

## МНОГОМЕРНОЕ МОДЕЛИРОВАНИЕ ДАННЫХ ДЛЯ СИСТЕМЫ МОНИТОРИНГА ДЕЯТЕЛЬНОСТИ УВО

ШУЛЬДОВА С.Г., ЛАПИЦКАЯ Н.В.  
БГУИР, Минск, Беларусь, shsg@bsuir.by

В современном мире обеспечение качества образования становится основным стратегическим направлением, определяющим уровень развития и потенциал общества. При этом качество образования можно рассматривать как критерий эффективности деятельности образовательного учреждения, соответствия реально достигаемых результатов нормативным требованиям, социальным и личностным ожиданиям [1].

Соблюдение нормативных требований обеспечивается наличием у учреждения высшего образования (УВО) сертификата об аккредитации, подтверждающего соответствие качества образовательных услуг образовательному стандарту. Это внешняя оценка, основанная на критериях результативности образовательного процесса, эффективности организационно-управленческой, научно-исследовательской и методической деятельности, качества ресурсного обеспечения деятельности учреждения образования, проводимая специализированными организациями.

Соответствие социальным и личностным ожиданиям выражается степенью удовлетворенности выпускников, работодателей и преподавателей, и оценивается, как правило, путем анкетирования, а также отслеживанием карьерного роста выпускников.

Внутренняя оценка качества образования предполагает наличие у учреждения высшего образования собственной системы контроля. Такие системы предполагают проведение мониторинга, который рассматривается как систематическая процедура сбора, обработки, анализа, оценки и интерпретации информации о состоянии объекта; прогноза его дальнейшего функционирования и развития и выработки коррекционных мер [2]. Объект в данном случае представляет собой совокупность процессов по осуществлению образовательной, научно-исследовательской и учебно-методической деятельности УВО, и для реализации процедур мониторинга необходима разработка методов и моделей их обеспечивающих.

Поскольку качество – многомерная категория, для его описания целесообразно использование многомерной модели – представление данных в виде гиперкуба  $G = \langle M, D \rangle$ , в ячейках которого находятся анализируемые количественные показатели (меры), а оси представляют собой измерения – множество объектов (элементов измерения) одного или нескольких типов, организованных в виде иерархической структуры и обеспечивающих информационный контекст типизированного показателя. Пусть  $D = \{D_1, D_2, \dots, D_n\}$  – множество измерений,  $M = \{M_1, M_2, \dots, M_m\}$  – множество показателей в гиперкубе  $G$ , тогда  $M_i = \{m_{ij}\}$  – множество показателей в точке (ячейке куба), описывающих определенный факт, а  $m_{ij} = f_j(d_1, d_2, \dots, d_n)$ , где  $f_j$  – функция, определяющая в каждой точке  $n$ -мерного пространства значение анализируемого показателя,  $i=1..m, j=1..n$ .

Значения, связанные с измерением, характеризуют какое-либо классификационное свойство (атрибут) сущностей предметной области, то есть если задано множество атрибутов  $A = \{a_1, a_2, \dots, a_p\}$ , то  $\forall A \in D$ .

В каждом измерении гиперкуба задается множество ключевых (Key Members) элементов измерения (факт-координат) – однородных элементов с наибольшей степенью детализации (гранулярности). Во множестве этих элементов может быть задано разбиение на непересекающиеся подмножества, каждому подмножеству соответствует укрупненная координата измерения, совокупность которых составляет вышестоящий уровень иерархии, и т. д. На самом верхнем уровне иерархии вводится единственная корневая координата «Все» («All»), символизирующая совокупность всех элементов нижестоящего уровня. В общем

случае, для одного измерения может быть задано несколько иерархий, основанных на общем множестве факт-координат:  $l_i^k = \{d_{ij}^k\}$ , где  $i$  – порядковый номер измерения в кубе;  $k$  – порядковый номер уровня иерархии в измерении;  $j$  – порядковый номер элемента измерения на  $k$ -м уровне иерархии;  $L_i = \{l_i^k\}$  – множество уровней  $i$ -го измерения.

