

Моделирование предметных областей пользователей при использовании облачных технологий

В работе приводятся модели и методы формализованного описания предметных областей пользователей, анализа и классификации их на пользователей облачных баз данных (ОБД) и пользователей локальных БД (ЛБД), проектирования канонической структуры ОБД.

Ключевые слова: облачные технологии; предметная область пользователей облачных баз данных; функция подобия объектов предметной области; объектная модель требований пользователей; объектная каноническая структура облачной базы данных.

Парадигма облачных вычислений (cloudcomputing) в настоящее время активно используется при построении АИУС различного класса и назначения¹. Использование облака в качестве информационно-технологической среды распределенной обработки данных, предоставляющей на условиях аренды повсеместный сетевой доступ пользователям к пулу конфигурируемых по требованию ресурсов (серверам, системам хранения данных, приложениям, программному обеспечению, базам данных, информации и сервисам) провайдера облака, обеспечивает следующие важные преимущества для организации:

- существенное сокращение капитальных и операционных затрат на приобретение дорогостоящих средств вычислительной техники, лицензионного системного и прикладного программного обеспечения, их постоянное обновление и модернизацию, создание и содержание ИТ-служб и подразделений;
- доступность и мобильность: с помощью веб-сервисов доступ к информации и приложениям можно получать из любой

точки мира с помощью компьютеров и мобильных устройств (смартфонов, ноутбуков, планшетов), подключенных к сети Интернет;

- гибкость и высокая технологичность: необходимые ресурсы (информационные, вычислительные, программные) оперативно предоставляются, конфигурируются и освобождаются провайдером облака по требованию пользователей с минимальными эксплуатационными и временными затратами;
- экономичность: парадигма облачных вычислений основана на том, что требуемые ресурсы предоставляются пользователям как пакеты услуг по требованию и оплачиваются только за фактическое количество предоставленных ими функций. Основными услугами, предоставляемыми облачными технологиями, являются: SaaS (программное обеспечение как сервис), PaaS (платформа как сервис) и IaaS (инфраструктура как сервис). Разновидностями данных услуг являются такие, как STaaS (система хранения как сервис), DBaaS (база данных как сервис), INaaS (информация как сервис) и некоторые другие²;
- надежность: облака в виде полноценных центров обработки данных (ЦОД), включающих десятки и сотни серверов, разбросанных по сети Интернет, мощные системы хранения (СХД) и разнообразное ПО, создаются и поддерживаются специализированными организациями, укомплектованными штатом высококвалифицированных ИТ-специалистов. Создание таких ЦОД для многих организаций и предприятий, требующих значительных капитальных затрат, практически не представляется возможным.

В России облачные вычисления пока не получили столь широкого распространения, как на Западе. Объясняется это не только отсутствием понимания новой парадигмы, но и предельной осторожностью, связанной с рисками «перемещения» бизнес-процессов, программного и информационного обеспечения в облако и связанными с этим вопросами безопасности.

В зависимости от назначения различают 4 типа облаков³: частное, публичное, общественное и гибридное.

Частное облако (*private cloud*) представляется в виде ИТ-инфраструктуры, предназначенной для использования, как правило, одной организацией (предприятием). Частное облако может находиться в собственности, управлении и эксплуатации как самой организации, так и провайдера облака, либо их комбинации. Физически оно может размещаться как внутри организации, так и вне нее (на стороне провайдера облака).

Публичное облако (*public cloud*) представляет собой ИТ-инфраструктуру, предназначенную для свободного использования общественностью. Публичное облако может находиться в собственности, управлении и эксплуатации коммерческих, научных и правительственных организаций (или какой-либо их комбинации) и физически размещается на территории владельца – поставщика услуг. Пользователи не имеют возможности управлять данным облаком, и вся ответственность за его обслуживание возлагается на владельца ресурса.

