows to develop different versions of the algorithm according to the specific tasks to be solved with desired properties and actions, that it should include.

To solve the problem of optimization scenarios analysis of the budget process on an branched network based ontologies describe the structure of storing data in a hierarchical graph, offered the following modified breadth-first search algorithm that significantly reduces the time to find ways of building scenarios collect information on streaming branched network.

Let a given graph without weights with a dedicated source of top u . In this way $P(u, v)$ between vertices u and v is the set of arcs

$$(U, v_1), (v_1, v_2), \dots, (v_{n-1}, v) \tag{10}$$

Path length $d(u, v)$ denote the number of arcs in this way between vertices u and v . Breadth-first search finds the shortest path from vertex u to all other vertices.

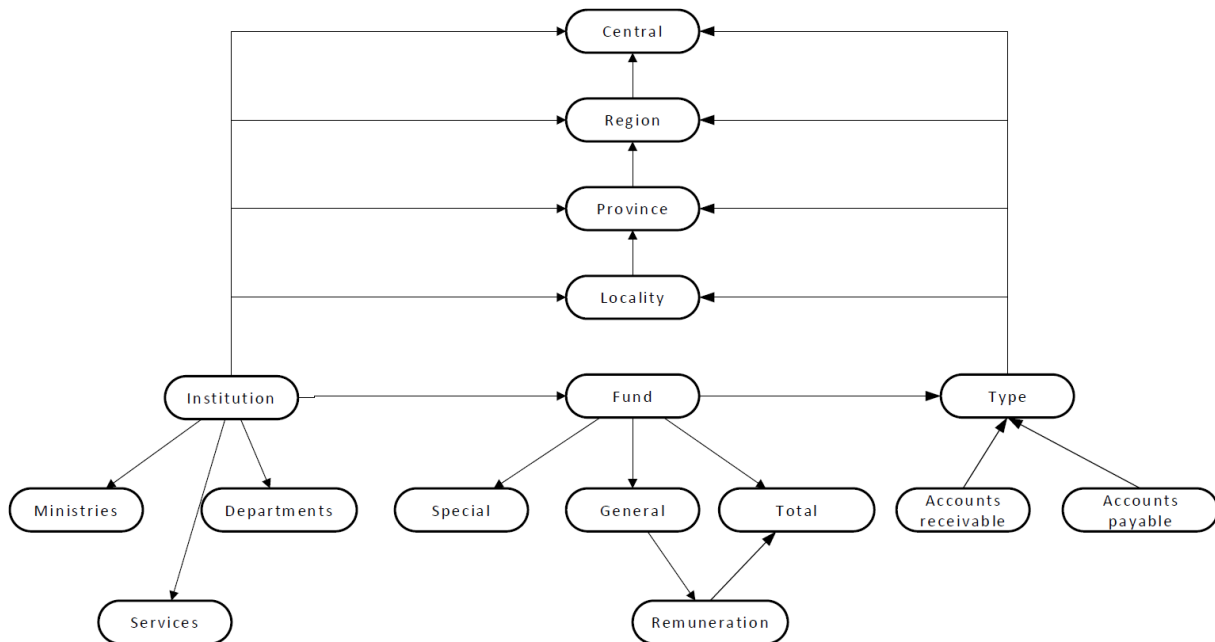


Figure 1. Structure branched network analysis of the budget process areas.

At the beginning of the algorithm, the distance to the top of the source

$$d(u) = 0, \tag{11}$$

for other peaks

$$d(v) = \infty, \forall v \neq u. \tag{12}$$

Computing core algorithm is circumvention of vertices adjacent to the selected vertex v , followed by the addition of not yet visited peaks in the set P. This operation is performed at each step for each vertex, such that $v \in V$.

The algorithm consistently clarifies the function $d(v)$. To take into account the features of construction problems scenario based on a hierarchical ontology introduce an additional function as the process of forming a scenario with desired properties and actions, that it should come in the form:

$F(v_j, v_i) = 1$, if v_j corresponds to the implementation of at least one of the actions that are a part of the script actions under construction.

$F(v_j, v_i) = 0$, if v_j is not responsible enforcement is not one of the actions that are a part of the script actions under construction.

Number of levels in the column is r , and the number of items in each level e .

The sequence of actions to assess the quality of the modeling of the analysis of information on the network can be as follows:

1. Selection on the lower layer of the vertices of the graph corresponding to the quality function.

2. Analysis of the links of the selected vertices with the vertices of the next level, which is located higher in the model hierarchy.

3. Saving the edges of the graph corresponding to the found relationships, for subsequent analysis and use in search operations.

