

Н. В. Васильев

К.т.н., начальник сектора, ПАО «Интелтех»,

О. В. Забродин

Инженер, ПАО «Интелтех»

Д. В. Куликов

Инженер, ПАО «Интелтех»

МЕТОД PROCESS MINING ДЛЯ АНАЛИЗА РЕГЛАМЕНТОВ ДОКУМЕНТООБОРОТА

АННОТАЦИЯ. Предложен метод анализа журналов событий системы электронного документооборота, основанный на методологии *Process Mining* (глубинный анализ процессов), позволяющий осуществить реконструкцию проекции потока управления, проекции ресурсов и проекции данных рабочих процессов обработки документов на предприятии. Метод основывается на анализе журналов действий над документами. Описывается компонент анализа в составе редактора процессов, реализующий приведенные методы реконструкции проекций моделей процессов.

КЛЮЧЕВЫЕ СЛОВА: анализ процессов, документооборот, бизнес-процесс, принятие решений, альфа-плюс алгоритм, жизненный цикл документа, анализ журналов событий

Введение

В современном мире подавляющий объём информации существует в электронном виде. Для хранения, обработки и управления документами на цифровых носителях широкое распространение получили системы электронного документооборота (СЭД), представляющие собой автоматизированные многопользовательские системы, сопровождающие процесс создания и перемещения по организации документов.

Большинство современных СЭД строится на основе процессного подхода, в соответствии с которым документооборот промышленного предприятия или государственного ведомства представляется в виде формализованного множества описаний последовательности-выполняемых сотрудниками операций над документами.

Однако, внедрение подобной системы приводит к чрезмерному увеличению нагрузки на аналитиков и сотрудников служб обеспечения. Это связано с необходимостью формализации процессов движения документов на предприятии. Особенно сложной эта задача становится при изменении структуры организации, штат-

ной численности или при переориентации деятельности предприятия. Вследствие описанных структурных изменений имеющаяся модель процессов теряет актуальность. Порой степень несоответствия модели процессов приводит к необходимости разработки моделей процессов «с нуля».

Актуальность темы работы обусловлена тем, что в настоящий момент в составе СЭД отсутствуют средства, позволяющие автоматизировать процесс создания и корректировки моделей процессов обработки документов на предприятии.

Большинство современных СЭД поддерживают журнализацию действий пользователей. В последние годы в зарубежных [1] и отечественных [2] работах получило развитие направление, носящее название «Глубинный анализ процессов» (*Process Mining*), которое позволяет на основе журналов событий информационных систем реконструировать схемы рабочих процессов (*work flow*), реализуемые пользователями. Однако в настоящее время данные средства в системах электронного документооборота пока еще не получили широкого распространения.

В работе предпринята попытка разработки инфраструктуры глубинного анализа процессов в структуре реализованной коллективом авторов СЭД «Цера». Предложенные модели и методы были реализованы в прототипе модуля, позволяющего помочь интеграторам СЭД при решении задачи создания и актуализации бизнес-процессов документооборота предприятия.

В статье приводится общая характеристика современных систем документооборота, приведен разработанный метод глубинного анализа процессов в СЭД и описан разработанный прототип модуля анализа процессов, реализующий предложенный метод.

1. Характеристика современных систем электронного документооборота

За последние 20 лет концепция электронного документооборота получила свое развитие от идеи сканирования и централизованного хранения графических образов документов до идеи управления документами и их карточками включая создание, регистрацию, подписание и сдачу в архив. Необходимость решения задачи маршрутизации документов внутри организации между исполнителями привела к внедрению в СЭД технологии рабочих (бизнес) процессов (БП). Русский термин «системы электронного документооборота» является некорректным, так как основным объектом хранения СЭД выступают не документы, а регистрационно-учетные карточки.

Документ при этом может храниться в базе данных СЭД, файловой системе или в бумажном виде в деле. В этом отношении англоязычный термин *EDRMS (Electronic Document Record Management Systems)* является более правильным.

Современные СЭД условно можно разделить на 2 класса: документ — ориентированные (*doc-flow*) и процесс — ориентированные (*workflow*) СЭД. Основой как первого так и второго класса СЭД является подход на основе рабочих процессов. Пример автоматизации деятельности сотрудников на примере процесса согласования документа представлен на рис. 1.

Организация рабочих процессов в *doc-flow* — СЭД идет от документа. Для каждого документа, прошедшего систему, создается задача (экземпляр рабочего процесса). Рабочие процессы имеют в *doc-flow* — СЭД *общий характер*, и, как правило, соответствующее название «Согласование», «Утверждение», «Рассмотрение». В то время как в *workflow* — СЭД с одним экземпляром рабочего процесса могут быть связаны несколько документов. Процессы в этом случае носят специфический характер и название: «Обслуживание заявки на подключение клиента», «Проведение сделки», «Аттестация персонала» и пр. Безотносительно типа СЭД задачи размещаются на сервере баз данных. Задачи характеризуются статусом (выполнена/в процессе/просрочена), прикрепленными документами, маршрутом движения документа, списком пользователей-исполнителей, а также временными параметрами.