Как правило, многомерная модель представляет собой множество локальных кубов для решения частных задач [3], которые в данном случае заключаются в анализе показателей по видам деятельности: учебная, учебно-методическая и научно-исследовательская.

Рассмотрим некоторые критерии научно-исследовательской деятельности кафедры и соответствующие им количественные показатели:

- научный потенциал кафедры – количество докторов, кандидатов наук, % от общего количества сотрудников кафедры;

- публикации: монографии, статьи в научных изданиях, материалы конференций и т.п., в том числе перечня ВАК и зарубежные – количество публикаций, количество публикаций каждого вида; количество публикаций каждого сотрудника;

- изобретательская и патентно-лицензионная работа – количество патентов, количество объектов права промышленной собственности, количество патентов каждого сотрудника;

- участие преподавателей в конференциях, конгрессах – количество конференций, в которых приняли участие сотрудники, количество сотрудников, принявших участие, количество конференций по видам;

- подготовка научных кадров высшей квалификации через аспирантуру и докторантуру – количество докторантов и аспирантов, количество защищённых диссертаций, количество присвоенных научных степеней и научных званий;

- наличие договоров о международном сотрудничестве, участие преподавателей в международных научных проектах – количество договоров, количество проектов, количество участвующих в них сотрудников.

Тогда множество интегрированных показателей или групп мер, характеризующих научно-исследовательскую деятельность, за отчетный период  $M_n = \{\text{Количество публикаций, Количество сотрудников, Количество защищенных диссертаций, Количество конференций, Количество патентов, Количество аспирантов и докторантов; Количество международных проектов}\}$ .

Одним из методов проектирования куба, является использование ER-модели предметной области, при этом целесообразно руководствоваться следующими положениями [4]:

- связь «многие-ко-многим» (М:М) ER-модели соответствует гиперкубу. Если в исходной ER-модели имеется несколько М:М-связей, то в результате замены их гиперкубами получается многокубовая модель, в которой отдельные кубы взаимосвязаны;

- для установлении иерархий в измерениях используются функциональные (1:М) зависимости;

- измерению может соответствовать несколько сущностей, в этом случае родительские сущности становятся частью измерений в виде уровней иерархий;

- в случае, если какой-либо атрибут М:М-связи представляет интерес в плане использования его в качестве измерения гиперкуба, целесообразно выделить этот атрибут в самостоятельное измерение (вырожденное), позволяя проводить по нему многомерный анализ данных.

Связь типа «многие-ко-многим» соответствует таблице фактов схем «звезда» или «снежинка» реляционной базы данных, служащей для поддержки многомерного представления содержащихся в ней данных, при этом показателям таблицы фактов соответствуют атрибуты связи. В свою очередь, группе мер куба соответствует факт, а измерениям куба – таблицы измерений реляционной БД.

На рисунке 1 представлен фрагмент схемы нормализованной реляционной БД, содержащей только анализируемые атрибуты. Таблицы «Сотрудники» и «Публикации

сотрудников» являются таблицами фактов, которые соответствуют 3 группам мер:  $M_1$ =Количество сотрудников,  $M_2$ =Количество публикаций,  $M_3$ =Количество защищенных диссертаций.

