

СЕКЦИЯ № 6

**Информационные технологии
в образовании и производстве**

Автоматизация сбора информации об электронных компонентах при проектировании микропроцессорных систем

В докладе рассматривается проблема сбора информации об электронных компонентах при проектировании электронных схем микропроцессорных систем, исследование и разработка методов автоматизации сбора и анализа данной информации, принципы организации программной системы на основе модели с применением нечеткого классификатора.

В настоящее время при разработке аппаратного обеспечения микропроцессорных систем, таких как микропроцессорные системы сбора и анализа данных, системы и сети передачи информации и управления, может применяться множество сложных электронных компонент различных производителей. Такие компоненты, представленные в основном элементами, выполненными в виде микросхем и микромодулей, имеют большой набор характеристик с перекрывающимися диапазонами и различными возможностями, например, производительностью, поддержкой тех или иных стандартов или технологий. Кроме этого, при проектировании микропроцессорных систем, для обеспечения производства конкурентоспособных электронных изделий, необходимо учитывать цену и доступность необходимого количества электронных компонент, а также наличие их аналогов.

Разработчики при проектировании архитектуры системы и подборе электронных компонент вынуждены анализировать значительный объем информации, полученной, в основном, с Интернет-сайтов производителей или поставщиков электронных компонент. При этом, перед разработчиками стоит задача оптимального выбора компонент для обеспечения соответствия проектируемой системы выбранному набору общих критериев, которыми может являться цена, производительность, набор выполняемых функций и т.п. Данный этап проектирования микропроцессорных систем требует от разработчиков больших затрат времени для анализа имеющегося рынка электронных компонент, учета влияния выбранных компонент на соответствие разрабатываемой системы заданным критериям, а также взаимосвязи между выбираемыми компонентами.

Таким образом, актуальной является задача автоматизация процесса анализа имеющейся информации о множестве электронных компонент для обеспечения принятия решения по выбору оптимальных с учетом заданных критериев.

Для проектирования электронных схем используются системы автоматизированного проектирования (САПР), такие как, PCAD, Altium Designer и другие. В таких САПР отсутствует возможность получения характеристик электронных компонент проектируемой электронной схемы из внешних источников, например из сети Интернет. Поэтому, при решении задачи автоматизации процесса анализа, существует также необходимость реализации возможности получения информации об используемых компонентах электронной схемы из САПР для анализа оптимальности используемых решений.

Множество электронных компонент может рассматриваться как сложная, слабо структурированная система с большим числом входов. Задачу анализа характеристик и классификации электронных компонент можно формулировать как решение задачи классификации при многокомпонентном критерии. При этом параметры компонент могут быть представлены характеристиками, часто имеющими нечисловую форму представления, например, список поддерживаемых интерфейсов связи. При этом могут существовать критерии, отражающие предпочтения разработчика, не имеющие четкого числового представления, например, критерий выбора производителя компонент.

Таким образом, выбор компонентов относится к решению задачи принятия решения в нечеткой среде, так как в начале имеется неполное или нечеткое представление о конечном наборе компонент.

Данная задача может быть реализована при помощи адаптивной системы анализа и классификации информации в нечеткой среде. Для разработки методов решения задачи и модели системы могут быть применены элементы теории нечетких множеств и нейросетевых алгоритмов.

Для исследования и оптимизации методов и алгоритмов решения задачи анализа и классификации электронных компонент в зависимости от заданных критериев разрабатывается модель в системе математического моделирования MATLAB с использованием пакета моделирования процессов Simulink. Модель использует методы и алгоритмы принятия решения на основе элементов теории нечетких множеств с использованием пакета FuzzyLogic. Критерии классификации формируются экспертом на основе нечетких отношений предпочтения.

Структура программного обеспечения системы принятия решения приведена на рисунке 1.

Модуль обработки информации обеспечивает структурный и синтаксический разбор (парсинг) Интернет-страниц для дальнейшей классификации характеристик на множества по заданным нечетким критериям.

Система использует модули анализа параметров и классификатора, реализованные в виде отдельных динамически подключаемых библиотек, для обеспечения возможности подключения модулей реализующих различные алгоритмы с целью сравнительного анализа эффективности разрабатываемых методов реализуемых в модели.

Модуль интерфейса пользователя реализован в виде Web-интерфейса. Это позволяет организовать одновременный сетевой доступ к системе для нескольких пользователей, что необходимо при использовании завершённой системы на предприятиях.

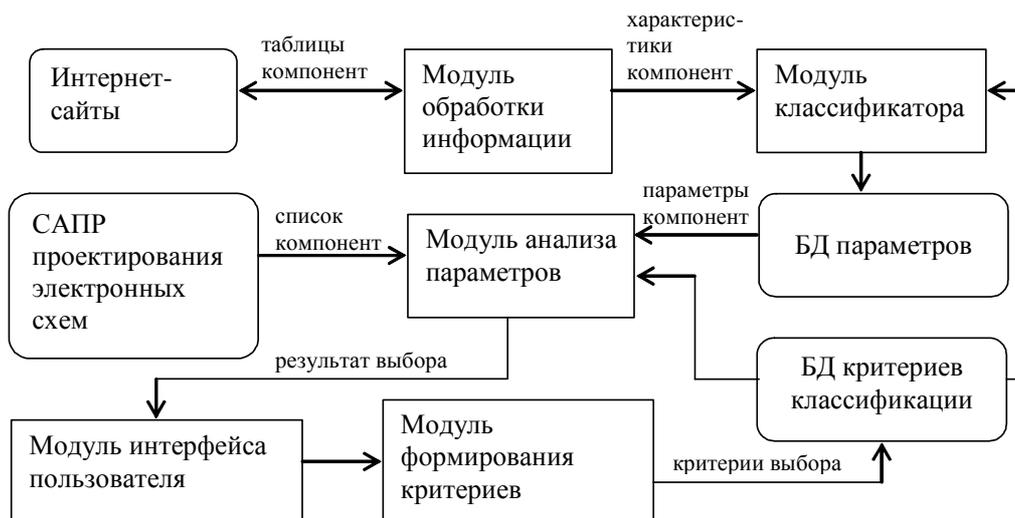


Рис. 1. Структурная схема программного обеспечения автоматизированной системы сбора и анализа информации об электронных компонентах

Таким образом, разрабатываемая система позволяет автоматизировать процесс анализа имеющейся информации о множестве электронных компонент при разработке архитектуры микропроцессорной системы, а также отслеживать актуальную информацию об электронных компонентах на этапе проектирования электронных схем в САПР.

Задачей развития системы является исследование, разработка и оптимизация алгоритмов для модели анализа и классификации на основе элементов теории нечетких множеств и нейросетевых алгоритмов, а также анализ эффективности примененных алгоритмов классификации по критерию соответствия заданным параметрам классификации.

И.К. Будникова,
М.В. Симонов
Казанский государственный энергетический университет
г. Казань, ул. Красносельская, д. 51
E-mail: ikbudnikova@yandex.ru

Информационные технологии для обработки структурированной информации

Для обработки структурированной информации применяется один из видов многомерного статистического анализа – кластерный анализ. Реализация кластерного анализа в системе Statistica [1] позволяет существенно сократить время обработки массивов биллинговой информации, автоматизировав этот процесс.

С помощью этого типа анализа решаются такие важные задачи, как классификация и группировка объектов по множеству их атрибутов. Актуальность задачи заключается в том, что результаты исследования могут быть использованы при разработке новых тарифов и предложений в мобильной связи, а также статистического анализа и прогнозирования качества работы Call-центров.

Методы кластерного анализа можно разделить на две группы: иерархические и неиерархические. Каждая из групп включает множество подходов и алгоритмов.

Иерархические методы предполагают последовательное объединение объектов в кластеры по степени их близости друг к другу или, напротив, последовательное разбиение совокупности объектов на все более мелкие кластеры. В этом случае кластерное решение представляет собой иерархическую структуру вложенных друг в друга кластеров. Иерархические методы кластерного анализа используются при небольших объемах наборов данных. Преимуществом иерархических методов кластеризации является их наглядность. Иерархические алгоритмы связаны с построением дендрограмм. Дендрограмма представляет собой вложенную группировку объектов, которая изменяется на различных уровнях иерархии.

Неиерархические методы позволяют находить и идентифицировать "сгущения" объектов в пространстве переменных.

Решением задачи кластерного анализа является разбиение, удовлетворяющее некоторому критерию оптимальности. Этот критерий может представлять собой некоторый функционал, выражающий уровни желательности различных разбиений и группировок. Этот функционал часто называют целевой функцией. Например, в качестве целевой функции может быть взята внутригрупповая сумма квадратов отклонений.

Меру близости (сходства) объектов удобно представить как обратную величину от расстояния между объектами. В многочисленных изданиях посвященных кластерному анализу описано около 100 различных алгоритмов вычисления расстояния между объектами. Кроме термина "расстояние" в литературе часто встречается и другой термин - "метрика", который подразумевает метод вычисления того или иного конкретного расстояния. Наиболее распространенный способ в случае количественных признаков является так называемое "евклидово расстояние" или "евклидова метрика".

В данной работе проведено исследование, в котором рассматривается применение разного сочетания метрик и алгоритмов кластерного анализа к детализации телефонных переговоров.

Биллинг - это зафиксированные параметры событий, произошедших с объектом исследования за определенный период времени. Такими объектами могут быть:

- телефон, где событием является входящее или исходящее соединение;
- сотовые телефоны, где событием являются SMS сообщения;
- интернет соединения, где событием является соединение с поставщиком услуги;
- другие объекты.

Для телефонного номера зафиксированными параметрами события являются дата, время и длительность соединения, номер телефона, с которым произошло соединение и т.п. Детализации телефонных переговоров в зависимости от компании имеют разные структуры, в которые часто вводят сложные правила и зависимости, из-за чего невозможно создать инструментарий, позволяющий автоматически обрабатывать все возможные типы структур.

На основе алгоритма кластеризации в работе выделены следующие основные этапы: изменение исходных данных, принятие решений, анализ полученных результатов [2]. В общем случае, все эти этапы взаимосвязаны, и решения принятые на каждом из них взаимно обуславливают друг друга.

Проведена кластеризация исследуемых объектов с применением различных комбинаций метрик в программном комплексе Statistica. В зависимости от режима анализа информация выдается на экран двумя способами отображения: графом и диаграммой длительности телефонных переговоров.

Выбирая между иерархическими и неиерархическими методами, следует обратить внимание на следующие моменты. Неиерархические методы обнаруживают более высокую устойчивость по отношению к выбросам, неверному выбору метрики, включению незначимых переменных в базу для кластеризации и пр. При этом исследователь должен заранее фиксировать результирующее количество кластеров, правило остановки и, если на то есть основания, начальный центр кластера. Последнее условие существенно отражается на эффективности работы алгоритма. Если нет оснований искусственно задать это условие, рекомендуется использовать иерархические методы. Отметим существенное свойство для обеих групп алгоритмов: не всегда правильным решением является кластеризация всех наблюдений. Возможно, наиболее оптимально будет сначала очистить выборку от выбросов, а затем продолжить анализ. Можно также не задавать очень высокий критерий остановки (можно делать остановку, к примеру, когда кластеризовано более 90% наблюдений).

Среди всех подсистем современных центров обслуживания вызовов (Call-центры) отдельную важность имеют те, которые реализуют функциональность сбора и обработки статистической информации, предоставления отчетности о функционировании комплекса оборудования и осуществляют прогнозирование поведения системы в целом в тех или иных ситуациях.