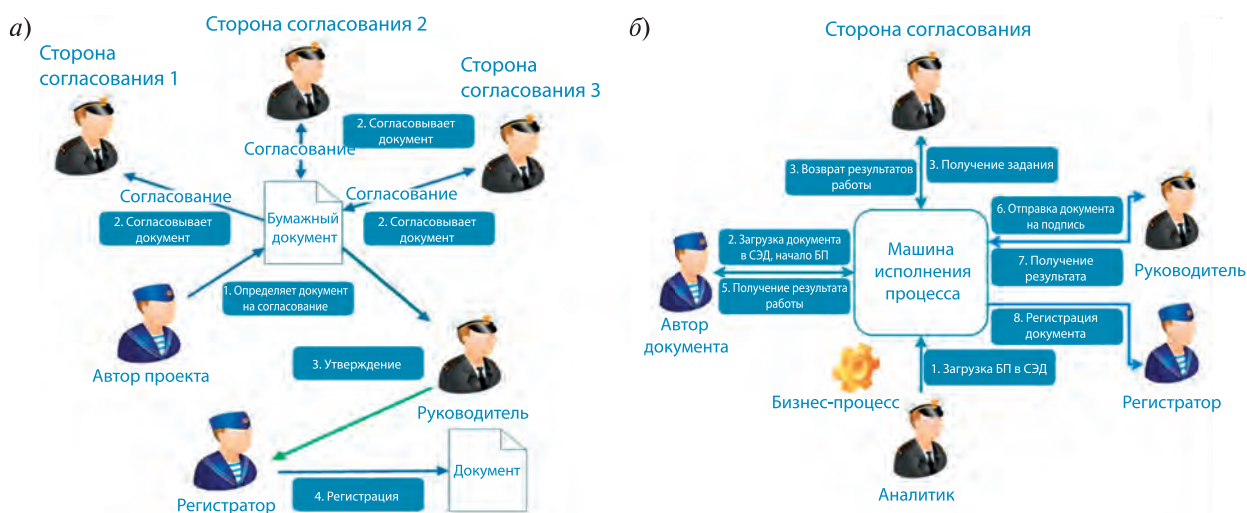


Рис. 1. Пример согласования документа в ручном режиме (а) и средствами системы электронного документооборота (б)



Рис. 2. Проекция модели рабочего процесса

Как правило, в СЭД каждый документ характеризуется типом, а тип в свою очередь моделью жизненного цикла. Жизненный цикл определяет, какие стадии и в каком порядке может проходить документ. Например, практически все документы проходят стадии разработки, согласования и утверждения, а также списания в архив. Для специфических типов документов могут выделяться специфические стадии.

Современные СЭД (см. [3, 4]) строятся на основе реляционных баз данных. Более близкая к СЭД концепция документ — ориентированных *NoSQL* — баз данных не получила пока широкого распространения. Идеологически, в составе СЭД можно выделить набор сервисов:

сервис справочников, предназначенный для хранения условно-постоянной информации, используемой пользователями СЭД при работе с документами;

сервис пользователей, предназначенный для управления пользователями и разграничения прав доступа. Он отвечает за авторизацию и аутентификацию пользователей системы по доступу к папкам и файлам, для чего использует как механизмы клиентской составляющей системы, так и встроенные механизмы безопасности базы данных;

сервис поиска и индексации, предназначенный для реализации механизмов полнотекстового поиска. Сервис производит периодическую индексацию таблиц документов и справочников с сохранением индекса в специальной таблице или на диске;

объектные сервисы, реализующие базовые операции создания, чтения, обновления и удаления (*CRUD*) над объектами, включая документы, справочники, задачи и задания рабочих процессов. Для объектов — документов это включает установку и получение свойств и потоков содержимого файлов;

сервис каталогов, осуществляющий доступ к объектам, размещаемым в иерархии папок путем добавления или удаления объектов из папки. Папки могут содержать другие папки и документы;

сервис исполнения рабочих процессов, служащий для создания новых экземпляров рабочих процессов документооборота, генерации и назначения заданий исполнителям, поддержки и контроля выполнения созданных экземпляров. Данный сервис использует для своей работы все перечисленные выше сервисы.

Как показано на рис. 2 описание рабочего процесса документооборота может быть представлено как набор из 4 проекций-перспектив.

Перспективе «*управление потоком*» соответствует маршрут движения документа между исполнителями (схема рабочего процесса).

Перспективе «*данные*» соответствует документ, над которым выполняется экземпляр процесса, а также набор дополнительных переменных процесса (переменных управления).

Перспективе «*ресурсы*» соответствует набор ролей и исполнителей, которые могут выполнять действия над документом в узлах схемы рабочего процесса.

Перспективе «*операции*» соответствует список элементарных действий, совершаемых исполнителями с документом в рамках задания. Например, скачать документ, подписать документ, создать новую версию, перенести в другую папку и пр.

В существующей схеме создание и загрузку рабочих процессов в СЭД в виде файлов *BPMN* (*Business process management notation*) производит администратор с помощью редактора (см. [5]). Подобная схема в силу субъективности и неточности имеет недостатки. В работе далее предлагается метод, который позволяет частично их устранить.

