УДК 004.8:004.9

Исаков Степан Алексеевич, студент магистратуры,
ФГБОУ ВО «Амурский государственный университет»

Научный руководитель:
Ерёмин Илья Евгеньевич, д-р техн. наук, профессор,
ФГБОУ ВО «Амурский государственный университет»

АЛГОРИТМ АВТОМАТИЗИРОВАННОГО ПОСТРОЕНИЯ И ВАЛИДАЦИИ МОДЕЛЕЙ БИЗНЕС-ПРОЦЕССОВ В НОТАЦИИ BPMN 2.0

Аннотация. В статье рассматривается алгоритм автоматизированного построения моделей бизнес-процессов в нотации BPMN 2.0 на основе текстовых регламентов с применением больших языковых моделей. Предложен двухуровневый подход к обработке данных: семантическое извлечение сущностей языковой моделью отделено от синтеза диаграммы через промежуточную каноническую модель процесса. Разработан контур детерминированной валидации графовой структуры, позволяющий выявлять логические тупики, разрывы связей и некорректные шлюзы. Внедрение алгоритма обеспечивает синтаксическую корректность моделей и сокращает трудоемкость работы аналитиков.

Ключевые слова: Моделирование бизнес-процессов, нотация BPMN 2.0, большие языковые модели, промежуточная модель, валидация графов, автоматизированное проектирование.

Проектирование и регулярная актуализация моделей бизнес-процессов составляют основу системного анализа при внедрении и развитии корпоративных информационных систем. Традиционно построение процессных схем выполняется аналитиками вручную на базе регламентов, инструкций и нормативных документов компании. Данный этап характеризуется высокой трудоемкостью: по оценкам специалистов, до шестидесяти процентов предпроектного времени расходуется на формализацию и ручную отрисовку графических моделей [1]. При этом ключевой сложностью остается существенный семантический разрыв между естественным языком текстовых описаний и строгой формальной спецификацией процессных стандартов.

Развитие методов машинного обучения и больших языковых моделей открыло практические возможности для автоматизированного извлечения сведений из неструктурированной документации. Тем не менее попытки прямой генерации исполняемых файлов стандарта BPMN 2.0 с помощью нейросетей сталкиваются с принципиальным ограничением: вероятностный характер генерации последовательностей текста неизбежно приводит к искажению топологии графа. Модели формируют фиктивные идентификаторы, допускают пропуски промежуточных шагов, создают изолированные узлы и неразрешимые логические тупики, что делает итоговый файл непригодным для верификации и практического использования. Стандарт BPMN 2.0 опирается на формальный аппарат системного анализа и теории направленных графов [2], требующий строгой математической и структурной корректности.

Для преодоления обозначенной проблемы разработан алгоритм, основанный на разделении процесса синтеза диаграмм на два взаимосвязанных уровня. На первом уровне осуществляется лингвистический анализ естественно-языкового описания, в рамках которого выявляются операции, ответственные исполнители, развилки и граничные условия. На втором уровне формируется инженерная модель процесса, где каждое действие преобразуется в

типизированный элемент, сопоставляется с конкретной дорожкой исполнителя и включается в топологическую структуру ориентированного графа потоков управления.

Связующим звеном выступает промежуточная каноническая модель. Текст регламента не транслируется в графический формат напрямую: предварительно строится нормализованное описание последовательностей, связей и условий ветвления в машинно-читаемой форме. Каждому шагу присваивается идентификатор, наименование операции, роль исполнителя и ссылки на смежные узлы. Принципиальной особенностью промежуточного формата является сохранение сквозной привязки созданных элементов к номерам исходных предложений регламента, что обеспечивает полную прозрачность и возможность верификации логики человеком [1].

В рассматриваемом алгоритме роль нейросетевых компонентов ограничена извлечением семантических конструкций из исходного текста: выявляются субъекты действий, целевые операции и условия ветвления. Полученные фрагменты преобразуются в нормализованные записи промежуточной модели по заданной схеме вывода. При обнаружении в тексте неоднозначных формулировок или неполных указаний на исполнителей система регистрирует соответствующее предупреждение, исключая произвольную генерацию несуществующих параметров процесса.

Сформированное промежуточное представление преобразуется в топологический граф и передается в контур детерминированной валидации. Проверка осуществляется последовательно по четырем направлениям (рисунок 1): контроль структуры промежуточных данных, анализ связности графа процесса, проверка синтаксических правил BPMN и верификация параметров геометрической раскладки.

