

УДК 004.62

**Ю.Н. Возовиков**

Одесский национальный политехнический университет, г. Одесса  
кафедра системного программного обеспечения  
E-mail: [yuri\\_email@mail.ru](mailto:yuri_email@mail.ru)

## ТЕХНОЛОГИЯ СОЗДАНИЯ И СОПРОВОЖДЕНИЯ МЕХАНИЗМА УПРАВЛЕНИЯ МАТЕРИАЛИЗОВАННЫХ ПРЕДСТАВЛЕНИЙ

*Предлагается технология создания и сопровождения периодически подключаемых материализованных представлений (МП). Увеличение эффективности использования МП достигается, за счет корректировки состава МП при изменениях в автоматизированных системах, расширения количества запросов, для которых использование МП было не эффективным, сокращения издержек на поддержание МП в течение времени, когда некоторые из них будут неэффективными.*

**Ключевые слова:** информационные системы, материализованные представления, запрос.

### Общая постановка проблемы

Одним из путей решения повышения производительности в автоматизированных системах (АС), где используются реляционные базы данных (РБД), является применение материализованных представлений [1,2]. В большинстве существующих АС наблюдается определенная периодичность выполнения различных задач. Данная периодичность отражается в периодичности появления различных запросов. Постоянное включение МП, как это предлагалось ранее [3,4], снижает эффективность работающих МП.

С течением времени в АС происходят изменения. Данные, полученные в течение периода исследования системы, и расчёты, выполненные на их основе, постепенно устаревают. Предварительно рассчитанная эффективность применения некоторого МП может уменьшиться. Для поддержания эффективной работы механизма управления МП прежде всего следует выявить факторы, влияющие на понижение эффективности.

Предлагается следующая классификация этих факторов.

1. Структурные изменения в БД.
2. Структурные изменения в АС.
3. Организационные изменения, связанные с фактором времени.
4. Появление новых и удаление существовавших ранее задач (видов деятельности).
5. Изменение общей интенсивности решения задач и их периодичности.

Некоторые из перечисленных факторов могут быть в определенной степени взаимосвязаны. Например, структурные изменения в АС могут быть следствием появления или удаления определенных задач или видов деятельности.

Приведенная классификация позволяет разработать конкретные способы выявления указанных факторов, а также реакцию механизма управления МП на их появление.

В основу информационной технологии поддержки актуальности механизма управления МП положено проведение периодического анализа потока входных запросов к АС, выборочный анализ запросов, а также административное управление системой в некоторых ситуациях.

Рассмотрим возможные способы устранения негативного влияния выявленных факторов.

### Структурные изменения в БД

Из множества запросов  $Q$  выделяем подмножество запросов, выполняющих структурные изменения в БД

$$Qt = \{q_j \mid Ty_j = tt\} \tag{1}$$

Разобьём структурные изменения в БД, влияющие на эффективность работы механизма МП, на два вида. К первому отнесём удаление таблиц, удаление полей таблиц, переопределение полей. Ко второму – создание новых таблиц, добавление полей в существующие таблицы.

Реакция на структурное изменение первого вида заключается в поиске и удалении МП, использующих соответствующие таблицы и поля.

#### Удаление $Mpi$

Сформируем множество  $Qt1 \in Qt$ , содержащее запросы типа DROP TABLE, ALTER TABLE – DROP COLUMN, ALTER TABLE – ALTER COLUMN.

Пусть некоторое  $Mpi$  реализует запрос  $q_i$ . Запишем условие удаления некоторого  $Mpi$  при удалении базовой таблицы

$$q_i \in Qs \wedge \forall (T_{ir} \in T_{qr} \wedge T_{ir} = T_j \wedge T_j \in T \wedge q_j \wedge q_j \in Qt1), \tag{2}$$

где  $T_{ir}$  - таблица, используемая в запросе  $q_i$ , а  $T_j = T_{ir}$ , таблица, удаляемая в запросе  $q_j$ .

Условие удаления некоторого  $Mpi$  при удалении или переопределении колонки базовой таблицы:

$$q_i \in Qs \wedge \forall (T_{ir} \in T_{iqr} \wedge f_i \in T_{ir} \wedge T_{ir} = T_j \wedge f_j \in T_j \wedge T_j \in T \wedge q_j \wedge q_j \in Qt1), \tag{3}$$

где  $T_{ir}$ ,  $f_i$  - таблица и её столбец, используемые в запросе  $q_i$ , а  $T_j = T_{ir}$ ,  $f_j$  - таблица и её столбец, который удаляется или переопределяется в запросе  $q_j$ .