2. Метод глубинного анализа процессов в системе электронного документооборота

Привлечение глубинного анализа вносит в привычную схему развертывания бизнес-процессов документооборота коррективы. На пред-

приятие ставится система с минимально необходимым набором процессов и пользователи выполняют привычные действия в ручном режиме. Например, при согласовании документа вручную указывают все согласующие инстанции, которые должен пройти данный документ. В процессе согласования факты выполнения всех действий заносятся в журнал. После обработки нескольких однотипных документов журнал становится «полным», что позволяет реконструировать предполагаемый процесс обработки документа. То есть после рецензии полученного процесса аналитиком и внесения изменений процесс может быть загружен в систему и движение документа по организации будет автоматизировано.

Согласно [1–8] для обеспечения реконструкции схемы процесса журнал событий должен иметь как минимум четыре атрибута:

действие (*activity*) — действие, выполненное пользователем например, «подпись документа», «наложение резолюции»;

время регистрации (*timestamp*) — момент времени, когда произошло события;

идентификатор последовательности событий (*case id*) — идентификатор последовательности действий над определенным документом;

ресурс (*resource*) — исполнитель, или инициатор действия пользователь или внешняя информационная система).

В рамках СЭД отдельный экземпляр рабочего процесса ассоциируется с документом, поэтому трассы могут быть выявлены по идентификатору документа, который соответствует идентификатору последовательности (*case id*) рассмотренного журнала. Общая схема метода реконструкции рабочих процессов документооборота показана на рис. 3.

Необходимость шагов 1,2 обусловлена тем, что с одним типом документов в журнале могут быть связаны несколько рабочих процессов, поэтому важно определить признак, по которому трассы одного процесса отличаются от трасс другого. Это может быть сделано исходя из положения, что каждый документ имеет свой *тип*, а тип в свою очередь характеризуется *жизненным циклом*. Жизненный цикл (ЖЦ) документа — тип поведения документа от момента формирования до момента передачи в архив (на хранение) или уничтожения. Жизненный цикл может быть описан в форме графа, в котором вершинами являются стадии жизненного цикла, а ребра-

ми — переходы между стадиями. При выделении трасс рабочего процесса можно исходить из принадлежности действий к одной стадии ЖЦ. К одному рабочему процессу могут быть отнесены трассы от момента начала до момента окончания стадии ЖЦ.

Жизненный цикл каждого типа документа должен быть описан. В разработанном прототипе для этих целей был использован справочник «Жизненный цикл документа», связанный со справочником «Типы документов».

После выделения множества трасс может быть проведена реконструкция перспективы «Поток управления» рабочего процесса одним из алгоритмов *Process Mining* (шаг 3 на рис. 3). При разработке прототипа был использован альфа-плюс алгоритм [9]. Общая схема реконструкции проекции «Поток управления» показана на рис. 4.

Алгоритм на первом этапе на основе журнала событий строит матрицу пар отношений между событиями. Выделяется 4 типа отношений:

прямая преемственность ($a >_L b$) — шаблонное отношение, наблюдающееся, когда в журнале событий присутствует хотя бы одна трасса, в которой событие *b* следует сразу же за событием *a*;



Рис. 3. Схема процесса реконструкции рабочих процессов документооборота

причинность ($a \rightarrow_L b$) — отношение наблюдается в журнале только когда есть хотя бы одна трасса, где ($a >_L b$) и нет ни одной трассы, в которой ($a \not>_L b$). То есть можно говорить, что причиной появления события b в журнале служит событие a ;

несвязность ($a \#_L b$) — шаблонное отношение, наблюдаемое, когда в журнале ($a \not>_L b$) и ($b \not>_L a$);

параллельность ($a \parallel_L b$) — отношение, наблюдаемое в журнале, если ($a >_L b$) и ($b >_L a$).

