

УДК 004.5

А. Г. Додонов, А. В. Бойченко

Институт проблем регистрации информации НАН Украины
ул. Н. Шпака, 2, 03113 Киев, Украина

Разработка сценариев аналитической деятельности

Рассмотрены формы и методы аналитической деятельности в корпоративных информационно-аналитических системах, предложен подход к описанию аналитической деятельности на основе сценарного подхода, исследованы варианты сценариев аналитической деятельности, предложена методика формирования сценариев, основанная на пошаговой детализации описания процесса аналитической деятельности.

Ключевые слова: аналитическая деятельность, информационное взаимодействие, корпоративная информационно-аналитические системы.

Введение

В эпоху бурного развития информационного общества задача поиска путей повышения эффективности аналитической деятельности в корпоративных информационно-аналитических системах приобретает чрезвычайную актуальность. Аналитикам приходится оперировать беспрецедентными по своим размерам, многоплановости, динамичности и темпам роста информационными ресурсами, а это заставляет разработчиков информационных систем искать пути к усовершенствованию процессов и технологий обработки данных.

Наличие современных средств хранения, извлечения и визуализации данных само по себе не решает основных задач корпоративной информационно-аналитической системы. Это невозможно без мощной аналитической составляющей, позволяющей выделить из гигантского объема данных, поступающих из разнообразных источников, сведения, необходимые и достаточные для успешного реагирования на поступающие вызовы.

Описание аналитической составляющей должно быть эффективным. Это значит, что расходы на его проведение должны быть наименьшими при оптимальной глубине анализа и его комплексности. Поэтому использование современных методов и средств для повышения эффективности аналитической деятельности является чрезвычайно важным. Для проектирования аналитической деятельности це-

лесообразно использовать методы анализа бизнес-процессов, методы моделирования программного обеспечения, методы сетевого планирования и управления, которые прошли апробацию при выполнении многих сложных проектов [1, 2].

В настоящее время для спецификации аналитической деятельности используется целый ряд различных методик. Наиболее распространенные методики базируются на использовании специально разработанных нотаций [3–5]:

1) UML — унифицированный язык объектно-ориентированного моделирования, используется в парадигме объектно-ориентированного программирования. Является неотъемлемой частью унифицированного процесса разработки программного обеспечения, хотя может использоваться и в других областях системного анализа [2];

2) методология Gantt/PERT (Program Evaluation and Review Technique) — использует последовательную сеточную логику и средневзвешенные оценки продолжительности операций для расчета продолжительности проекта. Многопользовательские диаграммы отображают сетевую модель в графическом виде как множество вершин (узлов), соответствующих работам, связанных линиями, представляющими зависимости предшествования между работами. Этот граф, называемый сетью типа вершина-работа или диаграммой предшествоющего, является наиболее распространенным изображением сети на сегодняшний день, в частности именно такой тип диаграмм поддерживает ряд пакетов программ планирования и контроля выполнения работ. Самой распространенной из таких пакетов является MS Project, который входит в состав MS Office [7];

3) IDEF0 — методология функционального моделирования и графическая нотация, предназначенная для формализации и описания бизнес-процессов. Особенностью IDEF0 является ее акцент на соподчиненность объектов, что значительно облегчает понимание предметной области. В IDEF0 рассматриваются логические, а не временные связи между работами;

4) BPMN (Business Process Modeling Notation) — графическая нотация моделирования деловых процессов. BPMN интересен наличием средств трансляции BPEL/BPEL4WS (Business Process Execution Language for Web Services), на основе которых уже могут функционировать сетевые приложения корпоративной информационно-аналитической системы (КИАС) [1].

В последнее время все большее распространение получает подход к описанию программных комплексов на основе сценариев [8–10].

В работе [6] определяется, что сценарий является описанием конкретных случаев использования системы, связанных во времени, и содержит описание среды, данных (контекста), пользователей (актеров) и реакции системы, с определением условий их начала и завершения. Таким образом, сценарии являются более подробной формой спецификации в сравнении с вариантами использования (Use Cases), которые являются лишь представлением действий пользователя и реакции системы.