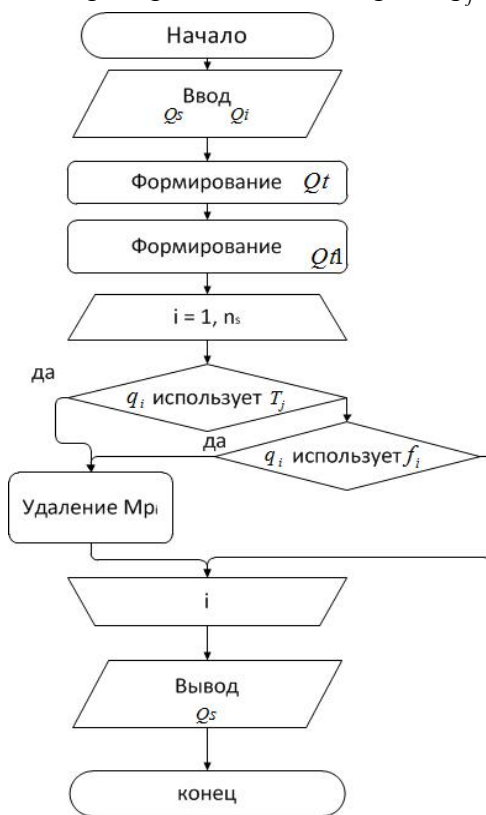


Рисунок 1 — Алгоритм удаления  $Mpi$  при удалении базовой таблицы.

### Выявление новых запросов

Изменения второго рода, очевидно, предусматривают последующую работу с новыми таблицами и полями. Исследования запросов использующих эти таблицы и поля ранее не проводилось. Поэтому предлагается включение механизма выборочного исследования запросов, использующих новые поля и таблицы.

Сформируем множество  $Qt2 \in Qt$ , содержащее запросы типа CREATE TABLE, ALTER TABLE – ADD COLUMN.

Условие формирования множества запросов  $Qinv1$ , которые подлежат дальнейшему исследованию при вводе новой таблицы:

$$Qinv1 = \{q_i \mid T_{ir} \in T_i \wedge T_{ir} = T_j \wedge q_j \in Qt2\}$$

Условие формирования множества запросов  $Qinv2$ , которые подлежат дальнейшему исследованию при вводе или переопределении колонки в таблице:

$$Qinv2 = \{q_i \mid T_{ir} \in Tq_i \wedge f_i \in T_{ir} \wedge f_j \in T_j \wedge T_{ir} = T_j \wedge f_i = f_j \wedge q_j \in Qt2\}$$

