УДК 004.021

А. М. Тележкин

СИСТЕМА САМПО+ ДЛЯ СОЗДАНИЯ И АНАЛИЗА ИСТОРИЧЕСКОЙ БАЗЫ ДАННЫХ

Представлена методология создания исторических баз данных, предназначенных для более точной оценки необходимых ресурсов для инициируемых проектов. Приведена процедура создания такой базы с помощью системы САМПО+.

Ключевые слова: *историческая база данных проекта, сбор и анализ проектных метрик.*

Введение. Компании, работающие в сфере программного обеспечения, достигают определенного уровня зрелости, когда накапливается некоторый объем информации об уже выполненных в компаниях проектах [1]. Причем эти данные не всегда используются для оценки ресурсов при запуске следующих проектов. Сбор и накопление информации по реализованным проектам в исторической базе данных (ИБД) обеспечивает возможность использования имеющегося опыта для более точного планирования и контроля новых проектов.

Историческая база данных о выполненных проектах содержит такие характеристики, как: предметная область проекта, модель жизненного цикла, используемые технологии, затраченные ресурсы, риски, бюджет, цели проекта, даты начала и окончания этапов проекта и т.д., а также различные процессные и продуктовые метрики, связанные с исполнением проектов в конкретной организации. Использование ИБД позволяет на основе анализа найденных проектов-аналогов произвести оценку ресурсов, необходимых для успешного завершения инициируемого или выполняемого нового проекта.

Проект-аналог — это проект, который, по мнению эксперта, может служить в качестве основы для оценивания соответствующих характеристик исследуемого проекта, а в предельном случае и заимствования модели процесса разработки программного изделия.

Система САМПО+. Система поддержки создания ИБД компаний (САМПО+), разрабатывающих программное обеспечение (далее — Система), предназначена для снижения трудозатрат и оценки необходимых ресурсов. Система позволяет на основе ИБД сформировать модель базы данных для конкретной организации, а также произвести объективную оценку выполнимости инициируемых проектов. В этом качестве Система служит инструментом для снижения риска незавершения нового проекта.

Наряду со своим прямым назначением Система может использоваться в качестве средства повышения уровня профессиональных знаний сотрудников компании, а также в учебном процессе при подготовке специалистов по разработке программного обеспечения.

Пользователем Системы является эксперт в области управления процессом разработки программных изделий, владеющий знаниями, плохо поддающимися формализации. В Системе используется методология моделирования и принятия решений (прозрачные технологии) на основе алгоритмических сетей, разработанная в СПИИРАН [2].

Базу прозрачной технологии в системе САМПО+ составляет алгоритм вычисления оценок, предложенный в начале 1970-х гг. акад. Ю. И. Журавлёвым [3]. Основная идея этого алгоритма заключается в том, что для определения класса какого-либо объекта используются не отдельные характеристики, а их совокупность (ансамбль).

Важной особенностью Системы является хранение ссылок на документы, из которых взята информация, что связано с необходимостью обеспечения возможности контроля и повторного анализа вводимых значений характеристик.

Методология, на основе которой в Системе определяется экспертное множество характеристик, формирующих ИБД, была предложена и испытана в СПИИРАН и включает следующие аналитические процессы:

вербальный — формирование, исходя из экспертного множества, исходных множеств источников, характеристик и проектов на основе анализа методической и специальной литературы, а также бизнес-процессов компании;

количественный — формирование уточненных множеств источников, характеристик и проектов на основе анализа обеспеченности каждого из них проектной информацией и представления об их значимости для поиска проектов-аналогов;

качественный — формирование рекомендуемых множеств источников, характеристик и проектов на основе экспериментов, подтверждающих достаточность сформированного в ходе предыдущих процессов множества характеристик, для решения задачи распознавания (под задачей распознавания в данном случае понимается поиск проекта-аналога).

Подробно проблемы формирования ИБД, а также методология их построения на примере системы САМПО+ рассмотрены в работе [4]. В настоящей статье рассматриваются уточнения к методологии, а также особенности формирования и исследования ИБД в системе САМПО+ по опыту ее практического применения в организации “Exigen Services” (Санкт-Петербург) в ходе анализа базы выполненных проектов.

Режимы Системы. Система обеспечивает поддержку режимов формирования ИБД, исследования ИБД и режима использования собранной информации в ИБД.

