

**ФУНКЦИИ И УСЛУГИ В СИСТЕМИТЕ ЗА АНАЛИТИЧНА
ОБРАБОТКА НА ДАННИ****FUNCTIONS AND SERVICES IN THE SYSTEMS FOR ANALYTICAL
DATA PROCESSING****Ivanka Valova***Bulgarian Academy of Sciences***Abstract***This article studies some OLAP functions and services are used by many software products.***Keywords:** OLAP functions, Drill down, Drill up, Currency Translation, Selecting Constants.**ВЪВЕДЕНИЕ**

Системите за аналитичната обработка на данни в реално време (OLAP) се използват за анализ на големи обеми оперативни и исторически данни. Процесорът OLAP позволява обработка на многомерни модели в съответствие с различни бизнес перспективи. Складът на метаданни (фиг. 1) позволява информацията, намираща се в обекти с метаданни за OLAP-системата да бъде показана в съответната работеща система или да бъде използвана в дадена операция. За целите на планирането се създават специални приложения- BPS. Сферите на приложение варират от ръчно въвеждане на данни до сложни сценарии по планиране. За целите на планиране на отчетни функции във фоновата среда се използва програмен инструмент наречен Агент за отчети [1].

Специални процеси за анализ, като добив на данни-DM могат да бъдат реализирани чрез т.нар. Инструмент за анализ на процеси (APD). Чрез използването на APD, наличната информация може да бъде обработена, с цел генериране на нова информация. Продуктът на SAP Business Information Warehouse (Хранилище на бизнес информация) използва OLAP технология за анализ на данни. Аналитичната обработка на данни в реално време определя BW-системата като система за поддръжка и вземане на решения. OLAP технологиите позволяват на ръководителите, вземащи решения да анализират бързо и интерактивно многомерните модели, в съответствие с техните бизнес съображения. Доставчиците на данни (InfoProviders)

позволяват показването на съответните данни. Данните се съхраняват в инфокубове, които са оптимизирани за четене на данни. Инфокубовете и наборите MultiProviders трябва да бъдат разглеждани за предпочитани набори от данни [2].

Процесорът OLAP, се намира между крайния потребител и базата данни. Той позволява на многомерно форматираните данни да бъдат налични чрез използването на специални интерфейси за отворен анализ. Процесорът OLAP е оптимизиран за анализ и генериране на отчети на много големи набори от данни. Потребителите имат възможност да искат не-планирани индивидуални аспекти на бизнес данни чрез използване на Business Explorer.

Фигура 1 показва статуса и заданията на OLAP процесора в рамките на процеса за обработка на данни, изпълняван като анализ на многомерни модели:

Заявките са основата за всички анализи, извършвани в средата на OLAP система. За да бъде дефинирано формално запитване към многомерен модел, заявката определя следното:

- Структурата – аналог на работния лист;
- Филтърът, който оказва въздействие върху структурата [1];
- Пространство за навигация (свободни характеристики).

OLAP-системата разполага с поредица от функции за анализ и навигация за целите на форматирането и оценяването на данни за дадена компания. Потребителят има възможност да формулира индивидуални запитвания въз основата на набори от данни

(InfoProviders). Впоследствие, потребителят може да преглежда и оценява



Фиг. 1. OLAP система

данните от различни перспективи по време на изпълнението на заданието. Целият набор от функции, необходими за изтеглянето, обработката и форматирането на данните се осигурява от процесора.

2. ПРЕГЛЕД НА OLAP ФУНКЦИИТЕ И УСЛУГИТЕ

2.1 АГРЕГАЦИЯ

За да се изчисли стойността на ключови показатели, данните от ИнфоДоставчик (Info Provider) трябва да бъдат агрегирани. Използват се различни формули за изчисление и агрегиране на характеристики. За всеки ключов показател, една характеристика може да бъде агрегирана чрез използване на различни правила.

Процесорът OLAP в BW-системата извършва следните операции:

1. Първо се изпълнява стандартна агрегация, която може да се зададе в дефиницията на показателя. Възможни видове агрегация са сумиране (SUM), минимум (MIN), и максимум (MAX). Минимум и максимум може да се зададе, например, за ключови показатели - дата.

2. Агрегацията на избрана характеристика се извършва след стандартната агрегация - агрегация по изключение. Възможни агрегации по изключение в наличност за потребителя са средна стойност, брояч, първа стойност, последна стойност, минимум, максимум, стандартно отклонение, сумиране, дисперсия, средно отклонение и др. Случаите, в които се прилага агрегация по изключение включват, например, съхраняване на не-кумулятивни стойности, които не могат да бъдат сумирани по час, или броячи които броят количеството на характеристики за дадена специфична характеристика.

