OPTIMIZATION OF ATM TELECOMMUNICATION NETWORKS

Leonid Hulianytskyi, Andrii Baklan

Abstract: ATM network optimization problems defined as combinatorial optimization problems are considered. Several approximate algorithms for solving such problems are developed. Results of their comparison by experiments on a set of problems with random input data are presented.

Keywords: network, ATM, optimization, combinatorial optimization, local search, simulated annealing, genetic algorithm

ACM Classification Keywords: G.2.1 Combinatorics: Combinatorial algorithms

About Network Optimization Problems

Telecommunication and information technologies play a fundamental role in the development of society and economics nowadays. Requirements for telecommunication networks are constantly increasing and, therefore, Broadband Integrated Services Digital Network (B-ISDN) conception has appeared. Evolution of this conception has resulted in appearing of new network technology, universality of which makes it extremely attractive, – Asynchronous Transfer Mode or ATM for short.

The appearance and expansion of this technology along with its high potential make the development of methods for solving ATM network optimization problems important. One of the most important problems is a problem of optimal choice of bandwidth of transmission links for different kinds of traffic for which no effective methods of solution have been developed to present day.

ATM technology is a high-speed, broadband transmission data communication technology based on packet switching (ATM packets are also called cells) and multiplexing technologies and used to carry integrated heterogeneous information, such as data, voice, and video information.

The main requirement for telecommunication networks is the requirement for increasing in their bandwidth together with decreasing in their cost. Availability of high-performance and relatively inexpensive personal computers, workstations, commercial software, expansion of distributed computing – these all demand higher bandwidth at lower cost, available on both local and metropolitan networks. Thus a challenge of providing easy-to-manage broadband services on demand and at an affordable price is arisen.

Different classes of service are used to accommodate transmission of different traffic types in optimal ways, and ATM optimizes traffic flow performance through these various classes of service, which can be allocated on a per-connection basis by using ATM Quality of Service (QoS) settings.

The basic Quality of Service (QoS) parameters (or traffic parameters) that can be negotiated on an ATM network include the following:

- Cell Transfer Delay (CTD);
- Cell Delay Variation (CDV);
- Cell Loss Ratio (CLR);
- Maximum Burst Size (MBS);
- Peak Cell Rate (PCR);
- Sustainable Cell Rate (SCR).

ATM supports several different service categories (kinds of traffic):

- Constant Bit Rate (CBR);
- Variable Bit Rate (VBR);
- Available Bit Rate (ABR);
- Unspecified Bit Rate (UBR).

All these service categories were introduced to attain the ability to transfer heterogeneous traffic, adequate network resources dispatching for each traffic component, more network flexibility and usability. The introducing of service categories increases the advantages making ATM suitable practically for an unlimited range of applications. ATM service categories make it possible for users to choose specific combinations of traffic and performance parameters.

When designing new or analyzing existent telecommunication networks the following problems arise:

- a problem of optimal choice of bandwidth of existent transmission links;
- a problem of optimal choice of routes for transmission and optimal flows distribution;
- a combined problem of optimal choice of bandwidth of existent transmission links, optimal choice of routes for transmission and optimal flows distribution;
- a problem of analysis of survivability indices;
- a problem of network structure synthesis.

One of the most important problems is a problem of optimal choice of bandwidth of transmission links, whose description can be done in the following way. A network structure consisted of switches linked by transmission links is defined. Each transmission link is associated with its length. For each ordered pair of switches the traffic volumes that must be transferred over the network are given. Moreover, an aggregate data flow is given for each transmission link. The bandwidth of each transmission link is proportional to the bandwidth of the basic transmission link. Specific costs per unit length for different bandwidth transmission links are also given. It is necessary to choose such a number of the basic transmission links allocated for traffic for each transmission link that the cost of the network would be minimal and the QoS constraints would be met.

When finding a solution to the problem of optimal choice of bandwidth of transmission links for different kinds of traffic, we must take into account both different kinds of traffic and different QoS parameters for them. There are four different kinds of traffic in ATM technology: CBR, VBR, ABR and UBR, for each of them there are certain QoS parameters negotiated. Having regard to very strong requirement for QoS parameters for CBR traffic, constant bandwidth is allocated for this kind of traffic in each transmission link. VBR and ABR traffics have common bandwidth, which distributes among them in the following way: VBR traffic occupies the greater part of bandwidth and are served by switches with higher relative priority by FIFO procedure, and if there are no VBR cells standing in a queue of a switch, ABR cells are transferred. Finally, the rest of the bandwidth is used for UBR traffic transmission, QoS parameters being not negotiated [1].

