

Проблемы интеграции и доступа в гетерогенных распределенных базах данных

В.С. РАМАЗАНОВА, Н.П. АЗАНОВ

Казахский национальный университет им. аль-Фараби, Алматы

e-mail: azanovnp@zmail.ru

Аннотация

В работе исследуются проблемы интеграции и доступа в гетерогенных распределенных базах данных на основе программных продуктов различных производителей.

Внедрение корпоративных сетей и средств хранения информации привело к развитию распределенных информационных систем, реализуемых на основе разнообразных баз данных (БД), управляемых как современными, так и "морально"устаревшими системами управления БД. Это привело к появлению гетерогенного информационного пространства и соответственно гетерогенных распределенных БД [13]. Организация поиска информации в таких базах данных требует, в первую очередь, решить задачи интеграции и доступа к различным типам данным, находящимся в БД различных производителей. Проблема также состоит в сложности доступа к распределенной информации в гетерогенных БД.

Определим виды гетерогенности. Литвин В. [4] делит гетерогенность данных на физическую и семантическую. Физическая гетерогенность подразумевает различия в представлении данных и может выражаться в различии типов данных, реализаций моделей данных, языков описания процедур, триггеров, языков запросов, манипулирования и определения данных. Кроме этого, могут различаться и сами модели данных. Данные могут быть представлены не только в БД, но и в электронных таблицах, почтовых файлах и т.д. Семантическая гетерогенность проявляется в различиях в наименованиях данных, значениях и логических структурах.

Кудашев Е.Б. [17] добавляет, что гетерогенность присуща не только данным, но и системам. Могут быть определены следующие разновидности гетерогенности информационных систем, которые содержат и обрабатывают данные:

1. аппаратная гетерогенность, или гетерогенность аппаратного обеспечения;
2. гетерогенность программного обеспечения, которую можно, в свою очередь, разделить на:
 - интерфейсную гетерогенность;
 - гетерогенность прикладных программ;
 - гетерогенность операционных систем.

Такое разнообразие видов гетерогенности создает существенные проблемы доступа к данным, технологиям и информационным системам, что, в свою очередь, негативно влияет на процесс поиска в распределенных гетерогенных БД.

Кашников А. и Лядова Л. [15] отмечают, что проблема физической гетерогенности решается посредством введения стандартов взаимодействия (ODBC, DAO, OLE DB, ADO, ADO.NET). Достоинством универсальных механизмов является возможность применения одних и тех же средств доступа к разным типам источников, поэтому приложения легко модифицировать, если необходима замена СУБД. Но за универсальность приходится платить невозможностью доступа к уникальной функциональности, специфичной для конкретной СУБД, снижением производительности. В области семантической гетерогенности проводится большое количество исследований, но до сих пор не появилось какого-либо стандарта. Для решения этой проблемы были разработаны прототипы реляционных языков (MSQL, IDEAL). Многие специальные функции этих языков нуждаются в выразительной мощности, выходящей за пределы существующих реляционных систем, диалектов SQL. Требуются мультиреляционные операции и логика исчисления предикатов более высокого порядка, тогда как SQL основан на реляционной алгебре и логике предикатов первого порядка.

Черняк Л. [21] уточняет понятия консолидации и интеграции. Консолидация является более простой формой объединения информации, когда работа с данными ограничена составлением отчетов и нет потребности в применении серьезных аналитических методов. Интеграция отличается тем, что она представляет пользователю единообразный взгляд на разнородные источники данных, что предполагает общую модель и общее отношение к семантике, с тем чтобы обеспечить возможность для доступа к данным, а в случае необходимости разрешать конфликтные ситуации.

Он выделяет отличие в интеграции данных от интеграции информации в отсутствие семантики.