Все новые запросы будут представлены множеством

$$Qinv = Qinv1 \cup Qinv2 \tag{4}$$

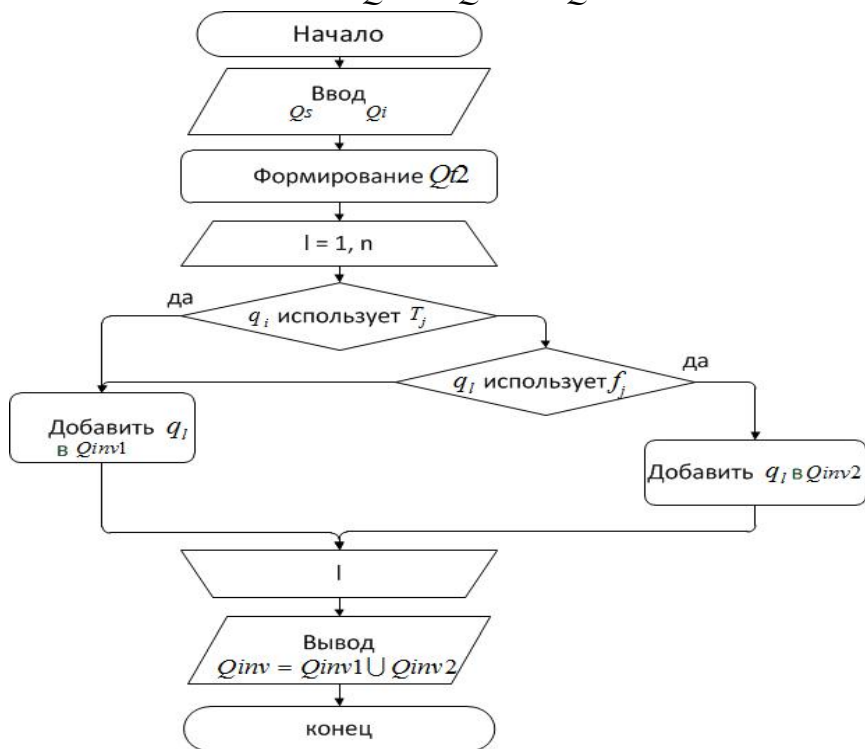


Рисунок 2 — Алгоритм формирования множества новых запросов.

### Структурные изменения в АС

Структурные изменения в АС, связанные с образованием новых и удалением старых рабочих мест, не могут быть однозначно определены в рамках работы с БД. Поэтому здесь можно рекомендовать проведение организационных мероприятий.

#### Удаление рабочего места

Если ответственному лицу, например, администратору БД, известно, что удаление рабочего места связано с прекращением определенного вида деятельности, то по его команде имеется возможность внести изменения в список используемых МП. Для этого предлагается из всего множества МП удалить те таблицы, которые соответствуют запросам, сформировавшимся только на удаляемом рабочем месте.

Представим запрос в виде  $q = \langle tx, T_q, Ty, \langle tm, d \rangle, w \rangle$ ,

где  $w$  - источник запроса (рабочее место, с которого был отправлен запрос). В данном случае нас интересуют только запросы типа SELECT, для которых были созданы МП. Множество таких запросов было определено как  $Q_s$ .

Будем считать, что  $w_d$  - удаляемое рабочее место. Сформируем множество запросов, которые были посланы с этого рабочего места

$$Q1 = \{q_i \in Q_s \mid w_i = w_d\} \quad (5)$$

В это множество могли попасть запросы, которые совпадают по виду с найденными, но поступали в АС от других рабочих мест и исследовались ранее. Определим множество этих запросов

$$Q2 = \{q_i \in Q1 \mid \exists q_j \in Q_s \mid tx_i = tx_j\}$$

Тогда множество запросов, для которых следует удалить МП, будет определено как разность множеств  $Q1$  и  $Q2$

$$Qd = Q1 / Q2 \quad (6)$$

### Добавление рабочего места

Добавление нового рабочего места, связанного с появлением нового вида деятельности, может быть поводом к выборочному исследованию запросов, направляемых в АС с этого рабочего места.

Пусть  $W_{new}$  множество новых рабочих мест. Представим множество запросов типа SELECT, которые поступают в АС от новых рабочих мест в виде

$$Q1 = \{q_i \mid w_i \in W_{new}\}$$

Это множество может содержать запросы, которые совпадают по виду с запросами, поступавшими в АС и от других рабочих мест и исследовались ранее.

$$Q2 = \{q_i \in Q1 \mid \exists q_j \in Q_s \mid tx_i = tx_j\}$$

Тогда множество запросов, для которых следует проводить исследование, будет определено как разность множеств  $Q1$  и  $Q2$

$$Q_{new} = Q1 / Q2 \quad (7)$$

### Общий подход к корректировке настроек механизма управления МП

Факторы 3, 4 и 5 (параграф 1) в основном предусматривают проведение исследования системы. При существенных организационных изменениях в системе можно рекомендовать проведение исследований незамедлительно. Если удаление некоторых задач и их связь с МП хорошо прослеживается ответственным лицом, например, администратором БД, то можно удалить соответствующие МП в «ручном режиме». В остальных случаях проблемы будут решаться в результате периодического анализа АС.

Для поддержания эффективности применения МП на высоком уровне предлагается проводить корректировку ранее принятых решений на основании продолжения наблюдения за системой. Это также позволит определить периодичность в распределении запросов, которую, ранее не представлялось возможности обнаружить. Например, перед внедрением механизма МП в некоторую ИС проводились наблюдения за её работой в течение двух месяцев. Однако набор и интенсивность задач, решаемых в ИС, зависит от сезона года. Очевидно, что такой период не мог быть определён за период наблюдения за работой системы на стадии проектирования МП.