Известно, что планирование мощностей центра является обязательным этапом при его создании и требует специального программного обеспечения. В это время закладываются его возможности на долгосрочную перспективу, однако подобные механизмы могут использоваться не только в стратегических целях, но и в более узких временных масштабах. Сложность организации систем прогнозирования поведения напрямую зависит от логики обслуживания вызовов центром и поступающих на него потоков запросов.

Решение этих задач дает менеджерам центра средства для эффективного управления и позволяет организовать работу системы так, чтобы удовлетворить большинство запросов клиентов.

Кластерный анализ позволяет существенно снизить трудозатраты в процессе обработки значительных массивов информации и решить задачи классификации и выделения похожих групп объектов в соответствии с требованиями качества.

Литература

1. Боровиков В.П. STATISTICA. Искусство анализа данных на компьютере. Для профессионалов. – СПб.: Питер, 2003.
2. Халафян А.А. STATISTICA 6. Статистический анализ данных – М.: Бином-пресс, 2007.

В.С. Дороганов,
К.И. Ким
Кузбасский государственный технический университет имени Т.Ф. Горбачёва
650055 г. Кемерово, ул. Дзержинского, д. 9
E-mail: doroganovV@mail.ru

**Интеллектуальная геоинформационная система
для прогнозирования времени прибытия общественного транспорта
и поиска оптимального маршрута движения в улично-дорожной сети города**

Прогнозирование прибытия транспортных средств в условиях городской среды не является тривиальной задачей. Необходимость данного алгоритма появилась в проекте системы, дающий возможность пользователям посредством мобильного телефона отслеживать местоположение автобусов (общественного транспорта) в реальном времени, строить маршруты передвижения и возможные пересадки, а так же позволяет прогнозировать время прибытия на нужную остановку. Данный проект позволит уйти от традиционного статического расписания общественного транспорта, и даст возможность пользователю рационально планировать время на дорогу.

Очевидно, что основываться только на статическое расписание автобуса и его текущую скорость не является оптимальным. В условиях современного города, где на тысячу человек приходится более 200 автомобилей, проблема транспортных «пробок» не является ни для кого секретом. Мало кто не видел километровые пробки в час пик, или же не проезжал 100 метров за 30 минут. Влияние пробок на «пользователей» очевидны:

- водителям приходится больше времени выделять на «комьют», то есть дорогу от дома до работы;
- пробки сказываются на общественном транспорте, даже малейший затор на дороге создаст отставание от графика.

Чем больше отставание от графика, тем менее актуальным становится расписание автобусов. Как известно от подобных изменений в расписании страдают пассажиры. Автобус может прийти на остановку как раньше, так и позже, что редко является радостной новостью для пассажиров. В час-пик и вовсе ожидание требуемого автобуса может затянуться на 20-30 минут. Отсутствие необходимой информации о времени прибытия автобуса частенько приводило к так называемому закону «подлости»: «отходишь от остановки – приезжает автобус», «садишься на более долгий или дорогой автобус – подходит твой».

Основной задачей было поставлено решение проблемы прогнозирования прибытия общественного транспорта, определение оптимального маршрута для пассажиров общественного транспорта, отображение ситуации на дорогах в плане пробок. Согласно приказу Минтранса [1] все транспортные средства отвечающие за перевозку пассажиров должны быть оснащены навигационным оборудованием (датчиками Глонасс/Gps). Данный приказ был издан для снижения аварийности и экономии ГСМ. Наличие уже установленных датчиков позволило начать разработку системы с охватом сразу всех транспортных средств с возможностью накопления статистических данных о положении дел в городе.

Представляемая система, создающаяся для мобильных приложений, позволит в режиме реального времени отслеживать местоположение автобусов, рисовать маршруты, показывать возможные места пересадок. Одним из главных достоинств системы станет гибрид динамического расписания и возможности расчёта пересадок в удаленных остановках. Это позволит более быстро добираться до нужной точки.

Система строится по модульному принципу, для облегчения развертывания по регионам, а так же для повышения отказоустойчивости. Немаловажным аспектом является то, что большая часть решения строится на облачных технологиях (Microsoft Windows Azure), обусловленных как экономическими выгодами, так и снижением трудозатрат на разработку и масштабирование системы. Данная система является интеллектуальной и адаптивной, способной подстраиваться под ситуацию в городе за короткое время.

В основе системы лежит связка базы знаний статистической информации о каждом сегменте города с множеством дополнительных критериев, срезов и база данных с текущей информацией. Показания средней скорости по временным срезам и локация города накапливается в базе знаний, с учётом поправочных коэффициентов для оперативного реагирования на ситуацию в городе. По запросу пользователя специальный модуль системы, основываясь на знаниях о скорости на дорогах и данных полученных с датчиков Глонасс/GPS, рассчитает время прибытия нужного автобуса на заданные остановки и определяет оптимальный вариант по различным критериям (прим.: «раньше сесть в точке А» или «раньше прибыть в точку Б»). Из-за большого количества клиентских платформ было принято решение о создании закрытого API для кросс-платформенного клиент-серверного взаимодействия. Так же нахождение в облаке увеличивает доступность, и являющейся единой точкой входа для всех видов клиентов.

Данная система будет весьма полезной людям, использующим общественный транспорт для передвижения по городу. Система будет способствовать популяризации общественного транспорта, позволит не опаздывать на нужный автобус, незначительно снизить нагрузку на дорожную сеть города, а так же искать оптимальные пути пассажирам.

Так же не маловажным свойством системы является накопление знаний, которые применимы для решений задач логистики городских компаний. Бесплатность приложения позволит всем желающим получить к нему доступ независимо от платформы.

Литература

1. Приказ Минтранса РФ от 9 марта 2010 года №55 «Об утверждении перечня видов автомобильных транспортных средств, используемых для перевозки пассажиров и опасных грузов, подлежащих оснащению аппаратурой спутниковой навигации ГЛОНАСС или ГЛОНАСС/GPS»

Анализ обобщенных взаимосвязей методов автоматизированной сборки

Существует достаточно классификаций методов автоматизированной сборки, однако, к настоящему времени нет единого подхода, позволяющего объединить их. Выделим взаимосвязи методов автоматизированной сборки, проанализировав классификационные группы с позиции обобщенной модели функционирования сборочной системы [1], и представим их графически с помощью обобщенной схемы взаимосвязей методов автоматизированной сборки (рис. 1).

Первая группа основана на разделении устройств сборки по степени функциональной универсальности и включает три направления. Первое направление – промышленные роботы и робототехнические устройства – представляет собой универсальные средства автоматизации, позволяющие в наибольшем объеме воспроизводить функции оператора-сборщика. Второе направление – специализированные средства автоматизации, воспроизводящие отдельные функции оператора. Третье направление – взаимодействующие универсальные средства и специализированные средства автоматизации, сочетающие преимущество универсальности первых и качество отработки отдельных операций вторых. Это направление, как и первое, представляет собой область роботизации сборочных операций. Устройства рассмотренной классификационной группы могут быть представлены с помощью элементов обобщенной модели функционирования [1]. Для наглядного представления взаимосвязей соответствующих направлений используется принцип функциональных преобразований над объектами сборки (присоединяемой, базовой деталями или сборочным узлом) при их переходе с одного этапа сборочной операции на другой.

К следующей классификационной группе относятся устройства, отличающиеся характером функциональных связей. Связи элементов в обобщенной модели функционирования сборочной системы делят на две группы: информационные и энергетические. Информационные связи обеспечивают представление информации в виде сигналов, необходимых для формирования программных траекторий или управляющих воздействий. Информационные связи определяют характер поиска согласованного положения сопрягаемых поверхностей, и позволяют выделить следующие типы сборочных устройств: первые, для которых характерно прохождение информационного канала непосредственно через измерительный элемент, конструктивно представляющий собой датчик, и вторые, информационный канал которых датчик не включает.

В первом случае имеет место прямая, во втором – косвенная информационная связь. Энергетические связи осуществляют передачу, преобразование, накопление энергии требуемого для функционирования элементов сборочного устройства. Для реализации энергетических связей может потребоваться введение не относящихся к структуре автоматизированной сборочной системы специальных устройств, (энергетических модулей) служащих для передачи, преобразования, накопления электромагнитной и различных видов потенциальной энергии (энергии сжатого газа, жидкости и т.п.) – прямая энергетическая связь. Однако реализация энергетических связей может осуществляться без применения подобных устройств, если осуществляется передача кинетической энергии – косвенная энергетическая связь.

В зависимости от характера энергетических и информационных связей выделяют следующие группы устройств: активные корректирующие устройства, осуществляющие управление движением соединяемых деталей только при наличии внешних прямых энергетических и информационных связей; пассивные корректирующие устройства, осуществляющие управление движением соединяемых деталей при наличии косвенных энергетических и информационных связей.

По результатам обзора существующих классификационных групп устройств автоматизированной сборки, в процессе анализа функциональных связей структурных компонент соответ-

ствующих групп, на основании разработки обобщенной модели функционирования сборочной системы была получена обобщенная схема взаимосвязей методов автоматизированной сборки, наглядно отображающая состав и взаимосвязи структурных компонент, присущих различным сборочным устройствам, операционные переходы в процессе сборочной операции, а так же представлены функциональные переключатели (управляемые ключи), условно отображающие возможности перехода от одного метода автоматизированной сборки к другому.