Рассмотрим интерпретированную схему классификаций технологий интеграции данных, разработанную Клаусом Дитрихом:

1. Common Data Storage (Общие системы хранения). Осуществляется за счет слияния данных из разных систем хранения данных в одну общую. Сегодня мы бы объединили эти два уровня в один и назвали бы его виртуализацией систем хранения.
2. Uniform Data Access (Унифицированный доступ к данным). На этом уровне осуществляется логическая интеграция данных, различные приложения получают единообразное видение физически распределенных данных. Такая виртуализация данных имеет свои несомненные достоинства, но гомогенизация данных в процессе работы с ними требует значительных ресурсов.
3. Integration by Middleware (Интеграция средствами ПО промежуточного слоя). ПО этого слоя играет посредническую роль, его составляющие способны к выполнению отдельных предписанных им функций, в полном объеме интеграционная задача решается во взаимодействии с приложениями.
4. Integration by Applications (Интеграция средствами приложений). Обеспечивает доступ к разным источникам данных и возвращает пользователю обобщенные результаты. Сложность интеграции на этом уровне объясняется большим разнообразием интерфейсов и форматов данных.

5. Common User Interface (Общий пользовательский интерфейс). Дает возможность единообразного доступа к данным, например, с помощью браузера, но при этом данные остаются неинтегрированными и неоднородными.
6. Manual Integration (Интеграция вручную). Пользователь сам объединяет данные, применяя различные типы интерфейсов и языки запросов.

Приведенная схема Дитриха интересна тем, что позволяет связать вместе интеграцию данных с интеграцией информации – по мере продвижения снизу вверх простые атомарные данные обретают семантику.

Таким образом, автор утверждает, что в настоящее время ни о какой реальной интеграции информации речи пока быть не может: все, что существует сегодня, сводится лишь к консолидации данных на том или ином уровне абстракции.

Елмагармид А., Русинкевич М. и Снетч А. [1] выделяют три категории моделей интеграции данных: модель с глобальной схемой, интегрированные базы данных и подход на основе языка мультибаз данных.

Кент В. [3] пишет, что вне зависимости от выбранных технологий и метода интеграции данных, остаются вопросы, связанные с их смысловой интерпретацией и различиями в представлении одних и тех же вещей. Приходится разрешать несоответствие схем данных и несоответствие самих данных. В несоответствии схем данных он выделяет структурные и семантические проблемы:

1. Различие в типах данных. Некоторый домен в одном источнике может представляться числом, в другом – строкой фиксированной длины, в третьем – строкой переменной длины.
2. Различие в единицах измерения. В одной БД указана величина в сантиметрах, в другой – в дюймах. В этом случае существует отображение 1:1.
3. Различие в множестве допустимых значений. Один и тот же признак может определяться разными наборами констант. Например, выполнение задания одним источником может оцениваться по четырехбалльной шкале (неудовлетворительно, удовлетворительно, хорошо, отлично), другим – по трехбалльной (–, ±, +), третьим – по столбальной. Отображение не является 1:1, оно может быть многозначным, может не иметь обратного, может зависеть от сторонних данных (например, 30 по математике соответствовать «удовлетворительно», а по русскому языку – «неудовлетворительно»).
4. Различие «домен-отношение». Домен в одной БД (например строковое значение) соответствует таблице в другой БД (записи из таблицы-справочника).
5. Различие «домен – группа доменов». В одном источнике адрес записывается одной строкой, в другом – отдельные поля для улицы, дома, строения, квартиры.
6. Различие «данные-схема». Данные одной БД соответствуют схеме (метаданным) другой. В одной БД «инженер» – значение атрибута «должность» отношения «работник», в другой «инженеры» – отношение, содержащее некоторых работников, в то время как «бухгалтеры» содержит других.

7. Отсутствующие значения. В каком-то из источников может отсутствовать информация, имеющаяся в большинстве других.

Типы несоответствия собственно данных:

1. Различие формата данных. «ул. Манаса, 18-1» или «Манаса, д.18, стр.1»; «8(777)-234-45-32» или «8-777-234-45-32»
2. Различие в представлении значений. Например, некая организация может быть записана в отдельных источниках как «Карагандинский металлургический комбинат», «КМК», «ОАО КМК».
3. Потеря актуальности данных одним из источников. Например, смена фамилии при замужестве: в одной БД записана новая фамилия, в другой старая, и они не совпадают.
4. Наличие ошибок операторского ввода (или ошибок распознавания бланков) в отдельных источниках данных. Сюда относятся механические опечатки, ошибки восприятия на слух сложнопроизносимых имен/названий, отсутствие единых стандартов транскрипции с иностранных языков.
5. Намеренное внесение искажений с целью затруднить идентификацию сущностей.

Рассмотрим проблемы физической интеграции данных.