Общественное облако (*community cloud*) – вид инфраструктуры, предназначенный для использования конкретным сообществом потребителей из организаций, имеющих общие задачи. Общественное облако может находиться в кооперативной (совместной) собственности, управлении и эксплуатации одной или более из организаций сообщества или третьей стороны (или какой-либо их комбинации), и оно может физически существовать как внутри, так и вне юрисдикции владельца

Гибридное облако (*hybrid cloud*) – это комбинация из двух или более различных облачных инфраструктур (частных и публичных). Часто такой тип облака применяется, когда внутренняя ИТ-инфраструктура частного облака не справляется с текущими задачами и часть мощностей и ресурсов перемещается на публичное облако (например, большие объемы статистической информации). Поэтому актуальной задачей, стоящей в настоящее время перед организациями – создателями облаков, является интеграция частных и общественных облаков⁴.

Для обеспечения пользователей требуемой информацией наиболее важными являются услуги по предоставлению доступа к информации (INaaS) и базам данных (DBaaS). Реализовать такие услуги при наличии проектных решений можно быстро, получив при этом значительный эффект. В то же время отсутствие формализованных моделей и методов анализа и структуризации информации предметных областей пользователей облачных вычислений не позволяет принимать обоснованных решений о возможности создания для отдельных групп пользователей частного облака, определения оптимального состава и структуры «облачной» БД.

Разработка блок-схемы и алгоритма анализа и структуризации предметных областей пользователей облачных технологий

В настоящем разделе рассмотрены методы кластерного анализа предметных областей пользователей, обеспечивающие классификацию их на пользователей, для которых целесообразно использование частного облака, и пользователей, для которых целесообразно

создание собственной (локальной) БД (ЛБД). Целесообразность определяется наличием заданной степени общности между предметными областями рассматриваемого множества пользователей. В дальнейшем для группы пользователей частного облака с использованием объектно-ориентированного подхода и объектно-ориентированной модели данных (ООМД) осуществляется построение и анализ объектных моделей требований пользователей, обобщенной объектной модели предметной области и объектной канонической структуры «облачной» БД (ОБД). Полученные результаты могут в дальнейшем использоваться при синтезе оптимальных логических и физических структур объектно-ориентированной ОБД. Предложенные методы основаны на результатах работ, полученных авторами в области анализа и синтеза оптимальных структур локальных, распределенных и объектно-ориентированных БД⁵. Блок-схема этапов анализа и структуризации предметных областей пользователей облачных технологий приведена на рисунке.

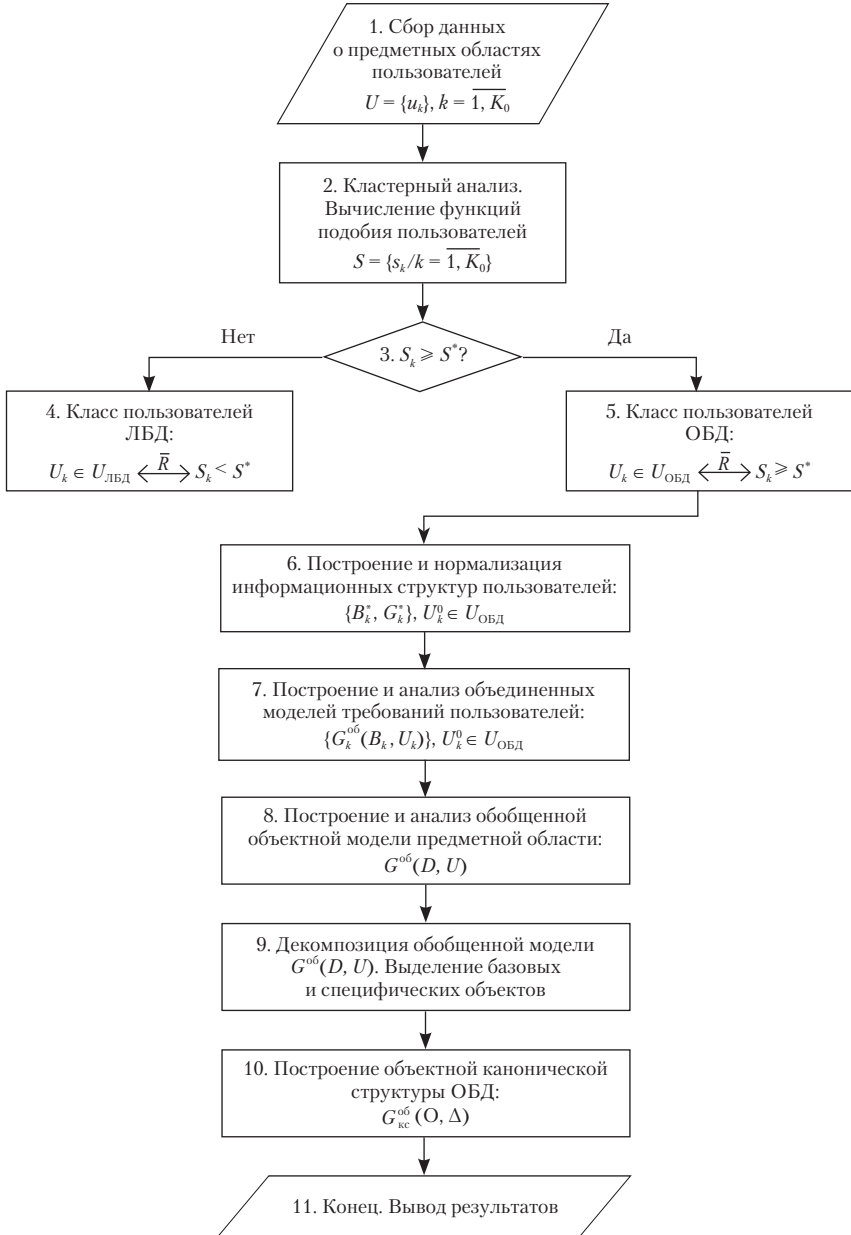
Рассмотрим методы кластерного анализа для классификации предметных областей пользователей. Предлагаемые методы основаны на процедурах анализа степени общности и оценке функции подобия предметных областей заданного множества пользователей.