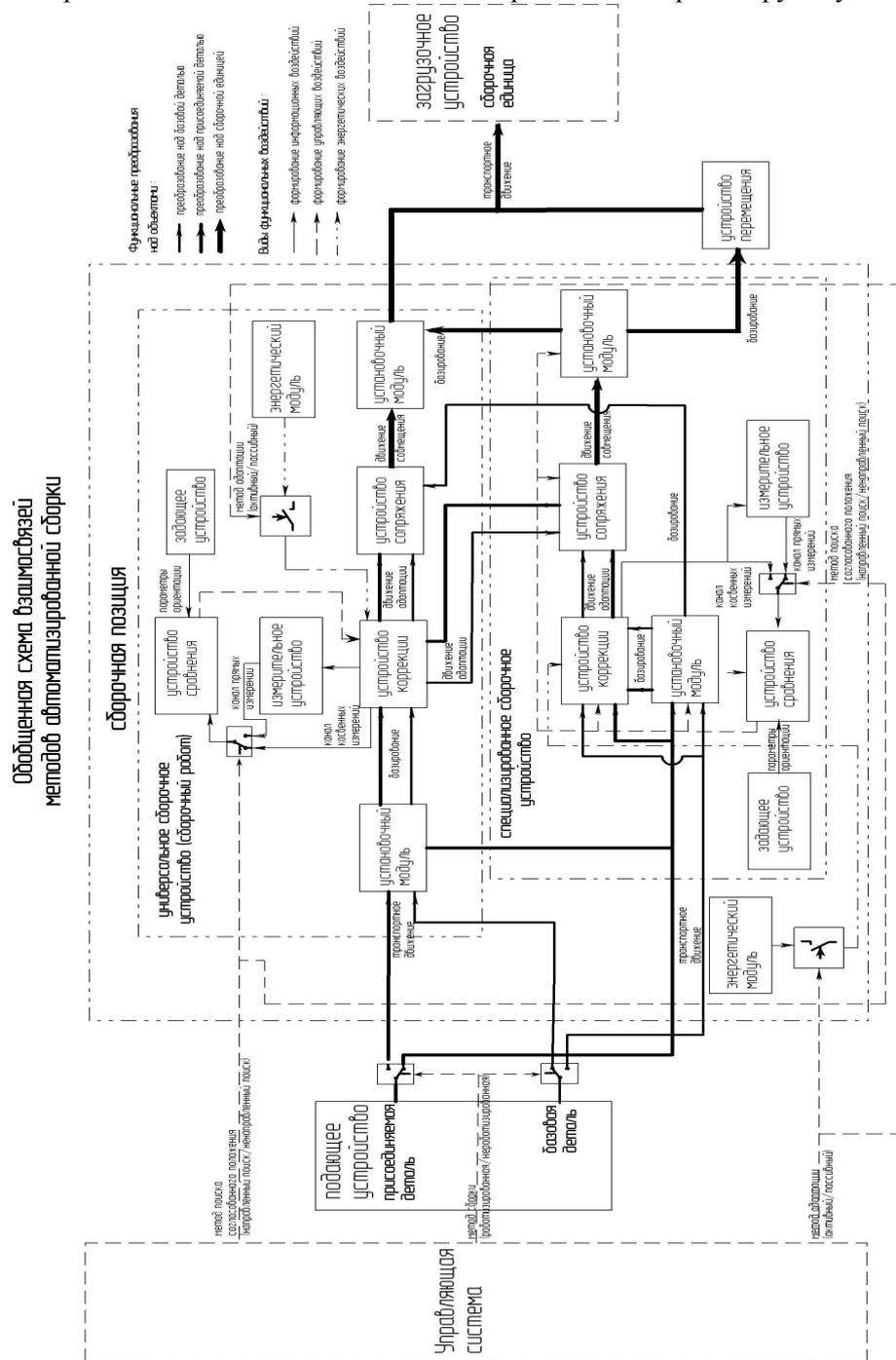


Рис. 1. Обобщенная схема взаимосвязей методов автоматизированной сборки

Литература

1. Кузнецова, С.В. Обобщенная модель функционирования автоматизированной сборочной системы [Текст] / С.В. Кузнецова, Е.П. Тетерин // Наука и современность – 2011: сборник материалов XIII Международной научно-практической конференции: в 2-х частях. Ч. 2 / Под общ. ред. С.С. Чернова. – Новосибирск: Издательство НГТУ, 2011 – С. 206-210.

Разработка базы данных по характеристикам надежности механических элементов

Принятые в настоящее время методики расчетной оценки надежности радиоэлектронной аппаратуры (РЭА), приведенные в ОСТ 4Г 0.012.242 [1], основаны на использовании данных справочника «Надежность ЭРИ» [2], в котором содержатся математические модели интенсивностей отказов электрорадиоизделий (ЭРИ) и численные значения их коэффициентов. Однако в состав РЭА, наряду с ЭРИ входят и механические элементы (МЭ), отказы которых так же приводят к отказам РЭА. Расчет интенсивностей отказов МЭ может быть проведен по данным РМ 25 446 [3], но приведенная в нем номенклатура МЭ достаточно скудна, а модели настолько примитивны, что не учитывают ни физико-химических свойств материалов, ни геометрии МЭ, ни условий их применения.

В этом плане большой интерес представляют модели, приведенные в американском стандарте NSWC-06/LE10 [4], разработанного специалистами Кардерокской дивизии ВМФ США. Следует отметить, что модели, приведенные в нем достаточно сложны, что обуславливает высокую трудоемкость «ручных» расчетов. Поэтому, наряду со стандартом этим подразделением выпускается программа MechReal. Однако, применение этой программы, равно как и модулей «Reliability prediction» систем RAM Commander, Relex, BlockSim и др., осложняется тем, что в их базах данных (БД) отсутствуют физико-химические характеристики материалов, а содержатся только «усредненные» значения. Это приводит к существенным погрешностям при оценке интенсивностей отказов МЭ и ограничивает возможность обоснования эффективности мероприятий, направленных на повышение надежности [5].

Поэтому, при создании новой (4.12) версии программного комплекса (ПК) АСОНИКА-К в его базовую версию (систему АСОНИКА-К-СЧ) было решено ввести возможность расчета надежности РЭА с учетом МЭ. Т.к. основу таких расчетов составляют исходные данные, то в настоящее время проводятся исследования моделей интенсивностей отказов МЭ стандарта NSWC-06/LE10 [4] с целью классификации их параметров и коэффициентов по следующим признакам: переменные проектирования (геометрические размеры, режимы применения и т.д.); физико-химические характеристики материалов (твердость, теплопроводность, и др.); эмпирические коэффициенты; физические константы (абсолютная температура, постоянная Больцмана и др.).

В результате этих исследований будет создана концептуальная модель БД по характеристикам надежности МЭ и физико-химическим свойствам материалов, интегрированная в справочную часть БД системы АСОНИКА-К-СЧ [6]. Использование такой БД, в отличие от БД зарубежных программных средств, позволит проводить расчеты надежности МЭ не по «усредненным», а по значениям, нормированным в ГОСТах, ТУ и др.

Литература

1. ОСТ 4Г 0.012.242-84. Аппаратура радиоэлектронная. Методика расчета показателей надежности.
2. Справочник «Надежность ЭРИ». - М.: МО РФ, 2006.
3. РМ 25 446-87. Изделия приборостроения. Методика расчета показателей безотказности.
4. NSWC-06/LE10. Handbook of Reliability prediction Procedures for Mechanical Equipment.
5. Zhadnov V. Methods and means of the estimation of indicators of reliability of mechanical and electromechanical elements of devices and systems. / Reliability: Theory & Applications, Vol. 2, No 4, 2011. - p. 94-102.
6. Жаднов В.В., Сарафанов А.В. Управление качеством при проектировании теплонагруженных радиоэлектронных средств: Учебное пособие. - М.: СОЛОН-ПРЕСС, 2012. - 464 с.

А.А. Орлов,
Л.В. Антонов

Муромский институт Владимирского государственного университета
602264 г. Муром, Владимирской обл., ул. Орловская, д. 23
E-mail: alexeyalexorlov@gmail.com, levantonov@yandex.ru

Обзор и анализ существующих информационных систем управления животноводческим хозяйством

В настоящее время животноводческое хозяйство в Российской Федерации находится в упадочном состоянии. Информационно-техническая база значительно устарела, так как на большинстве комплексов последний виток модернизации был завершен в конце 80-ых годов. Современные предприятия отрасли хоть и используют информационные технологии, также являются убыточными, так как процесс производства не оптимизирован и затруднен большим количеством неверных управленческих решений. Ошибки возникают из-за того, что современные информационные системы, которые внедрены в процесс производства животноводческой продукции, не могут быть интегрированы друг с другом, а следовательно персонал предприятия должен выполнять большой объем ручной работы по перенесению данных из одной системы в другую. При использовании промежуточного информационного ресурса, в виде рабочего персонала, данные искажаются и решения принимаются неверно. Но даже при условии верного перенесения информации перед менеджером встает трудная задача ее анализу, так как большой объем малосвязанных данных со всего предприятия нужно оценить в короткий промежуток времени. Ограниченность времени связана со спецификой производства, где животные являются основным источником производства, а продукция - скоропортящаяся. В связи с этим растет интерес к системам, призванным увеличить производительность предприятия, эффективность и оперативность принимаемых решений, повысить качество продукции, оперативно отслеживать состояние и поведение животных, что позволит избежать внепланового снижения производственных мощностей и падежа скота.

Целью работы является анализ информационных систем присутствующих в настоящий момент на рынке, с целью выявления их достоинств и недостатков, влияющих на качество выпускаемой продукции и здоровье животных в условиях специфики управления животноводческим предприятием.

С 2000 года объемы поставок оборудования и программного обеспечения для животноводческих комплексов из-за рубежа постоянно увеличивались. Ведущими игроками на российском рынке являются шведская «ДеЛаваль», немецкая «Вестфалия-Сёрдж», датская S. A. Christensen & Co: на их долю приходится 87,5 % всех поставок в стоимостном выражении. Хорошие позиции занимают фирмы «Лада-Сервис» (Латвия) и «Бау-Матик» (США/Бельгия). Доля поставок остальных производителей 1-2 %. Особую группу поставщиков составляют производители из бывших союзных республик.

Таблица 1

Сравнительный анализ существующих и разрабатываемой систем

	ALPRO	Dairy Plan 5	Cattle Code	«Ферммаш»	Автоматизированная система управления стадом (ВИЭСХ и БИМ)	Разрабатываемая система
Измерение индивидуальных надоев	+	+	+	+	+	+
Регистрация поедаемости корма	+	+	+	+	-	+

Определение активности животного	-	+	-	-	-	+
Ведение календаря ветеринарных мероприятий	+	-	-	-	+	+
Контроль биологического состояния животных	+	-	+	+	-	+
Селекция стада	-	-	-	+	-	+
Визуализация мониторинговых данных	-	-	-	-	-	+
Экспертная поддержка принятия решения	-	-	-	-	-	+

Как видно из таблицы 1 все существующие на рынке системы являются узкопрофильными, направленными на решение лишь одной конкретной задачи и не предоставляют пользователю полного спектра решений. Использование нескольких систем в совокупности бессмысленно, так как они не совместимы друг с другом. Перечисленные программные продукты не удовлетворяют запросам потребителей и являются неэффективными при сборе и анализе большого объема разноплановой информации со всего производственного комплекса.

В связи с этим возникает необходимость разработки универсального программного продукта, позволяющего проводить оперативный анализ данных на предприятии и выявлять проблемные ситуации на ранних стадиях их появления, а также прогнозировать их. Функции интеллектуального анализа информации должны предлагаться реализовать в специальном экспертном модуле, который позволит оперативно обрабатывать данные, а также исключить управленческие ошибки, что позволит увеличить объем и качество производства.

Экспертный анализ возможно реализовать на основе системы включающей обширную базу знаний и хранящей большое количество оценок, мнений и заключений экспертов области о когда-либо возникавших проблемных ситуациях. При сопоставлении текущей информации с контентом базы знаний экспертный алгоритм определяет, существовала ли подобная ситуация в прошлом или нет. Если проблема была успешно решена в прошлом, система сообщает менеджеру о наличии проблемной ситуации и предлагает варианты для ее нейтрализации. Менеджер принимает вариант системы поддержки принятия решений или сообщает системе о том, что возникшая проблема принадлежит к другому классу и ранее не встречалась. Экспертная система фиксирует эту информацию и накапливает её с годами, обучаясь. Можно резюмировать, что чем больший поток информации стекается в экспертную систему из различных подразделений предприятия и чем дольше система на данном предприятии находится, тем достовернее будут предлагаемые ей альтернативы и эффективнее работа системы в целом.

Внедрение технологии экспертной системы на животноводческом предприятии позволит увеличить оперативность обработки данных, сократить производственные и экономические риски и поддерживать качество продукции на высшем уровне.

Информационная система автоматизированного управления финансовой стратегией рекламных кампаний в сети Интернет

Целью проекта «Информационная система автоматизированного управления финансовой стратегией рекламных кампаний в сети Интернет» является реализация задачи автоматизированного управления и формирования стратегии рекламной деятельности в сети Интернет.

В настоящее время объем рынка рекламы постоянно увеличивается и характеризуется появлением новых форм рекламы. Одной из наиболее перспективных и целенаправленных форм рекламы является реклама в сети Интернет, а именно контекстная реклама в поисковых системах и на популярных сайтах. Многие компании в сети предлагают услуги по размещению рекламы ориентированной на определённый круг лиц и область интересов пользователя. Аудитория популярных ресурсов российского сегмента Интернета давно превосходит цифру в 1 миллион уникальных посетителей в сутки. Среди рекламодателей существует плотная конкуренция. Стоимость размещения рекламной ссылки формируется по рыночным принципам и, зачастую, совершенно не рентабельна для заказчика. Возникает необходимость управления финансами по индивидуальной для каждого заказчика логике.

