DOI: 10.31319/2519-8106.2(51)2024.317498
UDC 004.65

Yalova Kateryna, Candidate of Technical Sciences, Associate Professor, Head of the Department of the Systems software
Ялова К.М., кандидат технічних наук, доцент, кафедра програмного забезпечення систем
ORCID: 0000-0002-2687-5863
e-mail: yalovakateryna@gmail.com

Babenko Mykhailo, Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of the Systems software
Бабенко М.В., кандидат технічних наук, доцент, кафедра програмного забезпечення систем,
ORCID: 0000-0003-1013-9383
e-mail: mvbab@ukr.net

Ismailov Vitaliy, PhD student, Department of the Systems software
Ісмаілов В.В., аспірант кафедри програмного забезпечення систем
email: ivv15081996@gmail.com

Dniprovsky State Technical University, Kamianske
Дніпровський державний технічний університет, м. Кам'янське

OLAP HYPERCUBES AS A TOOL FOR ANALYZING MULTIDIMENSIONAL HIGHLY STRUCTURED DATA

ГІПЕРКУБИ OLAP ЯК ЗАСІБ АНАЛІЗУ БАГАТОВИМІРНИХ СИЛЬНО СТРУКТУРОВАНИХ ДАНИХ

The article is devoted to the analysis of OLAP technology for multidimensional data analysis in relational data storage. The key components of OLAP technology are described: hypercube, dimension, measure, dimension table, and fact table. The models for transforming relational data structures into multidimensional models are also discussed. The authors propose a generalized algorithm for applying OLAP technology to multidimensional data analysis, which includes the following steps: acquiring input data, defining measures and dimensions, constructing a multidimensional data model, creating a query structure, and processing the resulting data. The paper presents the results of designing a hypercube based on statistical data from a national multi-subject test, stored in the tables of a relational data warehouse. To demonstrate the principles of using a hypercube for multidimensional data analysis, the article showcases the application of the following OLAP operations: slicing, dicing, rotating, drill-down, and roll-up.

***Keywords**: multidimentional data analysis, OLAP, ROLAP, hypercubes, relational data storage, snowflake data model.*

Статтю присвячено аналізу технології OLAP з точки зору багатовимірного аналізу даних, збережених у реляційних сховищах даних. Технологія OLAP забезпечує збереження і представлення інформації для її комплексного аналізу за рахунок узагальнення даних великих масивів, структурованих за принципом багатовимірності. Для застосування OLAP технології для проведення багатовимірного аналізу даних та їх візуалізації необхідно перетворити вхідні дані до специфічної архітектури сховища даних OLAP, визначивши категорії-виміри, які використовуються для реалізації багатовимірних запитів. В статті описано основні об'єкти технології OLAP: гіперкуб, вимір, показник, таблиця вимірів і фактів; наведено опис моделей перетворення реляційної структури даних до багатовимірної моделі. Багатовимірний куб або гіперкуб даних — це абстрактне представлення підмножин даних у декількох вимірах одночасно, опера-

ції з яким можуть бути застосовано для різностороннього аналізу багатовимірних даних. Вимірами гіперкубу є числові значення, числові значення, координати яких представляються індивідуальними значеннями атрибутів, що містяться в таблиці фактів. Показниками гіперкубу є чисельні значення, які зіставляються з числовими значеннями стовпців таблиці фактів.

Авторами запропоновано узагальнений алгоритм застосування технології OLAP для здійснення багатовимірного аналізу даних, який полягає у виконанні таких кроків, як: отримання вхідних даних, визначення показників і вимірів, формування багатовимірної моделі даних, формування структури запитів, обробка отриманих даних. Наведено результати проєктування гіперкубу на основі статистичних даних результатів національного мультипредметного тесту, збережених у таблицях створеного реляційного сховища даних. З метою демонстрації принципів використання гіперкубу для аналізу багатовимірних даних в статті наведено результати застосування операцій: slice, dice, rotation, drill-down, roll-up.

***Ключові слова**: аналіз багатовимірних даних, OLAP, ROLAP, гіперкуб, реляційне сховище даних, модель даних «сніжинка».*

Problem’s formulation

In today's information technology landscape, the need for efficient data processing methods is increasingly important. Databases (DBs) and their management systems have become integral to the storage and processing of data in industrial enterprises, banking institutions, educational establishments, commercial organizations, and other institutions. Relational Database Management Systems (RDBMS) store data in normalized tables optimized for transaction processing systems. However, queries involving multiple tables in RDBMS can be relatively slow [1]. In contrast, OLAP (On-Line Analytical Processing) technology significantly simplifies data analysis by utilizing multidimensional cubes for data representation. OLAP technology enables the storage and presentation of information in a manner conducive to comprehensive analysis by aggregating large datasets, structured according to the principles of multidimensionality. To effectively apply OLAP technology for multidimensional data analysis and visualization, it is necessary to transform the input data into an OLAP-specific data warehouse architecture [2]. This involves defining categories and dimensions to support multidimensional queries.