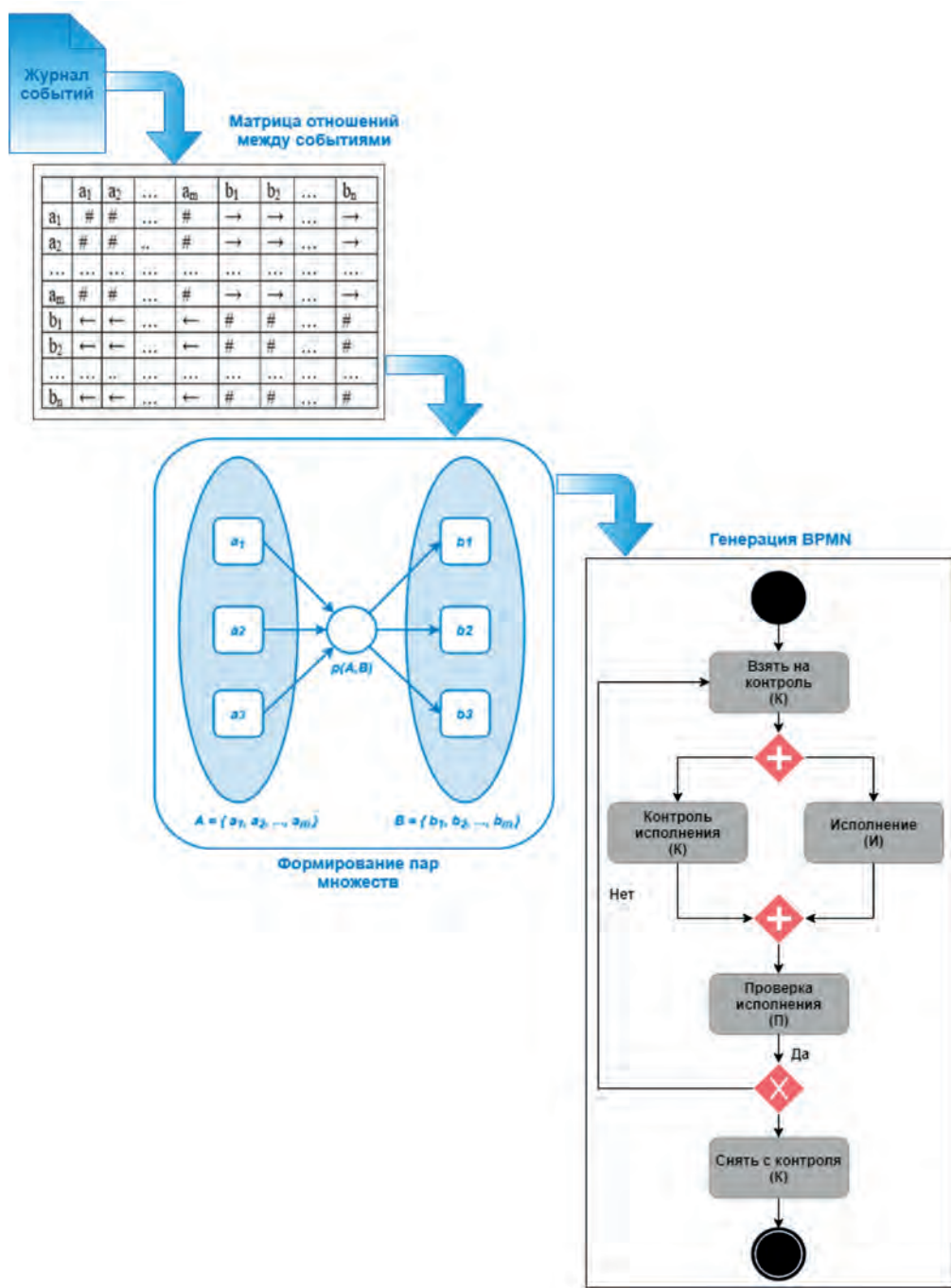


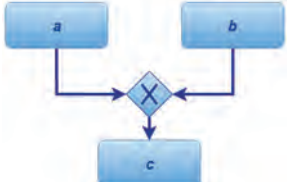




Рис. 4. Схема реконструкции модели потока управления

Таблица 1

Условия для генерации фрагментов нотации BPMN на основе альфа-алгоритма

| Фрагмент в нотации BPMN | Условие |
|--|--|
| <p>Последовательное выполнение</p>  | <p>присутствует пара (A, B) $a \in A, b \in B$, не выполняются другие условия</p> |
| <p>Условное разделение</p>  | <p>присутствует пара (A, B) $a \in A, b \in B, c \in B$</p> |
| <p>Условное слияние</p>  | <p>присутствует пара (A, B) $a \in A, b \in A, c \in B$. не выполняются другие условия</p> |
| <p>Параллельное разделение</p>  | <p>$a \in A_1, a \in A_2, b \in B, c \in C$ и присутствуют пары отношений (A_1, B) (A_2, C)</p> |
| <p>Параллельное слияние</p>  | <p>$a \in A_1, a \in A_2, b \in B, c \in C$ и присутствуют пары отношений (B, A_1) (C, A_2)</p> |

На втором этапе множество событий разбивается на пары множеств (A, B) по следующему принципу: каждый элемент множества A связан с каждым элементом множества B , при этом внутри A и B события связаны отношением ‘#’ (несвязность).

Базовая версия альфа-плюс алгоритма использует в качестве целевого представления сеть Петри. Переход к нотации BPMN на третьем этапе был осуществлен при помощи таблицы преобразования (табл. 1).

Проекция «Операции» (см. рис. 2) поддерживается на уровне системы электронного документооборота и может в себя включать такие элементарные действия над документом как: изменение, просмотр, подписание, блокировка, создание и удаление версии. Исполнение каждого действия сопровождается занесением записи в журнал. Кроме этого в модель операций должны быть включены действия исполнителей с данными регистрационных карточек. Предполагается, что в качестве одного действия

результатирующего журнала рассматривается набор действий по изменению атрибутов карточки, выполняемых последовательно одним пользователем.