Taking into account the above-mentioned, problem description for ATM networks can be formulated in the following way. An ATM network structure consisted of switches linked by transmission links is defined. Each transmission link is associated with its length. For each ordered pair of switches CBR, VBR and ABR traffic volumes that must be transferred over the network are given. Moreover, an aggregate data flow is given for each transmission link. The bandwidth of each transmission link is proportional to the bandwidth of the basic transmission link. Specific costs per unit length for different bandwidth transmission links are also given. It is necessary to choose such numbers of the basic transmission links allocated for CBR traffic, common VBR and ABR traffic, and the share of common VBR and ABR traffic allocated for VBR traffic for each transmission link that the cost of the network would be minimal and the following QoS constraints would be met: CLR and CTD for CBR traffic, CLR and CTD for VBR traffic, CTD for ABR traffic.

The problem of optimal choice of bandwidth of transmission links of ATM network is a new combinatorial optimization problem, to which no effective methods of solution have been developed. Before now only one approach to the problem is proposed – the method of successive analysis and screening of candidate solutions [2]. But this approach don't allow improving solutions found (finding several solutions) and also has essential computational complexities when there is an increase in problem size. Therefore, development of approximate algorithms is expedient to solve the problem.

Statement of the problem of optimal choice of bandwidth of transmission links for CBR traffic as well as VBR and ABR traffics is considered in [2]. Modified statement of the problem under consideration in the form of a combinatorial optimization model is presented below.

Statement of Problem

As it is mentioned above, constant bandwidth is allocated for CBR traffic in each transmission link, independently of flow distribution. This allows solving the problem of optimal choice of bandwidth of transmission links for CBR traffic independently of VBR and ABR traffics.

Therefore, we will consider two subproblems: optimal choice of bandwidth of transmission links for CBR traffic and optimal choice of bandwidth of transmission links for VBR and ABR traffics.

1) Optimal choice of bandwidth of transmission links for CBR traffic.

Let $G = (V, E)$ be an undirected graph that represents a network structure where $V = \{v_1, v_2, \ldots, v_n\}$ is a set of vertices (switches), $E = \{e_1, e_2, \ldots, e_m\}$ is a set of edges (transmission links). Each transmission link e_k is associated with its length l_k, $k = \overline{1, m}$. A CBR traffic matrix $H^{(0)} = \left\| h_{ij}^{(0)} \right\|_{i,j=\overline{1,n}}$ is also given, where $h_{ij}^{(0)}$ is the traffic volume (Mbit/s) that must be transferred from a switch i to a switch j. Moreover, an aggregate CBR traffic data flow vector $f^{(0)} = \left(f_1^{(0)} \quad f_2^{(0)} \quad \ldots \quad f_m^{(0)} \right)^T$ is given, where $f_k^{(0)}$ is an aggregate CBR traffic data flow in a transmission link e_k, $k = \overline{1, m}$. The bandwidth of each transmission link e_k is proportional to the bandwidth of the basic transmission link μ (Mbit/s): $\mu_k^{(0)} = x_k^{(0)} \mu$, where $x_k^{(0)} \in \{1,2,\ldots,N\}$ is the number of the basic transmission links, $k = \overline{1, m}$.

Specific costs per unit length for different bandwidth transmission links $C = \{c_1, c_2, \ldots, c_N\}$ are also given. That is the cost of a transmission link e_k with the number of the basic transmission links $x_k^{(0)}$ and length l_k equals to $c_{x_k^{(0)}} l_k$, $k = \overline{1, m}$.

It is necessary to choose such numbers of the basic transmission links allocated for CBR traffic $x^{(0)} = \left(x_1^{(0)} \quad x_2^{(0)} \quad \ldots \quad x_m^{(0)} \right)^T$ that the cost of the network would be minimal:

$$\sum_{k=1}^{m} c_{x_k^{(0)}} l_k \to \min \tag{1}$$

provided that the following constraints would be met:

$$CLR^{(0)} \le CLR_{set}^{(0)}, \tag{2}$$

$$CTD^{(0)} \le CTD_{set}^{(0)}, \tag{3}$$

$$f_k^{(0)} < x_k^{(0)} \mu, \; k = \overline{1, m}, \tag{4}$$

$$x_k^{(0)} \in \{1,2,\ldots,N\}, \; k = \overline{1, m}, \tag{5}$$

where $CLR^{(0)}$ is the mean probability of losses (the loss ratio) for CBR cells; $CTD^{(0)}$ is the mean transfer delay for CBR cells; $CLR_{set}^{(0)}$ is the specified loss ratio for CBR cells; $CTD_{set}^{(0)}$ is the specified mean transfer delay for CBR cells.