Цель разработать автоматизированную систему для управления рекламной кампанией в области ценовой политики в виде удобного для пользователя и простого для быстрого старта использования приложения. Продукт должен быть расположен на выделенном сервере с круглосуточным доступом в Интернет, клиенты должны подключаться к нему через web браузер.

Проект направлен на разработку системы автоматизированного управления кампанией контекстной рекламы по средствам планирования расходов и выполнения запрограммированной стратегии поведения пользователя сервиса по направлению ценовой политики. Выступая промежуточным звеном между поставщиком услуг контекстной рекламы (например Яндекс.Директ) и рекламодателем система позволяет настроить робота для управления ставкой за целевой переход посетителя по рекламируемой ссылке. Широкий спектр аналитики и статистики позволяет получать подробные отчёты об эффективности действий робота и прогнозировать расходы.

Основной задачей поставленной в проекте является разработка системы состоящей из двух компонентов: серверная часть для взаимодействия с API рекламной площадки и реализации заложенной пользователем бизнес логики; модуль взаимодействия с клиентом (заказчиком).

Первый модуль (ядро) представляет собой реализацию WSLD или RESTFull клиента для подключения к API исполнителя по защищённым соединениям (SSL) основанного на сертификатах x509 подписанных компанией оператором для агентства от имени которого работает клиент. Содержит инструменты для работы с Базой Данных и систему реализации пользовательской бизнес логики. Модуль функционирует по расписанию в виде псевдо демона. Для взаимодействия с клиентским модулем используется RESTfull интерфейс.

Второй модуль позволяет подключаясь к RESTfull ядра формировать источник для подключения пользователей. Обработывая запросы браузера, возвращает html страницы для отображения. Большая часть функционала реализована с использованием ajax.

Проектируемое приложение представляет из себя продукт категории SaaS (Software as a Service), что подразумевает основной доход от подписки пользователей на услугу предоставляемую сервисом. Это означает, что структура доходов будет формироваться из взимания периодических платежей, а рост доходов в такой ситуации сильно привязан к росту числа подписчиков программного обеспечения.

Целевая аудитория услуги это индивидуальные предприниматели и бизнес структуры разного уровня. Подписчиками SaaS могут быть любые структуры в обязанности которых входит работа с рекламой.

Разрабатываемая система позволит собрать в единый интерфейс информацию от разных поставщиков рекламы с использованием программного API(см. рисунок 1).

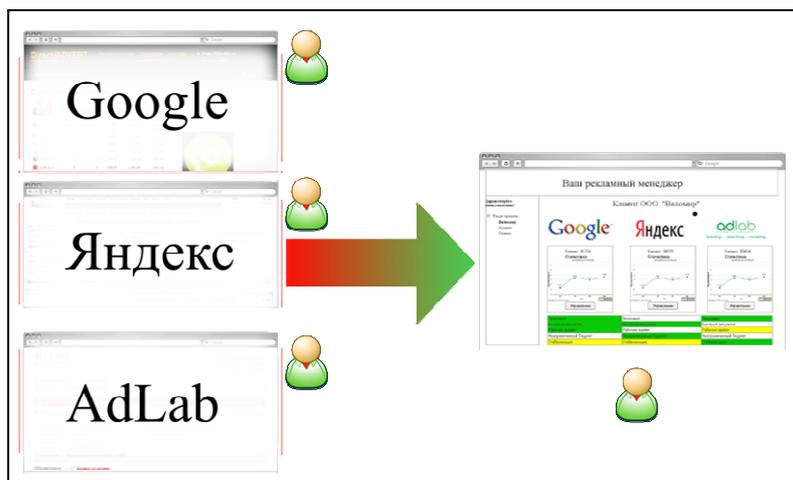


Рис 1. Единый интерфейс для различных поставщиков рекламы

Управление компанией во всех системах сведена к одному интерфейсу и выполняется одновременно, для них доступны как общие настройки, так и индивидуальные. Это позволяет обеспечить синхронную реакцию на изменения.

На рисунке 2 описана техническая схема нововведения. Между клиентом и Провайдером контекстной рекламы появляется новое звено — сервер, который содержит набор программного обеспечения, которое берёт на себя регулярный обмен информацией с провайдером и корректировку стратегии поведения клиентов исходя из действий клиента в интерфейсе (см. рисунок 1). Клиент производит первичную настройку рекламных кампаний, допустимые коридоры изменений, на основании событийной модели программирует реакцию сервера. Сервер работает в непрерывном режиме и имеет техническую возможность очень быстро реагировать на изменения рыночной ситуации.

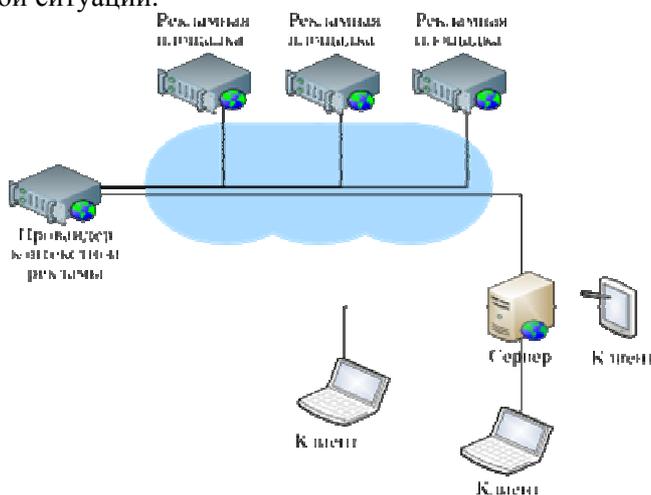


Рис 2. Техническая схема проекта

Разрабатываемая система использует собственную логику принятия решений, а её оператором с нашей стороны (со стороны исполнителя) является профессиональный SEO специалист, который будет постоянно обучать и настраивать робота.

Таким образом, планируется использовать индивидуальный подход к клиентам, им будет позволено менять логику работы системы, но они всегда будут подстрахованы поддержкой профессионального специалиста. Наличие в разрабатываемом приложении открытого публичного API будет позволять на его основе разрабатывать любые дополнительные инструменты аналитики и принятия решений.

Р.А. Симаков,
Е.В. Пугин

*Муромский институт Владимирского государственного университета
602264 г. Муром, Владимирской обл., ул. Орловская, д. 23
E-mail: roman.simakov@gmail.com, egor.pugin@gmail.com*

Способы построения процессоров запросов в СУБД

Процессор запросов является одной из главных составляющих любой СУБД. При построении, организации и оптимизации процессора запросов необходимо помнить о базовых принципах, относящихся к работе аппаратного обеспечения и к выполнению запросов:

1. Дисковый ввод/вывод доминирует в стоимости выполнения запроса, поэтому надо избегать частых обращений к дискам. В противоположность этому более интенсивное использование оперативной памяти позволяет ускорить работу системы.

2. Случайный ввод/вывод более дорог, чем последовательные операции, до тех пор, пока операции ввода/вывода не будут закешированы.

3. Следует уменьшать объём данных, передаваемых между операторами. Для этого необходимо применять проекции, группировки и предикаты, как можно раньше (предполагается, что предикаты относительно дешёвы).

Процессор запросов должен взаимодействовать с менеджером ресурсов (управление параллельной обработкой на локальном узле), а также сетевой подсистемой, чтобы организовывать выполнение в параллельном режиме на локальном и на других узлах сети. В таких системах оптимизатору запросов необходимо уметь работать с учётом распределённой структуры и обработки [2].

Попадания данных и инструкций в кэш является одной из наиболее сильно влияющих на производительность причин. Необходимо учитывать это, как и при исполнении запросов (сокращение числа вызовов функций, сбрасывающих конвейер процессора; конвейерная обработка порций данных внутри СУБД), так и при организации структуры данных (размеры чанков и т.д.).

Существует несколько методов, применяемых при проектировании процессоров запросов:

1. Конвейерное исполнение запроса. Операторы получают значения ячеек или кортежей на входах и вычисляют результат. Они требуют меньшего объёма памяти для хранения состояния в противоположность материализующим операторам, у которых часто размер результата превосходит объём оперативной памяти, и системе приходится сбрасывать его на диск.

2. В СУБД с вертикальным хранилищем, а также в СУБД, которые работают с многомерными массивами [1] удобно в качестве индексов использовать битовые карты. Основная идея заключается в том, чтобы отделить сканирование индекса от сканирования данных. Также логические операции в данном случае выполняются не над данными, а над битовыми картами, что во много раз ускоряет выполнение этих операций, т.к. не требует чтение самих данных, а только их битовых карт.

3. Использование механизмов сжатия данных. Оно уменьшает размер данных и увеличивает производительность операций ввода/вывода путём уменьшения затрат времени на поиск, так как данные хранятся ближе друг к другу. Сжатие данных снижает время передачи данных и увеличивает вероятность попадания в буфер. Существует несколько видов сжатия данных (null suppression, dictionary encoding, run-length encoding (RLE), bit-vector encoding).

В большинстве СУБД используется pull (от англ. тянуть) модель процессора запросов. Это означает, что после получения окончательного плана от оптимизатора запросов, происходит последовательное выполнение операторов по дереву от узлов к корню. Инициатором обработки в этом случае является процессор запросов. При таком подходе данные запрашиваются системой и «протягиваются» через цепочки операторов. Часто такая обработка организуется с использованием итераторов и принципа конвейерной обработки. Процессор запросов в данном случае имеет точный детерминированный план выполнения.

Второй моделью является push (от англ. толкать) процессор запросов (push executor). Обычно такая модель используется в peer-to-peer и сетевых процессорах запросов. Они обрабатывают данные, сильно распределённые по сети. При таком подходе используется гибкое планирование (flexible scheduling) выполнения запроса. Это позволяет центральному процессору обрабатывать различные части плана, когда он сталкивается с задержкой при ожидании определённого источника данных.

Для сравнения различных способов построения процессора запросов составим сводную таблицу характеристик в различных СУБД:

1. Тип СУБД (реляционная - Р, многомерная - М, хранилище пар ключ-значение - Х).
2. Организация данных (горизонтальная - Г, вертикальная - В).
3. Распределённость.
4. Многопоточное выполнение запроса.
5. Push/pull модель или смешанная (mixed).
6. Конвейерная обработка и использование итераторов.
7. Архитектура распределённой обработки (shared memory - SM, shared disk – SD, shared nothing – SN, shared everything - SE) [3].
8. Сжатие данных.

Таблица 1

Сводная таблица характеристик процессоров запросов в различных СУБД

СУБД	Характеристики							
	1	2	3	4	5	6	7	8
PostgreSQL	Р	Г	-	-	pull	+	SM	?
Oracle database	Р	Г	+	+	pull	?	?	?
Firebird	Р	Г	-	-	pull	?	?	?
MySQL	Р	Г	+	-	mixed	?	SN	?
MS SQL Server	Р	Г	+	+	pull	?	?	?
Vertica	Р	В	+	+	pull	+	SN	+
SciDB	М	В	+	+	pull	+	SN	+
MonetDB	Р	В	?	+	pull	+	?	+
Sybase IQ	Р	В	+	+	pull	?	SE	+

Построение процессора во многом зависит от остальной архитектуры СУБД и от способа организации данных. В качестве перспективных направлений по дальнейшему увеличению производительности можно выделить следующие: программная векторизация вычислений и адаптация аппаратной векторизации, эффективное сжатие и обработка сжатых данных, а также повышение вероятности попадания данных в кэш процессора.

