

ӘОЖ 004.65

Гумарова А.А., МИСҒ-21

Ғылыми жетекші: Камалова Г.А., ф.-м.ғ.к.

Қуанғалиев Т.Г., т.ғ. магистрі

Ілиясов Т.Ж., т.ғ. магистрі

Жәңгір хан атындағы Батыс Қазақстан аграрлық-техникалық университеті, Орал қ.

КӨП ӨЛШЕМДІ ДЕРЕКТЕР ҚОРЫНЫҢ СИПАТТАМАСЫ ЖӘНЕ МОДЕЛІ

Андатпа

Мақалада көп өлшемді талдауды қолдана отырып деректерді өңдеу аспектілері қарастырылған. Көп өлшемді модельдің негізгі ұғымдары келтірілген, OLAP жүйелерінің жіктелуі қарастырылған.

Түйін сөздер: OLAP сервер, көп өлшемді OLAP текшесі, «Звезда» құрылымы, «Снежинка» құрылымы, OLAP клиент.

Көп өлшемді мәліметтердің шығуы және OLAP (On-line Analytical Processing - мәліметтердің жедел талдауы) қазіргі таңда информациялық-аналитикалық жүйелер құрастыруының кең таралған тұжырымдамаларының бірі. OLAP терминін 1993 жылы E. F. Codd енгізді. Кодд бойынша, көп өлшемді концептуалды ұсыныс (multi-dimensional conceptual view) басқару объектісін бағдарлаушы қызметшінің табиғи көзқарасы болып табылады.

OLAP мәні - мәліметтің талдауы үшін бастапқы түрі көп өлшемді куб түрінде көрсетіледі, сонымен қатар онымен манипуляция жасауға қамтамасыз етілген және қажетті ақпараттық резервтерді - есептемені беру болып табылады. Сонда соңғы пайдаланушы әр түрлі өлшемдерде мәліметтерді (айғақтар) автоматты түрде жинақтайтын және есептеу формасы мен есептеулерді интерактивті басқаруға мүмкіндік беретін кубты көп өлшемді динамикалық кесте түрінде көреді. Бұл операциялардың орындалуы OLAP машинасымен қамтамасыз етіледі.

Мұндай жүйелер өңдеуге бағытталған және алдын ала сұраныспенен жасақталған. Мұндай жүйелердегі мәліметтер өздігінен тәртіпте сирек жүргізіледі. Әдетте жүктеме сыртқы көздерден орындалады: оперативті ДҚ , электронды кестелер. Кубтар үшін бастапқы және агрегаттық мәліметтер реляциялық және көп өлшемді деректер қорларында сақтала алады.

Мәліметтердің орналастыру тәсілі бойынша OLAP машинасының классификациясы.

OLAP мәліметтерін сақтау үшін қолданылады:

Cache (OLAP - серверлер) сияқты арнайы көп өлшемді ДҚБЖ. Осы жағдайда MOLAP(Multidimensional OLAP) туралы айтылады. Бастапқы және агрегаттық мәліметтер нақты тағайындауларға сәйкес көп өлшемді жергілікті кубтарда сақталады. Әр түрлі өлшемдердегі мәліметтерді талдайтын күрделі сұраныстарды орындағанда көп өлшемді ДҚБЖ реляциялыққа қарағанда үлкен өнімділікті қамтамасыз етеді.

Дәстүрлі реляциялық ДҚБЖ - ROLAP (Relational OLAP) - бастапқы деректер жазық кестелерде сақталады. Агрегаттық мәліметтер сол ДҚ-да қызметтік кестелерде жайғастырыла алады. Реляциялық ДҚ-дан көп өлшемді кубқа деректерді өзгерту OLAP-тың құралының сұранысы бойынша болады. ROLAP аспаптары реляциялық ДҚБЖ-мен метамәліметтер деңгейінде байланысады, ол мәліметтердің көп өлшемді тұрақты құрылымын жасамауына мүмкіндік береді. Арнайы мәліметтер құрылымдарының қолдану - "Звезда" (star) схемасы және "Снежинка" (snowflake), реляциялық мәліметтердің көп өлшемді талдауын жасай алады.