Formulas for $CTD^{(0)}$ and $CLR^{(0)}$ obtained in [2] in our notation are the following:

$$CTD^{(0)} = \frac{1}{H_{\Sigma}^{(0)}} \sum_{k=1}^{m} \frac{f_k^{(0)}}{x_k^{(0)}\mu - f_k^{(0)}}, \tag{6}$$

$$CLR^{(0)} = \frac{1}{m} \sum_{k=1}^{m} CLR_k^{(0)}, \tag{7}$$

$$CLR_k^{(0)} = P_0 \left(\frac{f_k^{(0)}}{\mu} \right)^{x_k^{(0)}} \frac{1}{x_k^{(0)}!} \left(\frac{f_k^{(0)}}{x_k^{(0)}\mu} \right)^{N_k^{(0)}}, \tag{8}$$

$$P_0 = \left[\sum_{t=0}^{x_k^{(0)}} \left(\frac{f_k^{(0)}}{\mu} \right)^t \frac{1}{t!} + \left(\frac{f_k^{(0)}}{\mu} \right)^{x_k^{(0)}} \frac{1}{x_k^{(0)}!} \sum_{t=1}^{N_k^{(0)}} \left(\frac{f_k^{(0)}}{x_k^{(0)}\mu} \right)^t \right]^{-1}, \tag{9}$$

where $H_{\Sigma}^{(0)} = \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^{(0)}$ is the aggregate CBR traffic volume; $N_k^{(0)}$ is the size of a buffer of ATM switch for CBR traffic cells.

2) Optimal choice of bandwidth of transmission links for VBR and ABR traffics.

Let $G = (V, E)$ be an undirected graph that represents a network structure where $V = \{v_1, v_2, \ldots, v_n\}$ is a set of vertices (switches), $E = \{e_1, e_2, \ldots, e_m\}$ is a set of edges (transmission links). Each transmission link e_k is associated with its length l_k, $k = \overline{1,m}$. Both a VBR traffic matrix $H^{(1)} = \left\| h_{ij}^{(1)} \right\|_{i,j=\overline{1,n}}$ and an ABR traffic matrix $H^{(2)} = \left\| h_{ij}^{(2)} \right\|_{i,j=\overline{1,n}}$ are also given, where $h_{ij}^{(1)}$ and $h_{ij}^{(2)}$ are respectively the VBR and ABR traffic volumes (Mbit/s) that must be transferred from a switch i to a switch j. Moreover, both an aggregate VBR traffic data flow vector $f^{(1)} = \left(f_1^{(1)} \quad f_2^{(1)} \quad \ldots \quad f_m^{(1)} \right)^T$ and an aggregate ABR traffic data flow vector $f^{(2)} = \left(f_1^{(2)} \quad f_2^{(2)} \quad \ldots \quad f_m^{(2)} \right)^T$ are given, where $f_k^{(1)}$ и $f_k^{(2)}$ are respectively an aggregate VBR and an aggregate ABR traffic data flows in a transmission link e_k, $k = \overline{1,m}$. The bandwidth of each transmission link e_k is proportional to the bandwidth of the basic transmission link μ (Mbit/s): $\mu_k = x_k \mu$, where $x_k \in \{1,2,\ldots,N\}$ is the number of the basic transmission links, $k = \overline{1,m}$.

Specific costs per unit length for different bandwidth transmission links $C = \{c_1, c_2, \ldots, c_N\}$ are also given. That is the cost of a transmission link e_k with the number of the basic transmission links x_k and length l_k equals to $c_{x_k} l_k$, $k = \overline{1,m}$.

