

**СИСТЕМА ПОДДЕРЖКИ СОЗДАНИЯ БАЗ ИСТОРИЧЕСКИХ ДАННЫХ КОМПАНИЙ, РАЗРАБАТЫВАЮЩИХ ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ****В. П. Морозов, Е. А. Калугина, А. М. Тележкин (Санкт-Петербург)**

Неправильная оценка ресурсов, необходимых для успешного завершения проекта разработки программного обеспечения (ПО), является одной из важнейших причин последующего превышения сроков, бюджета, отсутствия требуемого качества [1, 2].

Значимость правильной оценки, относительно короткий промежуток времени, отводимый на ее получение, неопределенный характер используемой информации обуславливают необходимость привлечения формальных методов и автоматизации процесса оценивания.

В работе [3] задачу оценки ресурсов предлагается рассматривать как совокупность задач, последовательно реализующихся в трех основных этапах:

1. Этап выбора проекта-аналога (формирование «поля», на котором, согласно [4], предполагается использовать прогнозную модель).
2. Этап формирования прогнозных значений множества характеристик иницируемого проекта (прогнозная модель).
3. Этап собственно оценки возможности реализации иницируемого проекта на основе полученного прогноза (модель принятия решения).

Очевидно, что уже на первом этапе эксперты имеют возможность проводить предварительную оценку необходимых ресурсов, однако в качестве неперемного условия его реализации необходима база данных, содержащая информацию о проектах, ранее выполнявшихся в организации. Хотя, в соответствии со стандартами разработки программного обеспечения [5–7], начиная с определенного уровня зрелости, организации должны иметь базу данных процесса, последние создавались как базы метрических данных, ориентированные на решение задач оперативного контроля над ходом выполнения проектов, и для решения задачи оценки не пригодны. Базу данных, содержащую информацию, обеспечивающую поиск проектов, аналогичных иницируемому, и позволяющую по значениям характеристик найденных проектов прогнозировать значения характеристик иницируемого проекта, будем называть базой исторических данных организации.

В ходе создания базы исторических данных для компании «Эксиджен Сервисис», Санкт-Петербургским институтом информатики и автоматизации РАН (СПИ-ИРАН) была предложена и апробирована методология, включающая следующие этапы:

1. Формирование множества источников, в которых может/должна содержаться информация о выполняемых/выполненных проектах, их содержании, особенностях и ходе выполнения – исходное множество источников.
2. Формирование и структуризация множества характеристик, содержащих информацию о выполняемых/выполненных проектах, их содержании, особенностях и ходе выполнения – исходное множество характеристик.
3. Формирование множества выполненных проектов, обеспеченных источниками информации и перечня источников для каждого проекта – исходное множество проектов.
4. Уточнение исходного множества характеристик, исходя из степени их обеспеченности проектной информацией и представления об их значимости для поиска проектов-аналогов – уточненное множество характеристик.
5. Уточнение исходного множества проектов, исходя из степени их обеспеченности информацией уточненного множества характеристик – уточненное множество проектов.

6. Уточнение исходного множества источников, исходя из степени их использования в уточненных множествах проектов и характеристик – уточненное множество источников.
7. Проведение испытаний, подтверждающих достаточность уточненного множества характеристик – рекомендуемые множества источников, характеристик и проектов.

Как показывает практика, в процессе создания баз исторических данных имеют место ряд условий и ограничений, главными из которых являются:

- ◆ неопределенный характер информация о значениях характеристик, описывающих, по мнению лица, проводящего оценку, иницилируемый проект;
- ◆ квалификация лиц, проводящих оценку, достаточно высока – как правило, это специалисты, являющиеся экспертами в решаемых вопросах;
- ◆ размерность задач, решаемых в процессе создания и эксплуатации базы, позволяет опираться на опыт и интуицию эксперта.

Перечисленные условия определили выбор процедур и методов, используемых для обеспечения поддержки создания базы исторических данных. Прежде всего, речь идет о поддержке этапа, связанного с формированием рекомендуемого множества характеристик. Суть исследований, проводимых на данном этапе, сводится к следующему. Из уточненного множества проектов выделяется некоторое «контрольное» множество, для которого в оставшейся базе ищутся проекты-аналоги. Проекты контрольного множества выступают в качестве иницилируемых. Для выбранного «иницилируемого» проекта эксперт задает множество характеристик, которые наиболее полно отражают его особенности. По значениям этих характеристик с помощью формального метода по критерию максимальной близости ансамблей значений формируется множество проектов, выступающих в качестве возможных аналогов иницилируемому. Если, по мнению эксперта, найденные проекты действительно являются аналогами иницилируемого, осуществляется переход к следующему проекту контрольного множества. Если нет, эксперт меняет состав характеристик, либо их значение, и поиск проводится вновь. Итерации завершаются либо нахождением проекта аналога, либо констатацией факта необходимости расширения уточненного множества характеристик. Испытания проводятся до исчерпания контрольного множества. Таким образом, по результатам испытаний уточненное множество характеристик может быть дополнено новыми характеристиками, значения некоторых характеристик могут быть уточнены, часть характеристик отброшена как несущественная. Полученное таким образом множество характеристик рекомендуется обследуемой компании в качестве основы для создания базы исторических данных.