The main advantages of OLAP as a tool for multidimensional data analysis include [3]:

– increase in data processing speed compared to standard database queries due to preliminary aggregation of the input data set;

– the ability to perform data processing in real time, which allows for timely management decisions;

– ease of visualization and analysis with different parameters due to the multidimensionality of hypercubes and operations with them

– the ability to quickly and easily create pivot tables without the need to use database query languages;

– the integration property of OLAP technology provides the ability to further use the obtained data in information systems and dashboards created in Power BI, Tableau, or IBM Cognos.

Despite several disadvantages, such as resource intensity, design complexity and low flexibility, OLAP technologies remain a powerful tool for multidimensional data analysis for large-scale structured data.

Analysis of recent research and publications

The term «OLAP» was coined by Edgar F. Codd in 1993, when he published his article titled «Providing OLAP to User-Analysts: An IT Mandate» where he also described 12 laws of analytical data processing [2]. The most famous scientists who made a significant contribution to the further development of the OLAP idea are:

– R. Kimball, who proposed to implement dimensional modeling, representing data in the form of facts and measurements.

– J. Dorrian, R. Earle, who led the development of Essbase (Extended Spreadsheet Database), one of the first Multidimensional OLAP systems;

– The development of Business Intelligence (BI) technologies necessitates the search for effective means of processing multidimensional data to make timely and informed management decisions [4]. In this context, OLAP technologies are successfully used to achieve these goals. D. Lemire, S. Zuboff, B. Inmon devoted their scientific works to the topic of optimizing queries to relational data storage and the use of OLAP in BI and decision support.

Since the 2000s, OLAP has been integrated into information systems that have internal data storages, and software tools such as QlikView, Tableau, and Power BI provide effective tools for interactive analysis of multidimensional data using pre-aggregated OLAP cubes.

The following scientists have devoted their scientific works in the field of multidimensional data processing and application of OLAP technology: Korobko A.V., Penkova T.G., Semchenkov S.Y., Vakhitov A.R., Demchenko A.A., Samochadin A. V. The task of building hypercubes based on data from relational and object-oriented databases is considered in the works of Fisun N.T., Gorbany G.V., Bondarenko A.V., Bessarabov N.A., etc. The problems of OLAP-systems development and application of hypercubes for multidimensional analysis and visualization of database query results are considered in the works of such foreign scientists as: V. Kamel, R. Djiroun, M. Maliappis, D. Kremmydas, M. Fotache, C. Strimbei, O. Boutkhoum and others.

The latest advances in cloud technologies and the transfer of corporate data to the clouds have become the basis for the development of OLAP technologies and their integration into cloud services. Modern platforms such as Google BigQuery, Snowflake, and Amazon Redshift contain integrated OLAP functionality, providing the ability to multidimensional analyze data stored in the clouds of large datasets [5]. Combining the principles of cloud data storage with OLAP increases the potential of real-time data processing, improving its performance.

Formulation of the study purpose

The purpose of the article is to analyze the OLAP technology and its application for analysis of multidimensional data by means of hypercube operations. The objectives of the study are:

– developing of an algorithm for transferring a domain to relational data storage;
– defining a test data set for conducting experiments with the OLAP hypercube;
– designing the architecture of a relational data storage of the «snowflake» type;
– defining the dimensions and indicators of the hypercube;
– creating a hypercube and filling it with data;
– conducting experiments on processing multidimensional data using hypercube operations.

Presenting main materials

The hypercube provides a structure for storing and organizing access to multidimensional data and is widely used in data analysis, particularly in OLAP technology [6]. Multidimensional data consists of several characteristics, or dimensions, each of which adds a new level of granularity to the data. This data is often represented in the form of a multidimensional table or a hypercube. A hypercube is an abstract representation of subsets of data across multiple dimensions simultaneously [7]. The dimension *D* of a hypercube *G* refers to the numerical values whose coordinates are represented by the individual attribute values contained in the fact table [8]. Dimensions can be organized hierarchically, and a set of dimensions forms the axes of the hypercube. A common example is the time dimension, which is often structured hierarchically as week → month → quarter → year. Measures *F* are numerical values used for performing multidimensional data analysis, typically mapped to numeric columns in the fact table of a relational database (RDB). Measures commonly contain aggregated data produced by operations like SUM, COUNT, and AVG. The correct identification of dimensions and measures, to be represented as part of the relational model of the hypercube, is crucial for its effective use in multidimensional data analysis [9].