Құрамалы вариант - HOLAP (Hybrid OLAP), ДҚБЖ-ның бірнеше түрін сәйкестендіреді. ДҚБЖ-ның екі типін қосатын варианттардың бірі - ол көп өлшемді ДҚБЖда агрегаттардың, ал мәліметтер бұйымын (көлем болатын ең үлкен) реляциялықта сақтауы болып табылады. OLAP кубасын құрастыру реляциялық және көп өлшемді мәліметтердің негізінде OLAP құралының сұранысы бойынша орындалады.

Келесі классификация – OLAP машинасының орналасуы. Бұл белгі бойынша OLAP өнімдері OLAP - сервері және OLAP – клиенттерге жіктеледі:

- серверлік OLAP құралдарында агрегаттық мәліметтердің есептеулер және сақтауы жеке процесстермен – сервермен орындалады. Клиент қосымшасы серверлерде сақталатын көп өлшемді кубтардың сұраныс нәтижесінен кейін алады. Кейбір OLAP - серверлер тек қана реляциялық

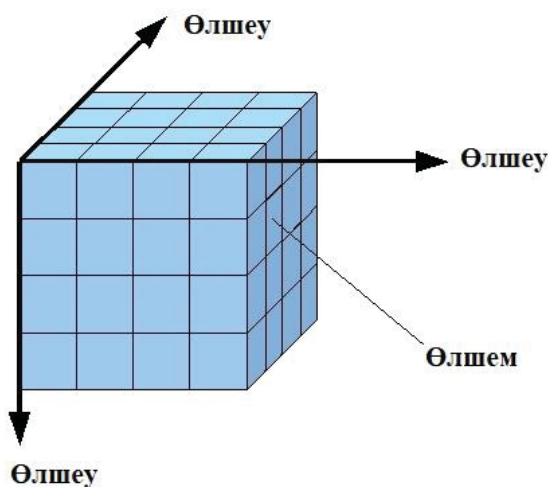
базаларындағы мәліметті сақтайды, кейбір - тек қана көп өлшемді. Көпшілік қазіргі OLAP - серверлер мәліметті сақтауды барлық үш әдістерін қолдайды: MOLAP, ROLAP және HOLAP.

OLAP - клиент өзгеше жасалған. Көп өлшемді куб пен OLAP - есептеу клиент компьютерінің жадтарында орындалады. OLAP –клиенттер сонымен бірге ROLAP және MOLAP жіктеледі. Бұл тұжырымдаманы MQEның сұранысты басқаратын орта жүзеге асырады.

Бұл аспаптар реляциялық ДҚБЖда тікелей қолданылатын немесе MOLAP аралық сервердің көмегіменен талданатын шектеулі функциясын көрсетеді. MQE аспаптары ДҚБЖ-ғы (тікелей немесе MOLAPсы сервері арқылы) мәліметтерді стол компьютерге немесе сақталатын, талданатын және жергілікті талданатын мәліметтердің куба түріндегі жергілікті серверге жібереді. MQE аспаптарының типті архитектурасы суретте көрсетілген.

OLAP қажетті ақпараттың анализін, оны көру мен қол жеткізетін құралдың жылдам жұмыс істеуін көрсетеді. Қолданушы мәліметтерді көп өлшемді кубтар (Cubes) түрінде ұйымдастыра отырып, мәліметтердің интуициялық түсінігін алады оларды Көп өлшемді координаталар жүйесінің өстеріне талданатын бизнес- процесстің негізгі атрибуттары қызмет көрсетеді. Мысалы, бұл сатылымдарға тауар, өлке, сатып алушының түрі бола алады. Өлшемдердің бірі ретінде уақыт қолданылады. Өстердің қиылысуларында - (Dimensions) өлшемдерде - (Measures) шара процессін санмен сипаттайтын мәліметтер болады. Бұл ақшалай өрнектелген немесе дана түріндегі тауар көлемі, қоймадағы қалдықтар және т.б. болуы мүмкін. Ақпаратты талдайтын қолданушы кубты түрлі бағытта кесе алады, жыл бойынша немесе апта бойынша берілген мәліметтерді алады және талдау процесінде оның ойына келген бәрнеше манипуляцияларды орындайды.