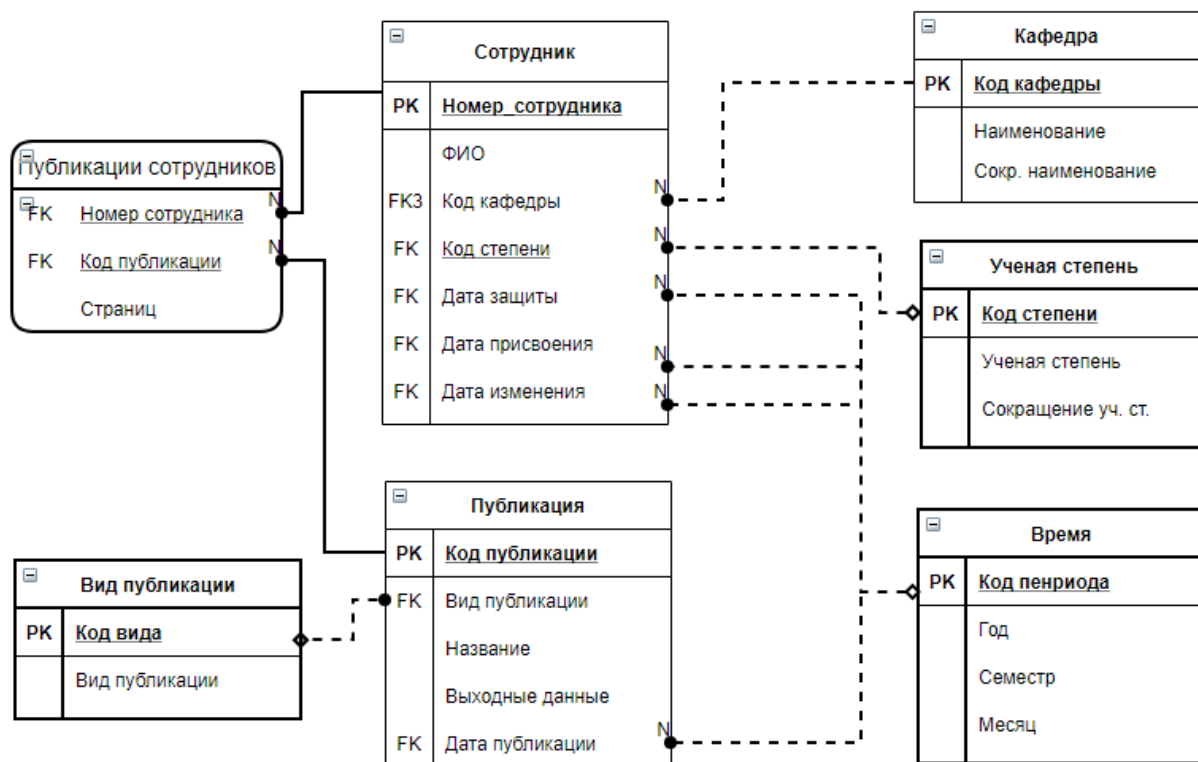


Рисунок 1 – Реляционная схема для поддержки многомерной модели

Измерение времени – особое измерение, поскольку правильное определение его атрибута гранулярности и иерархий влияет на корректность агрегирования мер. Исходя из задач анализа, атрибут гранулярности равен месяцу, более детальные данные для анализа не требуются.

Измерение «Сотрудники» относится к типу 2 медленно меняющегося измерения, так как ученая степень и звание сотрудника со временем могут измениться. С целью сохранения истории в таблицу добавлены поля: Дата защиты, Дата присвоения и Дата окончания.

Сущность «Вид публикации» трансформируется в уровень иерархии измерения «Публикации».

Измерениями куба будут:  $D_1$ =Кафедра,  $D_2$ =Ученая степень,  $D_3$ =Сотрудники,  $D_4$ =Публикация,  $D_5$ =Период времени.

Иерархии, полученные на основе функциональных связей:

- Номер сотрудника →Ученая степень →Кафедра;
- Номер публикации →Вид публикации.

Измерение времени содержит естественную иерархию Месяц→ Семестр →Год.

Отношение сопоставления  $R \subseteq D \times M$ ,  $(D_j, M_i) \in R$  определяется, если показатель  $M_i$  может быть проанализирован по измерению  $D_j$ . Графическое представление отношений для определенных выше показателей и измерений показано на рисунке 2.

Так как два показателя определены на одном и том же множестве измерений, то, следовательно, один из них содержится (включен) во множество другого, то есть  $M_3 \subset M_1$ .