Mathematically, the hypercube *G* can be described as:

$$G = \langle F, D \rangle, \tag{1}$$

where $F = \{f_1,\ldots,f_m\}$ — a set of hypercube measures that characterize the process under analysis; $D = \{d_1,\ldots,d_n\}$ — is a set of hypercube dimensions, each of which is an ordered set of values of a certain type.

The process of forming a hypercube involves selecting the necessary tables, setting up links between them, selecting data fields and their correlation with the terms of the domain. When designing a hypercube, the most difficult thing is to represent the knowledge of the domain as a set of relational tables of a certain architectural model suitable for multidimensional data analysis. Hypercubes are the basis of OLAP technology, and if a hypercube is represented as a relational data storage, ROLAP (Relational OLAP) technology is used. There are two types of tables in a ROLAP storage [10]:

– fact table is the main table of the data storage that contains information about objects or events for multidimensional analysis. The fact table must contain a composite primary key that combines the primary keys of the dimension tables. There are four types of facts: facts related to transactions and based on individual events (transaction facts); facts related to the state of domain objects (snapshot facts); facts related to elements of domain documents (line-item facts); facts that provide information about the occurrence of an event without details about it (event or state facts);

– dimension table — a table containing conditionally unchangeable data of the domain. Each dimension table must contain at least one descriptive field and a primary key. If the dimensions are represented as a hierarchical entity, then the fact table is supplemented with fields that point to descendants and ancestors.

The choice of a ROLAP data storage model for building a hypercube depends on the complexity of the domain and the structure of the final data samples. The input data in the ROLAP model is transformed into a multidimensional model through an intermediate metadata layer. When using ROLAP, the results of queries are stored in the same database as the input data. Aggregated data is stored in specially created service tables within the database. To create a ROLAP data storage, two models for organizing fact and dimension tables are typically used, differing in complexity and the level of data normalization. These models are [11]:

– Star schema is the most common data storage model for storing hypercube data. This model involves creating a set of tables and defining a central (single) fact table among them that is linked to several dimension tables. To implement the «star» model, it is necessary to convert the RDB tables into a form where each dimension of the cube is contained in a separate table, and a one-to-many relationship is used to implement the connection with the fact table.

– Snowflake schema is a normalized data storage model for storing hypercube data, where dimension tables can be presented in a hierarchical relationship, meaning they can have several subtables. The snowflake model is characterized by a situation where information about one dimension is contained in several related tables. In this case, a child table from the dimension hierarchy is linked to the fact table.

A generalized algorithm for constructing a hypercube for multidimensional analysis using ROLAP technology consists of the following steps:

1. Obtaining input data and analyzing their structure and interdependence.
2. Determination of the set of dimensions and measures for the hypercube, as well as the desired level of detail for the multidimensional analysis of input data.
3. Developing a ROLAP model by representing the input data as a set of tables in the «star» or «snowflake» model. Determine the structure, number and content of the fact and dimension tables.
4. Formation of the queries structure for multidimensional analysis and the application of analytical operations to the developed hypercube.
5. Processing and visualization of the results from data queries.

As an example of multidimensional data analysis using ROLAP technology, the authors propose to consider the statistical data of external independent testing results. To analyze multidimensional data, the set of dimensions is described as: $D = \{d_1,\ldots,d_n\}$, where $n = 3$. The dimension that describes territorial units has subdimensions and is given as: $d_2=\{d_{21},\ldots,d_{2k}\}$, where $k = 6$. The names of the dimensions are as follows:

1. Time: year.
2. Territorial units: region $\to$ district of region $\to$ settlement $\to$ district of settlement $\to$ educational institution $\to$ participant.
3. Disciplines in which testing was conducted.

The set of measures describes as: $F = \{f_1,\ldots,f_m\}$, where $m = 4$. Hypercube measures include:

1. Number of test participants.
2. Average score.

3. Maximum score received.
4. Number of participants who received the highest score.

It is difficult to visualize the created hypercube in full, but mathematically it is processed as a structure with n number of dimensions and m measures. To simplify the representation of the hypercube, it can be represented in three-dimensional form, as shown in Fig. 1, which presents a partial visualization of the constructed hypercube.