Примером может служить MSDL (Military Scenario Definition Language) [11] — нотация, разработанная Simulation Interoperability Standards Organization, которая позволяет создавать как военные сценарии, так и сценарии процесса моделирования. Военный сценарий включает в себя такую информацию как карты, меры контроля, особенности местности, погоды, связи и организации. Кроме того, во-

енный сценарий отражает результаты планирования миссий, в том числе целей плана, сценарий ситуации, и других данных, поддерживающих выполнение матрицы и связанных с ней программ планирования.

В табл. 1 приведены преимущества и недостатки использования сценариев.

Таблица 1.

Преимущества	Недостатки
<ul style="list-style-type: none"> • Сценарии являются интуитивно понятным и удобным способом учета требований всех заинтересованных сторон, таких как дизайнеры и пользователи. • Сценарии особенно хорошо подходят для оперативного описания реактивных и телекоммуникационных систем. • Сценарии хорошо приспособлены к использованию в итерационной модели разработки. • Сценарии могут быть абстрагированные от основной структуры системы, если это необходимо. • Сценарии являются наиболее удобными для документации и коммуникации. • Сценарии могут быть использованы для автоматизированной генерации тестов для проверки на различных уровнях (спецификация, проектирование и внедрение). • Сценарии могут использоваться для разработки более подробных моделей и реализаций. 	<ul style="list-style-type: none"> • Поскольку сценарии являются разрозненными, трудно провести оценку полноты набора сценариев, особенно если сценарии не описаны на едином уровне абстракции. • Сценарии не в состоянии выразить нефункциональные требования. • Синтез компонентов сценариев остается сложной проблемой. <p>Использование сценариев приводит к обычным проблемам, связанным с отслеживанием с помощью других моделей, используемых в разработке процесса.</p> <ul style="list-style-type: none"> • Современные подходы, основанные на сценариях, еще не имеют достаточно высокого уровня зрелости.

Модель сценария аналитической деятельности

Под сценарием аналитической деятельности (далее — САД) будем понимать определенную последовательность действий, которую выполняет аналитик или группа аналитиков при решении аналитических задач в составе корпоративной информационно-аналитической системы.

Структуру САД можно получить путем последовательной декомпозиции САД вплотную к уровню микроопераций (МО). На нижнем уровне детализации выделяются МО, отвечающие наименее детализированным элементам аналитической деятельности, которые можно отобразить в модели.

Выделим следующие элементы САД (рис. 1):

- цель выполнения сценария;
- задание (программа) выполнения сценария;
- информационные ресурсы;
- программно-технические ресурсы;
- начало (инициализация) САД;
- последовательность исполняемых операций (собственно сценарий);
- аналитический результат выполнения САД.



Рис. 1 Модель сценария аналитической деятельности

Исполнители сценария, в свою очередь, характеризуются полученными ролями. В [1] приводятся основные роли, используемые при формировании САД:

- руководитель;
- руководитель группы;
- аналитик;
- эксперт;
- техник;
- секретарь;
- исполнители.

Следует выделить 2 вида начала САД:

- плановый;
- инициализированный в связи с возникновением проблемной ситуации.

Примером проблемной ситуации могут выступать: стихийное бедствие, техногенная катастрофа, экономический кризис, эпидемия [1].

Программа выполнения сценария может быть описана в виде алгоритма, и в свою очередь, включает выполнение определенных элементарных операций.

Операция сценария аналитической деятельности (осадка) — спланированное действие, направленное на выполнение определенной элементарной функции. Типовые операции, которые выполняются в КИАС, приведены в табл. 2.

Может быть использована следующая классификация САД.

Типы САД:

- организационно-технический;
- информационно-технический;
- информационно-аналитический;
- системно-аналитический.

Таблица 2.

№№	Код	Название	Содержание
1	СИ	Сбор информации	Запрос, прием по каналам связи, размещение в хранилище
2	ПОД	Первичная обработка данных	Выборка, группирование, оформление таблиц, графиков, диаграмм
3	АД	Анализ данных	Статистическая обработка, построение зависимостей, расчет прогнозных значений
4	ЭД	Экспертиза данных	Проверка отсутствия противоречий принятым фактам
5	ИД	Интерпретация данных	Выполнение аналитического исследования и оформление отчетного документа
6	ПЛ	Планирование	Составление, рассмотрение и оформление планов
7	ОТ	Отчетность	Составление, рассмотрение и оформление отчетов
8	К	Контроль	Контроль выполнения планов, отчетов, распоряжений, указаний, документов
9	ВР	Выработка сложных решений	Изучение проблемы, организация коллегиальной работы, принятие решений
10	ПД	Подготовка документов	Разработка концепции сложного документа, подготовка проекта, согласование, рассылка адресатам

Способ организации САД:

- локальный;
- распределенный.