Суретте бейнелелген үш өлшемді кубта шаралар ретінде тауарлар сомасы пайдаланған, ал өлшемдер ретінде - уақыт, тауар және дүкен. Өлшемдер топтаудың нақтылы деңгейлеріне көрсетілген: тауарлар дәрежелер бойынша топталады, дүкендер - елдер бойынша, операциялардың іске асыруды уақыты туралы мәліметтер - айларға.



Сурет 1. Куб мысалы

Мәліметтердің көп өлшемді моделі және деректермен манипуляция жасау операциясы

Мәліметтердің көп өлшемді үлгісінің негізгі ұғымдары болып табылады:

- мәліметтердің гиперкубі (Data Hypercube);
- өлшем (Dimension);
- таңбалар (Members);
- ұяшық (Cell);
- шара (Measure).

Мәліметтердің гиперкубінде бір немесе бірнеше өлшемдер бар және ұяшықтардың ретке салған жиыны болады.

Ұяшықта мәліметтер болуы мүмкін - шаралы немесе бос болады. . Ұяшықтың термині орынына кейде көрсеткіш термині қолданылады. Оның мәндері қалыптасуына байланысты көрсеткіш анықталады:

- айнымалы (Variable) - мұндай көрсеткіштердің мәндері қандай болмасын сыртқы көздерден бір рет жүргізіледі немесе программалы қалыптасады және содан соң анық түрде (Мбд) көп өлшемді деректер қорларында сақталады;

- формула (Formula) - мұндай көрсеткіштердің мәні кейбір алдын ала спецификацияланған формула бойынша есептеледі. Формула, ДҚ-ға көрсеткіш болатын түр үшін емес, бұл мәндер бойынша есептейтін формуламен сақталады.

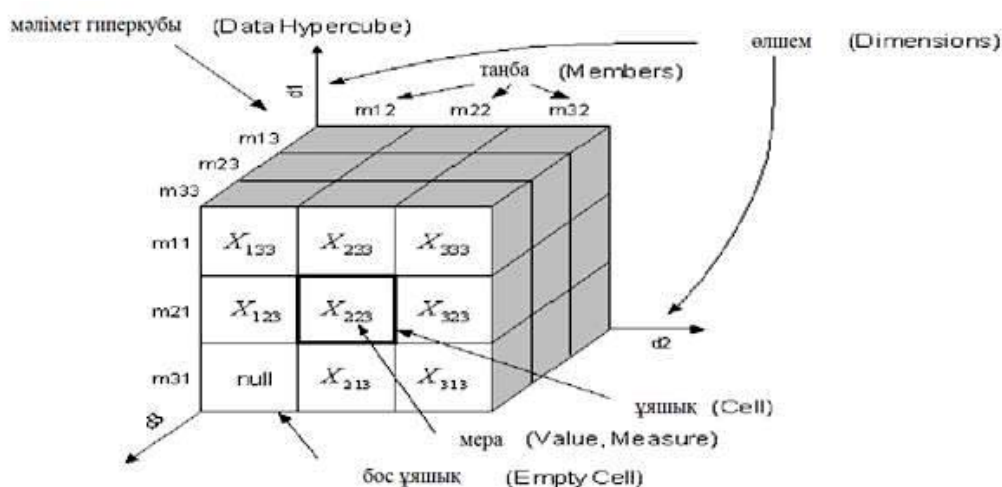
Әрбір ұяшықпен өлшем мәндерінің бір-ақ жиынымен анықталады – ол таңбалар. Өлшеммен гиперкубтің қырларының бір жасаушы таңбалардың жиынын түсінеміз. Өлшем - бұл біртепті мәліметтердің жиыны.

Географиялық өлшемдердің мысалдары болып табылады: қалалар, аудандар, өлкелер, ел және тағы басқалар. Мәліметтің көпөлшемді моделінде Өлшем гиперкуб ұяшығында болатын (көрсеткіштер) нақты мәндер үшін қолданылатын индекстер ролін ойнайды.