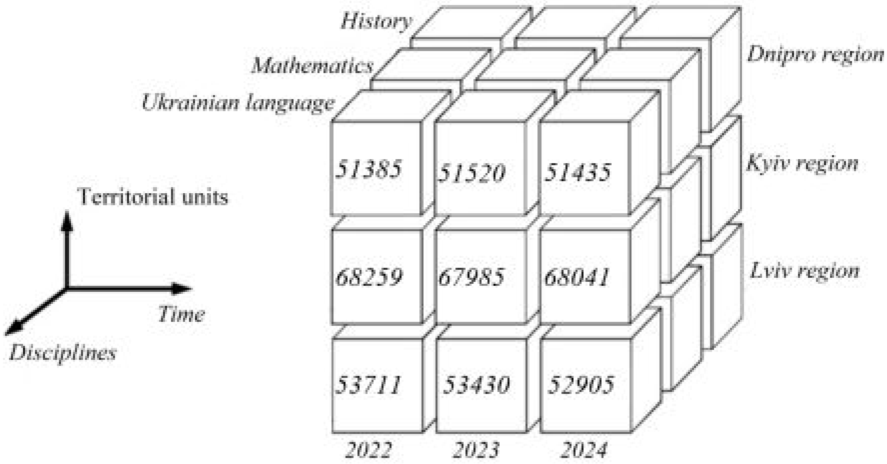

Fig. 1. Fragment of the hypercube for the results of external independent testing

The number of measure values in the hypercube depends on the number of years selected for analysis, the disciplines, and the values of the territorial distribution hierarchy. The total number of measure values in the hypercube can be calculated as the product of the number of elements in each i-th dimension:

$$G = d_1 \times d_2 \times ,..., \times d_n = \prod_{i=1}^{n} d_i \ . \tag{2}$$

To determine the possibility of joint analytical processing of the measures F and the dimensions D for the situation when the indicator f_i can be analyzed by the measurement d_j, the relation $R \subseteq F$ Ч D, $(f_i, d_j) \in R$ is set in the hypercube.

Since the set of dimensions D includes a hierarchical dependency within the territorial distribution dimension, it is advisable to use the «snowflake» schema as the relational data storage model, as illustrated in Fig. 2.

Unlike conventional RDB processing, where the output data is generated as a result of executing database queries, analytical operations can be used to perform multidimensional data analysis based on a hypercube [12]. The result R of each operation O with a hypercube can be described as:

$$R = O(d_1, d_2, ..., d_n, F) \ . \tag{3}$$

The main operations with a hypercube include:

1. Slicing: This involves fixing the values of a particular dimension, thereby reducing the dimensionality of the cube [13]. A slice is essentially a subcube that retains all other dimensions. The amount of data in the slice decreases depending on the number of values in the dimension being sliced. For example, when the i-th dimension (where i=1) is absent, the amount of data in the slice can be calculated using the following formula:

$$C_{slice} = d_2 \times ,..., \times d_n \ . \tag{4}$$

For a described hypercube, it is possible, for example, to create a C_{time_slice} for a certain value of the year of testing, which will reduce the dimensionality of the cube like:

$$C_{time\_slice} = \{d_2, d_3\}, \ \textit{where time}='2023'. \tag{5}$$

Fig. 3 shows a fragment of the C_{time_slice} data slice, the results of applying the slice operation to the proposed hypercube, where the time dimension is selected as the slice parameter.