Пыхалов А.В. и Шаройко О.В. [20] делят проблемы, возникающие при получении данных из различных БД на две части: проблемы, связанные с отсутствием стандартного API для доступа к различным СУБД; проблемы, возникающие при использовании в одном запросе данных из разных БД. Стандарты ODBC и OLE DB частично обеспечивают унифицированный доступ к различным СУБД (если не учитывать различия в SQL-диалектах, использующихся в разных СУБД). Более подробно они рассматривают проблемы, относящиеся ко второй части – одновременная работа с несколькими СУБД. Не существует стандартного решения для предоставления в совместное использование разнородных источников данных, а также, не существует OpenSource решений, позволяющих предоставить единый интерфейс для запросов одновременно к нескольким разнотипным СУБД.

Для решения этой проблемы предлагается создание ODBC драйвера по причине распространенности интерфейса ODBC и существования доступных шаблонов ODBC драйверов в рамках проекта UnixODBC, что существенно упрощает процесс реализации драйвера. Среди недостатков данного метода стоит отметить возрастание нагрузки на клиентский компьютер при анализе и выполнении запросов, а также возрастание нагрузки на канал связи с клиентом, так как происходит избыточная передача данных между клиентом и сервером.

Блажко А.А. и др. [10] отмечают, что в распределенных гетерогенных базах данных приходится решать трудоемкую задачу: «прозрачно» для пользователей согласовывать содержимое БД и одновременно управлять доступом пользователей или процессов к БД. Существует множество решений по согласованию ГБД с использованием протоколов

тиражирования на основе middleware-программных шлюзов или триггеров СУБД. Управление доступом к автономным БД всегда должно быть санкционированным, что в современных ИС реализуется на основе избирательной или полномочной избирательной или полномочной политик безопасности.

Вышеописанные задачи должны решаться в комплексе, иначе лучшее решение по согласованию ГБД с точки зрения производительности и целостности будет конфликтовать с правилами управления доступом в автономных БД.

Унификацию процесса управления предложено проводить на основе модели процесса гетерогенного тиражирования операций согласования ГБД в виде пятерки

$$\langle DS, SM, RS, AQ, AL \rangle,$$

где DS – граф описания структур распределенных БД; SM – множество элементов-описаний соответствия между логическими структурами автономных БД; RS – множество элементов-описаний схемы тиражирования операций согласования; AQ – упорядоченное множество элементов-описаний операций асинхронного согласования БД; AL – множество элементов-описаний операций доступа к БД с учетом идеологии безопасности, применяемой в ИС.

На основании предложенных формализаций упрощается автоматизация санкционированного управления процессом согласования в гетерогенных тиражируемых БД за счет создания алгоритмов: автоматического формирования множества элементов-описаний схем тиражирования операций согласования на основе множества описаний структур распределенных БД с удовлетворением применяемой в организации политики безопасности; генерации программного кода процедур согласования ГБД с использованием программных шаблонов *OTPL*; генерации команд управления доступом на основании множества описаний операций доступа к БД с учетом идеологии применяемой в ИС безопасности.

Антипин К.В. и др [8] предлагают решить проблему унифицированного доступа к гетерогенным источникам данным, как если бы они имели единое логическое и физическое представление посредством системы виртуальной интеграции BizQuery основанной на использовании технологий XML и UML.

Система разработана группой MODIS Института системного программирования РАН, которая на протяжении последних трех лет занимается вопросами исследования и разработки методов управления XML-данными. BizQuery обладает следующими основными чертами:

- обеспечение интегрированного доступа к нескольким источникам данных, каждый из которых может содержать реляционные данные, либо XML-данные;
- использование XML как для внутреннего представления данных, так и для представления результирующих данных запроса;
- представление глобальной схемы интегрированных данных как в терминах UML, так и в терминах XML;
- использование декларативных языков запросов UQL (разработан в ходе выполнения проекта [4]) и XQuery для формулировки запросов к интегрированным данным в терминах UML и XML соответственно;

- полнофункциональная обработка запросов, включающая оптимизацию запросов, их декомпозицию с выделением частичных запросов, адресованных индивидуальным источникам данных, и формирование окончательного результата путем выполнения соединений и трансформаций частичных результатов.