Пусть $U = \{u_k\}$, $k = \overline{1, K_0}$ – множество территориально распределенных пользователей организации, $D = \{d_j\}$, $j = \overline{1, n_0}$ – полное без избыточное множество информационных элементов рассматриваемой совокупности пользователей, где n_0 – количество элементов в полном множестве.

Пусть $D_k = \{d_{ki}\}$, $i = \overline{1, L_k^0}$ – множество информационных элементов, описывающих предметную область k -го пользователя, L_k^0 – количество информационных элементов в множестве.

Отношение каждого пользователя к полному множеству информационных элементов D формализованно представляется в виде матрицы смежности $B = \|b_{kj}\|$, проиндексированной по осям множеством пользователей $U = \{u_k\}$ и полным множеством информационных элементов $D = \{d_j\}$. Элементы матрицы $B = \|b_{kj}\|$ принимают единичные значения, если информационный элемент d_j требуется пользователю u_k , в противном случае $b_{ki} = 0$.

На основании введенных исходных данных определим объединенное множество информационных элементов $D^0 = \{d_j\}$, $i = \overline{1, N}$, $N = \sum_k L_k^0$, получаемое путем объединения множеств D_k и содержащее повторяющиеся информационные элементы, т. е. $D^0 = \bigcup_{k=1}^{k_0} D_k$. Далее определим множество $D_k^y : D_k^y = D^0 - D_k$, $k = \overline{1, K_0}$.



Степень общности информации предметных областей пользователей определяется методом последовательного получения и анализа попарных пересечений информационных множеств пользователей с соответствующими множествами D_k^y .

Пусть $D_k^p = D_k \cap D_k^y$ – множество общих элементов информационного множества отдельного пользователя D_k и подмножества пользователей D_k^y . Если множество D_k^p является не пустым, а мощность пересечения удовлетворяет заданной величине, то предметная область k -го пользователя имеет достаточную степень общности с предметными областями других пользователей, что позволяет рассматривать его в качестве пользователя ОБД.

Для получения количественной характеристики степени общности предметной области k -го пользователя с объединенной предметной областью остальных пользователей воспользуемся понятием меры подобия.

Пусть заданы множества D_k и D_k^y . Тогда мерой подобия называется отображение пересечения множеств $D_k^p = D_k \cap D_k^y$ на некоторое множество вещественных чисел, выраженное неотрицательной вещественной функцией S_k , удовлетворяющей условию $0 \leq S_k \leq 1$.