Литература

1. Overview of SCIDB: Large scale array storage, processing and analysis. Rogers J., Zdonik S., Simakov R., Velikhov P., Smirnov A., Knizhnik K., Soroush E., Balazinska M., DeWitt D., Heath B., Maier D., Madden S., Stonebraker M., Patel J., Brown P.G. Proceedings of the ACM SIGMOD International Conference on Management of Data 2010 International Conference on Management of Data, SIGMOD '10. Sep. "Proceedings of the 2010 International Conference on Management of Data, SIGMOD '10" sponsors: ACM SIGMOD. Indianapolis, IN, 2010. С. 963-968.

2. Пугин Е.В., Симаков Р.А. Распараллеливание математических вычислений на примере операторов линейной алгебры. Алгоритмы, методы и системы обработки данных: Электронный научный журнал /под ред. С.С.Садыкова, Д.Е. Андрианова.- Вып. 17.- Электрон. журн.- Муром: Муромский институт (филиал) ВлГУ, 2011. - Режим доступа: <http://amisod.ru>, свободный.-ISSN 2220-878X (Online)

3. Симаков Р.А., Смяткин М.А. Особенности архитектуры распределенной массивно-реляционной СУБД. Алгоритмы, методы и системы обработки данных: Электронный научный журнал /под ред. С.С.Садыкова, Д.Е. Андрианова.- Вып. 18.- Электрон. журн.- Муром: Муромский институт (филиал) ВлГУ, 2011. - Режим доступа: <http://amisod.ru>, свободный.-ISSN 2220-878X (Online)

М.А. Смяткин,
Р.А. Симаков

Муромский институт Владимирского государственного университета
602264 г. Муром, Владимирской обл., ул. Орловская, д. 23
E-mail: smyatkinmaxim@gmail.com, roman.simakov@gmail.com

Реализация распределенного хранилища с применением DHT алгоритмов

Как в современной науке, так и в промышленности существуют задачи, решение которых связано с большими и непрерывно возрастающими наборами данных. Очевидно, что в таких системах из-за роста объема хранимых и обрабатываемых данных необходимо, так или иначе, реализовать возможность масштабирования. При малом количестве поддерживаемых узлов система быстро перестанет отвечать на запросы пользователей из-за роста объемов данных, а при заведомо большом количестве узлов большинство вычислительной мощности будет потрачено впустую. Можно также использовать вертикальное масштабирование (увеличение производительности конкретных узлов, вместо добавления новых), однако такой подход рекомендовал себя как более дорогой, так как значительно дешевле построить быструю систему из большого числа дешёвых ЭВМ, чем такую же, но с использованием нескольких дорогостоящих.

Поэтому важную роль в масштабируемости системы играет возможность добавления в неё новых узлов. Кроме того, узлы могут и покидать систему, как по своей воле, так и вследствие возникновения непредвиденных ситуаций. Наивным решением проблемы обработки входов/выходов узлов в системе может быть введение центральных элементов, таких как узлы, хранящие адреса других узлов и данных, расположенных на них. Однако в высоконагруженной системе такие центральные элементы быстро становятся узким местом для производительности, а также представляют собой центральную точку отказа, снижая отказоустойчивость всей системы.

Одним из решений этой проблемы являются распределённые хэш-таблицы (DHT [1]). Алгоритмы семейства DHT предоставляют возможности добавления/удаления узлов из системы с поддержанием целостности и эффективности доступа к данным (логарифмическая сложность поиска). Стоит отметить, что использование этих алгоритмов эффективно только в достаточно больших системах со средней вероятностью выхода из системы узлов. Например, если рассматривать распределённое хранилище данных, то если объёмы данных не достаточно велики, значительно выгоднее хранить метаданные (индексы, таблицы маршрутизации и т. д.) на каждом узле. Однако при увеличении нагрузок объёмы метаданных растут очень быстро, и могут занимать значительную долю хранилища. Так же достаточно дорого поддерживать целостность метаданных, когда они дублируются на каждом из узлов. В таких случаях DHT, позволяющие исключить из системы нежелательное дублирование и центральные элементы за счет дополнительной нагрузки на сеть, становятся выгодным решением.

К таким системам, в которых применение DHT имеет смысл, относится распределённое хранилище, являющееся частью разрабатываемой авторами системы [2, 3]. На текущий момент, уже имеется первая реализация сетевой подсистемы, за основу которой взята популярная DHT Chord[1]. Реализация DHT уже продемонстрировала свою работоспособность, предоставляя масштабируемость, репликацию, целостность данных и отказоустойчивость. При добавлении или удалении нового узла в систему изменения происходят только на K узлах, находящихся в непосредственной близости от него (в хэш-таблице). Примеры обработки изменения структуры системы приведены на рисунке 1.

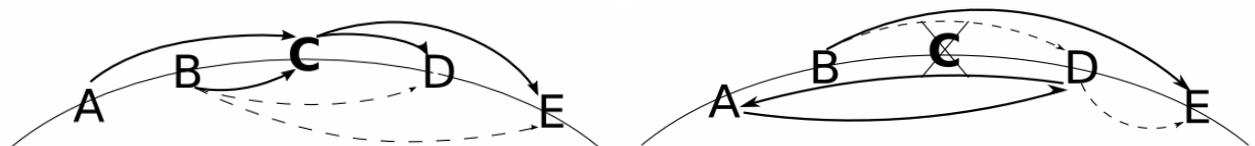


Рис. 1. Вход и выход узла D из системы с репликацией данных на двух дополнительных узлах.
Непрерывные стрелки – потоки передачи данных, пунктирные – обновление реплик

Однако существуют специфические задачи, решение которых требует дальнейших исследований. Ниже перечислены наиболее важные из них:

1. Уменьшение общего количества данных, передаваемых между узлами вследствие входов/выходов. В ряде случаев имеется возможность замедленного реагирования на входы/выходы узлов из системы. Это возможно либо если система сконфигурирована на высокую производительность, а не на поддержание целостности (т. е. потеря небольшого количества данных для клиентов не так важна, как потеря времени реагирования на запросы), либо, что более вероятно, если данные реплицируются более чем на один узел. В таком случае можно уменьшить или даже исключить перебросы данных в следующих ситуациях:

1.1. Перезагрузка узла за приемлемый промежуток времени.

1.2. Выход или вход в систему нескольких узлов (возможно, целых дата-центров) за короткий промежуток времени.

2. Распределение и балансировка нагрузки, основанные на статистических данных. В ряде случаев скорость выполнения запросов может быть значительно увеличена за счёт запроса данных непосредственно с узлов, находящихся по близости (например, в том же дата-центре). Этого можно добиться с применением специальных хэш-функций (пункт 4) или при помощи накопления статистических данных. Также, основываясь на статистических данных, имеет смысл повышение доли запросов, выполняемых на слабо загруженных узлах.

3. Адаптивное перераспределение хранилища и другие возможности саморегулирования. Целевыми применениями DHT были peer-to-peer системы, в которых каждый узел хранил набор автономной информации, которая редко имела зависимости с данными на других узлах. Однако в распределённых СУБД, например, каждый узел может хранить лишь какое-то подмножество данных и данные, как правило, имеют явные зависимости между собой. Поэтому очень важно конфигурировать систему с учётом этих зависимостей, чтобы уменьшить среднее количество узлов, участвующих в выполнении одного запроса.

4. Разработка специальных хэш-функций. Пространство идентификаторов большинства имеющихся DHT строится на основе широко известных хэш-функций, таких как sha-1. Однако, решение любой из трёх перечисленных выше проблем значительно упрощается при использовании хэш-функций, которые могут учитывать расположение узлов и наборы хранимых на них данных. Так, для репликации и эффективного распределения нагрузки важно физическое расположение узлов, а для перераспределения хранилища – важны данные, хранимые на узлах.

Решение этих проблем важно не только для конкретной разрабатываемой системы, но и для решения многих задач, затрагивающих работу с большими наборами данных в науке и промышленности, и именно на решение этих проблем нацелена дальнейшая работа.

Литература

1. Ion Stoica, et all. Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications. ACM SIGMOD, 2001.

2. М.А. Смяткин, Р.А. Симаков. Предпосылки создания распределенной массивно-реляционной СУБД. Алгоритмы, методы и системы обработки данных. 2011. №18.

3. Р.А. Симаков, М.А. Смяткин. Особенности архитектуры распределенной массивно-реляционной СУБД. Алгоритмы, методы и системы обработки данных. 2011. №18.

Е.А. Соколов,
С.Н. Серeda

*Муромский институт (филиал) Владимирского государственного университета
602264 г. Муром, Владимирской обл., ул. Орловская, д. 23
e-mail:center@mivlgu.ru*

Архитектура информационной системы управления вузом

Одной из актуальных задач информатизации для любого образовательного учреждения является внедрение автоматизированной информационной системы управления, позволяющей накапливать, анализировать и контролировать информацию обо всех бизнес-процессах организации. В различных ВУЗах эта задача решается либо локальной автоматизацией отдельных функций, либо внедрением комплексной информационной системы. На рынке программного обеспечения представлено значительное число коммерческих продуктов, направленных на автоматизацию работы ВУЗов, отличающихся функционалом, стоимостью и технологией реализации. Такие решения имеют более или менее успешные внедрения в российских ВУЗах, но не учитывают специфики работы конкретного ВУЗа, что вынуждает учреждения образования разрабатывать информационные системы своими силами.

Для решения задачи автоматизации работы ВУЗа в Муромском институте в 2012 году была разработана и успешно внедрена информационная система управления (СКАЛА). Структура информационной системы является масштабируемой и содержит базовые модули: учет показателей научной деятельности [1]; планирование учебной нагрузки преподавателей, кафедр и факультетов; контроль успеваемости студентов [2]; электронный документооборот [3]; формирование отчетов; управление показателями и документами системы менеджмента качества (СМК) института; администрирование и управление доступом.

Архитектура системы реализуется комплексом программных средств: серверная ОС Linux, web-сервер Apache, СУБД MySQL, язык программирования php. Обработка и хранение данных выполняется централизованно на высокопроизводительном многопроцессорном сервере института. Авторизованный доступ пользователей к информационным сервисам системы реализован по технологии «тонкий клиент» из любой программы-браузера и не требует установки какого-либо специального программного обеспечения, обеспечивая аппаратную и программную независимость, а также мобильность подключения с любого ПК, как в локальной сети института, так и из сети Интернет. Архитектура системы является масштабируемой, что позволяет в дальнейшем наращивать её функционал в виде новых информационных сервисов. Кроме того, благодаря простоте и гибкости установки и настройки система может быть внедрена в любом учебном заведении в короткие сроки.

Литература

1. *Соколов Е.А., Серeda С.Н.* Информационная система мониторинга показателей деятельности Вуза // Труды X Международной НТК «Новые информационные технологии и системы», Пенза, 27-29 ноября, 2012.- Пенза: ПГУ, с.324-326.
2. *Соколов Е.А., Серeda С.Н., Карнов А.В.* Автоматизация контроля успеваемости студентов // Материалы Всероссийской научной конференции «Информационно-телекоммуникационные системы и технологии», г. Кемерово, 20-22 сентября 2012 г. / Под редакцией проф. А.Г. Пимонова; Кузбас. гос. техн. ун-т им. Т.Ф. Горбачева.– Кемерово, 2012. с.241
3. *Соколов Е.А., Серeda С.Н.* Информационный сервис электронного документооборота ВУЗа // Современные проблемы науки и образования. – 2012. - №5. URL: www.science-education.ru/105-7031

Автоматизированный метод оценки надежности бортовых радиоэлектронных устройств со сложными алгоритмами работы

Важным аспектом при разработке бортовых радиоэлектронных устройств (БРЭУ) является обеспечение их надежности, особенно это актуально при проектировании аппаратуры военного и специального назначения. Одним из эффективных методов повышения надежности является введение возможности реконфигурации структуры аппаратуры. Такой метод повышения надежности широко применяется в разработке БРЭУ космических аппаратов, например, в устройствах телеметрии, электропитания и др. При этом методе в случае отказов отдельного компонента происходит перераспределение его функций между исправными компонентами либо подключение резервных, что позволяет сохранить работоспособность. Алгоритмы реконфигурации могут быть крайне разнообразны, так как они учитывают функциональные особенности БРЭУ. В устройствах питания обычно происходит перераспределение нагрузок между исправными блоками, в информационных устройствах зачастую требуется изменить маршрут обработки информации, подключить резервные вычислители и т.п.