Согласно тестам в независимости от общего времени выполнения, чистое время работы VizQuery (включающее оптимизацию и, если необходимо, выполнение кроссдоменных операций) относительно невелико. Это обеспечивается посредством перезаписи запросов и декомпозиции запросов, которое приводит к выделению наиболее ресурсоемкой части запроса источнику. Необходимо заметить, что общее время выполнения запроса можно снизить, вводя дополнительные индексы в источниках.

Стандарт языка запросов XQuery платформы XML воплощает функциональность, свойственную интегрирующей модели данных. Базовая модель данных этого языка поддерживает иерархические и реляционные структуры данных и, таким образом, обеспечивает возможности для интеграции XML-данных и данных, содержащихся в реляционных базах данных.

Перейдем к рассмотрению проблем семантической интеграции.

Белодед Б.В. и др. [9] решают проблемы семантической интеграции и доступа к БД на основе единой терминологии понятной пользователю. На сегодняшний день для описания и формализации предметных областей используются онтологии. В данном случае, онтология будет представлять собой иерархически структурированное множество терминов. Одним из ключевых моментов использования онтологии является концептуализация – первичная теоретическая форма, рассматриваемая независимо от словаря предметной области и конкретной ситуации. Основными компонентами онтологии являются классы (или понятия), отношения, функции, аксиомы, экземпляры [7].

Для реализации информационной системы предложенным способом были решены следующие задачи: построение базы метаданных для интеграции БД и онтологии; преобразование запроса пользователя в SPARQL запрос к онтологии; преобразование SPARQL запроса в SQL запросы к БД.

Данный подход предоставляет единый унифицированный интерфейс для получения данных. При этом конечному пользователю предоставляется возможность формулировать свой запрос в терминах предметной области.

Жучков А.В. [14] и другие также утверждают, что наиболее перспективным направлением решения проблемы интеграции и поиска семантически разнородной информации являются технологии семантического связывания на основе онтологий. В их работе под онтологией понимается эксплицитная (явная, вербализованная, формальная) спецификация некоторой концепции. При этом речь идет об онтологиях уровня предметной области (domain specific) или конкретной задачи (task specific).

В качестве инструмента для работы с онтологиями было разработано оригинальное программное обеспечение G-Ontology. Созданное ПО совместимо с созданным ранее ПО Gozelle, которое обеспечивает функциональность ИС в целом: поиск данных, создание и модификация данных; различные интерфейсы взаимодействия со сторонними источниками данных; работа с авторскими и распределенными наборами данных; работа с метаданными; различные фильтры и конверторы для работы с гетерогенными распределенными электронными библиотеками; различные; разнообразные поисковые средства.

Использование онтологий открывает новые возможности при создании современных электронных библиотек. Особенно ярко преимущества этой технологии проявляются в больших распределенных гетерогенных информационных системах.

Новицкий А.В. [19] продолжает тему использования онтологий для решения семантических проблем в электронных библиотеках. В процессе обзора семантической интероперабельности в Европейских проектах он выделяет ряд таких проблем, как:

- Процессор генерирования запросов SeRQL не знает, в каком классе или свойстве могут быть найдены слова.
- Существуют некоторые ограничения в подстроках соответствия SeRQL при использовании символа общности <*>.
- Каждый справочник имеет свою поисковую систему и использует свою грамматику метаданных для описания и индексации данных, в частности, она никогда не будет работать на других системах.

Огромное количество различных форматов, используемых для индексирования данных, является большим препятствием на пути к интеграции, и должны быть серьезно проанализированы.

Чтобы создать единый концептуальный слой, семантическая информация должна быть взята из базы данных, HTML-страниц, описательных текстов, метаданных и представлена в стандартном формате, с целью получения концептуального содержания информации, создав концептуальный маппинг. Как только концептуальный слой для данных и метаданных готов семантическая информация будет храниться в контейнере, основанном на RDF и онтологии.

Из платных программных продуктов выделяется СУБД Oracle, имеющая хорошо проработанные механизмы физической и начало семантической интеграции.

Корпорация Oracle [5,6] предлагает спектр решений, обеспечивающих возможность связи, которые позволяют компании прозрачно интегрировать различные системы и обеспечивать консолидированное представление компании в целом. Этими решениями являются Generic Connectivity и Oracle Transparent Gateways.