На основании сказанного корректировка механизма управлением МП может производиться в результате решения трёх задач.

1. Получение текущего значения эффективности каждого из используемых МП
2. Сравнительная оценка количеств запросов, для которых созданы и не созданы МП.
3. Накопление и анализ данных для возможности создания МП с большими периодами работы.

**Расчетные значения за предыдущий период наблюдения**

Пусть механизм МП был включен после предварительного наблюдения за системой в течение времени  $t_0$ . Будем считать, что за это время получено:

- множество всех запросов -  $Q$  и их количество -  $n = |Q|$
- множество разных запросов типа  $t_s$ , для которых были созданы МП -  $Q_s$  и их количество -  $n_s = |Q_s|$ ;
- множества запросов определенного вида, для которых были созданы МП -  $Q_i$  и их количество -  $n_i = |Q_i|$ ;
- множества запросов, обновляющих данные, используемые в  $Q_i$  -  $Q_{u_i}$  и их количество -  $n_{u_i} = |Q_{u_i}|$ ;

По завершению этапа предварительного наблюдения за системой были построены матрицы управления МП и получены расчетные характеристики:

- суммарного времени выполнения всех запросов  $q_i$  с учётом  $MP_i$  -  $S\tau_i$ ;
- относительной эффективности для каждого  $MP_i$  -  $\varepsilon\tau_i$ ;

а также определено количество запросов  $n\tau_i$ , попадающих в периоды включенного состояния  $MP_i$ .

**Оценка текущей эффективности МП, созданных ранее**

Для анализа работы МП предлагается метод скользящей оценки эффективности МП. Суть метода заключается в непрерывном отслеживании работы системы и периодическом расчете эффективности МП в заранее определённые моменты.

Прежде всего, следует определить минимальный интервал времени  $\Delta t$ , по истечению которого имеет смысл проводить анализ данных. Выбор  $\Delta t$  определяется следующими соображениями.

- Величина  $\Delta t$  не может быть слишком короткой, поскольку в этом случае изменения, происшедшие в системе, мало отразятся на оценке эффективности МП.
- Величина  $\Delta t$  не может быть слишком большой, т.к. радикальные изменения в ИС, происшедшие за этот период, могут не только понизить эффективность некоторых МП, но и сделать их применение совершенно бесполезным.

Исходя из приведенных соображений, максимальным значением может быть  $\Delta t = t_0$ , а минимальным – самый длинный период  $\tau_{\max}$  некоторого реализованного МП. Например,  $t_0$  - два месяца, а  $\tau_{\max}$  - неделя, тогда  $\Delta t = 7$  суток.

Оценка эффективности МП может выполняться разными способами зависимости от времени эксплуатации механизма МП.

1. Время эксплуатации МП не превышает времени  $t_0$  наблюдения за ИС.
2. Время эксплуатации МП превышает время  $t_0$ .

В первом случае могут сравниваться расчетные эффективности с реальными эффективностями для некоторых МП. Во втором – имеется возможность проанализировать изменения реальных эффективностей.

Рассмотрим первый случай.

Пусть  $Q\delta$  -- множество всех запросов, поступивших в ИС в течение времени  $\Delta t$ , а  $n\delta$  - их количество.

Определяем множество запросов  $Q\delta_i$  вида  $q_i$ , выполненных в течение времени  $\Delta t$  ( $q_i \in Q\delta$ ) и их количество -  $n\delta_i$ .

Если будет получено соотношение

$$\frac{n\delta_i}{n\delta} \ll \frac{n_i}{n}, \quad (8)$$

то следует пересмотреть целесообразность дальнейшего использования  $MP_i$ .

Найдём суммарное время выполнения запросов  $q_i$  поступивших в АС в течение  $\Delta t$

$$S\delta_i = \sum_{j=1}^{n\delta} d_{ij},$$

где  $d_{ij}$  – длительность выполнения некоторого запроса  $q_i$  в течение  $\Delta t$ .

Определим среднее время выполнение запросов  $q_i$ , поступивших в АС в течение  $t_0$

$$Sa_i = \frac{S\tau_i}{n\tau_i},$$

и среднее время выполнения  $q_i$  в течение  $\Delta t$

$$Sa\delta_i = \frac{S\delta_i}{|Q\delta_i|}.$$

Если значения  $Sa_i$  и  $Sa\delta_i$  существенно не отличаются

$$Sa_i \cong S\delta_i, \quad (9)$$

то можно считать, что  $MP_i$  продолжает работать эффективно.