На шаге 4 (см. рис. 3) для обеспечения реконструкции проекции «Ресурсы», т. е. модели исполнителей процесса в журнале должно содержаться поле «Исполнитель действия». Указанное поле может быть взято из справочника «Пользователи». Реконструкция роли исполнителя действия может быть осуществлена на основе связанных справочников «Должностные лица» (ДЛ), «Должности» (Д) и «Подразделения организации» (ПО). На основе указанных справочников формируется дерево, в котором промежуточные узлы — подразделения организации и должности, а листья — должностные лица (см. рис 5).

Назначение исполнителей заданий рабочих процессов может быть осуществлена на основе следующих эвристических правил:

если на множестве трасс действия всегда исполняет одно и то же должностное лицо, то роль может соответствовать только этому ДЛ;

в случае если исполнителем действия выступают разные должностные лица, в качестве роли может быть использован промежуточный узел дерева (наименьшее по численности подразделение или должность), включающий в качестве потомков всех указанных ДЛ.

Рассмотрим заключительный шаг 5 метода (см. рис. 3). В случае электронного документооборота, перспектива «Данные», описывающая основные атрибуты экземпляра процесса содержится в регистрационной карточке. Карточка, наряду с жизненным циклом описаны

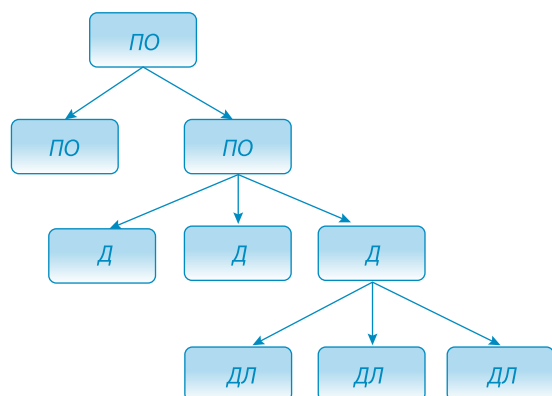


Рис. 5. Фрагмент дерева организационно-штатной структуры для реконструкции перспективы «Ресурсы»

в справочнике «Тип документа» куда заносится администратором СЭД. Для заполнения регистрационной карточки используются следующие базовые типы атрибутов: «Дата», «Дробное число», «Признак», «Справочник», «Строка», «Текст», «Целое число».

В разных действиях рабочего процесса документооборота исполнители работают с различными подмножествами атрибутов регистрационной карточки. Как было отмечено, под одним действием предполагается набор изменений реквизитов карточки, выполняемых последовательно одним пользователем.

Полнота описания перспективы «Данные» также обеспечивается заданием модели поведения исключаяющих шлюзов. То есть условий заданных на значениях реквизитов карточки в зависимости от которых срабатывают исключаяющие шлюзы, соответствующие условным переходам процесса (см. рис. 6). Каждый вариант прохождения исключаяющего шлюза представлен отдельной трассой в журнале событий. Таким образом, для каждого исключаяющего шлюза, полученного при реконструкции модели управления потоком по журналу должна быть построена пара предикатов вида:

$(\text{атрибут}_1 \text{ оп.сравн. значение}_1)$

ИЛИ $(\text{атрибут}_2 \text{ оп.сравн. значение}_2)$ ИЛИ...

где *оп.сравн.* — операции сравнения: «>», «<», «=».

Каждое из полученных выражений определяет вариант исполнения исключаяющего шлюза.

Данное построение может быть выполнено алгоритмом автоматического построения деревьев решений С4.5 (см. напр. [10]). Метки классов, соответствующих листьям дерева принятия решений соответствуют пары событий до и после условного перехода (см. рис. 7). Однако реализация для перечисленного набора типов данных налагает свои особенности:

для атрибутов типа «Признак» и «Справочник» алгоритм используется без модификаций. В процессе построения дерева решений для каждого возможного значения признака или записи справочника создается отдельное поддерево;

для атрибутов «Целое число» и «Дробное число» производится дискретизация. Для каждого численного реквизита определяется возможный размах значений с последующим разбиением на интервалы — по одному для каждого поддерева;

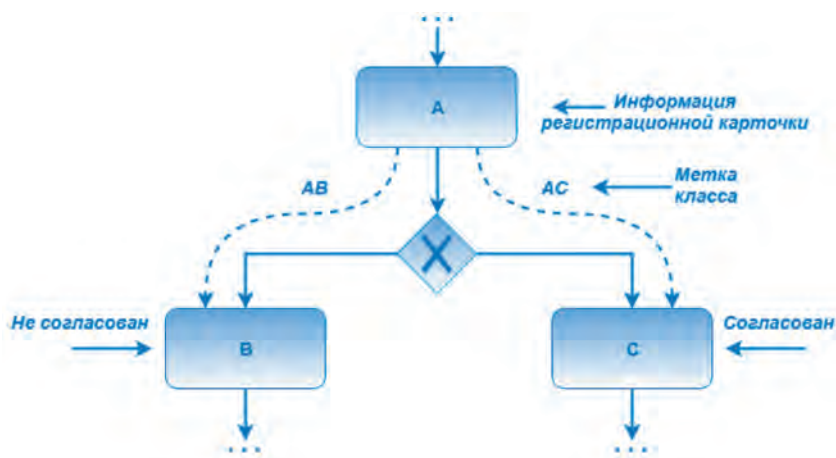


Рис. 6. Схема срабатывания условного перехода в рабочем процессе с меткой класса маршрута

на множестве атрибутов типа «Дата» вычисляются все возможные разности (целые числа), после чего задача сводится к классификации целочисленных значений;

на множестве значений реквизитов типа «Строка» и «Текст» выполняется индексирование (см. [11]) и последующее выделение типов резолюций методом кластеризации на основе дерева суффиксов *STC (Suffix tree clustering)*. Сформированные для каждого исключаящего шлюза классы резолюций соответствуют под деревьям узла.