It is necessary to choose such numbers of the basic transmission links allocated for common VBR and ABR traffic $x = \left(x_1 \quad x_2 \quad \ldots \quad x_m \right)^T$, and the share of common VBR and ABR traffic allocated for VBR traffic $x^{(1)} = \left(x_1^{(1)} \quad x_2^{(1)} \quad \ldots \quad x_m^{(1)} \right)^T$ that the cost of the network would be minimal:

$$\sum_{k=1}^{m} c_{x_k} l_k \to \min \tag{10}$$

provided that the following constraints would be met:

$$CLR^{(1)} \le CLR_{set}^{(1)}, \tag{11}$$

$$CTD^{(1)} \le CTD_{set}^{(1)}, \tag{12}$$

$$CTD^{(2)} \le CTD_{set}^{(2)}, \tag{13}$$

$$x_k^{(1)} < x_k, \ k = \overline{1,m}, \tag{14}$$

$$f_k^{(1)} + f_k^{(2)} < x_k \mu, \ k = \overline{1,m}, \tag{15}$$

$$f_k^{(1)} < x_k^{(1)} \mu, \ k = \overline{1,m}, \tag{16}$$

$$x_k \in \{1,2,..., N\}, \ k = \overline{1,m}, \tag{17}$$

$$x_k^{(1)} \in \{1,2,..., N\}, \ k = \overline{1,m}, \tag{18}$$

where $CLR^{(1)}$ is the mean probability of losses (the loss ratio) for VBR cells; $CTD^{(1)}$ and $CTD^{(2)}$ are the mean transfer delays for VBR and CBR cells respectively; $CLR_{set}^{(1)}$ is the specified mean probability of losses (the loss ratio) for VBR cells; $CTD_{set}^{(1)}$ and $CTD_{set}^{(2)}$ are the specified mean transfer delays for VBR and CBR cells respectively.

Formulas for $CTD^{(1)}$, $CTD^{(2)}$ and $CLR^{(1)}$ obtained in [2] in our notation are the following:

$$CTD^{(1)} = \frac{1}{H_{\Sigma}^{(1)}} \sum_{k=1}^{m} \frac{f_k^{(1)}\left(f_k^{(1)} + f_k^{(2)}\right)}{x_k \mu \left(x_k \mu - f_k^{(1)}\right)}, \tag{19}$$

$$CTD^{(2)} = \frac{1}{H_{\Sigma}^{(2)}} \sum_{k=1}^{m} \frac{f_k^{(2)}\left(f_k^{(1)} + f_k^{(2)}\right)}{\left(x_k \mu - f_k^{(1)}\right)\left(x_k \mu - f_k^{(1)} - f_k^{(2)}\right)}, \tag{20}$$

$$CLR^{(1)} = \frac{1}{m} \sum_{k=1}^{m} CLR_k^{(1)}, \tag{21}$$

$$CLR_k^{(1)} = P_0 \left(\frac{f_k^{(1)}}{\mu}\right)^{x_k^{(1)}} \frac{1}{x_k^{(1)}!} \left(\frac{f_k^{(1)}}{x_k^{(1)} \mu}\right)^{N_k^{(1)}}, \tag{22}$$

$$P_0 = \left[\sum_{t=0}^{x_k^{(1)}} \left(\frac{f_k^{(1)}}{\mu}\right)^{t} \frac{1}{t!} + \left(\frac{f_k^{(1)}}{\mu}\right)^{x_k^{(1)}} \frac{1}{x_k^{(1)}!} \sum_{t=1}^{N_k^{(1)}} \left(\frac{f_k^{(1)}}{x_k^{(1)} \mu}\right)^{t}\right]^{-1}, \tag{23}$$

where $H_{\Sigma}^{(1)} = \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^{(1)}$ and $H_{\Sigma}^{(2)} = \sum_{i=1}^{n} \sum_{j=1}^{n} h_{ij}^{(2)}$ are respectively the aggregate VBR traffic volume and the aggregate ABR traffic volume; $N_k^{(1)}$ is the size of a buffer of ATM switch for VBR traffic cells.

Experimental Results

For the purpose of making experimental investigation of the developed algorithms a program system implementing local search algorithm (LS) [3], iterated local search algorithm (ILS) [4], simulated annealing algorithm (SA) [5], G-algorithm [6], and genetic algorithm (GA) [7] was developed. This program system also allows generating test problems with random input data.

At the first stage the following parameter values of the algorithms were empirically specified. For ILS algorithm: maximum number of iterations $t_{\max}$ =30, maximum number of transitions $h_{\max}$ =20. For SA algorithm: maximum number of iterations $t_{\max}$ =20; maximum number of transitions $h_{\max}$ =10000; initial temperature value T_0 =50; temperature schedule coefficient r =0,925; number for equilibrium condition determination ε =0,01; number of passages k =2; number of transitions per passage v =35. For G-algorithm: maximum number of iterations $t_{\max}$ =40; number for equilibrium condition determination ε =10; number of passages k =2; number of transitions per passage v =35; initial value of parameter μ_0 =0; parameter γ =0,05. For GA: maximum number of generations $t_{\max}$ =500; number of individuals in initial population *K*=*20*; number of selected individuals for crossing over Q =10; parental gene inheritance probability P_c =0,1; probability of mutative change of gene P_m =0,5. Let radiuses of vicinities in LS, ILS and SA algorithms equal 1.