Уақытша өлшемнің мысалы күндер, айлар, тоқсандардың тізімі болып табылады. Географиялық өлшемнің мысалымен аймақтық объекттердің тізімі бола алады: халық пункті, аудандар, өлкелер, елдер және тағы басқаларлар.

Қолданушы мәліметке қол жеткізу үшін қажетті ұяшықтарға сәйкес келетін өлшемдердің мәндерін таңдауда бір немесе бірнеше ұяшықтар жолымен көрсетілуі керек. Өлшем мәндерін таңдауы процесі таңбаларды бекіту деп атайық, ал өлшемдердің таңдаулы мәндер жиынын - бекітілген таңбалар жиыны дейміз.

Айғақ (шара) субъекттерді ұсынады – талдауға қажетті белгісіз үлгі немесе оқиға. Мәліметтердің көп өлшемді үлгілерінің көпшілігінде айғақтар өлшем мәндерінің комбинацияларымен анықталады; айғақ тек қана мәндердің нақты комбинациясы үшін арналған ұяшық бос емес болғанда ғана бар болады. Бірақта, кейбір үлгілер айғақтарды ерекше қасиеттері бар бірінші класстың объекттері деп түсіндіреді. Сонымен бірге көп өлшемді үлгілердің көпшілігі әрбір айғақтарға әрбір өлшемнің деңгейі төмен болатындай және бір мәнді болуын талап етеді.



Сурет 2. Куб мысалы

$$D = \{d_1, d_2, \dots, d_n\} - \text{гиперкуб өлшемі}$$

$$M_{d_i} = \{m_{1_i}, m_{2_i}, \dots, m_{k_i}\}, i = 1, \dots, n - \text{өлшемнің көпшілік белгілері}$$

$$M = M_{d_1} \cup M_{d_2} \cup \dots \cup M_{d_n} - \text{гиперкубтің көпшілік белгілері}$$

$$D' \subseteq D - \text{өлшемнің көпшілік бекітілімі}$$

$$M' \subseteq M - \text{көпшілік бекітілім белгілері.}$$

Әрбір айғақ өлшемдер мәндерінің комбинациясы жасалатын анықталған деңгейге ие. Мәліметтер қоймаларында айғақтардың келесі үш түрлері болады.

Оқиғалар (event), кем дегенде ең үлкен айғақты, әдеттегідей нақты дүниенің оқиғаларын пішіндейді мұнымен әрбір айғақ зерттелетін құбылыстың нақты данасын ұсынады. Мысал ретінде

сатылымдар, Web беттерінде тышқанның шертулері немесе тауарлардың қоймасындағы қозғалысты жатқызамыз.

Лездік суреттер (snapshot) дәл осы уақыттағы объектінің күйін пішіндейді, мысалы, дүкенде немесе қоймадағы тауарлардың бар болу деңгейлері және Web қолданушыларының саны. Нақты дүние құбылысының бір данасы, мысалы, бобаның нақты банкісі бірнеше айғақтарда пайда бола алады.

Жиынтық лездік суреттер (cumulative snapshot) нақтылы уақыт мерзіміндегі ұйым қызметі туралы мәліметтерді құрайды. Мысалы, алдыңғы мерзім бойынша сатылымның жиынтық көлемін, ағымдағы айды қоса алғанда, өткен жылдың айлардың көрсеткіштеріменен оңай салыстыруға болады.

Бұл айырмашылықтар тек қана жобалау кезеңінде болады және саоңғы пайдаланушылардан толық бұрмеленген.

МДБ мәліметтерді өлшемдердің кез келген санға кестелерінің жалпылаулары болып табылған куб түрінде қарайды. Тиісті кубтар жиыны көп өлшемді деректер қорын (немесе мәліметтердің қоймасы) құрайды. Қазіргі құрал-сайман гиперкубтың 10-15 өлшемдерден артық болғанда өнімділіктің жетіспеуімен жиі қақтығысады. Өлшемдердің жаңа мәндерін толықтыра отырып, текшелерді оңай басқаруға болады.

Әр түрлі МДҚБЖ-да мәлімет ұйымының екі негізгі варианттары қолданылады:

- гиперкуб үлгісі;
- поликубтық үлгі.