Если

$$Sa_i < S\delta_i,$$

то следует провести анализ причин увеличения времени выполнения запросов. Такие причины могут быть следующими.

1. Замедлилась работа АС в силу различных причин, не связанных непосредственно с работой рассматриваемой БД.

2. Изменилась периодичность появления запросов  $q_i$ ;

3. Увеличилось количество запросов обновления  $MP_i$ .

Для проверки первой причины следует сопоставить эффективность применения  $MP_i$ , рассчитанную для периода  $t_0 - \varepsilon\tau_i$  с эффективностью  $MP_i$  полученной для периода  $\Delta t - \varepsilon\delta\tau_i$ .

Если

$$\varepsilon\tau_i \cong \varepsilon\delta\tau_i, \quad (10)$$

то не следует менять алгоритм управления  $MP_i$ .

В случае, когда

$$\varepsilon\tau_i > \varepsilon\delta\tau_i, \quad (11)$$

Рекомендуется выполнить новый анализ периодичности для запросов  $q_j$ .

Если решено провести новый анализ периодичности, то рекомендуется использовать все запросы, имеющиеся в распоряжении исследователя.

$$Q_{new} = Q \cup Q\delta \quad (12)$$

Следует учитывать, что корректировке должны в первую очередь подвергаться те МП, в управлении которых присутствуют «длинные» периоды, то есть для которых, возможно, было недостаточно данных для хорошей настройки.

Если время эксплуатации МП превышает время наблюдения за ИС, то во всех ранее приведенных оценках работы МП следует вместо периода наблюдения  $t_0$  брать равный по длительности период работы ИС, что позволит получить более точные результаты.

Если отмечено снижение эффективности для многих МП, а также существенно увеличилось количество запросов, для которых нет МП, необходимо выполнить полный анализ запросов системы для нового периода наблюдения.

**Анализ возможности применения новых МП**

Если обеспечить накопление данных о запросах в течение времени эксплуатации МП, то появляется возможность проанализировать добавление новых МП с большими периодами включения/ выключения. При отсутствии ориентировочной информации о возможной периодичности задач, решаемых в конкретной системе, можно рассмотреть периоды, основанные на общеизвестных календарных отметках времени таких как, месяц, квартал, сезон года. При этом предлагается связать с каждой временной отметкой, уточняющий атрибут «начало», «конец», «середина».

Для исследования новой периодичности следует решить две задачи.

1. Определение запросов, для которых следует искать новый период повторения.
2. Определение длительности нового периода повторения.

В качестве запросов – кандидатов для нового исследования следует выбирать все виды запросов, для которых ранее не создавались МП, а также запросы с низкой эффективностью соответствующих МП.

При определении длительности периода повторения  $\tau_{new}$  предлагается руководствоваться следующим отношением  $2\tau_{new} \leq t_0 + \Delta t$ ,

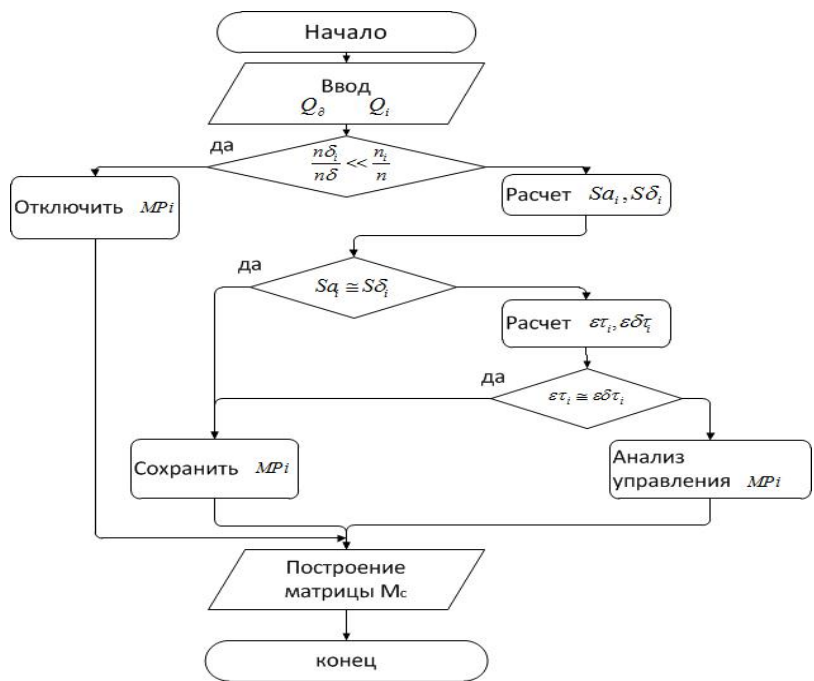


Рисунок 3 — Алгоритм корректировки настроек механизма МП.