Рассмотрим пример построения схемы принятия решений (см. табл. 2). Из регистрационной карточки было взято два атрибута «Срок исполнения» и «Поле резолюции» (X_1 и X_2). Первое поле относится к типу «Дата», а второе имеет тип «Текст». Для дальнейшего анализа было введено поле X_1'' и X_2'' . Первое представляет разницу между сроком выполне-

ния и текущей датой (10.10.2018 г.) в днях, а во втором хранится отношение резолюции к одной из групп: *positive* (положительное решение), *negative* (отрицательное решение), *null group* (без резолюции). Данное отношение было получено с использованием алгоритма ранжирования (см. [11] и табл. 2). Вариант прохождения исключаящего шлюза указан в виде столбца Y . На данном наборе данных было построено дерево принятия решений (рис. 8) по алгоритму $S4.5$. На каждом шаге алгоритм последовательно вычисляет энтропию и прирост информации (см. подробнее [10]) для каждого атрибута регистрационной карточки. Выбор атрибута для текущего узла дерева решения производится на основе критерия максимизации прироста информации. Энтропия приведенного фрагмента до разбиения равна 0,9852.

Энтропия при разбиении по атрибуту $X_1'' = 0,3935$.

Таблица 2

Реквизиты фрагмента регистрационной карточки для реконструкции дерева решений

| X_1 | X_2 | X_1'' | X_2'' | Y |
|------------|--|-------------------------|---------|-----------|
| 20.10.2018 | Изменить название пункта 2.2 | <i>neg</i> (изменить) | 9 | <i>AB</i> |
| 18.10.2018 | В приказе добавить подпись Иванова В. В. | <i>neg</i> (добавить) | 7 | <i>AB</i> |
| 15.10.2018 | Приказ одобрен. Для ознакомления | <i>pos</i> (одобрено) | 4 | <i>AC</i> |
| 09.10.2018 | Null | <i>Null</i> | -1 | <i>AB</i> |
| 11.10.2018 | Null | <i>Null</i> | 1 | <i>AC</i> |
| 12.10.2018 | На счет правок невозрожаю. | <i>pos</i> (невозрожаю) | 2 | <i>AC</i> |
| 01.10.2018 | Не смотрел | <i>Null</i> | -9 | <i>AB</i> |

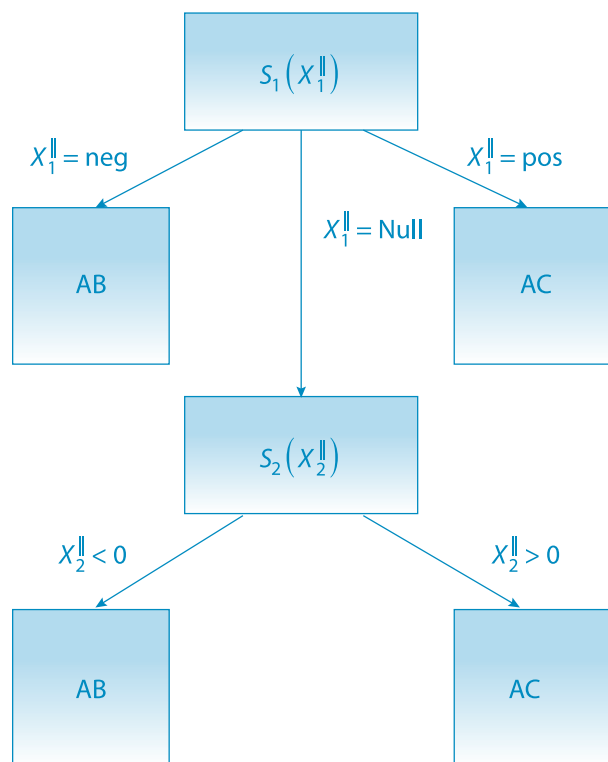


Рис. 7. Пример дерева принятия решений

Прирост информации по атрибуту $X_1^{||} = 0,5917$.

Энтропия при разбиении по $X_2^{||} = 0,6935$.

Прирост информации по $X_2^{||} = 0,2916$

Таким образом, на основе приведенного критерия на первом шаге для ветвления должен использоваться $X_1^{||}$.

Фрагмент построенного дерева показан на рис. 7.

Построенные предикаты в виде делегатов хранятся в описании рабочего процесса и ассоциируются с конкретным шлюзом.