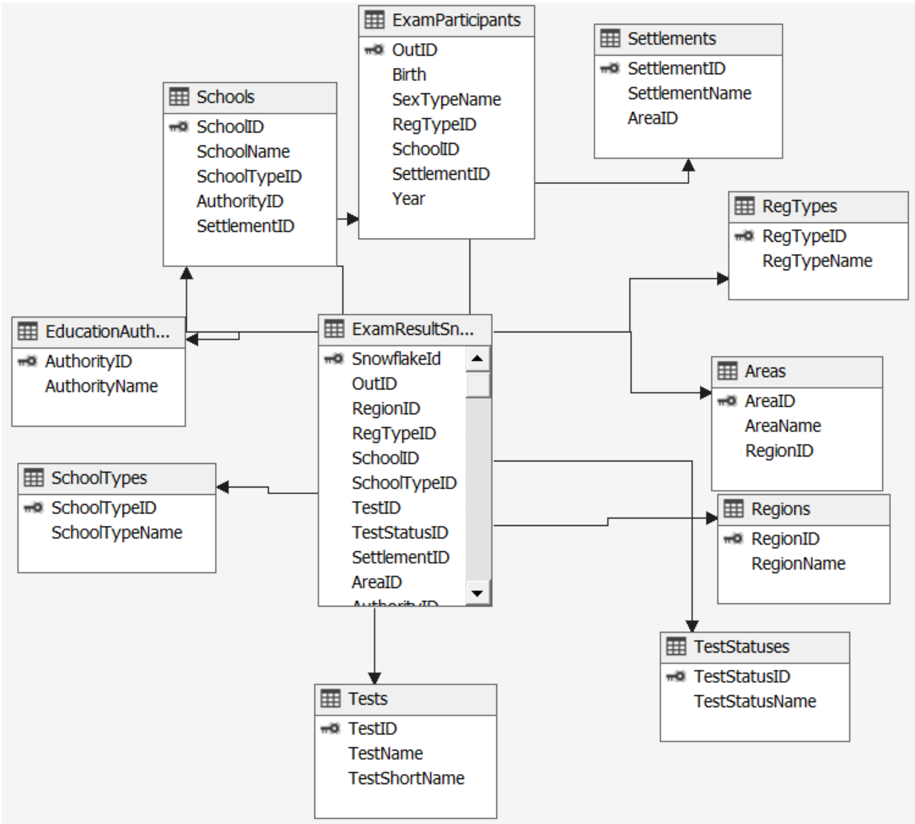

Fig. 2. Diagram of a ROLAP data storage hypercube

	Англійська мова	Біологія	Географія	Іспанська мова	Історія України	Математика	Німецька мова	Українська література	Українська мова	Фізика	Французька мова	Хімія
Вінницька область	6342	3100	2574	4	24268	26593	55	1541	28283	399	4	193
Волинська область	4592	2667	2196	1	20116	21882	57	1659	23156	271	4	210
Дніпропетровська область	13744	5885	2860	12	42176	48668	300	2584	51435	849	52	450
Донецька область	3323	1629	615	2	11771	13502	219	617	14152	210	19	134
Житомирська область	4807	2670	1981	2	19436	21108	74	1418	22686	280	8	162
Закарпатська область	4462	2749	1183	(null)	15768	17533	66	981	18975	163	9	156
Запорізька область	4749	1595	764	3	13129	15679	228	701	16489	198	22	138
Івано-Франківська область	5845	2957	2128	9	22317	24249	104	1932	25990	333	6	245
Київська область	8828	3537	1965	43	27517	31163	160	1841	32568	431	14	230
Кіровоградська область	3046	1995	995	4	12671	13794	44	841	14684	243	2	103
Луганська область	582	218	102	(null)	2357	2715	38	89	2831	34	5	18
Львівська область	15462	3822	3457	12	44986	50743	235	2609	52905	422	22	483
м.Київ	26073	4691	2509	65	53081	65635	924	2423	68041	1063	111	552
Миколаївська область	3661	2150	751	(null)	12350	14041	102	812	15006	319	7	140
Одеська область	13948	4008	2152	5	32597	38934	295	1748	41203	577	32	334
Полтавська область	5357	3239	1828	1	20061	22315	61	1280	23732	410	4	161
Рівненська область	4518	3047	2220	6	21545	23302	54	1477	24816	405	3	212
Сумська область	3561	2341	1120	2	13967	15602	80	904	16469	259	2	90
Тернопільська область	4162	1755	1834	11	16940	18391	79	1196	19500	190	5	238
Харківська область	11480	3782	1607	12	28993	34523	562	1387	36258	558	58	298
Херсонська область	1699	723	235	14	4921	5530	91	290	5782	111	1	33
Хмельницька область	5249	2576	1522	3	20903	22001	67	1509	24175	341	1	237
Черкаська область	4633	2736	1336	(null)	16270	18317	51	1199	19502	295	5	165
Чернівецька область	3177	2134	982	4	12094	13480	79	666	14441	168	48	140
Чернігівська область	3367	1957	1396	6	14401	15663	78	963	16706	273	12	121

Fig. 3. Fragment of a slice operation's result

2. Dice is an operation that allows the creation of a subcube using multiple dimensions [14]. For a given hypercube, it is appropriate, for instance, to create a subcube of C_{time_reg} based on a specific value of the testing year and the selected region like:

$$C_{time\_dice} = \{d_3\}\ where\ (time='2023')\ and\ \ (region='Дніпропетровська'). \qquad (6)$$

Fig. 4 shows the results of the dice operation when the hypercube dimensionality is reduced according to the condition specified in (6).