Субъект выполнения САД:

- индивидуальный;
- групповой;
- кооперативный;
- массовый.

Объект выполнения САД:

- обработка документа;
- информирование руководства;
- консультирование подчиненных;
- коллегиальное обсуждение.

Возможность формализации:

- операции САД не подлежат формализации;
- частично формализуемые;
- полностью формализуемые.

Вид результата:

- предоставление исходных данных;
- формирование концепции решения;
- разработка проекта решения;
- экспертное заключение;
- участие в коллективном решении;
- редактирование исходного документа.

САД можно представить как

$$S = \{I, Z, O, R, A, T\},$$

где I — инициатор операции; Z — субъект операции; O — объект сделки; R — получатель результата; A — порядок выполнения операции; R — результат операции; T — срок выполнения операции.

Каждый САД можно разложить на ряд микроопераций (элементарных операций), которые имеют следующие особенности:

- четкую определенность постановки задачи, метода ее решения;
- наличие технологических альтернатив решения;
- ситуационность использования.

Пример МО:

- запросить информацию;
- предоставить информацию;
- предложить вариант решения;
- проанализировать вариант решения;
- сравнить два информационных объекта;
- сравнить с тенденцией в прошлом;
- экстраполировать тенденцию развития событий;
- прогнозировать тенденцию развития событий на отдаленную перспективу;
- оценить последствия принятия решения;
- оценить риск
- отложить принятие решения;
- сообщить решение.

Функционирование САД предполагает наличие определенной среды, которое включает:

- лингвистическое обеспечение;
- информационное обеспечение;
- аппаратно-программное обеспечение;
- программное обеспечение;
- нормативно-правовое обеспечение;
- организационное обеспечение.

При моделировании САД результатом является определенная операция. Таким образом, сделка является основным элементом деятельности на самом нижнем уровне детализации, на выполнение которого требуется время, и который может задержать начало выполнения других операций. Момент окончания операции означает выполнение определенного САД. Каждая операция имеет сроки выполнения и требует назначения ресурсов для своего исполнения.

Методика формирования сценария аналитической деятельности

Предлагается следующая последовательность формирования сценариев:

- анализ существующих элементов КИАС;
- описание задач аналитической деятельности;
- определение исполнителей САД;

- построение списка отдельных операций;
- назначение исполнителей для каждого САД;
- задание для всех операций основных параметров (продолжительность, сроки выполнения);
- определение взаимной подчиненности вхождения отдельных операций САД;
- определение зависимостей предшествования операций (логических зависимостей);
- назначение ресурсов для выполнения САД;
- моделирование выполнения САД;
- анализ результатов моделирования.
- в случае необходимости возврат к первому шагу.

Программная составляющая среды для формирования совокупности САД представлена на рис. 2.

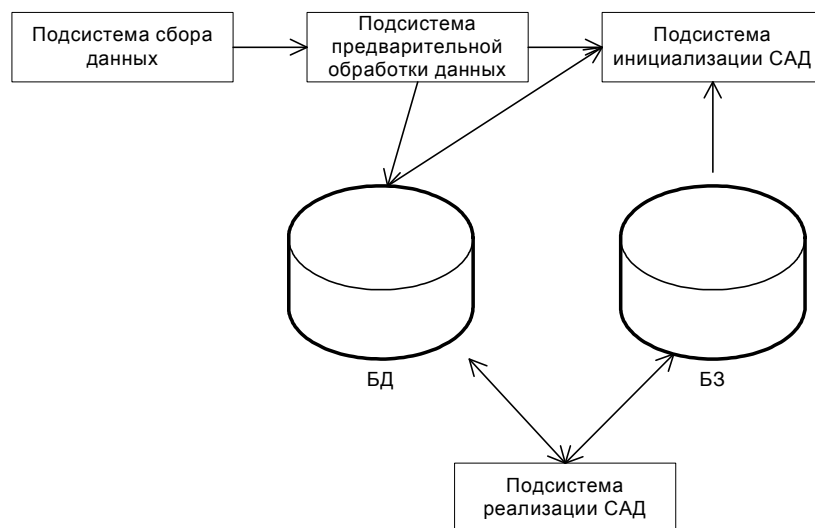


Рис. 2. Структура среды формирования совокупности САД

Подсистема сбора данных обеспечивает получение данных из различных источников в форме: доклада, сообщения, аналитического обзора, отчета, справки, сведения, телеграммы, телефонограммы.