Расчетное определение надежности реконфигурируемых БРЭУ производится грубыми оценочными способами с заведомым занижением полученных значений, что приводит к необходимости избыточных мер обеспечения надежности, то есть использованию более надежной элементной базы, дополнительного резервирования и т.д., а, следовательно, снижению качества и конкурентоспособности БРЭУ. Точную оценку вероятности безотказной работы (ВБР) для таких БРЭУ можно получить только при построении модели расчета по методу перебора гипотез, основанному на формуле полной вероятности. Однако для получения точного результата необходимо учесть все возможные сценарии функционирования БРЭУ на протяжении срока эксплуатации, то есть все последовательности и вариации отказов компонентов и реконфигураций, не приводящие к отказу БРЭУ, которых может быть достаточно много, что сильно затрудняет применение данного метода.

Повысить точность расчетов можно при использовании методологии имитационного моделирования, однако на практике она применяется редко, но для реконфигурируемых устройств это наиболее перспективный метод исследования надежности. Его применение позволяет создать модель, в которой будут учтены все особенности структуры устройства, любые виды зависимости интенсивности отказов и их комбинации во времени. Анализ методов и средств имитационного моделирования показал, что применение имитационного моделирования для расчета надежности ограничено сложностью создания и верификации модели. Так как в модели необходимо учесть все особенности функционирования БРЭУ, то для каждой конкретной задачи требуется повторять разработку и верификацию. Не смотря на крайнее разнообразие инструментов имитационного моделирования (одних языков имитационного моделирования более 3 тысяч) в инженерной практике не используются готовые шаблоны, пригодные для построения моделей БРЭУ со сложной структурой для определения показателей надежности.

В связи с этим была поставлена задача по созданию автоматизированного метода оценки надежности БРЭУ, применимого для устройств с реконфигурируемой структурой и сложными алгоритмами функционирования. Для этого необходимо разработать специализированные модели, способ верификации, средства проведения имитационного моделирования и обработки результатов. Для обеспечения адекватности результатов моделирования модель должна отражать действия реконфигурации на разных уровнях разукрупнения, учитывать зависимость интенсивности отказов от времени, историю функционирования и сложные критерии отказов. Для обеспечения требования применения модели к широкому и разнообразному классу БРЭА необходимо гибко описывать состав, структуру и критерии отказов. Такие требования можно обес-

печатать только при моделировании жизненного цикла исследуемого устройства от ввода в эксплуатацию до отказа, с учетом всех изменений происходящих в процессе функционирования.

На основе сформулированных требований была предложена форма представления устройства как множества отдельных компонентов и возможных событий. Параметрами модели компонента являются состояния и режим функционирования, также модель независимо определяет время, которое компонент пробудет в исправном состоянии при неизменном режиме работы с учетом истории эксплуатации, и информирует об изменениях своего состояния при перемотке модельного времени. При этом разрешается смена режимов и состояний компонента как вследствие случайных процессов (с некоторой интенсивностью) так и при некоторой ситуации в модели (как следствие действий реконфигурации). В качестве инструмента работы с моделями был разработан специализированный язык описания отказов и реконфигураций. Язык содержит в себе конструкции для описания состава моделируемого БРЭУ, алгоритмов реконфигурации и критериев отказов. Конструкции языка позволяют описывать самые разнообразные устройства с высокой достоверностью и за счет этого получать адекватную оценку показателей надежности.

Для работы с формальными моделями разработано программное средство (ПС), основными модулями которого являются компилятор формальной модели, модуль верификации и модуль выполнения имитационных экспериментов. Пользователь вводит формальную модель исследуемого устройства в формате языка описания отказов и реконфигураций, которая преобразуется в программную модель. В соответствии с практикой имитационного моделирования необходимо перед началом имитационного эксперимента провести верификацию модели на соответствие исследуемому устройству. Для верификации модели в ПС предусмотрена возможность проведения управляемого эксперимента. В этом случае пользователь сам определяет последовательность отказов компонентов и контролирует состояние модели после каждого отказа. В случае несоответствия реакции модели описанию работы поиск ошибки можно осуществить при помощи анализа лог-файла выполнения шага эксперимента, в котором содержится информация о том, какие условия каких событий реконфигурации были выполнены и какие именно действия с компонентами модели ими были проведены [1].

Экспериментальная проверка разработанных метода, моделей, алгоритма моделирования и программно-методических средств проводилась в два этапа. На первом проводилось моделирование стандартных структур, для которых известны точные аналитические формулы для расчета ВБР. По результатам моделирования были получены значения ВБР, сходящиеся к результатам аналитического расчета при увеличении числа экспериментов. При количестве имитационных экспериментов более 1000 относительная ошибка составляла менее одного процента. На основе этих экспериментов можно сделать вывод о правильности разработанных моделей и алгоритмов проведения имитационного эксперимента, так как такое поведение соответствует ошибке, обусловленной только погрешностью конечного числа экспериментов, устранить которую при имитационном эксперименте невозможно, однако при достаточном количестве экспериментов ею можно пренебречь в силу ее незначительности. На втором этапе проводился расчет ВБР сложных реконфигурируемых БРЭУ. При этом для верификации моделей привлекались специалисты с предприятий-разработчиков, которые экспертно подтверждали соответствие модели алгоритму реконфигурации. Во всех случаях результаты моделирования давали ожидаемые результаты, которые можно считать ближе к истинным по сравнению с аналитическими моделями, в которые заведомо были внесены допущения, приводящие к занижению ВБР. При этом разница в оценке ВБР предложенным и аналитическим методом превышала 15%.

Литература

1. Тихменев А.Н., Абрамешин А.Е., Жаднов В.В. Имитационное моделирование в оценке надежности электронных систем с реконфигурируемой структурой для космических аппаратов. / Электромагнитная совместимость и проектирование электронных средств: сб. науч. тр. // Под ред. Л.Н. Кечиева. - М.: МИЭМ, 2012. - с.13-23.

Визуально-образный язык геометрического моделирования в геометро-графической подготовке инженеров

Предметным языком геометрического моделирования является визуально-образный язык. Поэтому очень важно определить его роль и место в геометро-графической подготовке инженеров.

Визуально-образная информация непосредственно обрабатывается в оперативной памяти мозга в целом или значительными порциями, которые способны передать органы зрения человека. Информация, поступающая в виде имен одной из знаковых систем, попадая в оперативную память, вызывает из долговременной памяти соответствующий ей мысленный образ отдельного элемента “описываемой” модели. После ввода всей информации, “описывающей” данную модель осуществляется синтез отдельных элементов образа в единый мысленный образ. Обратный процесс, заключающийся в преобразовании мысленного образа во внешнюю информацию, понятную другим индивидам, представляет процесс моделирования, т.е. процесс создания модели, адекватной мысленному образу. Моделирование мысленного образа осуществляется посредством языка. Поскольку в основе мысленного образа преобладает визуально-образная информация, то наиболее адекватной формой представления его является визуально-образная модель (визуально-образный язык). Визуально-образная модель воспринимается естественным путем, также как процесс отражения реально существующих объектов. Визуально-образный язык является первичным по отношению к описательным. Это хорошо видно из анализа появления, развития и формирования визуально-образного языка у ребенка. Именно визуально-образный язык является первым языком (посредником) человеческого общества. Чем совершенней визуальный образ, тем он ближе к реально существующим параметрам объекта.

Можно сделать вывод, что визуально-образный язык является первоосновой языка вообще и первым естественным языком человеческого общества, в частности, а визуально-образное геометрическое моделирование является первой областью знания человечества, давшей начало другим её направлениям и обеспечивающей дальнейшее их развитие.

Наиболее ярко прослеживается роль визуально-образного языка в познавательном процессе. Как доказал Ж. Пиаже, наглядно-образное мышление характеризуется обратимостью осуществляемых в воображении операций и возможностью мысленно производить операции, неосуществимые в действительности. Другая важная особенность визуально-образного мышления – установление непривычных, даже «невероятных», сочетаний предметов и их свойств. В этом качестве визуально-образное мышление практически неотличимо от воображения и выступает как фактор креативности.

Известно, что визуально-образный геометрический язык играет исключительно важную интегративную роль в геометро-графической подготовке инженеров, главной целью которой является – поэтапное формирование у студентов пространственно-конструктивного мышления, предусматривающее целенаправленную подготовку, имеющую сквозной междисциплинарный характер.

Первоосновой в формировании пространственно-конструктивного мышления является овладение визуально-образным языком в курсе “Начертательная геометрия и инженерная графика” методом выполнения типовых и творческих задач по геометрическому моделированию. Следующий этап – совершенствование навыков владения визуально-образным языком на базе конструкторской подготовки профилирующих дисциплин. Завершающим этапом геометро-графической подготовки является дипломное проектирование. Следует отметить, что каждый этап формирования пространственно-конструктивного мышления характеризуется наличием разноразрядных по сложности и профессиональной направленности способов в достижении учебной цели (задачи) с использованием визуально-образного геометрического языка.

Литература

1. Рукавишников, В. А. Инженерное геометрическое моделирование как методологическая основа геометро-графической подготовки в техническом вузе: дис. ... д-ра пед. наук /В.И. Рукавишников – Казань, 2004. – 357 с.

Автоматизированная система поддержки принятия решений в задачах управления высшим учебным заведением

Текущее состояние дел в сфере высшего профессионального образования таково, что при увеличении конкуренции среди ВУЗов за ежегодно сокращающееся число абитуриентов, отсутствии притока молодых преподавателей и слабой мотивации имеющегося профессорско-преподавательского состава от высших учебных заведений по-прежнему требуется качественная подготовка выпускаемых специалистов. В подобных условиях разработка качественно новой информационной системы поддержки принятия решений, позволит снизить нагрузку на административный и профессорско-преподавательский состав, обеспечить оперативное принятие решений по организации учебного процесса, реализовать лично-ориентированный подход при работе со студентами, что в конечном итоге повысит как качество работы сотрудников учебного заведения, так и качество подготовки выпускаемых им специалистов.

Внедрение единой информационно-вычислительной системы поддержки принятия решений позволит высшим учебным заведениям повысить качество подготовки выпускаемых специалистов, улучшить условия труда сотрудников и выйти на качественно новый уровень, так как наличие подобной системы является важным критерием соответствия учебного заведения мировым стандартам.

Реализация данного проекта в высших учебных заведениях позволит повысить престиж самих учебных заведений, что очень важно для привлечения абитуриентов, а также существенно увеличить уровень заинтересованности работодателей в выпускаемых учебным заведением специалистах. Существующие тенденции к снижению федеральных дотаций, сокращению числа высших учебных заведений и ужесточению требований к организации учебного процесса вынуждают руководство учебных заведений обратить внимание на средства оптимизации и качественного улучшения деятельности ВУЗа в целом. Система поддержки принятия решений позволит снизить стоимость принимаемых организационных решений за счет упрощения информационных потоков и автоматизации обработки данных, обеспечить выполнение предъявляемых к высшим учебным заведениям России требований по организации учебного процесса.