	Англійська мова	Біологія	Географія	Іспанська мова	Історія України	Математика	Німецька мова	Українська література	Українська мова	Фізика	Французька мова	Хімія
Дніпропетровська область	13744	5885	2860	12	42176	48668	300	2584	51435	849	52	450

Fig. 4. Dice operation's example

3. Rotation (pivot) — the implementation of the C_{pivot} rotation operation leads to a change in the position of the cube measurements as:

$$C_{pivot} = \{d_1, d_2\} \longrightarrow \{d_2, d_1\} \tag{7}$$

Rotation changes the viewpoint of the data [12]. This operation leads to a swapping of the rows and columns of the result and is used to improve the convenience of data perception. For the proposed hypercube, the results of the rotation are shown in Fig. 5.

	Вінницька...	Волинська об...	Дніпропетровськ...	Донец...	Житомирськ...	Закарпатськ...	Запорізька...	Івано-Франк...	Київсь...	Кіровоградс...	Луганська...	Львівськ...	м.Київ	Миколаї...	Одеська...	Полтавс...	Рівненська...	Сумсь...
Англійська мова	6342	4592	13744	3323	4807	4462	4749	5845	8828	3046	582	15462	26873	3661	13848	5357	4518	3561
Біологія	3100	2667	5885	1629	2870	2749	1595	2357	3537	1995	218	3822	4691	2158	4008	3239	3047	2341
Географія	2574	2196	2860	615	1981	1183	764	2126	1965	995	102	3457	2509	751	2152	1828	2220	1120
Іспанська мова	4	1	12	2	2	(null)	3	9	43	4	(null)	12	65	(null)	5	1	6	2
Історія України	24268	20116	42176	11771	19435	15768	13129	22317	27517	12571	2357	44996	53081	12350	32597	20061	21545	13967
Математика	26593	21882	48668	13502	21108	17533	15678	24249	31163	13784	2715	50743	65535	14041	38934	22315	23382	15602
Німецька мова	55	57	300	219	74	66	228	104	160	44	38	235	924	102	295	61	54	80
Українська літ...	1541	1659	2584	617	1418	981	701	1932	1841	841	89	2609	2423	812	1749	1260	1477	904
Українська мо...	28283	23156	51435	14192	22686	18975	16489	25950	32988	14684	2831	52905	68041	15006	41203	23732	24816	16469
Фізика	399	271	849	210	280	163	198	333	431	243	34	422	1063	319	577	410	405	259
Французька м...	4	4	52	19	8	9	22	6	14	2	5	22	111	7	32	4	3	2
Хімія	193	210	450	134	162	156	138	245	230	103	18	483	552	140	334	161	212	90

Fig. 5. Fragment of a rotate operation's result

4. Drill-down (decomposition) defines the transition from aggregated data to a more detailed presentation [2]. This operation is utilized for a hierarchy of dimensions with subordinate values, where an additional level is incorporated into the dimensional hierarchy. Figure 6 illustrates the results of detailing the hypercube data with the «Territorial Distribution» detailing input parameter, where the «Region» dimension is expanded to the «District» level.

	Англійська мова	Біологія	Географія	Іспанська мова	Історія України	Математика	Німецька мова	Українська література	Українська мова	Фізика	Французька мова	Хімія
Дніпровський район	403	335	202	(null)	1984	2086	6	139	2293	31	2	22
Кам'янський район	472	411	186	1	2233	2408	3	156	2604	43	(null)	25
Криворізький район	172	226	83	(null)	1152	1171	4	91	1288	22	(null)	7
м.Дніпро	6240	1708	832	5	14713	17940	134	758	18701	319	37	167
м.Кам'янське	1038	581	264	1	3332	3836	18	210	4062	84	6	27
м.Кривий Ріг	2866	918	514	1	8826	10191	83	589	10664	168	5	106
Нікопольський район	810	453	186	(null)	2667	3063	27	174	3262	59	(null)	30
Новомосковський район	546	396	143	2	2141	2376	10	159	2567	28	(null)	20
Павлоградський район	691	442	238	2	2667	2992	11	149	3172	61	2	21
Синельниківський район	506	415	212	(null)	2461	2605	4	159	2822	34	(null)	25

Fig. 6. Fragment of a drill-down operation's result

5. Roll-up (aggregation) is the opposite of drill-down operation, defines the transition from a detailed data representation to an aggregated one [15]. The operation is used for a hierarchy of measurements with subordinate values, when one of the measurement values is replaced by another one of a higher hierarchy level. The consolidation operation, where the input parameter is a territorial unit, when information is combined from the entire hierarchy, returning the total values of indicators for the whole of Ukraine, as shown in Fig. 7.