3. Агрегацията по валута и мерни единици се изпълнява последна. Ако две цифри са агрегирани нееднакво с различни валути или мерни единици, системата ги маркира със знака '*'. Формулите се изчисляват след като цифрите са били напълно агрегирани. Съществуват три изключения от това правило:

➤ Когато времето за изчисление на показателя е избрано така, че изчислението на формулата трябва да се извърши преди агрегацията.

➤ Когато променлива на дадена формула се използва със заместване от стойност на даден атрибут в определен изчислен показател.

➤ При преизчисляване на валута, която е зададена в конструктора на заявки.

2.2 ФУНКЦИЯ ЗА ЛОКАЛНИ ИЗЧИСЛЕНИЯ

Функцията се използва за преизчисляване на единични стойности показани в даден отчет съгласно специфични критерии. Локалните изчисления включват само цифрите в изчислението, които се появяват в текущия отчет и по този начин, се заместват стандартните изчисления на OLAP процесора. Локалните изчисления променят само показването на стойностите. При изчисления, с използване на формули, системата не използва тези стойности променени за целите на показването, а използва оригиналните стойности, подадени от процесора.

Например, може да се създаде списък с предварително зададени параметри от даден отчет, който показва постъпленията от продажбата на съответните продукти. След това продуктите се подреждат съобразно постъпленията и им се присвоява позиция. Може да се използват локални изчисления за създаване на отчет с активно условие, който показва първите 5 продукта по постъпления. В резултат, общата сума се показва за всички продукти, тъй като условията не оказват влияние върху резултатите. В този случай, може локално да се изчисли сумата за първите 5 продукта:

Product	Revenue		Product	Revenue
Laptop-Backpack	\$ 77.063		Laptop-Backpack	\$ 77.063
Business Card Case	\$ 75.719		Business Card Case	\$ 75.719
CoffeeMug	\$ 71.438		CoffeeMug	\$ 71.438
Multifunctional-Pen	\$ 71.260		Multifunctional-Pen	\$ 71.260
Automatic umbrella	\$ 70.086		Automatic umbrella	\$ 70.086
Overall Result	\$ 742.344	➔	Overall Result	\$ 365.566

Фиг.2. Изчисляване на резултати като обобщение

OLAP функция	Операции в подробности
Навигация	<i>Drill down</i> - Задълбочаване в характеристика/ структурен елемент-детайлизиране. <i>Премахване на елемент</i> - напречен отрязък с данни. <i>Drill up</i> - консолидация и скриване на йерархични възли-обобщаване <i>Swap</i> -Преместване на елементи при задълбочаване
Филтри	Ограничаване (<i>slice</i> - <i>надлъжен отрязък с данни</i>) на характеристики до набори (единична стойност, обхват на стойност, йерархичен елемент, изключване)
Агрегация	<i>Стандартна агрегация</i> : Формулата за изчисление е зависима от ключов показател за агрегирането на единични стойности от резултат <i>Агрегация по изключение</i> (настройка за специална агрегация във връзка със специфична характеристика. Например, средна агрегация на салдо по сметка с позоваване на час на характеристика) <i>Локална (местна) агрегация</i> или местно изчисление (напр., изчисление на индивидуални стойности показано за целите на нормиране от цялостния резултат)
Разбивка (план)	Разбивка на характеристиките като ключ /идентификатор Показване/скриване на редове с резултати Смяна на позиция на йерархични възли (нагоре/надолу)
Избор и разбивка в зависимост от резултатите	Прагови стойности (Изключения)(Оцветяване в цвят на необичайна дисперсия средно отклонение в ключовите показатели) Условия: Ограничение на характеристики зависими от показатели съгласно дефинирани условия
Структуриране	Йерархично присвояване на стойности на характеристики със (<i>drill down</i>) задълбочаване за повече от един елемент („универсална йерархия за показване“)
Генерични и бизнес аналитични функции	Сортиране по отношение на характеристики и ключови показатели Изчислени ключови показатели и формули (дава се възможност за статистическо и математическо изчисление във основа на стойности на показатели, в допълнение към другите възможности) Преобразуване на валути Елиминирание на вътрешни бизнес обеми (бизнес елиминирание на вътрешни операции)
Концепции за оптимизиране на времето за изпълнение	Не-кумулятивни стойности Агрегирани данни OLAP кеш (регулируем в кеш режим (в зависимост от заявката))
Интегрирани допълнителни функции	Променливи за целите на параметризация и повишена повторна употреба на заявки Интерфейси отчет-отчет за целите на навигация в различни отчети Концепция за права на достъп за целите на управление на потребителските оторизации по отношение на данните за достъп