3. Реализация инфраструктуры анализа процессов в СЭД

Предложенный метод реконструкции проекций моделей процессов был реализован в виде компонента в составе редактора рабочих процессов документооборота «Цера» [12]. Отличительной особенностью разработанного модуля от уже существующих решений (см. напр. [13–15]) является ориентация на анализ процессов документооборота.

Инфраструктура *Process Mining* в СЭД показана на рис. 8. В процессе функционирования редактор обращается к базе данных и справоч-

никам документооборота. На первом этапе модулем построения модели управления потоком осуществляется процесс сегментирования журнала событий на основе выбранного пользователем типа документа и стадии его жизненного цикла. После чего производится реконструкция BPMN — графа рабочего процесса. На основании рассмотренной методики модуль построения модели ресурсов производит обращения к справочникам описания организационно-штатной структуры и последующее назначение ролей заданиям схемы рабочего процесса.

Дальнейшую работу выполняет модуль построения модели данных и решений, производящий на основе таблиц «Регистрационная карточка» и «Атрибут карточки», а также *json*-поля «образ атрибутов» таблицы «Журнал событий» производит создание форм заданий рабочего процесса, а также формирование предикатов для делегатов, обеспечивающих поведение исключających шлюзов.

4. Практическая апробация. Выводы

Практическая апробация разработанного модуля показала недостаточную эффективность альфа — плюс алгоритма для решения задач анализа реальных журналов СЭД. В случае наличия ошибок в журнале (дублирование или выпадение событий, ошибки ручного выполнения операций) авторами были получены слабо читаемые модели, объем ручных модификаций которых по доводке до рабочих процессов был значителен. В качестве дальнейших исследований авторы предполагают использовать алгоритм индуктивного анализа (*Inductive miner*), позволяющий, как и альфа-плюс алгоритм создавать бездефектное (*soundness*) описание процессов, но не столь чувствительные к ошибкам (см. напр. [15]). Была также определена зависимость результатов анализа от дисциплинированности исполнителей документов, на основе которых действий которых формируется исходный журнал. В случае несвоевременного указания значений реквизитов регистрационной карточки при выполнении действий наблюдалась некорректная проекция данных процесса. В качестве частичной меры преодоления указанного недостатка авторами предлагается привязка групп реквизитов регистрационной карточки к стадиям ЖЦ типа документа. Открытым остается также вопрос оценки степени деградации схемы рабочего процесса при организационно-штатных и нормативных изменениях на предприятии.