Проведенный анализ показал, что в условиях сформулированных ограничений решение поставленной задачи лежит исключительно в области методов имитационного моделирования, причем в качестве формальной основы используются методы статистических испытаний, теории принятия решений и теории распознавания, а в качестве неформальной – плохо формализуемые знания эксперта.

Разработанная в СПИИРАН система поддержки создания баз исторических данных, далее Система, предназначена для снижения трудозатрат, связанных с их созданием. Система позволяет на основе полученной информации сформировать модель данных для дальнейшей разработки специализированной базы, а также проводить оценку выполнимости проектов, иницилируемых в компании, для которой проводилось исследование. В этом качестве Система служит инструментом, обеспечивающим снижение риска неудачного завершения иницилируемых проектов.

Наряду с прямым назначением Система может использоваться в качестве средства повышения уровня профессиональных знаний сотрудников компании, а также в

учебном процессе, связанном с подготовкой специалистов в области управления разработкой программного обеспечения.

Система ориентирована на пользователя, являющегося экспертом в области управления процессом разработки программных изделий, и обеспечивает привлечение его знаний, плохо поддающихся формализации. Данная особенность Системы обуславливается использованием в ней методологии моделирования на основе алгоритмических сетей, а также методологии принятия решений на основе алгоритмических сетей (прозрачные технологии), разработанных в Санкт-Петербургском институте информатики и автоматизации РАН (СПИИРАН) [8-10].

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

Система реализована в среде MS Excel 2007 на языке Visual Basic for Applications и получила название САМПО (Система АвтоМатизированной ПОддержки).

Система поддерживает все этапы методологии создания базы.

С помощью системы САМПО СПИИРАН совместно с компанией «Эксиджен Сервисис» проведен анализ документации, создававшейся в компании в ходе выполнения проектов. На основании анализа сформированы исходные множества источников (1 этап), характеристик (2 этап) и проектов, введены данные более чем по 280 характеристикам и 70 проектам (3 этап).

На основании анализа введенных значений уточнены множества характеристик (4 этап), проектов (5 этап) и источников (6 этап).

В настоящий момент множество характеристик исследуются на способность решать задачу распознавания проектов-аналогов по заданным значениям характеристик иницируемого проекта (заключительный 7 этап).

### **Выводы**

На сегодняшний день в подавляющем числе стандартных производственных процессов организаций, разрабатывающих программное обеспечение, определено в лучшем случае множество характеристик отчета, формируемого по завершении проекта. Ни по своему составу, ни по смысловой семантике это множество недостаточно для решения задачи оценки ресурсов иницируемого проекта.

Использование организациями общих стандартов технологии разработки ПО определяет целесообразность разработки единой методологии создания баз исторических данных и систем, обеспечивающих её поддержку.

Специфика решаемой задачи обуславливает необходимость использования имитационного подхода и процедур, ориентированных на привлечение плохо формализуемых профессиональных знаний лица, оценивающего характеристики, что, в свою очередь накладывает ограничения на класс используемых формальных методов, на результаты, формируемые ими, а также на сложность операций по переработке информации, имеющей место в процессе оценивания.

Предлагаемая работа является попыткой решить проблему создания баз исторических данных за счет использования формального математического аппарата (теория принятия решений, теория распознавания образов), опирающегося как на опыт организации (документация по выполненным проектам), так и на личный опыт эксперта, осуществляющего оценку (знания эксперта, плохо поддающиеся формализации).

### **Литература**

1. **Гласс Р.** Факты и заблуждения профессионального программирования. Пер. с англ. СПб.: Символ-Плюс, 2007. 240 с.

2. **Брукс Ф.** Мифический человеко-месяц или как создаются программные системы. Пер. с англ. СПб.: Символ-Плюс, 1999. 304 с.
3. **Морозов В. П., Пунтиков Н. И., Тубольцева В. В.** Метод оценивания возможности выполнения проектов разработки программных изделий/XI Санкт-Петербургская международная конференция «Региональная информатика-2008», СПб.: 22–24 октября 2008 года.
4. **Бозм Б. У.** Инженерное проектирование программного обеспечения: Пер. с англ. М.: Радио и связь, 1985. 512 с.
5. **Паулк М., Куртис Б., Хриссис М. Б., Вебер Ч. В., Гарсия С. М., Буш М.** Модель зрелости процессов разработки программного обеспечения – Capability Maturity Model for Software (CMM): Пер. с англ. М.: Богородский печатник, 2002. 256 с.
6. ГОСТ Р ИСО/МЭК 12207-99 ИТ. Процессы жизненного цикла программных средств.
7. **Баранов С. Н., Домарацкий А. Н., Ласточкин Н. К., Морозов В. П.** Процесс разработки программных изделий. М: Наука, Физматлит, 2000. 176 с.
8. **Морозов В. П.** Методология и системы моделирования на основе алгоритмических сетей/X Санкт-Петербургская международная конференция «Региональная информатика-2006», 24–26 октября 2006 года: Труды конференции. СПб.: СПОИСУ, 2007. С. 112–118.
9. **Иванищев В. В., Марлей В. Е.** Введение в теорию алгоритмических сетей. СПб.: Изд-во СПбГТУ, 2000. 179 с.
10. **Морозов В. П.** Системы управления проектами на основе прозрачных технологий/IX Санкт-Петербургская международная конференция «Региональная информатика-2004», 22–24 июня 2004 года: Труды. СПб.: Наука, 2005. С. 255–261.