Англійська мова	Біологія	Географія	Іспанська мова	Історія України	Математика	Німецька мова	Українська література	Українська мова	Фізика	Французька мова	Хімія
167367	67571	40710	221	524144	596132	4095	32518	630245	8802	456	5243

Fig. 7. Roll-up operation's result

The analysis of multidimensional data by means of hypercubes is advisable to use when the number of records in the storage tables exceeds 5000. The total time T of data preparation and multidimensional cube construction for n iterations of analysis of a ROLAP storage represented as a «snowflake» model can be calculated as follows:

$$T_{total} = O(D \times F) + O(D \times L) + O(F \times D \times \log(F)) + O(\textit{complexity factor}), \tag{8}$$

where $O(D \times F)$ — time of data join operations from tables linked by a hierarchical relationship, $O(D \times L)$ — hypercube construction time, where L — number of each dimension levels, $O(F \times D \times \log(F))$ — data indexing time to improve the efficiency of data query execution,

O(*complexity factor*) — additional time, depending on the complexity of the model and the presence of nested dimensions and foreign keys.

Hypercube data processing performance depends on the correct construction of a multidimensional database model and the definition of primary and foreign keys.

Conclusions

This paper presents the results of substantiating the feasibility of using hypercube and ROLAP technology to analyze multidimensional high structured data. ROLAP technology facilitates the analysis of large volumes of data stored in relational data storage through the use of multidimensional models. The primary advantages of utilizing ROLAP cubes for multidimensional data analysis include high processing speeds for substantial datasets, the implementation of various types of numerical and statistical analyses, the capability for simultaneous multi-user access to data, and robust support for multidimensional analysis, which ensures a comprehensive examination of problematic aspects.

The use of hypercube operations, such as slice, dice, roll-up, drill-down, and pivoting, to analyze statistical data can significantly reduce the time required to conduct research compared to traditional queries to RDB. This is achieved through more efficient use of indexes and specialized data access mechanisms. ROLAP allows for multi-level analysis of information, which is especially useful when working with complex hierarchical structures. As an example, the study used data from the results of the national Ukraine multitest, a model of a relational data storage was developed, and a hypercube for data analysis was built.

However, the main disadvantage of using OLAP technologies, in particular ROLAP, is their increased sensitivity to the correct design of the data storage schema. Incorrectly defining indexes, foreign keys, or the schema of facts and measurements can lead to a significant decrease in system performance. In addition, the difficulty in analyzing dynamically changing data is another significant problem, as the structure of OLAP cubes does not always adapt to changes in real time.

References

[1] Samotyy, V., Dzelendzyak, U. (2016). Multivariate data analysis using OLAP cubes in billing system. Bullitin of LSU, 13, 32—38.

[2] Yalova, K., Yasyna, K. Konratenko, A. (2019). OLAP technology for the relational databases analysis. LOGOS, 8, 69—73.

[3] Kharkianen, O. V., Gladka, Y. A. (2019). Information support for product sales by intellectual data analysis methods. Modelling and information systems in economy, 2019, 209—214. doi: 10.33111/mise.98.20.

[4] Sarsimbaeva, S., Dimitrov, V. (2019) Multidimensional data analysis with OLAP. CEUR, 2464, 95—98.

[5] Kovacic, I., Schuetz. C., Neumayr, B., Schrefl, M. (2022). OLAP Patterns: A pattern-based approach tomultidimensional data analysis. Data & Knowledge Engineering, 138, 101948.

[6] Zhou, X.-H., Zhang, X.-M. (2017). The application of OLAP and Data mining technology in the analysis of book lending. Advances in Engineering Research, 118, 368—373.

[7] Leveraging OLAP for Advanced Data Analysis. URL: https://www.kandasoft.com/blog/leveraging-olap-for-advanced-data-analysis

[8] Yadav, A., Singh, B. (2023). Improve the performance of multidimensional data for OLAP by using an optimization approach. AIP Conference Proceedings, 2745. doi: https://doi.org/10.1063/5.0132474

[9] Mao, Y., Huang, S., Cui, S., Wang, H., Zhang, J., Ding, W. (2020). Multi dimensional data distribution monitoring based on OLAP. *Proceedings of the 2nd International Conference on Information Technology and Computer Application* (ITCA'20), pp. 1—15, Guangzhou, China.

[10] Francia, M., Marcel, P., Peralta, V., Rizzi, S. (2022) Enhancing Cubes with Models to Describe Multidimensional Data. Information System Frontiers, 24(1), 31—48. doi: 10.1007/s10796-021-10147-3.