В результате реализации проекта будет разработана информационно-вычислительная система поддержки принятия решений в задачах управления высшим учебным заведением. Применение данной системы в ВУЗах страны позволит автоматизировать обработку больших объемов информации, важной для принятия своевременных и продуманных управленческих решений, позволит повысить качество предоставляемых образовательных услуг, а также повысить престиж ВУЗа и выпускаемых специалистов. Использование данной системы может стать существенным конкурентным преимуществом как в борьбе за абитуриентов, так и в борьбе за привлечение дополнительных финансовых потоков, что особенно важно в период мирового финансово-экономического кризиса.

Подсистема планирования штатного расписания кафедр позволит планировать распределение штатных единиц профессорско-преподавательского состава исходя из реализуемых учебных планов специальностей. Применение подхода, связанного с анализом учебных планов, позволяет обеспечить равномерное распределение штатных единиц по кафедрам, оперативный расчет и прогнозирование, что немаловажно в условиях меняющихся требований к образовательным учреждениям.

Подсистема учета регистрируемых объектов интеллектуальной собственности позволит автоматизировать учет и обработку данных о зарегистрированных в учебном заведении объектах интеллектуальной собственности (изобретения, полезные модели и др.), а также ускорить принятие решений, касающихся научно-исследовательской работы в высшем учебном заведении.

Возможности подсистем анализа успеваемости студентов университета, количественной оценки научной и внеучебной деятельности студентов, учета социально незащищенных студентов позволят совершенствовать процессы принятия решений, принимаемых в отношении студентов. Необходимость данных подсистем обусловлена сложностью работы со студентами и невозможностью должного применения лично ориентированного подхода к студентам в связи с тем, что на одного преподавателя часто приходится до нескольких сотен студентов, а на одного административного служащего – и того больше. В подобных условиях применение систем поддержки принятия решений позволит не только заочно познакомиться со студентами, но и при общении с ними максимально эффективно решать поставленные задачи.

Оценка временных характеристик работы конвейера имитационной модели ядра ARM Cortex-M3 для случая однобайтных команд

Существенное развитие технологических процессов и элементной базы дало возможность активно разрабатывать и использовать системы сбора и обработки информации на основе микроконтроллеров. Одной из перспективных архитектур является Advanced Risc Machine (ARM) Cortex-M3 [1]. К преимуществам ее можно отнести незначительное энергопотребление и гибкую систему команд, которая поддерживает наборы инструкций Thumb и Thumb-2. Тактовые частоты для семейств микроконтроллеров на данном ядре находятся в диапазоне 24-120 МГц. Поэтому данные микроконтроллеры в основном применяются в целях получения информации и управления в низкочастотных трактах технических систем.

Для данной архитектуры существуют как с открытым исходным кодом, так и проприетарные средства разработки программного обеспечения, включая компиляторы, отладчики, про-файлеры. Но такие средства требуют физического наличия отладочной платы с микроконтроллером, либо целевой платы и внутрисхемного эмулятора.

В отдельную категорию можно выделить интегрированные среды разработки, которые дополнительно позволяют эмулировать некоторые внутренние модули микроконтроллера. Сюда можно отнести Texas Instruments CCS, Atmel Studio 6, IAR Electronic Workbench, Keil uVision IDE и другие. Однако, они не обладают достаточным функционалом для симуляции электрических схем. Поэтому разработка имитационной модели микроконтроллера на ядре Cortex-M3 и интеграция ее в системы симуляции электронных схем является актуальной задачей.

На основе подхода, описанного в [2, 3], была разработана имитационная модель ядра ARM-микроконтроллера на ядре Cortex-M3 и интегрирована в среду разработки Labcenter Proteus 7. Одной из важных задач является верификация поведения микроконтроллера в рабочем режиме, а также реакция на процедуры конфигурирования.

В данной работе рассматривается работа конвейера и временные диаграммы его работы для случая простых команд, которые выполняются в течение одного машинного цикла. Блок-схема микроконтроллера приведена на рисунке 1. Конвейер в модели находится в модуле ЦПУ и взаимодействует с менеджером шин и контроллером прерываний. Анализ функционирования конвейера и исполнителя команд показал, что их временные диаграммы формируются в зависимости от тактовой частоты и типа обращения к памяти. Транзакции доступа к памяти области «BITBAND» выполняются дольше, чем транзакции доступа к периферийным регистрам или модулям памяти. Это связано с дополнительной обработкой параметров исходной транзакции, которые дешифруются в транзакцию к целевому модулю.

Было проведено исследование временных характеристик работы конвейера имитационной модели ядра ARM Cortex-M3 для SPICE симуляции. Выявлены основные факторы, влияющие на временные диаграммы выборки, декодирования и исполнения команд.

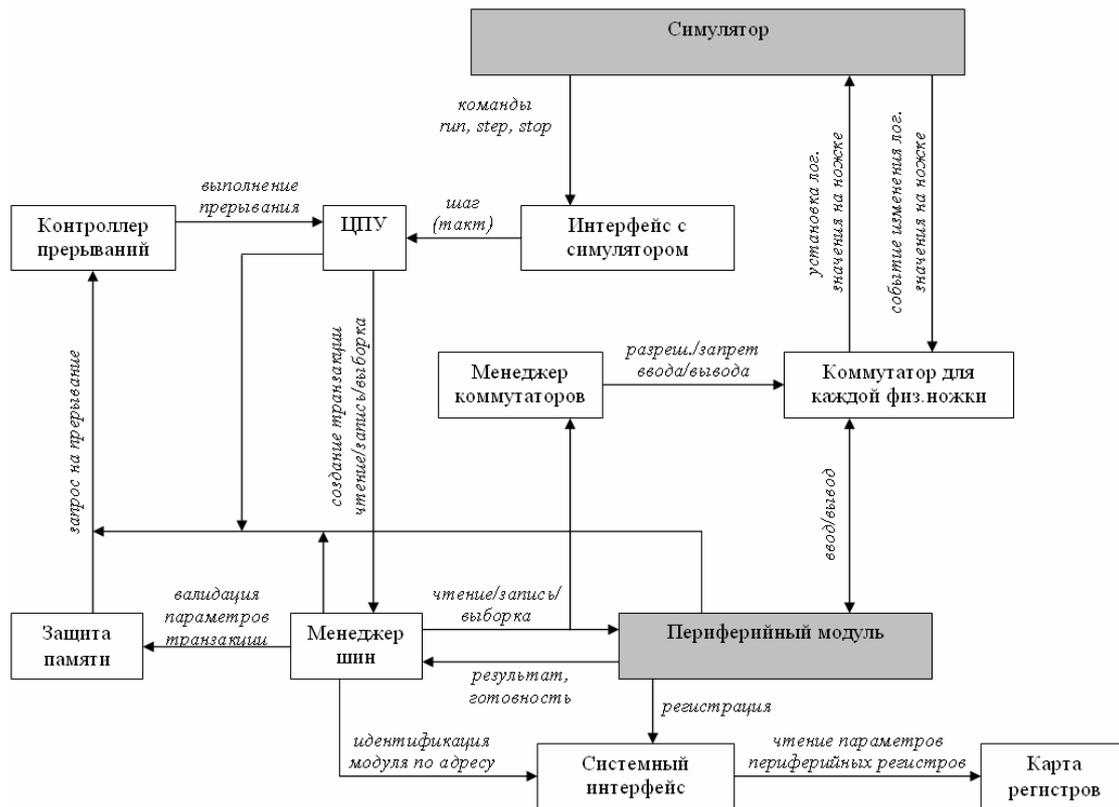


Рис. 1. Блок-схема микроконтроллера на ядре ARM Cortex-M3

Литература

1. Cortex™-M3 Technical Reference Manual. [Электронный ресурс]. URL: <http://www.arm.com>. (дата обращения: 05.10.2012)
2. Усачев М.В., Мольков Н.П. Обобщенная имитационная модель микроконтроллера смешанных сигналов для SPICE-симуляции // Методы и устройства передачи и обработки информации: межвуз. сб. науч. трудов. – Вып.12 / Под ред. В.В.Ромашова. – М.: «Радиотехника», 2010. - С. 75-78.
3. Усачев М.В., Мольков Н.П. Имитационная модель ядра ARM для SPICE симуляции микроконтроллерных систем управления. // Моделирование, идентификация, синтез систем управления: сборник тезисов пятнадцатой Международной научно-технической конференции. 9-16 сентября 2012., Национальная академия наук Украины. Институт прикладной математики и механики, Донецк, 2012. 199 с.

Автоматизация учета заявок в ЗАО «Электон»

Закрытое акционерное общество «Электон» является одним из лидеров в области производства наземного электротехнического оборудования для нефтяной промышленности в России и на международном рынке. Динамичное развитие предприятия, стремление к постоянному самосовершенствованию, ориентация на мировые стандарты качества – вот только некоторые из причин стабильного роста производства и сбыта продукции с торговой маркой "Электон".

В связи с тем, что предприятие работает с большинством клиентов удаленно, имеет смысл автоматизировать непосредственно процесс обработки заявок, а именно спроектировать и внедрить модуль учета заявок, что повысит экономическую эффективность предприятия.

Отслеживание заявок потребителей становится очень важным при частых и большого объема закупках. Для этой цели можно ввести возможность регистрации заявок на комплектующие, расходные материалы или услуги. Исходя из этого, в модуле можно выделить следующие типы заявок:

- ИТ заявка - предназначена для учета потребностей в покупке оборудования. То есть данная заявка вводится на основании запроса пользователей (его можно оформить документом: заявка пользователя) о необходимости каких либо комплектующих, услуг, расходных материалов, для того чтобы в будущем включить данную заявку в общую.
- Тендер - предназначен для организации тендера, в данной заявке регламентируются позиции, по которым будет проходить тендер на поставку. Данная заявка обычно распечатывается и отправляется всем участникам тендера.
- Счет потребителю - необходим для фиксации выставленного счета потребителю. В данном типе можно вводить суммы закупки.

Схема движения по этим видам заявок представлена на рисунке 1.

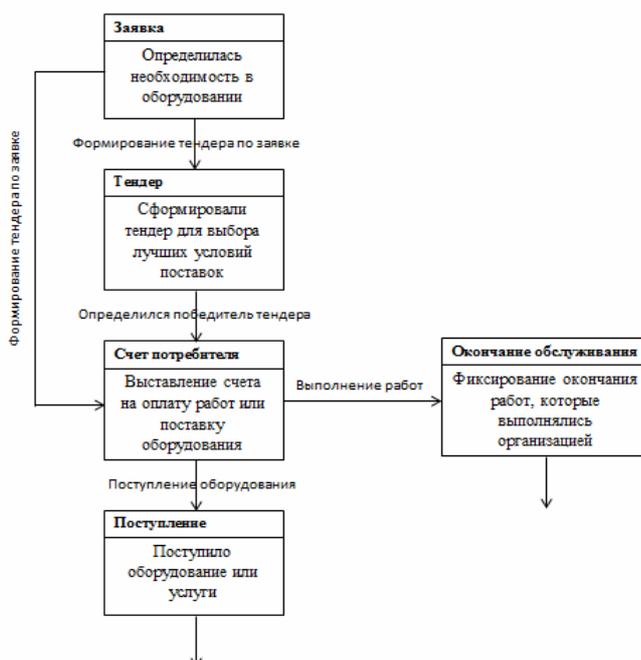


Рис.1. Схема движения по заявкам

Внедрение данного модуля учета заявок позволит значительно сократить время на согласование, особенно в организациях, имеющих территориально удаленные подразделения. Основная суть модуля заключается в автоматизации бизнес-процессов приема, обработки и отчетности по заявкам и обращениям клиентов.