At the second stage numerical experiments on the developed algorithms were made. For that a control set of 80 problem instances, consisted of 16 subsets of 5 problem instances of the same size, was generated. Note that the size of a problem is determined by the number of edges *m*.

For each problem instance each algorithm was executed 5 times and, as a result , the following values were found: total execution time *t* (ms), the best value of objective function f_* and improvement *q* (%), which is expressed by formula $q = (f_0 - f_*) / f_0 \cdot 100\%$, where f_0 is the value of objective function for initial candidate solution.

Average results of numerical experiments on two above-mentioned subproblems are presented in Table 1 for the first subproblem (1)–(9) (CBR traffic) and in Table 2 for the second subproblem (10)–(23) (VBR and ABR traffics).

The program system was developed in Object Pascal programming language in IDE Borland Delphi 7.0. Numerical experiments were run on personal computer with the following characteristics: CPU – AMD Athlon XP 1700+ , 1,47 GHz; RAM – DDR 333 MHz, 256 MB; operating system – Windows XP Professional.

Analysis of the results of the experiments has shown that for first subproblem instances a minimal improvement has been given by LS algorithm, other ones have found almost the same solutions, G-algorithm being the best. For second subproblem instances the best solutions have been found by ILS algorithm and G-algorithm, those solutions being almost the same.

Table 1. Average results of numerical experiments on the first subproblem (CBR traffic).

m	IL		ILS		SA		G-algorithm		GA	
	q, %	t, ms	q, %	t, ms	q, %	t, ms	q, %	t, ms	q, %	t, ms
15	31,54	212,20	43,29	1540,20	45,80	12498,00	45,08	8860,60	42,62	4917,20
17	28,54	258,20	44,11	2069,20	46,55	13032,60	45,12	9756,00	43,32	5596,00

20	36,48	450,80	46,15	2889,80	50,00	17735,80	48,89	13120,80	46,07	6609,60
21	15,15	220,00	26,45	2665,80	32,23	14354,60	30,62	11011,80	28,82	6786,00
23	40,74	625,00	53,36	3759,20	57,91	19776,40	56,72	16207,20	53,08	7767,40
25	13,69	308,40	24,40	3623,20	32,65	15636,40	31,09	13481,40	27,21	8217,80
27	31,19	709,00	50,16	5542,20	56,99	24939,60	54,60	18220,20	51,78	8123,60
29	54,04	1382,20	65,86	6673,60	72,37	25462,40	70,14	20828,00	65,28	8632,60
30	28,62	817,00	42,61	6303,20	48,94	28242,80	46,93	20451,40	44,65	8908,80
32	6,83	362,60	27,03	6541,40	40,30	17787,60	38,53	18793,00	36,18	10002,60
33	31,65	1138,00	44,58	8101,60	52,77	31260,80	50,77	23273,40	46,78	9770,20
35	13,24	1019,80	27,83	8884,40	36,74	28689,20	34,62	20770,00	33,28	10152,60
37	40,92	1798,20	56,31	10307,00	64,93	34583,60	62,80	28122,20	60,08	10854,00
40	5,77	619,00	19,61	10086,80	30,52	24569,00	27,44	20215,20	26,03	11813,00
45	10,88	1120,00	29,03	14124,20	44,62	36861,00	41,67	29396,20	38,99	12818,60
50	8,62	1091,60	23,04	15654,60	41,36	36672,40	37,69	31221,40	35,81	14803,00

Table 2. Average results of numerical experiments on the second subproblem (VBR and ABR traffics).