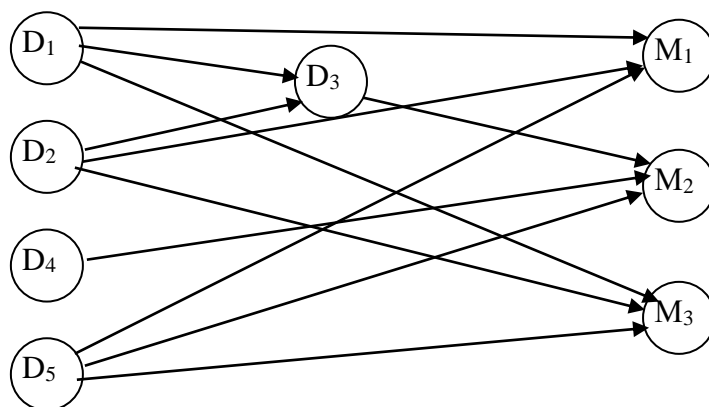


Рисунок 2 – Связи измерений и показателей

Таким образом, для анализа публикационной деятельности и научного потенциала кафедры получена 2-хкубовая модель, содержащая 5 измерений и 2 группы мер (рис.3):  $M_1 = \{\text{Количество кандидатов (докторов) наук, \% остепенённых сотрудников, количество защищенных диссертаций}\}$ ,  $M_2 = \{\text{Количество публикаций каждого вида, Количество публикаций сотрудника, Количество публикаций кандидатов (докторов) наук}\}$ .

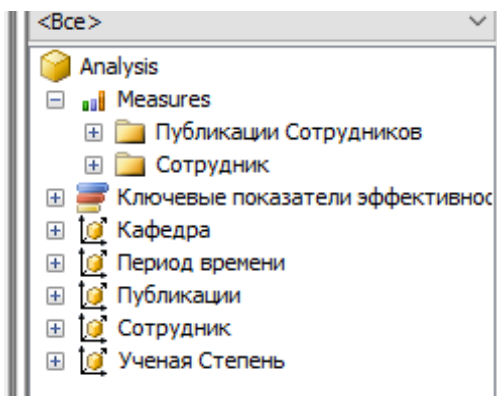


Рисунок 3 – Меры и измерения

Аналогично рассмотрев остальные показатели научной деятельности, получим многокубовую модель, применение к которой операций агрегации и детализации, среза и проекции, позволит оперативно проанализировать значимые показатели в различных разрезах и рассчитать агрегаты в динамике.

Источниками данных для анализа могут являться отчеты преподавателей за семестр, информационная система учреждения высшего образования, профили преподавателей в Google Академии.

Структура модели для системы в целом будет включать несколько секций с коллективными измерениями. Например, таковым является измерение «Сотрудники», так как сотрудники участвуют во всех видах деятельности.

Построение многомерной модели на основе преобразования реляционной схемы представляется наиболее целесообразным, так как в большинстве случаев источником данных для систем мониторинга является база данных информационной системы УВО.

Список литературы:

1. Григораш, О. В. Повышение эффективности управления качеством образовательного процесса / О. В. Григораш // Высшее образование в России. – 2013. – № 1. – С. 72–78.
2. Майоров, А.Н. Мониторинг в образовании [Текст] / А.Н. Майоров. – 3-е изд., испр. и доп. – М.: Интеллект-Центр, 2005.– 424с.

3. Зыкин, С.В. Автоматизация формирования представлений данных для их аналитической обработки / С.В Зыкин, А.Н. Полюянов // Вестник компьютерных и информационных технологий, № 4, 2010. – С. 3–9

4. Миронов, В.В. Построение OLAP-кубов на основе OLTP-ориентированной модели для анализа данных образовательного процесса / В. В. Миронов, Е. С. Макарова // Интеллектуальные технологии обработки информации и управления: сб. науч. тр. Междунар. молодеж. конф. Уфа: Аркаим, 2012. Т. 1. – С. 167–170.