Подсистема предварительной обработки данных строится, на базе лингвистического процессора (ЛП), выполняющего роль посредника между оператором и массивом данных, с которого необходимо получить нужную для активации САД информацию. Задачей ЛП является превращение полученных данных на набор семантических структур, является формальным отражением «содержания» полученной информации. Цель такого преобразования — обеспечить исходные данные для работы при запуске САД и наполнения БД.

Подсистема запуска САД обеспечивает обработку входящей информации, информации из БД и базы знаний и осуществляет запуск САД. Системы инициализации САД можно классифицировать по методам анализа входящей информации, периодичностью работы, реакцией, централизацией.

Подсистема реализации САД обеспечивает все необходимые средства для выполнения САД.

Процесс реализации САД можно рассматривать как ряд отдельных, но связанных между собой процессов:

- процесс запуска (инициации (активации)) САД;
- процесс планирования — определение и совершенствование целей и отбор лучшего из альтернативных планов действий, ведущих к достижению целей САД;
- процесс исполнения — координация использования человеческих и других ресурсов для выполнения САД;
- процесс контроля — регулярный мониторинг и оценка хода выполнения работ для определения отклонений от плана, при необходимости, корректирующих действий;
- процесс завершения — формальное принятие САД и оформление завершения должным образом.

Среда формирования САД может быть представлена в виде следующего алгоритма (рис. 3).

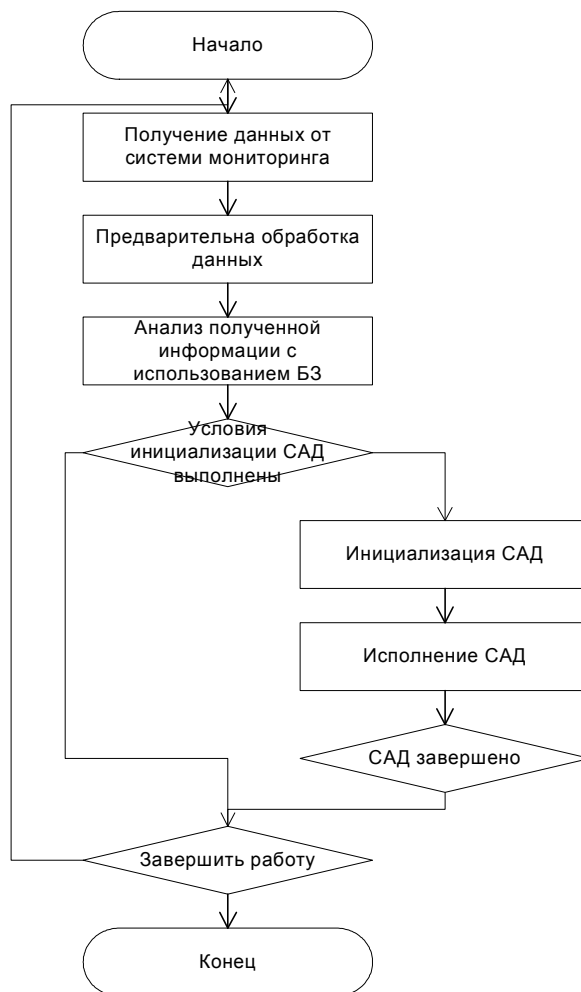


Рис. 3. Алгоритм формирования САД

Завершение работы алгоритма происходит в случае реализации среды формирования САД периодического типа.

Моделирование аналитической деятельности

Моделирование включает следующие этапы:

- подготовка САД;
- определение АРМов и их функций для выполнения задачи;
- передача требований по выполнению задачи группе администрирования и управления моделирующим комплексом (МК);
- конфигурирование моделирующего комплекса и передача постановки задачи функциональным подсистемам;
- создание внештатных ситуаций;
- мониторинг работы;
- реализация алгоритмов моделирования и выдача результатов;
- выполнение комплекса задач;
- сохранение результатов моделирования;
- анализ результатов моделирования;
- составление отчетной документации по результатам моделирования.