Литература

1. Костров, А. В. Методы и модели информационного менеджмента: учеб. пособие / Д. В. Александров, А. В. Костров, Р.И.Макаров, Е.Р. Хорошева; под ред. А.В. Кострова. – М. : Финансы и статистика, 2007. – 336 с.
2. Костров А.В. Основы информационного менеджмента: Учеб. пособие / Владим. гос. ун-т. – М.: Финансы и статистика, 2001. – 336 с. – ISBN 5-279-02314-0.

База данных по характеристикам надежности электрорадиоизделий

Надежность является одним из самых важных свойств современных радиоэлектронных средств (РЭС). От нее зависят показатели качества, эффективности, безотказности, живучести и другие важнейшие характеристики. РЭС может быть эффективным только при условии, что оно имеет высокую надежность. Для создания надежных РЭС инженеру необходимо рассчитывать надежность уже на ранних этапах проектирования [1].

Для расчета надежности РЭС необходимо знать параметры электрорадиоизделий (ЭРИ), составляющих их. Источником о номенклатуре разрешенных к применению ЭРИ служит ограничительный перечень, который обычно формируется МО РФ. В этом перечне содержится номенклатура только ЭРИ отечественного производства. Но, как известно, в РЭС могут входить и ЭРИ иностранного производства (ИП). Для этого предприятия, которые создают РЭС с использованием ЭРИ ИП должны подать соответствующую заявку и провести сертификацию таких ЭРИ. Однако, добавление ЭРИ ИП в перечень возможно только в том случае, если оно используется на нескольких предприятиях. Данный перечень должен пополняться регулярно, но обычно он пополняется раз в полгода.

Другим источником информации о характеристиках надежности ЭРИ является справочник «Надежность ЭРИ», который является официальным изданием МО РФ. Справочник содержит сведения о характеристиках надежности ЭРИ, применяемых при разработке (модернизации), производстве и эксплуатации аппаратуры, приборов, устройств и оборудования военного назначения. Справочник обычно обновлялся каждые два года, но, к сожалению, последние 6 лет обновлений не было [1].

Исходя из вышесказанного, возникает идея создания портала по характеристикам надежности ЭРИ с базой данных. Пользователь, пользующийся таким порталом, сможет получать достоверную информацию о параметрах и характеристиках надежности ЭРИ как отечественного, так и иностранного производства [2].

Для построения базы данных необходимо использовать такие программные средства, как Denwer, MySQL, MyPHPAdmin. Рассмотрим каждый программный продукт.

Denwer представляет собой пакет программ для создания веб сайта на локальном ПК под управлением операционной системы семейства Windows без необходимости выхода в интернет. Сразу после установки доступен полностью работающий веб-сервер Apache, работающий на локальном компьютере, на котором может работать неограниченное количество сайтов, что очень эффективно для разработки и отладки сценариев PHP без загрузки его файлов на удаленный сервер. Для запуска практически всех утилит «Денвера» используется приложение Run в подкаталоге /denwer (или /etc) корневого каталога установки «Денвера». При запуске создается виртуальный диск (по умолчанию Z:), где хранятся все файлы проектов. Denwer поддерживает работу со схемного флеш накопителя.

Особенностью, отличающей Denwer от других WAMP-дистрибутивов, является автоматическая правка системного файла hosts, являющегося локальным аналогом DNS-сервера, что позволяет обращаться к локальным сайтам, работающим под управлением Денвера, по именам, совпадающим с именем папки, расположенной в каталоге home пакета программ Denwer. Базовый пакет поставки включает в себя:

- Веб-сервер Apache с поддержкой SSI, SSL, mod_rewrite, mod_php.
- Интерпретатор PHP с поддержкой GD, MySQL, SQLite.
- СУБД MySQL с поддержкой транзакций (mysqld-max).
- Систему управления виртуальными хостами, основанная на шаблонах.
- Систему управления запуском и завершением.
- Панель phpMyAdmin для администрирования СУБД.

MySQL - это свободная система управления базами данных. Гибкость СУБД MySQL обеспечивается поддержкой большого количества типов таблиц: пользователи могут выбрать как таблицы типа MyISAM, поддерживающие полнотекстовый поиск, так и таблицы InnoDB, поддерживающие транзакции на уровне отдельных записей.

PhpMyAdmin - веб-приложение с открытым кодом, написанное на языке PHP и представляющее собой веб-интерфейс для администрирования СУБД MySQL. PHPMyAdmin позволяет через браузер осуществлять администрирование сервера MySQL, запускать команды SQL и просматривать содержимое таблиц и баз данных. Приложение пользуется большой популярностью у веб-разработчиков, так как позволяет управлять СУБД MySQL без непосредственного ввода SQL команд, предоставляя дружелюбный интерфейс.[3].

Рассмотрим создание базы данных с использованием этих программных продуктов. Для начала проектирования базы данных на локальном компьютере необходимо запустить пакет программ Denwer. После запуска данный программный продукт в автоматическом режиме произведет запуск и настройку всех необходимых программ для создания веб-ресурса. В нашем случае база данных является частью веб-ресурса.

Дальнейшее проектирование БД можно осуществлять в среде phpMyAdmin с применением, как визуальных средств создания БД, так и SQL команд. Создание базы данных осуществляется командой CREATE DATABASE db_name. db_name - имя базы данных. После создания базы данных необходимо создать необходимое количество таблиц, в которых будет храниться информация о каких-либо объектах. В нашем случае будет храниться информация о характеристиках надежности ЭРИ, их описание и название. Таблица создается командой CREATE TABLE table_name [(create_definition, ...)]. Table_name - имя таблицы. В скобках указывается информация о столбцах таблицы (имя и пр.) [4]. После успешного создания базы данных и таблиц в случае с БД по характеристикам надежности ЭРИ мы получили структуру БД. Название базы данных было выбрано Nadegnost. Имена таблиц задавались из такого расчета, чтобы администратор смог быстро найти в БД требуемую ему таблицу и внести некоторые изменения. В качестве примера можно использовать имена таблиц "Plis_osn" и "Plis_nad". Легко понять, что в первой таблице собрана информация об основных характеристиках ПЛИС, а во второй - характеристики надежности ПЛИС.

Для обеспечения взаимодействия базы данных с пользователем необходимо создать веб-страницу, где пользователь сможет получить информацию о необходимом элементе, а также, при наличии соответствующих прав, добавить информацию в базу данных. Веб-страница создается применением языка гипертекстовой разметки HTML [5]. С помощью этого языка создается дизайн страницы, и создаются необходимые формы для добавления информации в базу данных.

Создание базы данных для портала сложная задача, требующая знания многих программных продуктов. Но это необходимо для функционирования любого серьезного веб-ресурса, будь то интернет-магазин или портал по характеристикам надежности ЭРИ. База данных для портала позволит пользователям ресурса самостоятельно добавлять информацию в БД, а также просматривать ее в удобное для них время. При этом затрачивается намного меньше ресурсов на поиск информации, чем при использовании веб-ресурсов фирм-производителей.

Литература

1. Жаднов В.В. Информационные технологии в прогнозировании надежности электронных средств. / Информационные технологии в проектировании и производстве, № 1, 2012. - с. 20-25.
2. Цыганов П.А., Жаднов В.В. Информационный портал для специалистов в области надежности радиоэлектронных средств. / Инновационные информационные технологии: Материалы международной научно-практической конференции. // Под ред. С.У. Увайсова; Отв. за вып. И.А. Иванов, Л.М. Агеева, Д.А. Дубоделова, В.Е. Еремина. - М.: МИЭМ, 2012. - с. 337-340.
3. Кристиан Д. PHP и MySQL: создание интернет-магазина. - СПб.: Вильямс, 2011. - 632 с.
4. Голицына О.Л. Основы проектирования баз данных. - М.: Форум, 2012. - 415 с.
5. Фримэн Э. Изучаем HTML, XHTML и CSS. - СПб.: Питер, 2012. - 656 с.

Использование метода шинглов в сравнении документов

Сравнение документов – важная задача в развитии современной научной мысли. Для установления авторства работы важно знать, является ли научный документ подлинником, или он дубликат другого документа. Точно установленное авторство серьезной статьи, монографии или диссертации – путь к публикации в научных журналах, выступлению на конференциях, регистрации патентов и авторских прав на изобретения и к получению заслуженной научной степени. В научной среде встречаются случаи частичного или полного дублирования работ, в частности статей. Выявление таких фактов подчас является большой проблемой технического плана: не хватает аппаратной мощности и не совсем удовлетворительный способ поиска.

В связи с большим объемом текстов прямой поиск по словам – очень трудоемкая задача. Поэтому требуется использование более быстрых методов сравнения документов.

В работах [1, 2] был представлен метод сравнения документов с использованием шинглов. Шингл в переводе с английского языка – «чешуйка». На такие «чешуйки», шинглы, делятся оба рассматриваемых текста. Деление на шинглы может осуществляться по-разному. Они могут идти друг за другом (встык), могут перекрывать друг друга (внахлест), могут идти через некоторое расстояние друг от друга, пропуская несколько символов или слов. Шингл может содержать одну букву или одно слово, а может предложение или несколько предложений.

От величины шингла, используемого в системе сравнения документов, зависит качество и точность сопоставления работ. Чем больше шингл, тем больше разброс вероятности заимствования. Это ведет к неправильным результатам сравнения (вероятности совпадения). Если шингл маленький (1-2 символа), то точность сравнения может быть очень высокой, но вычислительные мощности, затрачиваемые на реализацию системы сравнения с такими шинглами, могут быть слишком велики, ведь одна из главных задач использования шинглов - снизить трудоемкость сравнения документов.



Рис. 1. Разбиение текста на шинглы внахлест

Уменьшение сравниваемого текста достигается путем кодирования шингла, превращением его в какое-то число, например, с помощью алгоритмов CRC, MD5. Все коды шинглов документа формируют список кодов, по которому проводится сопоставление. Для ускорения сравнения из шинглов может быть сформирован супершингл, тоже представляющий собой числовой код [3]. По кодам, полученным путем преобразования шинглов или супершинглов, системой сравнения проводится сопоставление документов и определяется вероятность их совпадения.

Документы, участвующие в сравнении лучше всего привести к «нормальной» форме: убрать пробелы, знаки препинания и даже приставки в словах, так как эти символы не несут решающей смысловой нагрузки.

Разные методы деления текста на шинглы могут сильно отличаться по результативности. Если шинглы идут друг за другом (встык), то их подвижка всего лишь одним-двумя символами может изменить весь текст до неузнаваемости для системы сравнения документов, так как подвинутые шинглы будут содержать уже совсем другие слова и символы. Этого эффекта можно избежать, если применять метод наложения шинглов (внахлест), либо метод выборочного взя-

тия шинглов. Вставка символов изменит лишь те шинглы, которые ей изменены. Другие шинглы не пострадают, и сравнение документов пройдет более точно, чем при использовании предыдущего метода.

В развитие метода шинглов заложен большой нереализованный потенциал, позволяющий многократно совершенствовать его путем проработки и добавления новых, неучтенных качеств сравнения текстов, так как проблема эта все еще актуальна в современной научной и образовательной средах.

Литература

1. Broder A., Glassman S., Manasse M., Zweig G. Syntactic clustering of the Web // Proc. of the 6th International World Wide Web Conference, April 1997.
2. Broder A. On the resemblance and containment of documents // Compression and Complexity of Sequences (SEQUENCES'97), IEEE Computer Society, 1998, p. 21-29.
3. Зеленков Ю.Г., Сегалович И.В. Сравнительный анализ методов определения нечетких дубликатов для WEB-документов // Труды 9-ой Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» RCDL'2007: сб. работ участников конкурса – Переславль-Залесский, 2007. – Т. 1. – С. 166-174.