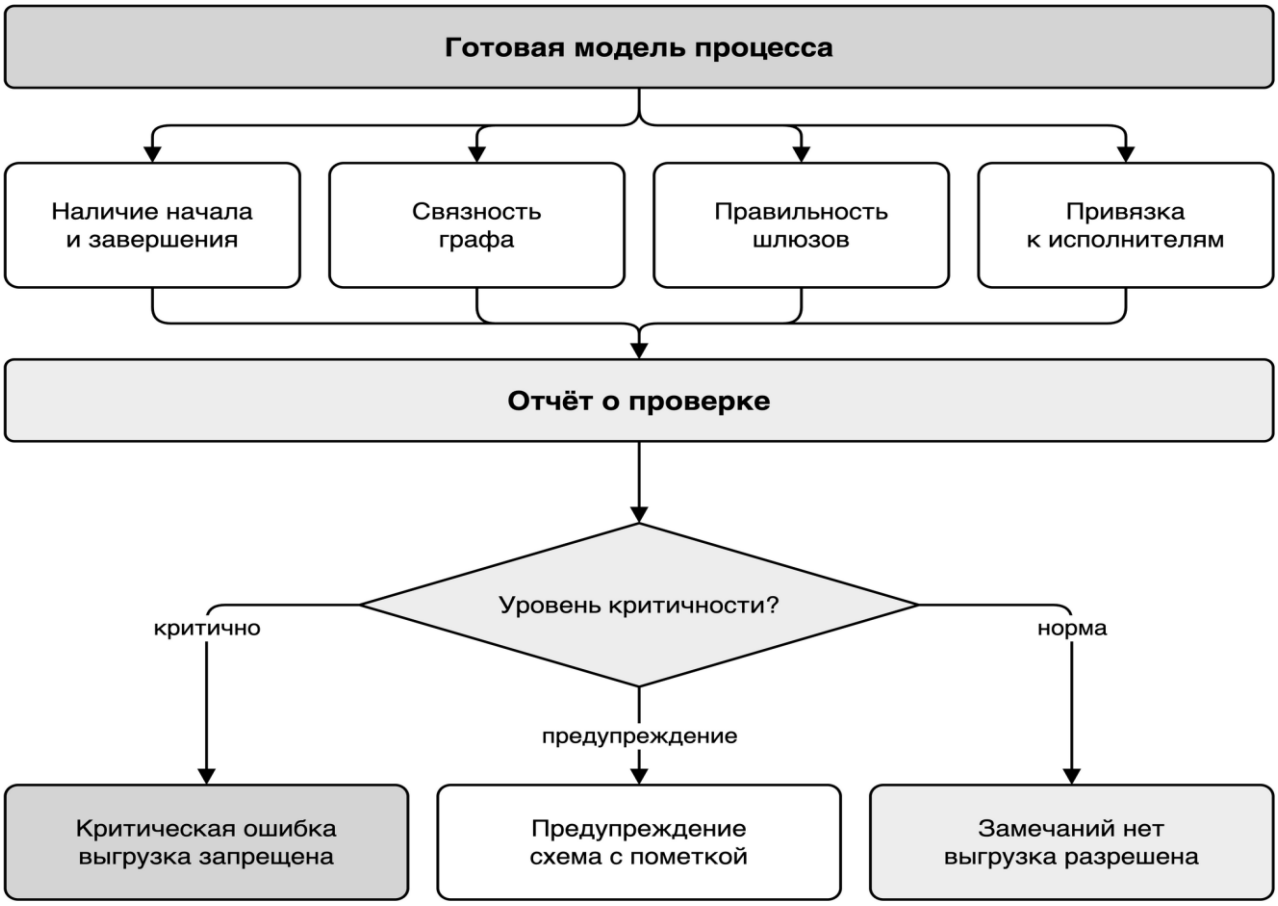


Рисунок 1. Контур структурной валидации и коррекции процессной модели

Графовый анализ модели базируется на обходе вершин от начальной точки для подтверждения фундаментальных свойств корректности структуры [2]. Контур контролирует обязательное наличие стартового события и корректное завершение всех ветвей конечными событиями, проверяет направленную достижимость каждой операции (что исключает появление изолированных узлов и смешение независимых процедур), а также валидирует правила построения шлюзов. Для исключающих развилок проверяется наличие не менее двух исходящих связей с сопоставленными условиями перехода, а для параллельных структур подтверждается корректность ветвления и слияния потоков управления.