Перед началом выполнения сценария администратор моделирующего комплекса должен выполнить следующие действия:

- конфигурирование функционального программного обеспечения — формирование профиля функционального ПО для каждого отдельного АРМа (набор задач, функций, процедур) из множества функциональных программных модулей (сервисов);
- назначение прав доступа к информационным ресурсам и ролей пользователей (должностных лиц), участвующих в решении комплексной задачи на МК;
- загрузку программного обеспечения в АРМы должностных лиц, участвующих в решении комплексной задачи;
- конфигурирование комплекса технических средств (серверов и АРМов должностных лиц), БД и синхронизация времени на всех АРМах моделирующего комплекса;
- выбор задачи (из списка комплексных задач) и ее активация;
- управление процессом моделирования комплексной задачи (пуск, прерывание, обмен сообщениями, ожидание подтверждения о получении/отправке сообщений);
- протоколирование действий АРМов должностных лиц в процессе решения комплексной задачи (действия операторов, системные нарушения, обмен сообщениями, нештатные ситуации);
- анализ журналов отчетности с результатами выполнения процедур и функций в процессе решения комплексной задачи (действия операторов, системные журналы, обмен сообщениями, нештатные ситуации).

Информация о типах параметров и результатах моделируемых операций уточняется таким образом, чтобы позволить сгенерировать значения параметров для простейших сценариев нормального использования этих операций и выполнить некоторые (далеко не полные) проверки корректности их результатов. По-

сколькx одни и те же типы данных в больших системах используются многократно, массовая генерация тестов может привести к существенному снижению трудозатрат на каждый тест.

Введенная информация должна пройти проверку на корректность. Такая проверка выполняется как за счет ее просмотра и дополнительного анализа, проводимого другими разработчиками, так и с помощью отладки сгенерированных тестов — они должны компилироваться и собираться, а все проблемы, возникающие при их выполнении, должны быть следствием ошибок в моделируемой системе, а не в тестах.

При конфигурировании МК результатом является определенное множество настроек. Отдельная настройка является основным элементом деятельности на самом нижнем уровне детализации, на выполнение которого требуется время, и который может задержать начало выполнения других операций. Момент окончания настроек означает выполнение данного этапа.

Роль и права доступа АРМов пользователей (должностных лиц МК) к информационным ресурсам устанавливает администратор МК перед началом моделирования на основании сценария. Топология и состав АРМ МК для решения комплексной задачи формируется в соответствии с должностными обязанностями и штатным расписанием.

Перед проведением конфигурирования необходимо упорядочить моделируемые операции и типы данных, связанные с ними так, чтобы операции, имеющие более сложные типы параметров, шли позже тех, которые имеют простые параметры и могут быть использованы для получения данных более сложных типов. При этом уточнение выполняется от простых операций к более сложным без необходимости часто переключаться на анализ других операций и может естественным образом многократно использовать на поздних этапах информацию, выявленную на ранних этапах.

Для реализации сценариев необходимо наличие:

- соответствия МК требованиям спецификации;
- загрузки функциональных программных модулей для каждого АРМа;
- конфигурации МК для выполнения обработки исходных данных моделирования;
- установки прав доступа пользователей АРМов и программных приложений к информационным ресурсам БД МК;
- подготовки операторов, участвующих в организации и проведении процесса моделирования;
- регламента информационного взаимодействия, обеспечивающего решение поставленной задачи.

Инструментом автоматизации разработки САД должна быть подсистема управления сценариями аналитической деятельности КИАС. Задачами этой подсистемы являются [1]:

1) генерация экранных форм интерфейса КИАС, которые являются отражением сущностей функциональности сценариев аналитической деятельности. Эти экранные формы отражают как запросы на получение результатов аналитической обработки информационных ресурсов КИАС, так и сами отображения результатов аналитической обработки в удобном для принятия решений виде;

2) генерация цепи соответствующих узлов обработки и преобразования данных КИАС (сервер(а) БД, сервер документооборота, сервер баз знаний, серверы приложений, Web-сервер, ГИС-сервер и т.д.), реализующих технологии и методы, обеспечивающие выполнение сценариев, и информационные технологии, с помощью которых реализуются механизмы взаимодействия между структурными компонентами КИАС.