В теории автоматической классификации используется ряд функций для вычисления мер подобия между объектами, представленных в таблице на стр. 80. Большинство функций подобия включает в свой состав величину p_{00} , которая определяет количество элементов, одновременно отсутствующих в обоих рассматриваемых множествах. Так как по определению $D^0 = \bigcup_k D_k$ и $D_k^y = D^0 - D_k$, то величина p_{00} всегда принимает нулевое значение. Таким образом, использование ряда функций подобия (S_2, S_5, S_7) приводит к вырожденным оценкам, не обеспечивающим адекватности сравнения. Анализ функций подобия S_1, S_3, S_4, S_6 показывает, что они обеспечивают сравнимость результатов. Среди них функция S_6 наиболее точно отражает степень общности между двумя множествами, так как позволяет учесть общие (p_{11}) и специфические (p_{10} и p_{01}) информационные элементы в обоих множествах.

Поэтому меру подобия предлагается вычислять с использованием выражения S_6 , где $n_0 = p_{11} + p_{10} + p_{01}$ – общее количество различных информационных элементов в рассматриваемых множествах, равное количеству элементов в полном множестве $D = \{d_i\}$; p_{11} – количество общих элементов в информационных множествах D_k и D_k^y .

Таблица

Код	Стандартная форма функции подобия	Вид функции в матричной модели
S_1	$\frac{p_{11}}{p_{11} + p_{10} + p_{01}}$	$\frac{p_{11}}{n_0}$
S_2	$\frac{p_{11} + p_{00}}{p}$	$\frac{p_{11}}{n_0}$
S_3	$\frac{p_{11}}{p}$	$\frac{p_{11}}{n_0}$
S_4	$\frac{2p_{11}}{2p_{11} + p_{10} + p_{01}}$	$\frac{2p_{11}}{p_{11} + n_0}$
S_5	$\frac{2(p_{11} + p_{01})}{p_{11} + p_{10} + p_{00}}$	$\frac{2(p_{11} + p_{01})}{p_{11} + n_0}$
S_6	$\frac{p_{11}}{p_{11} + 2(p_{10} + p_{01})}$	$\frac{p_{11}}{n_0 + p_{10} + p_{01}}$
S_7	$\frac{p_{11} + p_{00}}{p + p_{10} + p_{01}}$	$\frac{p_{11}}{n_0 + p_{10} + p_{01}}$

Величину p_{11} предлагается рассчитывать на основании характеристик матрицы $B = \|b_{kj}\|$ следующим образом: $p_{11} = \sum_{j=1}^{n_0} Z_j$, где $Z_j = 1$, если $\exists j: d_j \in D_k$ и $\sum_{k=1}^{K_0} b_{kj} \geq 1$; в противном случае $Z_j = 0$; p_{10} – количество элементов, принадлежащих множеству D_k , но отсутствующих в множестве D_k^y . Величина p_{10} рассчитывается как $p_{10} = \sum_{j=1}^{n_0} X_j$, где $X_j = 1$, если $\exists j: d_j \in D_k$ и $\sum_{k=1}^{K_0} b_{kj} = 1$; в противном случае $X_j = 0$; p_{01} – количество элементов, отсутствующих во множестве D_k , но принадлежащих множеству D_k^y , p_{01} определяется в виде $p_{01} = \sum_{j=1}^{n_0} Y_j$, где $Y_j = 1$, если $\exists j: d_j \notin D_k$ и $\sum_{j=1}^{n_0} b_{kj} \geq 1$; в противном случае $Y_j = 0$.

С учетом введенных обозначений функция подобия S_6 принимает вид:

$$S_k = \frac{\sum_j Z_j}{\sum_j Z_j + 2(\sum_j X_j + \sum_j Z_j Y_j)}.$$

Зададим на множестве пользователей $U = \{u_k\}$ отношение принадлежности R , которое определяется величиной меры подобия из множества $S = \{S_k\} : U_k \in U_{\text{ОБД}} \xleftrightarrow{\bar{R}} S_k \geq S^*$, где $U_{\text{ОБД}}$ – множество пользователей ОБД, S^* – критическая мера подобия.