ИЗЧИСЛЯВАНЕ НА ЛОКАЛНИ ЕДИНИЧНИ СТОЙНОСТИ:

Чрез единичните стойности, може да се нормализират данни на заявката за показател за различни резултати на този показател. Данните са показани като процент от резултата. Може да се създават списъци с предварително зададени параметри за единични стойности. Стойностите на характеристика са сортирани в съответствие с показателя, който е избран, и позицията която е зададена. Може локално да се изчисли максималните/минималните стойности на показател за дадена характеристика. С други думи, най-високата/най-ниска стойност (както постъпления) е показана за дадена характеристика (да я наречем ‘клиент’) за всички стойности на характеристика (клиент 1, клиент 2, ...).

Може да се броят всички стойности (където е необходимо и чрез изключване на нулевите стойности). Може локално да се изчисли средна величина на всички стойности (възможно и чрез изключване на нулевите стойности). Може също да се потиснат единичните стойности – показване само на стойности.

Изчисляване на резултатите локално:

С резултатите, както и с единичните стойности, може:

- да се показва максималната/минималната стойност на съответната област.

- да се брои количеството на стойностите на съответните области (където е необходимо и чрез изключване на нулевите стойности).

- да се изчисли средна величина на всички стойности на съответната област (възможно и чрез изключване на нулевите стойности).

- да се покаже сумата от стойностите на съответната област,

- да се скрие резултата или да се покаже резултата за най-високата/най-ниска стойност за съответната област.

С получените резултати, може да се използват функциите Стандартно отклонение и Дисперсия средно отклонение. Може да се сумират и закръглени стойности. Тази функция има смисъл, когато е създаден мащабен коефициент за показатели, защото може да има значителна разлика между сумата изчислявана по подразбиране и локално изчислената стойност на закръглените стойности.

2.4 ФУНКЦИЯ ЗА ВАЛУТНО ПРЕИЗЧИСЛЕНИЕ

Функцията валутно преизчисление позволява преизчислението на осчетоводените записи с данни в изходна валута в целевата валута или чрез повтарящо се трансфериране в няколко различни целеви валути. Тя е създадена като функционалност на SAP сървър за уеб приложения за валутно преизчисление. Валутното преизчисление прави възможно преизчислението на ключови показатели с валутни полета, които съществуват в сорс системата в различни валути, в една единствена валута в BW системата. Например, в местна или валута на фирмената група. Допълнителни възможности предоставя отчетът за курсови разлики. С него можете да сравнявате текущи обменни курсове с курсове валидни за датата на осчетоводяване и да изчислите въздействието на промените върху обменните курсове. Валутното преизчисление е основано върху видове валутни преизчисления, в които са установени корпоративни правила за преизчисляване. Той представлява комбинация от различни видове параметри (изходна валута и целева валута, вид обменен курс и часова база за преизчисление), които определят как се изпълнява обменния курс за даденото преизчисление.

1. Видът на обменния курс различава обменни курсове, които са валидни едновременно в една и съща времева рамка, например курс “купува”, курс “продава” или среден обменен курс.

2. Изходна валута е валутата, която трябва да бъде преизчислена. Изходната валута се определя от записа на данни или от даден ИнфоОбект (динамично). Целевата валута може да бъде определена във вида валута, доставена с даден ИнфоОбект, дефинирана от променлива или да не бъде избрана до времето на извършване на преизчислението.

3. Времевите референции за извършване на валутното преизчисление могат да бъдат фиксирани или променливи. Ако референцията е фиксирана, тогава моментът за определяне на обменния курс е независим от данните. Фиксирано време може да бъде текущата дата, фиксирана отчетна дата или времето от променливата (променлива към 0DATE). Ако е променлива, тогава времето за определяне на обменния курс идва от стойността заложена за времевата характеристика. В допълнение *отчетната дата на заявката* може да бъде избрана като времева референция. Тогава отчетната дата

която се използва е тази определена в свойствата на Конструктора на заявки.

Видът на валутното преизчисление се съхранява във форма за повторна употреба и е наличен за валутни преизчисления в правилата за обновяване за ИнфоКубове и в Business Explorer.