**Выводы**

1. Дано определение жизненного цикла механизма управления МП.
2. Определены основные факторы понижения эффективности механизма управления МП в процессе работы АС.
3. Предложен механизм корректировки состава МП и частичного исследования запросов в случае структурных изменений в АС и БД
4. Предложена технология непрерывного отслеживания за последовательностью запросов, поступающих в АС, позволяющая определить набор мероприятий по корректировке механизма управления МП.

**Список использованной литературы**

1. Компания Advanced Information Systems. Oracle 8. Энциклопедия пользователя: пер. с англ./ Компания Advanced Information Systems / ДиаСофт, К.: 1998. — 864 с.
2. Р. Виейра; Программирование баз данных Microsoft SQL Server 2005. Базовый курс.: пер. с англ. / Р. Виейра. Диалектика, М.: 2007. — 832 с.
3. Кунгурцев А.Б. Анализ возможности применения МП в ИС/ А.Б. Кунгурцев, Куок Винь Нгуен Чан // Тр. Одесского политехнического университета. Одесса, 2004. — 2(20). — С. 102-106.
4. Кунгурцев А.Б. Поиск закономерностей в распределении запросов для управления материализованными представлениями/ Кунгурцев А.Б., Возовиков Ю.Н.//Тр. Одесск. политехн. ун-та. Одесса, 2008. — 2(30). — с. 135— 140.
5. Кунгурцев А.Б. Управление материализованными представлениями в информационных системах/ Кунгурцев А.Б., Возовиков Ю.Н. // Восточ-Европ. журн. передовых технологий. — Харьков, 2010. — Вып.¼(43). — С. 18 — 21.

Надійшла до редакції:  
25.03.2013р.

Рецензент:  
д-р техн. наук, проф. Зорі А.А.

**Ю.М. Возовиков**

**Одеський національний політехнічний університет**

**Технологія створення і супроводу механізму управління матеріалізованих представлень.**

*Пропонується технологія створення і супроводу періодично підключаємих матеріалізованих представлень (МП). Збільшення ефективності використання МП досягається, за рахунок коригування складу МП при змінах в автоматизованих системах, розширення кількості запитів, для яких використання МП було б не ефективним, скорочення витрат на підтримку МП протягом часу, коли деякі з них будуть неефективними.*

**Ключові слова:** інформаційні системи, матеріалізовані представлення, запит.

**Yu.N. Vozovikov**

**Odessa National Polytechnic University**

**Technology of Development and Maintenance of Control Mechanism of Materialized Views.**

*There are various software methods of increasing the productivity of informational systems (IS) based on relational data bases (RDB). One of them presupposes the use of materialized views (MV). MV stores the result of a certain request in the data base (DB) and with the next entry of this request to the IS it allows getting the answer very fast. However the practical use of MV in a particular IS is impossible without preliminary study of this IS. As MV is a result of request which uses number of BD tables the refreshing of certain data in those tables leads to the necessity of refreshing MV too. If data are refreshed frequently the use of MV can decrease the effectiveness of IS instead of increasing it.*

*This paper offers periodical connection and disconnection of MV. The basis for this is the evident periodicity in solving different tasks for the most of organizations. For example, there is a period of admission, passing tests, and visiting hours in a university. In trading companies the periodicity of goods income, inventories, revaluation, sales out, seasonal variations in assortment and working hours is observed. The periodicity of solving tasks is reflected in the periodicity of requests coming to the system. The technology of periodically connected materialized views (MV) developed on the basis of an organization's schedule is offered. This approach related to the working hours /days of week /decades /months allows increasing MVs implementation effectiveness. The method of MV management can be used in many organizations where certain periodicity of tasks is observed.*

**Keywords:** information systems, materialized views, query.