Обнаруженные несоответствия дифференцируются по четырем градациям: критические дефекты (отсутствие начального события или ошибочные ссылки, блокирующие экспорт), структурные нарушения (недостижимые операции или незамкнутые ветви, требующие уточняющего запроса), предупреждения (неявное определение ролей или связывание операций исключительно по порядку следования предложений) и служебные уведомления (нормализация наименований и пространственное выравнивание элементов). При фиксации структурных нарушений алгоритм запускает процедуру локализованной обратной связи: формируется целевой повторный запрос с указанием конкретного логического разрыва и фрагмента исходного текста, что исключает необходимость пересчета всего процесса.

После подтверждения топологической корректности выполняется детерминированная трансляция в формат BPMN 2.0 XML. Генератор формирует структуру файла, включающую как семантическое описание процесса, так и графические координаты каждого символа. Для каждой выделенной роли создается отдельная горизонтальная дорожка (рисунок 2). Координаты элементов рассчитываются в соответствии с порядковыми уровнями графа, а линии связей маршрутизируются без взаимных пересечений, обеспечивая наглядность восприятия диаграммы аналитиком.

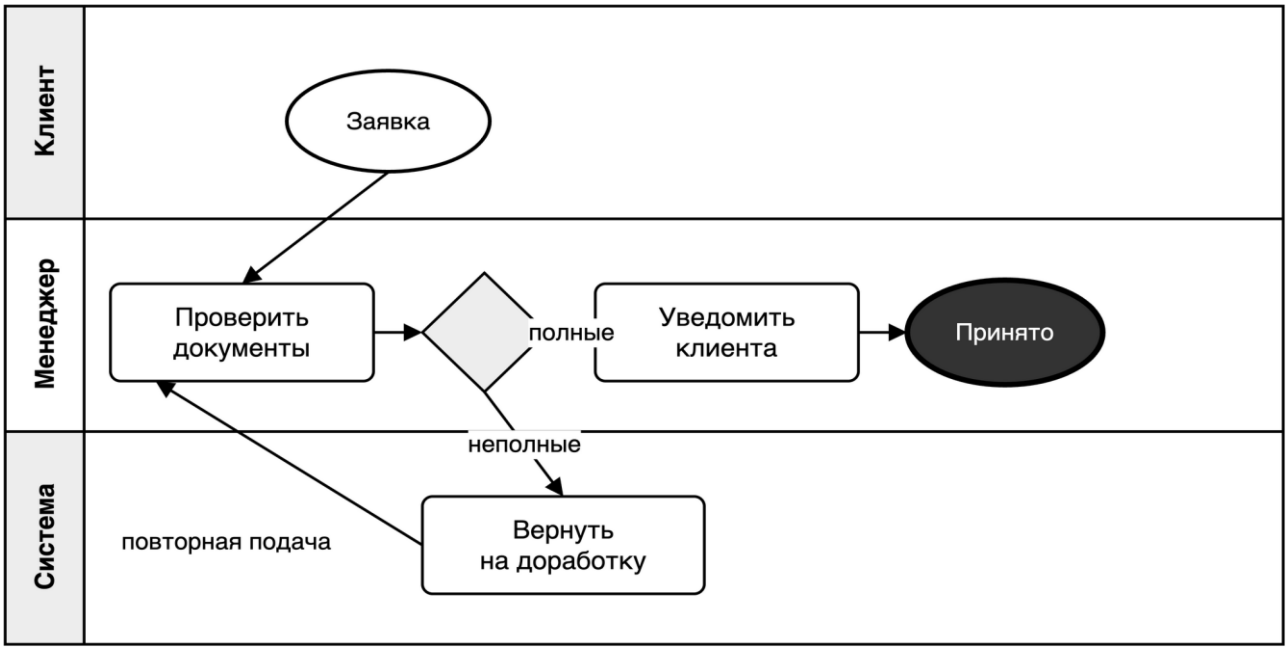


Рисунок 2. Форма построенной модели процесса в нотации BPMN 2.0

Разработанный алгоритм обеспечивает автоматизацию ключевого этапа предпроектного анализа, объединяя возможности современных языковых моделей по обработке естественного языка со строгой верификацией на основе методов системного и

графового анализа [2]. Предложенные решения составляют теоретическую и практическую основу диссертационного исследования и могут использоваться при разработке отечественных инструментальных средств проектирования программного обеспечения.

Список литературы:

1. Моделирование бизнес-процессов: учебник и практикум / под ред. О. И. Долгановой. М.: Юрайт, 2019. 289 с.
2. Моделирование систем и процессов: учебник / под ред. В. Н. Волковой, В. Н. Козлова. М.: Юрайт, 2019. 450 с.