

зультате МГК-анализа ее удалось понизить до 27, при этом текстурные признаки были брошены как неинформативные, а наибольшие нагрузки на первые главные компоненты обеспечили признаки кристаллов, связанные с их размерами. С помощью простого порогового критерия по значениям первой главной компоненты, удалось достоверно идентифицировать 95% пустых капель и 65% кристаллов. Принятие однозначного решения по нераспознанным каплям было затруднено специфическими помехами на изображениях, алгоритм подавления которых предстоит выработать.

Во второй задаче текстурные характеристики центральной части капель описывались с помощью 10 признаков Харалика. Классификатор, построенный на основе деревьев решений, позволил с точностью 85% отличать изображения капель с одиночными кристаллами от капель с фрагментами денатурированных белков. Относительно низкий процент корректно распознанных изображений в этом случае объясняется высоким разнообразием свойств образований в каплях.

#### **Библиографический список**

1. Wilson J. Towards the automated evaluation of crystallization trials // Acta Crystallographica Section D Biological Crystallography, 58, 2002, pp. 1907-1914.
2. Berry I.M., Dym O., Esnouf R.M., Harlos K., Meged R., Perrakis A., Sussman J.L., Walter T.S., Wilson J., Messerschmidte A. SPINE high-throughput crystallization, crystal imaging and recognition techniques: current state, performance analysis, new technologies and future aspects // Acta Crystallographica Section D Biological Crystallography, 62, 2006, pp. 1137-1149.

#### **Интеграции данных в интегрированной автоматизированной информационной системы вуза**

*О.А. Бубарева, Ф.А. Попов*  
*БТИ (филиал) АлтГТУ, г. Бийск*

Проблема интеграции данных особенно остро стоит на этапах проектирования и сопровождения комплексных ИС [1]. Решить проблему на этапе проектирования системы можно с помощью правильно построенной модели данных и формального описания предметной области. При этом необходимо решать задачи, связанные с обеспечением качественной информации, с поддержкой модификации моделей дан-

ных и информационных систем и с поддержкой часто меняющихся бизнес-процессов вуза.

В рамках данной статьи предлагается рассмотреть решение первой задачи на основе построения семантической модели ИАИС, включающей глоссарий, таксономию, онтологическое описание понятий и отношений между ними, расширяемый семантический базис и механизм изменения онтологического описания. В глоссарий входят термины предметных областей деятельности вуза и термины ИТ-области. Семантический базис представляет собой элементарную функциональность, которая может использоваться в различных подсистемах ИАИС, сюда входят управление доступом к данным, интерпретация понятий и отношений между ними, процедуры поддержки качества данных, а также механизм создания отчетов из различных систем на основе понятий. Онтологическое описание содержит онтологии предметных областей деятельности вуза, управления процессами и описание понятий ИТ-области, а также отношения между понятиями. Механизм изменения онтологических описаний реализуется с помощью инструмента создания понятий, отношений между понятиями, ограничений на атрибуты и создание, редактирование и удаление экземпляров понятий.

Интеграция данных обеспечивается единым хранилищем данных, куда поступает консолидированная информация из разных ИС, либо репликацией данных из одной системы в другую. Также данные должны проходить предварительную обработку, где оценивается их качество. Ввиду того, что репликация и хранилища предоставляют доступ к данным с задержкой по времени, можно использовать технологию интеграции данных по требованию - «на лету», т.е. интеграцию в реальном режиме времени, который подразумевает работу в режиме запрос-ответ, а ключевым в его реализации становится наведение связей между источниками и данными, которые можно запрашивать.

Для поддержки интеграции данных, обеспечения качества информации и для управления бизнес-процессами используется репозиторий метаданных, в котором хранятся описания всех понятий ИАИС, включая область управления процессами, а также описания физических и логических связей между источниками и системами.

Это позволяет в полной мере описать отношения между понятиями в гетерогенной информационной среде, где присутствуют СУБД с разной архитектурой. Одно понятие обычно имеет одну базовую таблицу или представление. Но полное описание понятия может включать несколько таблиц. Понятия отличаются друг от друга по имени; в рамках ИАИС имя понятия должно быть уникальным. Понятия, имея отноше-

ния проекции с источником данных, транслируют понятие в таблицы или представления, а атрибуты понятия в поля таблицы или представления.

Для поддержания интеграции данных используются описания контекстно-зависимых и контекстно-независимых связей между понятиями в репозитории метаданных. Он позволяет использовать в ИС понятия, созданные в другой ИС, поддерживать общую систему справочников для всех подсистем ИАИС, а так же общую систему отчетности ИАИС. При составлении отчетов пользователи используют определения понятий из репозитория, что позволяет ИАИС формировать произвольные, в том числе агрегированные отчеты из данных различных подсистем без привлечения программистов. Задача администратора репозитория метаданных корректно описать понятия и атрибуты, в том числе виртуальные, и связи между понятиями.

Рассмотренные методы интеграции данных были успешно применены при построении информационной системы поддержки учебно-организационной деятельности вуза. Это позволило не только объединить данные из разных ИС, но и повысить качество интегрированной информации, что обеспечило ее целостность и достоверность.

#### **Библиографический список**

1. Бубарева О.А., Попов Ф.А., Ануфриева Н.Ю. Решение проблемы интеграции данных при построении интегрированной автоматизированной информационной системы вуза // Международный журнал экспериментального образования. – 2011.– №5. – С. 95.

### **Модификация метода селекции рулетки для решения задачи Эйнштейна методами генетических алгоритмов**

*Д.С. Козлов*  
*АлтГУ, г. Барнаул*

Генетические алгоритмы имеют ряд недостатков, особенно заметных при решении задач большой размерности. Поиск решения задачи Эйнштейна в классическом варианте с 5 домами, 5 категориями и 15 условиями, в исполнении генетического алгоритма с ранговой или турнирной селекцией, почти всегда ведет к тупиковой ветви развития популяции и не дает правильный ответ на ключевой вопрос задачи.

В исследовании был применен непрерывный способ кодирования хромосом. Оценка жизнеспособности или целевая функция, была выбрана по количеству верных ответов на вопросы условий и может при-