Главное меню Системы представлено схемой, приведенной на рис. 1.

Рис. 1

Множество источников. Результатом анализа информации для создания ИБД является множество источников, в качестве которых могут быть названы знания эксперта, метрическая/корпоративная база данных компании, а также документы, создаваемые на основе шаблонов, и сами шаблоны.

Атрибуты источника: наименование; описание; множество, к которому принадлежит источник (исходное, уточненное, рекомендуемое); используемость во множестве характеристик; используемость во множестве проектов; дата и автор последнего изменения.

Множество характеристик. Данное множество позволяет описать любой проект, реализованный или выполняемый в компании, для последующего анализа и использования фактологических данных при инициации новых проектов.

Для формирования исходного множества характеристик используются, как правило, пять источников:

— общетехническая литература, содержащая информацию о выполнении проектов любых типов;

— специальная литература, содержащая информацию о выполнении проектов разработки программных изделий (CMM, CMMI, ISO и пр.);

— база данных компании, в которой собраны какие-либо процессные, продуктовые и проектные метрики;

— документы по проекту, или имущество проекта, которые формируются во время работы над проектом, а также по его завершении;

— мнения экспертов.

Для уменьшения пространства поиска проектов-аналогов предлагается структурировать исходное множество характеристик, а именно, выделить в нем три подмножества: „категории“, „интегральные характеристики“ и „терминальные характеристики“.

Атрибуты характеристики: уникальный номер; наименование; описание; тип (интегральная, вычисляемая, число, текст, шкала, дата); описание области задания; значение по умолчанию; документы, из которых характеристика может быть извлечена (заключение-мнение эксперта, корпоративная база данных, документы проекта); уникальный номер родителя; уникальные номера потомков; формула, по которой вычисляется значение характеристики; множество, к которому принадлежит характеристика (исходное, уточненное, рекомендуемое); используемость в проектах; дата и автор последнего изменения.

Множество проектов. Данное множество содержит все проекты организации, которые были выполнены, а также выполняются в настоящее время.

Атрибуты проекта: наименование; содержание; перечень источников информации; местонахождение источников; значения характеристик; множество, к которому принадлежит проект (исходное, уточненное, рекомендуемое); обеспеченность информацией; дата и автор последнего изменения.

Организация работы Системы. Добавление, редактирование и удаление источников, характеристик и проектов исходного, уточненного и рекомендуемого множеств организовано по единому сценарию. При входе в соответствующий режим Системы в выпадающем списке предлагается выбрать одно из множеств, которое будет редактироваться (исходное, уточненное или рекомендуемое). В зависимости от выбора в таблицу данных загружается необходимое множество.

Ввод значений, характеризующих конкретное множество, осуществляется по другой схеме. Пользователю предлагается таблица, в которой указываются наименования проектов, характеристик и их значения. При этом возможны следующие действия пользователя.

1. В графе наименований характеристик выбирается одна из характеристик, после чего открывается форма ввода значений, и пользователь указывает значения этой характеристики по всем проектам.

2. В графе наименований проектов выбирается один из проектов, после чего открывается форма ввода значений, и пользователь последовательно указывает значения характеристик для данного проекта.

3. Выбирается произвольная ячейка, содержащая значения характеристик по конкретному проекту, после чего открывается форма ввода значений, и пользователь указывает значение конкретной характеристики для конкретного проекта.

Режимы исследования. В Системе заданы четыре режима исследования накопленной информации — источников, характеристик, проектов и функциональной пригодности.

В режиме *исследования источников* пользователь работает с таблицей, в которой содержится информация о том, сколько раз данный источник был использован в том или ином проекте, а также процентное и абсолютное число „вхождений“ каждого из источников во все проекты.

В режиме *исследования характеристик* пользователь работает с таблицей, в которой содержится информация о значениях характеристик по каждому проекту, а также процентное и абсолютное число „вхождений“ каждой из характеристик во все проекты.

В режиме *исследования проектов* пользователь работает с таблицей, в которой содержится информация о том, сколько источников было использовано в каждом проекте, а также процентное и абсолютное число „вхождений“ источников в каждый проект.