Може да се определят в правилата за обновяване показатели или полета с данни за които дадено валутно преизчисление трябва да бъде извършено при актуализирането. *При някои специални случаи е възможно да се извърши валутно преизчисление използвайки потребителски подпрограми в правилата за трансфер и обновяване.*

2.5 ЕЛИМИНИРАНЕ НА ВЪТРЕШНИ СТОПАНСКИ ОПЕРАЦИИ

Тази функция позволява да се елиминира вътрешни стопански операции.

Например, приходите от продажби направени между два разходни центрове, които лежат в една и съща област на йерархията вече не се показват в отчета.

За всеки ключов показател, може да се създаде показател с референция, чиято (вътрешна) стойност да бъде елиминирана. *Трябва да се създаде правило, което да определи кога една стойност се счита за вътрешна.* Правилото се определя чрез двойки от характеристики (изпращач и получател, където всяка двойка се отнася към съответната характеристика. Стойността се счита за вътрешна тогава, когато стойността на характеристиката на изпращача принадлежи и връща избрания вариант на характеристиката на получателя. Стойността „не определена” винаги означава външна. Дори ако няколко двойки от характеристики са били вече определени, все още може да се определи дали стойността е вътрешна или не, ако правилото е валидно за всички или само за една двойка характеристики. Елиминиране чрез използване на йерархии може да бъде визуализирано в отчета. Вътрешната стопанска операция се елиминира на йерархично ниво под което са определени съответните двойки характеристики.

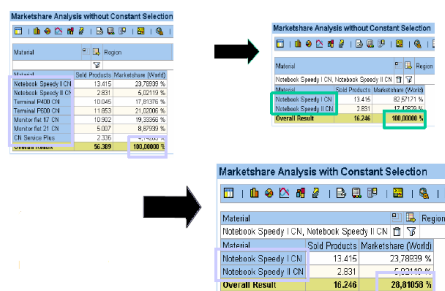
2.6 ФУНКЦИЯ ЗА ИЗБИРАНЕ НА КОНСТАНТИ

В конструирането на заявката, може да се използва подбор за да се определи, кои данни да се покажат при работата на отчета. След това може да се промени подбора по време на работа използвайки навигиране и филтри – така може да се ограничи допълнително подбора.

Функцията *избиране на константи* позволява да се маркира даден избор независим от другите филтри. Това означава, че навигирането и филтрирането нямат ефект върху избирането по време на работа. По този начин, можете лесно да се избират дименсии, които не могат да се променят за остатъка от работния процес.

Функцията *избиране на константи* може да бъде използвана за определяне например на пазарен индекс. Ако не е необходимо да се виждат абсолютните приходи от продажби в един списък с продукти (търсения продукт е в разбивката), а просто нормализираните приходи от продажби (които са свързани) със специфична група продукти. Използвайки функцията *избиране на константи*, може да се избират приходите от продажби от специфична група продукти, като константа за разбивката.

Може да се свържат приходите от продажби на индивидуален продукт с груповите приходи от продажби и например да се определят приходите от всеки индивидуален продукт, като част от приходите за групата.



Фиг. 3. Функцията позволява, изчисляване на индекс на пазара за отделни продукти[3]

Избирането на константи може да бъде използвано за създаване на списък на бавнооборотни стоки.

Може да се определи една заявка базира

на на МултиДоставчик в който са показани всички продукти, които не са били продадени в определен период от време и от които са били продадени само малки количества. С други думи, стойностите на характеристиката също трябва да бъдат показани в заявката, за която няма данни за извършена транзакция или има само малко количество данни за избрания период.

Може да се определят цели избирания като константи за структурните компоненти и клетки. Избирането остава неподвижно в навигирането независимо от всички филтри. Може да се дефинират компоненти от избирания, които са единични ИнфоОбекти и техните филтриращи стойности, като константи. Избирането остава константно само по отношение на дадения ИнфоОбект при навигацията и филтрирането. Трябва да се има предвид, че не може да се избира цял ограничен ключов показател като константа, а само неговите ИнфоОбекти.

ЗАКЛЮЧЕНИЕ

Статията изследва най-често използваните функции и услуги в системите за аналитична обработка на данни.

REFERENCE

- [1] Valova I., M. Noirhomme-Fraiture, B. Zhechev, V. Valov, Formalization of the OLAP theory, International Scientific Conference Computer Science'2008, Kavala, Greece, 756-761
- [2] Valova I., M. Noirhomme-Fraiture, Processing of Large Data Sets: Evolution, Opportunities and Challenges, PCaPAC'08, Ljubljana, Slovenia, 2008, 198-200.
- [3] www.sap.com