Отношение R принадлежности пользователей к классу $U_{\text{ЛБД}}$, для которого целесообразно проектирование локальных БД, определяются как $U_k \in U_{\text{ЛБД}} \xleftrightarrow{\bar{R}} S_k < S^*$.

Соотношение $S_k \geq S^*$ обычно выполняется в случае наличия пользователей, решающих близкие по функциональному назначению и характеристикам задачи управления и переработки информации.

Таким образом, в результате выполнения процедур кластерного анализа формируется множество пользователей «облачной» БД (ОБД) $U_k^0 \in U_{\text{ОБД}}$ и множество пользователей локальных БД (ЛБД) $U_k^0 \in U_{\text{ЛБД}}$.

Разработка методов построения объектной канонической структуры ОБД

В данном разделе статьи для группы пользователей $U_k \in U_{\text{ОБД}}$ предложены модели и методы построения внешних моделей, обобщенной внешней модели и объектной канонической структуры ОБД на основе методов объектно-ориентированного проектирования.

Использование данных методов неслучайно. Они наиболее адекватно отражают технологию облачных вычислений и предоставляемых услуг, поскольку предполагают инкапсуляцию в одном объекте как данных, так и методов (процедур) их обработки. Это позволяет создавать БД открытой архитектуры, с присущими ей свойствами переносимости, мобильности и интероперабельности; сократить сроки, стоимость и трудоемкость разработки БД за счет распараллеливания процесса создания структур БД и прикладных программ между коллективами разработчиков, а также возможности повторного использования наследованных объектов; обеспечить простоту сопровождения БД и приложений за счет использования принципа прозрачности информации, при котором обеспечивается сокрытие в объекте как процедур обработки данных, так и самих данных; свойство полиморфизма позволяет использовать одни и те же процедуры путем отправки сообщений для разных объектов, различающихся только переменными вызовов.

Введем ряд необходимых определений.

Под объектом предметной области будем понимать некоторую совокупность информационных элементов и методов (процедур) их обработки, а также отношений между ними, составляющих единое целое с точки зрения семантических и процедурных аспектов предметной области.

Под объектной моделью требований пользователей будем понимать информационно-функциональную структуру, формируемую в результате выполнения операций наложения на графы информационных структур пользователей технологических составляющих, описывающих процедуры поиска и обработки данных.

Под обобщенной объектной моделью предметной области будем понимать информационно-функциональную структуру, формируемую в результате операций наложения объектных моделей требований пользователей.

Под анализом обобщенной объектной модели предметной области понимается процесс сведения многообразия зафиксированных в ней объектных моделей требований пользователей к базовым (типовым) и специфическим объектам, в результате которого осуществляется построение объектной канонической структуры БД.

Проектирование объектных моделей требований пользователей осуществляется в 2 этапа.

На первом этапе с помощью предложенных в работах⁶ структурных методов осуществляется построение и нормализация информационных структур пользователей.

На втором этапе осуществляется отображение на сформированные информационные структуры требований пользователей по обработке данных.

Исходными данными для реализации этого этапа являются:

а) формализованные описания нормализованных информационных структур пользователей, задаваемые в виде матриц смежности B_k^* и орграфов G_k^* , на которых выделены групповые и информационные элементы, ключи и атрибуты групп данных;

б) формализованные описания требований пользователей по обработке данных.

Формализованные описания требований k -го пользователя по обработке данных задаются с помощью:

- множества процедур обработки данных $F_k = \{f_r^k / r = \overline{1, R_k}\}$, где f_r^k – r -я процедура k -го пользователя;
- матрицы использования l -го информационного элемента ($l \in I_k \subseteq L$) r -й процедурой ($r \in R_k \subseteq R$) $W_k = \|w_{rl}^k\|$. Элемент $w_{rl}^k = 1$, если l -й информационный элемент используется r -й процедурой; в противном случае $w_{rl}^k = 0$;

– структуры поиска требуемых для обработки данных, задаваемой в виде дерева поиска данных на графе информационной структуры G_k^* .

Отображение требований обработки данных на графах информационных структур пользователей производится следующим образом.