В режиме *исследования функциональной пригодности* базы данных пользователь работает с таблицей множества характеристик, с помощью которой он может построить свое подмножество характеристик и определить проекты, аналогичные инициируемому. Работа в данном режиме продемонстрирована на рис. 2 в виде снимка с экрана компьютера.

Рис. 2

Заключение. В результате исследований было сформировано множество из 360 различных характеристик, которые достаточно полно описывают процесс разработки программного обеспечения для стандартной компании.

Система САМПО+ использовалась при анализе характеристик для создания БД, собранных компанией “Exigen Services” в ходе выполнения проектов и при их завершении. Исследование проводилось в 2009—2010 гг., в общей сложности было рассмотрено 342 проекта, из них в Систему включен 71 проект.

Система также использовалась при анализе базы данных автоматизации документооборота НП „Объединение подземных строителей“ (Санкт-Петербург).

Система САМПО+ создана в среде MS Excel 2007/2010 и содержит 6200 строк программного кода на языке Visual Basic for Applications.

Статья подготовлена по результатам работы, выполненной при финансовой поддержке ведущих университетов Российской Федерации (субсидия 074-U01 — Университет ИТМО).

СПИСОК ЛИТЕРАТУРЫ

1. *Баранов С. Н., Домарацкий А. Н., Ласточкин Н. К., Морозов В. П.* Процесс разработки программных изделий. М.: Наука — Физматлит, 2000. 176 с.
2. *Морозов В. П.* Поддержка принятия решений, ориентированная на знания эксперта // Тр. XII Санкт-Петербург. междунар. конф. „Региональная информатика (РИ-2010)“, 20—22 окт. 2010 г. СПб: СПОИСУ, 2011. С. 69—73.
3. *Журавлев Ю. И., Никифоров В. В.* Алгоритмы распознавания, основанные на вычислении оценок // Кибернетика. 1971. № 3. С. 1—11.
4. *Тележкин А. М.* Создание исторических баз данных при помощи системы САМПО+ // Тр. Юбилейной XIII Санкт-Петербург. междунар. конф. „Региональная информатика (РИ-2012)“, 24—26 окт. 2012 г. СПб: СПОИСУ, 2013. С. 84—90.

Сведения об авторе

Александр Михайлович Тележкин — аспирант; СПИИРАН, лаборатория информационных технологий в системном анализе и моделировании; E-mail: telezhkin@gmail.com

Рекомендована СПИИРАН

Поступила в редакцию
10.06.14 г.

УДК 004.056

А. В. ФЕДОРЧЕНКО, А. А. ЧЕЧУЛИН, И. В. КОТЕНКО

ПОСТРОЕНИЕ ИНТЕГРИРОВАННОЙ БАЗЫ УЯЗВИМОСТЕЙ

Представлены результаты исследования открытых баз данных уязвимостей и описание процесса их интеграции для применения в системах оценивания защищенности компьютерных сетей. Предлагаются модель процесса формирования и структура интегрированной базы уязвимостей, а также описание и анализ разработанного прототипа.

Ключевые слова: *анализ защищенности, базы данных уязвимостей, системы мониторинга безопасности.*

Введение. В настоящее время существует большое количество баз данных (БД) уязвимостей, как открытых для общего доступа, так и закрытых, используемых в коммерческих продуктах. Они применяются в различных системах безопасности, сканерах уязвимостей и других средствах обеспечения комплексной защиты компьютерных систем. Однако применение таких баз данных в системах оценивания защищенности компьютерных сетей в режиме реального времени недопустимо вследствие низкой скорости поиска записей уязвимостей для оперативной обработки событий, нарушающих информационную безопасность [1—4]. Также следует отметить, что формирование БД уязвимостей не стандартизовано и производится несогласованно, что влияет на точность обнаружения уязвимостей в используемом программно-аппаратном обеспечении.

Для увеличения объема уникальных записей уязвимостей и списков программно-аппаратных продуктов, соответствующих этим уязвимостям, предлагается объединение открытых баз уязвимостей. Реализация данного процесса предусматривает разработку методики интеграции баз уязвимостей и проектирование структуры интегрированной базы для адаптации ее к быстрому поиску записей уязвимостей. Конечное использование формируемой интегрированной базы уязвимостей подразумевает ее эксплуатацию в системах оценивания защищенности компьютерных сетей, анализ которых должен проводиться в режиме, близком к