Поликубтық үлгіні қолдайтын жүйелер (мысал Oracle Express Server болып табылады) МДҚ-да бірнеше анықталған әр түрлі көлемді және әр түрлі өлшемді гиперкубтер анықталуы мүмкін. Олардың өлшемдері ретінде. Гиперкубтық үлгіде барлық көрсеткіштер өлшемдердің бір жиынымен ғана анықталуы керек болатынын есептейді.

Көп өлшемді деректер қорларының OLAP серверлері көп өлшемді мәліметтерді қысылған түрінде сақтауға ие. Бұл бар кеңістіктерді тиімді қолдануға барынша мүмкіндік беретін физикалық мәліметті сақтаудың динамикалық таңдау және қысу технологиясы негізінде іске асады. Тығыздап қапталған мәліметтер (яғни бос ұяшықтарда кіші кубтың бір бөліктері орналасатын мәліметтер) қиылған мәліметтерден бөлек сақтала алады. Кеңістікті қолдануды оптимизациялай отырып, мәліметтерді физикалық сақтау құралдарымен көрсетілетін, мәліметтердің үлкен көлемін талдауға мүмкіндік беретін OLAP серверлері талаптарды минимумға дейін қысқартады. Сонымен бірге бұл компьютердің оперативті жадысына үлкен мәліметтерді тасымалдауға мүмкіндік береді, осыдан атқарылатын енгізу-шығару операцияларының саны өнімділіктің жоғарылауына байланысты қысқартылады.

OLAP негізінде көп өлшемді деректер қорының серверлері төменде аталған негізгі аналитикалық операцияларды орындай алады.

Консолидация (агрегация) - бұл мәліметтерді талдауда қолданылатын анализдің арнайы техникасы. Қолданушы мәліметтегі бұйымның бағытымен жоғары (up) агрегатталғанға (down) өтуге мүмкіндігі бар. Ол мұндай агрегациялайтын операциялар қосады: мәндердің (орам) қарапайым жинақталуы немесе аралық мәліметті қосатын күрделі өрнектердің қолдануы бар есептеу.

Агрегация деңгейінің таңдау белгілері: егер қолданушыдан сұрайтын болсақ, оған тәптіштеу қандай деңгейі қажет десек, ол ойлап тұрмастан максималды мүмкіндігі бар деп жауап берер еді. Дегенмен мынаны ескеру керек: бұл шешім қанша тұрады және тәптіштеудің әрбір жаңа деңгейде мәліметтердің бар болуының мүмкін болатын экономикалық эффектті анықтауға тырысу керек. Мысалы, агрегация деңгейі ретінде Жылды таңдасақ, сіз жалпы тенденцияны талдау мүмкіндігіне ие боласыз және оның даму динамикасын болжай аласыз. Агрегация деңгейі ретінде Айды немесе Аптаны таңдасақ, сіз онымен қоса нақты уақыттағы нақты заттарға сұранысты болжай аласыз. Бұл маусымдық ауытқулар, өз қоймасын ұтымдылау қалыптастырып және маусымды жеңілдіктер және сатылымдарды тиімдірек құрастыруда саясат жүргізуге, іздеп табуға мүмкіндік береді. Егер жүйеге маркетингте шығындар туралы мәлімет енгізілсе, әрбір нақты маркетинг шарасынан эффектлі бақылауға мүмкіндігі пайда болады. Детализацияның әрбір келесі деңгейіне өткенде және жаңа мәліметтер көздерін қосқанда жоғарылап кетуі мүмкін.

Бәсеңдейтін талдау (түсіру, тәптіштеу). Қарастырылатын консолданған мәліметтер үшін кейбір ақпараттарды көрсететін операция, кері консолидация. Мысалы, талдауды Аймақ деңгейінен бастаса, қолданушы нақты бөлімше немесе менеджердің жұмысы туралы нақты мәліметті алғысы келеді.