Решение Generic Connectivity является опцией базы данных, использующей драйвер ODBC или OLEDB для прозрачного обращения к любой системе, разработанной другими производителями и соответствующей техническим требованиям ODBC или OLEDB. Она решает потребности доступа к данным во многих хранилищах данных, для которых у Oracle отсутствуют решения межсетевого шлюза (gateway solutions), например, к MySQL, Foxpro, Access, dBase и нереляционным целевым объектам типа Excel.

Oracle Transparent Gateways обращается к системам других производителей, используя их "родные" интерфейсы. Созданы Transparent Gateways ко многим источникам, среди которых можно назвать Sybase, DB2, Informix, Microsoft SQL Server, Ingres и Teradata.

Прозрачность местоположения может быть также распространена на объекты, постоянно находящиеся в не-Oracle системах. Пользователи могут создать синонимы для объектов в хранилищах данных других производителей и обращаться к ним, не имея

необходимости определить их физическое местоположение. Такая прозрачность устраняет необходимость прибегать к услугам разработчиков приложений для такой настройки приложения, чтобы оно могло обратиться к данным из различных систем других производителей, тем самым, уменьшая усилия по разработке и увеличивая мобильность приложений. Приложения могут быть построены поверх согласованного интерфейса Oracle и для систем Oracle и систем других производителей.

Для гладкой способности к взаимодействию между несопоставимыми системами требуются преобразования SQL, преобразования словаря данных и преобразования типа данных, даже если не-Oracle системы базируются на стандартах SQL. И Generic Connectivity, и Oracle Transparent Gateways имеют возможность преобразовать диалект одной системы к диалекту другой.

Рассмотрим SQL Server Integration Services (SSIS). SSIS может извлекать (а также выгружать) данные из различных источников, включая OLE DB, управляемые источники (ADO.NET), ODBC, плоские файлы, Excel, и XML, с помощью специального набора компонент, которые называются адаптерами (adapters). SSIS также содержит средства Text Mining. Text Mining, также известная как систематизация текстов, представляет собой определение связей между видами бизнеса и текстовыми данными (словами и фразами). Этот метод позволяет находить в текстовых данных ключевые понятия и, опираясь на них, автоматически идентифицировать весь текст как представляющий определенный интерес.

Основным недостатком этих продуктов является их высокая цена и недоступность на всех платформах.

Таким образом, анализ литературы в области исследований интеграции и доступа к распределенной информации в гетерогенных базах данных позволяет прийти к выводу, что многие СУБД поддерживают механизмы физической интеграции и доступа к данным, но большое количество вопросов семантической интеграции все еще остаются открытыми. Стандарты ODBC и OLE DB частично обеспечивают унифицированный доступ к различным СУБД (если не учитывать различия в SQL-диалектах, используемых в разных СУБД). Иллюзию единого физического и логического представления помогает достичь использование технологии XML. Основной средой разрешения проблем семантической гетерогенности чаще всего являются веб-системы на основе онтологий. Тем не менее эти проблемы являются весьма обширными. Исследования и разработки программного обеспечения в данной области востребованы и являются актуальными на протяжении уже более четверти века.

Авторы выражают глубокую признательность У.А. Тукаеву за помощь в выполнении данной работы.

Список литературы

- [1] *Elmagarmid A., Rusinkiewicz M., Sheth A.*, Management of Heterogeneous and Autonomous Database Systems. - San Francisco: Morgan Kaufmann Publishers, 1998.
- [2] *Hathi K.*, Введение в SQL Server 2005 Integration Services. // Microsoft - 2005.