Использование некоторой r -й процедурой обработки данных l -го элемента формально представляется на графе G_k^* петлей на группе d_j^r , в которую он входит, что свидетельствует об обработке данного группового элемента.

Деревья поиска, требуемые для обработки данных, отображаются дополнительными дугами на графе G_k^* .

Таким образом, формализованно объектная модель требований k -го пользователя представляется в виде мультиграфа с одним типом вершин и двумя типами дуг $G_k^{об}(D_k, U_k)$, где $D_k = \{d_l^k / l = \overline{1, L_k}, L_k \subseteq L\}$ – множество информационных элементов (включая ключи и атрибуты данных), выявленные в результате выполнения структурных методов, описанных в работе В.В. Кульбы и др.⁷; $U_k = U_k^{эл} \cup U_k^{пр}$, где $U_k^{эл}$ – множество дуг, характеризующих структуру взаимосвязей между информационными элементами (группами данных, ключами и атрибутами), а $U_k^{пр}$ – множество дуг, характеризующих технологию обработки данных для k -го пользователя в виде реализации совокупности методов (процедур) поиска и непосредственной обработки данных, включая петли и непосредственно дуги.

Основные характеристики объектной модели требований $G_k^{об}$

1. Вектор технологических весов вершин $Z_\mu^k = \{z_{lk}^\mu\}$, где z_{lk}^μ – технологический вес d_l -й вершины графа $G_k^{об}$.

2. Вектор технологических толщин дуг $Z_\eta^k = \{z_{ll',k}^\eta\}$, где $z_{ll',k}^\eta$ – технологическая толщина дуги на графе.

Технологический вес $z_{lk}^\mu = \text{const}$, где $\text{const} \in \{0, 1, 2, \dots, N\}$, N – множество целых чисел означает степень использования информационного элемента d_l множеством процедур обработки данных. Чем больше значение z_{lk}^μ для некоторого d_l , тем более важным является элемент d_l в процессах обработки данных.

Технологическая толщина дуги (d_p, d_l) $z_{ll',k}^\eta = \text{const}$ (где $\text{const} \in \{0, 1, 2, \dots, N\}$, N – множество целых чисел) означает степень использования дуги (ll') в процессах поиска требуемых данных. Чем больше значение $z_{ll',k}^\eta$ для некоторой дуги (d_p, d_l) , тем более часто используется данная дуга в путях доступа к требуемым для обработки данным.

Определение технологических весов вершин осуществляется путем суммирования значений матрицы по столбцам (информационным элементам): $z_{lk}^\mu = \sum_{\forall r \in R_k} w_{rl}^k$.

Определение технологических толщин дуг осуществляется путем суммирования дуг, входящих в множество U_k^{np} :

$$z_{ll',k}^{\text{н}} = \sum_{l,l' \in U_k^{\text{np}}} (d_l, d_{l'}).$$

Вычислив основные характеристики графа $G_k^{\text{об}}$, можно перейти к другому формализованному представлению объектной модели требований пользователей – взвешенному графу $G_k^{\text{взв}}(D_k^{\text{взв}}, U_k^{\text{взв}})$, где каждой вершине $d_l \in D_k^{\text{взв}}$ и дуге $(d_b, d_{l'}) \in U_k^{\text{взв}}$ приписаны соответствующие им веса.

Построение обобщенной объектной модели предметной области осуществляется после построения объектных моделей для всех требований пользователей путем последовательного наложения мультиграфов $G_k^{\text{об}}(D_k, U_k)$ друг на друга. Разработанная процедура основана на совмещении идентичных информационных элементов независимо от уровня их размещения на графах $G_k^{\text{об}}(D_k, U_k)$.

Результатом выполнения процедуры наложения является упорядоченный по уровням иерархии мультиграф обобщенной модели предметной области $G^{\text{об}}(D, U)$ с одним типом вершин $D = \{d_l / l = \overline{1, L}\}$, соответствующих множеству информационных элементов, и двумя типами дуг: $U^{\text{эл}}$ – множество информационных взаимосвязей между элементами $d_l \in D$ и $U^{\text{пр}}$ – множество процедурных (технологических) взаимосвязей между информационными элементами, $U = U^{\text{эл}} \cup U^{\text{пр}}$.