Бұрылыспен бөліктеу. Бұл операциялар (өлшемдермен манипуляция жасау) әртүрлі көзқарастармен алынған мәліметтердің берілуін алуға мүмкіндік береді. Мысалы, мәліметтердің бір

кесігі әрбір қала бойынша көрсетілген типтің жылжымайтын мүлік объектілерінің сатылғандағы табыстары туралы барлық хабарларды бейнелеу. Басқа кесік қалалардың әрқайсыларындағы серіктестіктердің бөлімшелердің табыстары туралы мәліметтерді дайындау алады.

«Срез» құрастыру. Қолданушыны өлшем мәндерінің потенциалды мүмкін болатын комбинациялары аз қызықтырады, ол бірдей барлық мәліметтер гиперкубымен жұмыс істемейді. Гиперкубті өлшемдер бір немесе көп мәнді бекітудің нәтижесінде жасалған ішкі жиын «Срез» (Slice) деп аталады.

«Вращение» операциясы. Өлшемдерді (әдетте екі өлшемді мәліметтердің берілуінде қолданылады) көрсету ретінің өзгеруі «Вращения» деп аталады. Бұл операциялар формадағы деректерді көрнекілеу мүмкіндігін қамтамасыз етеді, бұл олардың қабылдауы үшін өте жайлы.

ҚОЛДАНЫЛҒАН ӘДЕБИЕТТЕР ТІЗІМІ

1. Чаусов В., Амириди Ю. Классификация аналитических систем. Три года спустя // Журнал «Банки и технологии» – 2002. – №6.
2. Послед Б.С. Borland C++ Builder 6. Разработка приложений баз данных. – Спб.: ДиасофтЮП, 2003. – 320 с.
3. Microsoft Office PerformancePoint Server [Электронный ресурс] – URL: <http://www.microsoft.com/business/performancepoint/productinfo/previousversions.aspx>
4. Codd E.F. Providing OLAP (online analytical processing) to useranalysts: An IT mandate. E.F. Codd and Associates, 1993.
5. Intersoft Lab [Электронный ресурс] – URL: <http://www.iso.ru/>
6. Бергер А. Microsoft SQL Server 2005 Analysis Services. OLAP и многомерный анализ данных. СПб.: БХВ-Петербург, 2007.
7. Гудков А. С. Использование префиксных деревьев при построении систем анализа данных: автореф. дисс. ... канд. физ.-мат. наук. М., 2006.
8. Заботнев М. С. Разработка методов и средств анализа многомерных баз данных с неполной информацией: автореф. дисс. ... канд. техн. наук. М., 2006.
9. Ишенин П. П. Инструментальные средства построения комплексов моделей и аналитических приложений в OLAP-технологии: дисс. ... канд. техн. наук. Красноярск, 2006.
10. Кудрявцев Ю. А. Алгоритмы эффективной обработки MOLAP-кубов: дисс. ... канд. физ.-мат. наук. М., 2009.

РЕЗЮМЕ

Анализ современного состояния OLAP-технологий позволяет говорить о важнейших перспективах их развития. Многомерная обработка информации становится необходимой составляющей любого хранилища данных. В то же время широкие подходы к реализации таких систем и отсутствие адекватной интеграции форматов хранения и алгоритмов обработки данных заставляют искать новые подходы к описанию и разработке OLAP-систем.

RESUME

The analysis of the current state of OLAP technologies allows us to talk about the most important prospects for their development. Multidimensional information processing becomes a necessary component of any data warehouse. At the same time, broad approaches to the implementation of such systems and the lack of adequate integration of storage formats and data processing algorithms force us to look for new approaches to the description and development of OLAP systems.

ӘОЖ 004.92.075.8

Мухамедяр Ұ.Е., МИСФ-21

Ғылыми жетекші: **Қасымова А.Х.**, доцент, п.ғ.к.

Жәңгір хан атындағы Батыс Қазақстан аграрлық-техникалық университеті, Орал қ.

ҚАШЫҚТАН БАСҚАРЫЛАТЫН РОБОТТАР МЕН МАНИПУЛЯТОРЛАР

Андатпа

Қазіргі тағда әлем ғылыми зерттеулердің өрістерін, компьютерлік технологиялардың, информатика мен басқару теорияларының соңғы жетістіктеріне негізделген заманауи роботтарды