- [3] *Kent W.*, Solving Domain Mismatch and Schema Mismatch Problems with an Object-Oriented Database Programming Language. // Proceedings of the International Conference on Very Large Data Bases, 1991.
- [4] *Litwin W.*, From database systems to multidatabase systems: Why and how. // British National Conference on Databases, Cambridge Press, 1988.
- [5] *Makushkin B.*, Oracle 10G Development: доступ из Oracle к гетерогенным базам данных. // Alba Spectrum Technologies, Москва, 2007.
- [6] Oracle Database 10g: интеграция и способность информации к соединению. // Материалы конференции Oracle OpenWorld - San-Francisco, 2004.
- [7] *Yeung A.K.W., Hall G.B.*, Spatial database systems: Design, implementation and project management. - Dordrecht: Springer, 2007.
- [8] *Антупин К.В., А.В., Гринев М.Н., и др.*, Оперативная интеграция данных на основе XML: системная архитектура BizQuery. // Труды Института системного программирования РАН, Т.5, 2004. - С. 157-174.
- [9] *Белодед Б.В., Глоба Л.С., Терновой М.Ю., Штогрин Е.С.*, Единый интерфейс доступа к гетерогенным базам данных. // Материалы 21-й международной крымской конференции «СВЧ-техника и телекоммуникационные технологии» - Севастополь, 2011. - С. 47-50.
- [10] *Блажко А.А., Ибаа С., Альсаффади Т.Д.*, Санкционированное управление процессом согласования в гетерогенных тиражируемых базах данных. // Електротехнічні та комп'ютерні системи. Науковий журнал. №72, 2009. - С. 195-197.
- [11] *Вовченко А.Е., Калиниченко Л.А. и др.*, Краткий обзор технологий интеграции информационных ресурсов (на примере интеграции баз данных). // Материалы спецсеминара для студентов факультетов ВМК и Мехмата МГУ. Москва, 2009.
- [12] *Ганопольский Р.М., Кепещук Д.Б.*, Представление знаний в гетерогенных распределенных базах данных на примере ИНГРИС ТюмГУ. // Вестник кибернетики ТГУ. №5, 2006. - С. 70-76.
- [13] *Дорофеев А.Н.*, Моделирование и анализ конфликтов транзакций в информационных системах на основе гетерогенных баз данных. Дисс. на сиск. ... канд. техн. наук, 2003. - 155 с.
- [14] *Жучков А.В. и др.*, Интеграция и поиск информации в гетерогенных динамических информационных массивах с помощью онтологий. // Труды 6-й Всеросс. науч. конф. "Электронные библиотеки: перспективные методы и технологии, электронные коллекции" - RCDL 2004. Пушино, 2004. - С. 82-85.
- [15] *Кашников А., Лядова Л.*, Интеграция гетерогенных источников данных на основе рекурсивной декомпозиции. // International Journal "Information Technologies & Knowledge". Vol.5, №3, 2011. - С. 274-284.

- [16] *Коголовский М.Р.*, Методы интеграции данных в информационных системах. // Институт проблем рынка РАН. - Москва, 2010.
- [17] *Кудашев Е.Б., Марков С.Ю., Попов М.А.*, Использование гетерогенной пространственной информации при решении задач устойчивого развития территорий. // Журнал электронные библиотеки. - Т.14, Вып. 3., 2011.
- [18] *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.
- [19] *Новицкий А.В.*, Обзор некоторых направлений интеграции гетерогенных ресурсов в электронных библиотеках. // Труды 11-й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» - RCDL'2009. - Петрозаводск, 2009. - С. 350-356.
- [20] *Пыхалов А.В., Шаройко О.В.*, Методика работы с БД в гетерогенном окружении // Материалы XIV Всероссийской научно-методической конференции "Телематика 2007". - Санкт-Петербург, 2007.
- [21] *Черняк Л.*, Интеграция данных: синтаксис и семантика. // Открытые системы. 2009, №10.
- [22] *Шокин Ю.И., Федотов А.М., Жижимов О.Л., Мазов Н.А.*, Интегрированная распределенная система (ИРИС) Сибирского отделения РАН. // Выездное заседание научно-координационного совета по целевой программе «Информационно-телекоммуникационные ресурсы СО РАН». - Иркутск: ИДСТУ СО РАН, 2003.

V.S. Ramazanova, N.P. Azanov, The problems of integration and access to heterogeneous distributed databases. The Bulletin of KazNU, ser. math., mech., inf. 2011, №4(71), 72 – 81.

This paper investigates the problem of integration and access to heterogeneous distributed databases, based on software products from different manufacturers.

В.С. Рамазанова, Н.П. Азанов, Ықпалдаудың және гетерогендік бөлініс дерекқорларына жетімділік келеңсіз жағдайлары. ҚазҰУ хабаршысы, мат., мех., инф. сериясы 2011, №4(71), 72 – 81.

Мақалада ықпалдаудың және семантикалық гетерогендігі деректері бойынша қарастырылады.