[11] Cuzzocrea, A., Maio, C., Fenza, G., Loia, V., Parente, M. (2016) OLAP analysis of multidimensional tweet streams for supporting advanced analytics, *Proceedings of the 31st Annual ACM Symposium on Applied Computing* (SAIC'16), pp. 992—999, New York, USA. doi: https://doi.org/10.1145/2851613.2851662.

[12] Li, X., Shen, Q., Yang, T. (2024). Design and optimization of multidimensional data models for enhanced OLAP query performance and data analysis. Applied and Computational Engineering, 69, 161—166. doi: https://doi.org/10.54254/2755-2721/69/20241503.
[13] Neha, Garg, K. (2017). Efficient information retrieval using multidimensional OLAP cube. International Research Journal of Engineering and Technology, 4(7), 2885—2889.
[14] Appah, B., Amos, D. (2018). Multidimensional data model for health service decision making data. International Journal of Computer Science Engineering Techniques, 3(3), 1—6.
[15] Hu, R. (2017). The design of OLAP-based muti-dimensional analysis system for hospital information. Scholars Journal of Engineering and Technology, 5(2), 76—80.

Список використаної літератури

1. Самотий В. В. Дзелендзяк У. Ю. Багатовимірний аналіз даних у білінговій системі з використанням OLAP кубів. *Вісник ЛДУ*. 2016. №13. С. 32—38.
2. Ялова К. М., Яшина К. В., Кондратенко А. І. Технологія OLAP для аналізу реляційних баз даних. *LOGOS*. 2019. №8. С.69—73.
3. Харкянен О. В., Гладка Ю. А. Інформаційна підтримка збуту продукції методами інтелектуального аналізу даних. *Моделювання та інформаційні систем в економіці*. 2019. №98. С. 209—214. DOI: 10.33111/mise.98.20.
4. Sarsimbaeva S., Dimitrov V. Multidimensional data analysis with OLAP. *CEUR*. 2019. 2464. P. 95—98.
5. Kovacic I., Schuetz C., Neumayr B., Schrefl M. OLAP Patterns: A pattern-based approach tomultidimensional data analysis. *Data & Knowledge Engineering*. 2022. Vol. 138. P.101948.
6. Zhou X.-H., Zhang X.-M. The application of OLAP and Data mining technology in the analysis of book lending. *Advances in Engineering Research*. 2017. Vol. 118. P. 368—373.
7. Leveraging OLAP for Advanced Data Analysis. URL: https://www.kandasoft.com/blog/leveraging-olap-for-advanced-data-analysis (дата звернення: 30.08.2024).
8. Yadav A., Singh B. Improve the performance of multidimensional data for OLAP by using an optimization approach. *AIP Conference Proceedings*. 2023. Vol. 2745. doi: https://doi.org/10.1063/5.0132474.
9. Mao Y., Huang S., Cui S., Wang H., Zhang J., Ding W. Multi dimensional data distribution monitoring based on OLAP. *Proceedings of the 2nd International Conference on Information Technology and Computer Application* (ITCA'20), Guangzhou, China, 2020. pp. 1—15.
10. Francia M., Marcel P., Peralta V., Rizzi S. Enhancing Cubes with Models to Describe Multidimensional Data. *Information System Frontiers*. 2022. 24(1). P. 31—48. doi: 10.1007/s10796-021-10147-3.
11. Cuzzocrea A., Maio C., Fenza G., Loia V., Parente M. OLAP analysis of multidimensional tweet streams for supporting advanced analytics, *Proceedings of the 31st Annual ACM Symposium on Applied Computing* (SAIC'16), New York, USA, 2016. pp. 992—999. doi: https://doi.org/10.1145/2851613.2851662.
12. Li X., Shen Q., Yang T. Design and optimization of multidimensional data models for enhanced OLAP query performance and data analysis. *Applied and Computational Engineering*. 2024. Vol. 69, P. 161—166. doi: https://doi.org/10.54254/2755-2721/69/20241503.
13. Neha, Garg K. Efficient information retrieval using multidimensional OLAP cube. *International Research Journal of Engineering and Technology*. 2018. Vol. 4(7). P. 2885—2889.
14. Appah B., Amos D. Multidimensional data model for health service decision making data. *International Journal of Computer Science Engineering Techniques*. 2018. Vol. 3. Issue 3. P. 1—6.
15. Hu R. The design of OLAP-based muti-dimensional analysis system for hospital information. *Scholars Journal of Engineering and Technology*. 2017. Vol. 5. Issue 2. P. 76—80.

Надійшла до редколегії 14.10.2024