m	IL		ILS		SA		G-algorithm		GA	
	q, %	t, ms	q, %	t, ms	q, %	t, ms	q, %	t, ms	q, %	t, ms
15	75,68	412,60	76,02	2251,20	75,88	9151,20	76,04	5874,60	75,93	4099,80
17	68,82	478,80	71,96	2457,80	72,22	11201,80	71,99	7410,60	71,86	4819,00
20	76,78	787,20	80,18	5013,00	80,28	12269,80	80,05	9711,80	79,14	5834,40
21	67,98	799,00	69,37	4474,80	68,82	13934,00	69,17	8910,80	68,44	5407,60
23	80,57	1021,40	82,93	4867,00	82,95	15246,00	83,13	10126,40	80,95	6060,80
25	59,20	983,40	68,92	8231,80	69,92	16930,40	68,90	12808,60	68,37	6291,00
27	76,12	1434,20	76,35	7036,00	75,42	19414,00	76,39	12283,80	73,61	7514,60
29	84,63	1732,40	84,48	6982,00	83,63	21020,20	84,64	15260,00	80,44	7885,20
30	73,95	1974,00	74,31	9041,20	73,31	22744,60	73,66	15125,80	70,96	8107,60
32	68,62	1946,60	75,68	12577,80	75,38	22801,00	75,53	17947,80	71,79	8530,40
33	74,65	2157,20	75,50	13861,60	74,27	25254,40	75,33	17607,40	71,04	8766,60
35	61,85	2303,40	66,22	14619,20	63,83	27962,20	66,10	17360,60	61,60	9077,20
37	81,30	2794,00	81,48	14088,20	79,92	26670,20	81,50	20229,40	75,54	9882,20
40	65,13	3052,20	66,33	17221,00	63,67	29652,60	65,84	20607,60	60,25	10288,60
45	69,07	3943,40	70,36	20485,80	67,34	34727,60	70,31	24819,80	61,97	11516,60
50	67,47	4708,80	72,81	27754,00	70,02	42056,20	72,59	28853,80	62,42	12770,20

Bibliography

1. *Назаров А.Н., Симонов М.В.* АТМ: технология высокоскоростных сетей. М.: ЭКО-ТРЕНДЗ, 1997.
2. *Зайченко О.Ю.* Оптимізація характеристик мереж з технологією АТМ // Системні дослідження та інформаційні технології. – 2002. – №3. – С.57– 73.
3. *Сергиенко И.В.* Математические модели и методы решения задач дискретной оптимизации – К.: Наук. думка, 1988.
4. *Lourenço H. R., Martin O., Stützle T.* Iterated local search // Handbook of Metaheuristics: International Series in Operations Research & Management Science, vol. 57 (Eds. F. Glover and G. Kochenberger). – Norwell: Kluwer Academic Publishers, MA, 2002. – P. 321-353.
5. *Aarts E., Korst J.* Simulated annealing and Boltzmann machines. – New York: John Willey and Sons, 1989.
6. *Гуляницкий Л.Ф.* Решение задач комбинаторной оптимизации алгоритмами ускоренного вероятностного моделирования // Компьютерная математика. – Киев: Институт кибернетики им. В.М. Глушкова НАН Украины, 2004. – №1. – С. 64-72.
7. *Reeves C.R., Rowe J.E.* Genetic Algorithms: Principles and Perspectives. A Guide to GA Theory. – Boston: Kluwer Academic Publishers, 2002.

Authors' Information

Leonid F. Hulianytskyi – V.M.Glushkov Institute of Cybernetics of NAS of Ukraine, Glushkova ave. 40, Kyiv, 03680, Ukraine; e-mail: lh_dar@hotmail.com

Andrii M. Baklan – V.M.Glushkov Institute of Cybernetics of NAS of Ukraine, Glushkova ave. 40, Kyiv, 03680, Ukraine; e-mail: a_baklan@yahoo.com

---

ANALYZING THE DATA IN OLAP DATA CUBES[*]

Galina Bogdanova, Tsvetanka Georgieva

Abstract: OLAP applications provide a possibility to data analysis over large collections of historical data in the data warehouses, supporting the decision-making process. This paper presents an application that creates a data cube and demonstrates the effectiveness of the applying the OLAP operations when it necessary to analyze the data and obtain the valuable information from the data. It allows the analysis of factual data that is daily downloads of folklore materials, according to dimensions of interest.

Keywords: data cube, online analytical processing, multidimensional expressions

ACM Classification Keywords: H.4.2 Information Systems Applications: Types of Systems – Decision support

1. Introduction

Decision-support functions in a data warehouse, such as online analytical processing (OLAP), involve hundreds of complex aggregate queries over large volumes of data. It is not feasible to compute these queries by scanning the data sets each time [9]. The data cubes are structures designed to provide quick access to the data in data warehouses. The cube definition is determined from the requirements, which the users analyzing the data have

---

[*] Supported partially by the Bulgarian National Science Fund under Grant MM-1405/2004