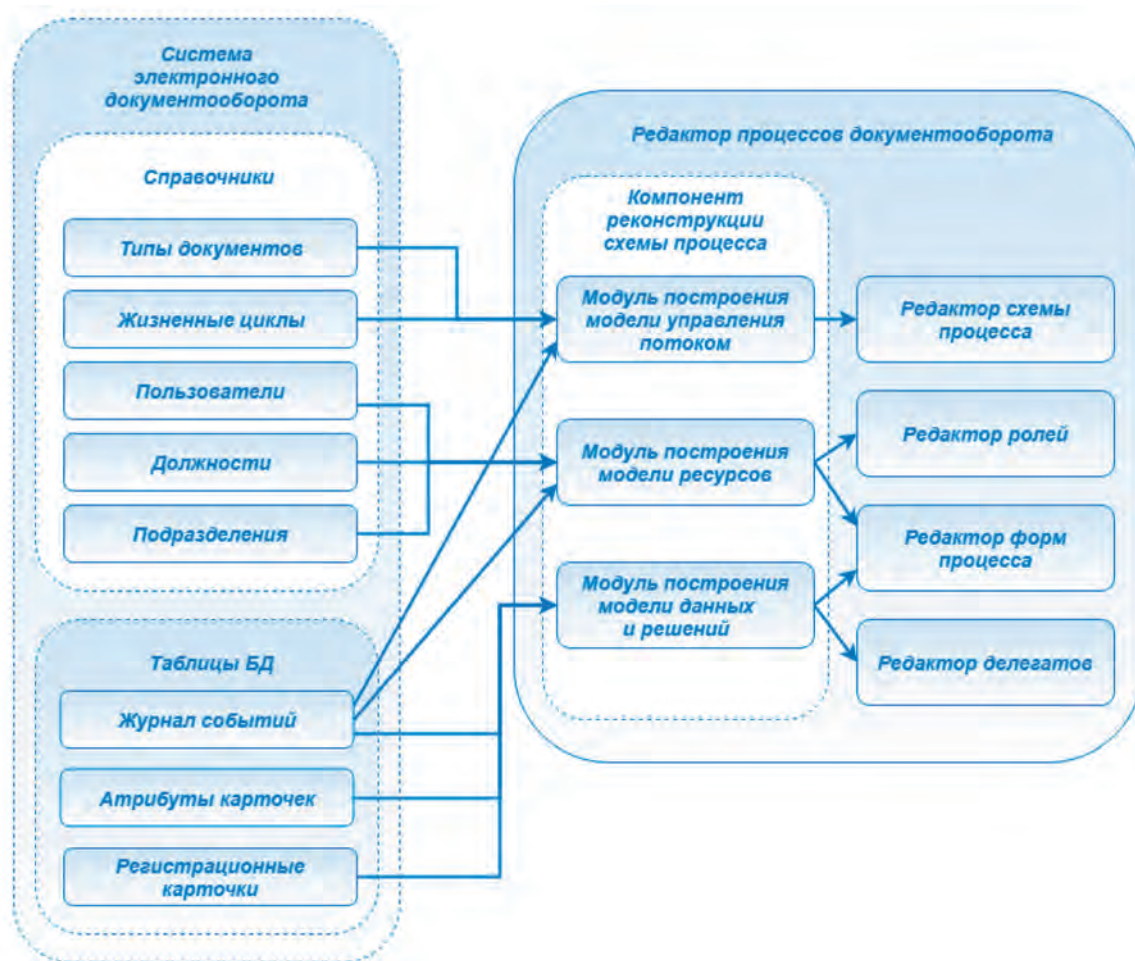


Рис. 8. Инфраструктура Process Mining в системе электронного документооборота

ЛИТЕРАТУРА

1. W.M.P. van der Aalst *Process Mining: Discovery, Conformance and Enhancement of Business Processes*. — Berlin: Springer-Verlag, 2011. — 352 p.
2. Барсегян А. А. Анализ данных и процессов: учеб. пособие / А. А. Барсегян, М. С. Куприянов, И. И. Холод, М. Д. Тесс, С. И. Елизаров. — СПб.: БХВ-Петербург, 2009. — 512 с.
3. Романченко Е. В. Основные тенденции развития СЭД в России / Е. В. Романченко // *Современные технологии делопроизводства и документооборота*. — 2015. — № 8. [Электронный ресурс]: Режим доступа: <http://e.deloprost.ru/article.aspx?aid=419473>.
4. Мокрый В. Ю. Системы электронного документооборота: учебное пособие / В. Ю. Мокрый. — СПб.: Инфо-да, 2018. — 48 с.
5. Михеев А. Г. Системы управления бизнес-процессами и административными регламентами на примере свободной программы RunaWFE: учеб. пособие / А. Г. Михеев. — М.: Альт Линукс, 2011. — 178 с.
6. W.M.P. van der Aalst, A.J.M.M. Weijters L. Maruster *Workflow Mining: Discovering process models from event logs*. In *IEEE Transactions on Knowledge & Data Engineering*, vol. 16, no.5, pp. 1128–1142, 2004.
7. W.M.P. van der Aalst and B.F. van Dongen. *Discovering Workow Performance Models from Timed Logs*. In Y. Han, S. Tai, and D. Wikarski, editors, *International Conference on Engineering and Deployment of Cooperative Information Systems (EDCIS2002)*, volume 2480 of *Lecture Notes in Computer Science*, pp. 45–63. Springer Verlag, Berlin, 2002.
8. W.M.P. van der Aalst, A.J.M.M. Weijters, and L. Maruster. *Workflow Mining: Discovering Process Models from Event Logs*. *IEEE Transactions on Knowledge and Data Engineering*, 16(9):1128–1142, 2004.
9. A. K. Alves de Medeiros, B. F. van Dongen, W.M.P. van der Aalst, and A.J.M.M. Weijters. *Process Mining: Extending the α -algorithm to Mine Short Loops*. BETA Working Paper Series, WP 113, Eindhoven University of Technology, Eindhoven, 2004.

10. **Паклин Н. Б.** Бизнес-аналитика: от данных к знаниям / Н. Б. Паклин, В. И. Орешков. — СПб: Питер, 2009. — 624 с.

11. **Марманис Х., Бабенко Д.** Алгоритмы интеллектуального интернета. Передовые методики сбора, анализа и обработки данных/Пер. с англ. — СПб.: Символ-Плюс, 2011. — 480 с.

12. Свид. 2017663083 Российская Федерация. Свидетельство об официальной регистрации программы для ЭВМ. Подсистема защищенного электронного документооборота «Цера» / Васильев Н. В., Компанец А. Н., Сопин Д. С.; заявитель и правообладатель ПАО «Интелтех» (RU). — № 2017663083; заявл. 06.10.17; опубл. 24.11.17, Реестр программ для ЭВМ. — 1 с.

13. B. F. van Dongen, A. K. Alves de Medeiros, H.M.W. Verbeek, A.J.M.M. Weijters, and W.M.P. van der Aalst. The ProM Framework: A New Era in Process Mining Tool Support. In G. C. and P. Darondeau, editors, ICATPN, volume 3536 of Lecture Notes in Computer Science, pp. 444–454. Springer, 2005.

14. B. F. van Dongen and W.M.P. van der Aalst. EMiT: A Process Mining Tool. In Jordi Cortadella and Wolfgang Reisig, editors, ICATPN, volume 3099 of Lecture Notes in Computer Science, pp. 454–463. Springer, 2004.

15. Leemans S.J.J., Fahland D., van der Aalst W. M.P. (2013) Discovering Block-Structured Process Models from Event Logs — A Constructive Approach. In: Colom JM., Desel J. (eds) Application and Theory of Petri Nets and Concurrency. PETRI NETS2013. Lecture Notes in Computer Science, vol 7927. Springer, Berlin, Heidelberg.