Основные характеристики обобщенной объектной модели предметной области

1. Вектор информационных весов вершин графа $G^{\text{об}}$: $Z_v = \{z_l^v\}$, где z_l^v – информационный вес вершины $d_l \in D$. Информационный вес для некоторой вершины d_l равен целому положительному числу $(0, 1, 2 \dots N)$ и характеризует степень потребности множества пользователей в данном элементе. Чем больше значение z_l^v , тем более важным и необходимым является элемент d_l для удовлетворения информационных потребностей пользователей.

2. Вектор технологических весов вершин графа $G^{\text{об}}$: $Z_\mu = \{z_l^\mu\}$, где z_l^μ – технологический вес d_l -го элемента (вершины), $z_l^\mu \in \{0, 1, \dots, N\}$.

3. Вектор информационных толщин дуг графа $G^{\text{об}}$: $Z_\theta = \{z_{ll'}^\theta\}$, где $z_{ll'}^\theta$ – информационная толщина дуги $(d_l, d_{l'})$. Информационная толщина дуги равна целому положительному числу $(0, 1, 2 \dots N)$ и характеризует степень семантической (смысловой) связности элементов d_l и $d_{l'}$ в заданной предметной области. Чем больше значение $z_{ll'}^\theta$, тем более семантически связаны (ассоциированы) элементы

d_l и $d_{l'}$ что подтверждено подмножеством пользователей в их информационных требованиях.

4. Вектор технологических толщин дуг графа $G^{об}$: $Z_\eta = \{z_{ll'}^\eta\}$, где $z_{ll'}^\eta$ – технологическая толщина дуги $(d_l, d_{l'})$, $z_{ll'}^\eta \in \{0, 1, 2 \dots N\}$.

Далее на основе анализа обобщенной объектной модели предметной области осуществляется построение объектной канонической структуры ОБД путем сведения многообразия объектных моделей требований пользователей, зафиксированных в обобщенной объектной модели, к базовым (типовым) и специфическим объектам.

На первом этапе анализа осуществляется переход от мультиграфа $G^{об}(D, U)$ к взвешенному графу $G^{взв}(D, U)$ обобщенной объектной модели предметной области. С этой целью для каждой вершины графа $G^{об}$ вычисляются их средние веса, которые рассчитываются по формуле

$$z_l^{cp} = \frac{z_l^v + z_l^h}{2}, \forall d_l \in D,$$

а для каждой связи – средние веса толщин дуг:

$$z_{ll'}^{cp} = \frac{z_{ll'}^v + z_{ll'}^h}{2}, \forall (d_l, d_{l'}) \in U.$$

Взвешенный граф $G^{взв}(D, U)$ представляется в виде орграфа с одним типом вершин и дуг, каждая из которых имеет определенный средний вес. Графу $G^{взв}(D, U)$ ставится в соответствие взвешенная матрица смежности $B = \|b_{ij}\|$. Элементы b_{ij} матрицы B , лежащие не на главной диагонали, равны положительным вещественным числам, соответствующим весам толщин дуг. Элементы b_{ii} матрицы B , лежащие на главной диагонали, равны положительным вещественным числам и соответствуют средним весам вершин графа $G^{взв}$.

На втором этапе анализа решается задача декомпозиции взвешенного графа $G^{взв}$ на ряд подграфов, соответствующих базовым (типовым) и специфическим объектам предметной области. Данная задача решается при следующих ограничениях:

а) на количество информационных элементов в одном объекте (N);

б) на количество процедур (методов) обработки данных в одном объекте (M);

в) на возможность включения отдельных информационных элементов в состав одного объекта. Данное ограничение формаль-

но можно представить в виде матрицы смежности $\bar{B} = \|\bar{b}_{ij}\|$, элементы которой: $\bar{b}_{ij} = 1$, если информационные элементы d_i и d_j семантически несовместимы в составе одного объекта, и $\bar{b}_{ij} = 0$, если появление элементов d_i и d_j в составе одного объекта допустимо;

г) на уровень типовости, при котором конструируемые объекты (как подграфы графа $G^{\text{взв}}$) можно отнести к базовым объектам заданной предметной области (Y).