Таким образом, подсистема разработки САД реализует требования к функциональности сценариев аналитической деятельности и генерирует подмножество узлов обработки и преобразования данных КИАС с обеспечением их взаимодействия, необходимых для реализации сценариев.

Выводы

Работа посвящена вопросам разработки методики описания, анализа и генерации сценариев аналитической деятельности на базе современных технологий моделирования бизнес-процессов.

Целью применения данной методики является улучшение качества и уменьшение стоимости и времени разработки аналитической составляющей современных корпоративных информационно-аналитических систем.

В работе исследованы основные требования к описанию аналитической деятельности и особенности ее реализации с применением современных вычислительных средств.

Разработана модель описания сценария аналитической деятельности, описывающая разработчиков САД, исполнителей, начало его выполнения, программу выполнения сценария, ресурсы, используемые при выполнении сценария (информационные, технические, вычислительные и др.) и результат выполнения САД.

Предложена методика формирования САД, основанная на пошаговой детализации описания процесса аналитической деятельности, позволяющая структурировать этапы САД на основные компоненты и обеспечить направленность деятельности на достижение всего комплекса целей аналитической деятельности.

1. *Андерсон Б.* Бизнес-процессы. Инструменты совершенствования / Б. Андерсон // РИА «Стандарты и качество». — М., 2007.

2. *Бойченко А.В.* Засоби аналізу макроекономічних показників / А.В. Бойченко, В.Р. Сенченко // Реєстрація, зберігання і оброб. даних. — 2005. — Т. 7, № 1. — С. 43–51.

3. *Додонов О.Г.* Методологія побудови корпоративних інформаційно-аналітичних систем / О.Г. Додонов, О.В. Коваль, В.Р. Сенченко // Реєстрація, зберігання і оброб. даних. — 2007. — Т. 9, № 4. С. 60–75.

4. *Додонов О.Г.* Методологія створення Національного реєстру електронних інформаційних ресурсів / О.Г. Додонов, О.В. Нестеренко, А.В. Бойченко // Реєстрація, зберігання і оброб. даних. — 2005. — Т. 7, № 3. — С. 88–97.

5. *Формування, інтеграція та використання інформаційних ресурсів органів державної влади / Додонов О.Г., Нестеренко О.В., Бойченко А.В., Бойченко О.А.* // Реєстрація, зберігання і оброб. даних. — 2002. — Т. 4, № 3. — С. 69–75.

6. *Плэтт В.* Добывание, анализ и обработка стратегических разведывательных данных / В. Плэтт. — М.: Эт Сетера Паблшинг. — 2006. — 288 с.

7. *Сенченко В.Р.* Питання побудови моделі функціональної спеціалізації експертів інформаційно-аналітичних систем / В.Р. Сенченко // Реєстрація, зберігання і оброб. даних. — 2003. — Т. 5, № 4. — С. 51–63.

8. *Andreas Tolk.* A Systems Engineering Process Supporting the Development of Operational Requirements Driven Federations [Электронный ресурс] / Andreas Tolk, Thomas G. Litwin, Robert H. Kewley. — Режим доступа: www.informs-sim.org/wsc08papers/155.pdf. — Проверено 05.12.2010.

9. *Hillel Kugler.* Testing Scenario-Based Models [Электронный ресурс] / Hillel Kugler, Michael J. Stern, E. Jane Albert Hubbard. — New York University. — New York, NY (USA). — Режим доступа: <http://www.cs.nyu.edu/~kugler/fase07.pdf>. — Проверено 05.12.2010.

10. *Shahen Naidoo.* Standardisation of Command and Control Systems and Simulator Interfaces, and the Definition of Scenarios. Council for Scientific and Industrial Research [Электронный ресурс] / Shahen Naidoo, Willem H. le Roux. — Режим доступа: http://researchspace.csir.co.za/dspace/bitstream/10204/3150/1/Naidoo_2007.pdf. — Проверено 05.12.2010.

11. *Simulation* Interoperability Standards Organization (SISO). — Standard for Military Scenario Definition Language (MSDL): SISO-STD-007-2008. — October 14, 2008.

Поступила в редакцию 07.12.2010