4. Repetition of iterations to the exit to the search source at the top level.

5. Consolidation of all collected edges and vertices of the graph, which is built on the basis of the quality function, that is, the request for information search.

6. The ordering of all paths on a constant graph in accordance with the function of quality.

7. Analysis of the scenario model based on the selection of the shortest paths with the minimum number of edges on the graph.

The structure simplified algorithm will take the following form:

1. Set $k=r, i=0, j=0$.

2. $v_i = a_{ki}$ for $i=1, \dots, e$.

$v_j = a_{k-1}$ for $j=1, \dots, e$.

3. $i = i + 1$.

4. $j = j + 1$.

5. Determine $F(v_j, v_i)$.

6. If $F(v_j, v_i) = 1$, then go to 10, otherwise go to 7.

7. If $j \neq e$ then go to 4, otherwise go to 8.

8. If $i \neq e$ then go to 5, otherwise go to 9.

9. If $k \neq 2$ then go to 11, otherwise go to 12.

10. $E(v_j, v_i)$ found.

11. $k = k-1$ go to 2.

12. The building process is completed

The work developed a simplified version of the algorithm passes to count layers (into layers) below the top of the last layer of the first. Obtained at the previous iteration result identifies one of the curves $e(v_j, v_i)$, which is the basis for finding a new arc on last layer on top v_j . If any layer except the bottom, you can not find $e(v_j, v_i)$, which complies with the quality features

$$F(v_j, v_i) = 1. \quad (13)$$

Then the process is repeated with the transition to a single layer and find your way back repeated. This excludes from consideration arc, poorly chosen. If any condition layer generally can not be done, it means that the scenario can not be built on this column. Then it is necessary to make changes and rebuild ontology graph that describes it, and then to re-build script on a new ontology.

As a result of actions described algorithm we get a script that meets these criteria, in accordance with the basic ontology that describes the graph. The algorithm reflects the sequential process of identifying a sequence of arcs connecting the vertices of the bottom layer to the top of the parameters of life.

When you search for information on loved ones in the form requested, we use an existing model script, which significantly reduces construction time by reducing the number of layers in question, but does not preclude the need to revise the ontology and its restructuring.

Conclusions

In the paper the problem of optimization scenarios streaming collection of information on an branched network of the example of the problem of building and further optimization scenarios to analyze the budget process on an branched network. Shown use structured approach to the ontology as considered most appropriate to the task of presenting the structure as a graph. Moved assessment of the effect of the use of partial information previously obtained by constructing scenarios collecting information flow on the network, described the count, with further refinement information. An formalized description of the hierarchical structure of the system. An example structure interaction problems for ontology elements.

The use of an integrated approach based on finding the shortest path in the graph (SSSP) and model ontology as a graph that realizes the possibility of using algorithms based on traversing the vertices in the hierarchy, i.e. layers.

An approach to the construction of a modified breadth-first search algorithm that significantly reduces the time to find ways of building scenarios collect information on streaming branched network. Described in the article optimization algorithm approach for constructing scenarios collect information on flow tested an branched network in the development of pilot systems, financial analysis of regional budgets. The algorithm allows to develop a set of software tools that provides plenty full and complete optimization problem solving scenarios searching and collecting information flow on an branched network. One of the promising areas of application of the algorithm is used to build its capacity information and analytical systems.

Referenses

1. Chernyak, Leonid. Big Data - a new Theory and Practice//Open system. DBMS. - М .:Open system, 2011. - № 10. -ISSN1028-7493
2. Chernov, V.A. The Economic Theory analysis: Textbook // V.A. Chernov .- Moscow: Prospect, 2017 .- 384 p. - ISBN 978-5-392-24867-4
3. Christopher J. Manning, Prabhakar Rahavan, Heinrich Schütz. Introduction Consumer Information Search (trans. With Eng.) - М .: ООО 'Y.D. Williams' 2011 – p.504.
4. Miller G. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. The Psychological Review, 1956. 63: 81-97.
5. Thorup, Mikkel. "Undirected Single-Source Shortest Paths with Positive Integer Weights in Linear Time." Journal of the ACM 46, no. 3 (May 1, 1999): 362-394
6. Moore, Edward F. "The Shortest Path Through a Maze," International Symposium on the Theory of Switching, 285-92, 1959.
7. Lee, C Y. "An Algorithm for Path Connections and Its Applications." IEEE Transactions on Electronic Computers 10, no. 3 (September 1961): 346-65.

Надійшла: 09.08.2019

Рецензент: д.т.н., професор Вишнівський В.В.