Уровень типовости задается проектировщиком ОБД и определяется как интегральная характеристика информационной и технологической связности объектов.

Значения ограничений N , M и Y выбираются таким образом, чтобы, по крайней мере, отдельная вершина графа $G^{\text{взв}}(D, U)$ представляла из себя отдельный объект предметной области.

Для решения поставленной задачи может быть использован алгоритм, приведенный в работе⁸. В результате его выполнения осуществляется преобразование:

$$G^{\text{взв}}(D, U) \rightarrow G_{\text{кс}}^{\text{об}}(O, \Delta),$$

где $G_{\text{кс}}^{\text{об}}(O, \Delta)$ – граф объектной канонической структуры ОБД, вершинами которого $O = \{O_{\varepsilon}/\varepsilon = \overline{1, \varepsilon_0}\}$ являются объекты предметной области, а дугами $\Delta = \{\delta_{\varepsilon\varepsilon'}/\varepsilon, \varepsilon' = \overline{1, \varepsilon_0}\}$ – связи (или отношения) между объектами.

Характеристиками графа $G_{\text{кс}}^{\text{об}}$ являются интегральные характеристики объектов и связей между ними. На сформированной объектной канонической структуре ОБД выделено множество базовых $O_{\text{баз}}$ и специфических $O_{\text{спец}}$ объектов.

Заключение

Таким образом, предложенные в работе методы и алгоритмы позволяют классифицировать пользователей на группу пользователей частного облака и пользователей локальных БД, сформировать внешние модели, обобщенную внешнюю модель и объектную каноническую структуру облачной БД. В дальнейшем данные результаты используются при синтезе логических и физических структур ОБД облачных технологий. Разработанные методы и алгоритмы использовались при разработке ряда проектов информационной инфраструктуры Евразийского патентного ведомства Евразийской патентной организации.

- ¹ Cloud Computing: Principles, Systems and Applications / N. Antonopoulos, L. Gil- lam (eds.) L.: Springer, 2010. 379 p.; *Батура Т.В., Мурзин Ф.А., Семич Д.Ф.* Облачные технологии: Основные модели, приложения, концепции и тенденции развития // Программные продукты и системы. 2014. № 3 (107). С. 64–72.
- ² Cloud Computing...; *Батура Т.В., Мурзин Ф.А., Семич Д.Ф.* Указ. соч.; *Mell P., Grance T.* The NIST Definition of Cloud Computing: Recommendations of the National Institute of Standards and Technology. NIST, 2011.
- ³ *Mell P., Grance T.* Op. cit.
- ⁴ *Черняк Л.* Интеграция – основа облака // Открытые системы. СУБД. 2011. № 7.
- ⁵ *Мамиконов А.Г., Ашимов А.А., Кульба В.В., Косяченко С.А., Сиротюк В.О.* Оптимизация структур данных в АСУ. М.: Наука, 1988. 256 с.; *Кульба В.В., Ковалевский С.С., Косяченко С.А., Сиротюк В.О.* Теоретические основы проектирования оптимальных структур распределенных баз данных. М.: СИНТЕГ, 1999. 660 с. (Серия «Информатизации России на пороге XXI века»); *Кульба В.В., Микрин Е.А., Сиротюк В.О., Сиротюк О.В.* Модели и методы проектирования оптимальных структур объектно-ориентированных баз данных в автоматизированных информационно-управляющих системах. М.: ИПУ РАН, 2005. 103 с.
- ⁶ *Мамиконов А.Г., Ашимов А.А., Кульба В.В., Косяченко С.А., Сиротюк В.О.* Указ. соч.; *Кульба В.В., Ковалевский С.С., Косяченко С.А., Сиротюк В.О.* Указ. соч.
- ⁷ *Кульба В.В., Ковалевский С.С., Косяченко С.А., Сиротюк В.О.* Теоретические основы проектирования оптимальных структур распределенных баз данных. М.: СИНТЕГ, 1999. 660 с. (Серия «Информатизации России на пороге XXI века»).
- ⁸ *Кульба В.В., Микрин Е.А., Сиротюк В.О., Сиротюк О.В.* Указ. соч.