

НАСТРОЙКА ЭКСПЕРТНОЙ СИСТЕМЫ АНАЛИЗА ДОКУМЕНТООБОРОТА НА ОСОБЕННОСТИ ОРГАНИЗАЦИИ

Усманова И.В., Коровина Л.В., Соколова О.Г.

ФГБОУ ВПО «Пензенский государственный университет», г. Пенза, Россия (440026, г. Пенза, ул. Красная, 40),
e-mail: inoup@pnzgu.ru

В работе рассмотрена и обоснована необходимость оптимизации и рационализации документооборота для обеспечения продуктивной работы организации. Анализ и оптимизацию документооборота предлагается произвести с применением экспертной системы (ЭС), в которой используется многоуровневая модель представления знаний. Модель основана на применении аппарата семантических сетей и позволяет прогнозировать показатели эффективности работы организации, зависящие от особенностей движения документов. Перечислены основные компоненты разработанной многоуровневой модели представления знаний, показан механизм формирования компонентов, обеспечивающих адаптацию системы к особенностям предметной области и специфики организации. Предложен алгоритм настройки системы логического вывода ЭС, основанный на применении метода Байеса. Предложенный алгоритм является инвариантным к предметной области и может быть применен для широкого спектра задач в области разработки систем управления документооборотом.

Ключевые слова: метод Байеса, анализ документооборота, логический вывод, вероятность, семантическая сеть.

SETTING THE EXPERT ANALYSIS OF DOCUMENT ON FEATURES OF ORGANIZATION

Usmanova I.V., Korovina L.V., Sokolova O.G.

FGBOU VPO «Penza State University», Penza, Russia (440026, Penza, Street Krasnaya, 40), e-mail: inoup@pnzgu.ru

In work need of optimization and document flow rationalization for ensuring productive work of the organization is considered and proved. The analysis and optimization of document flow is offered to be made with use of the expert system (ES) in which the multilevel model of representation of knowledge is used. The model is based on use of the device of semantic networks and allows to predict indicators of overall performance the organizations depending on features of movement of documents. The main components of the developed multilevel model of representation of knowledge are listed, the mechanism of formation of the components providing adaptation of system to features of subject domain and specifics of the organization is shown. The algorithm of control of system of a logical conclusion ES based on application of a method of Bayes is offered. The offered algorithm is invariant to subject domain and can be applied to a wide range of tasks in the field of development of systems of management by document flow.

Keywords: Bayess method, document flow analysis, logical conclusion, probability, semantic network.

В современных рыночных условиях повышение эффективности функционирования организации достигается не только за счет привлечения дополнительных финансовых средств, увеличения штатной численности, внедрения новых технологий в производственную сферу, но и благодаря рационализации и оптимизации документооборота. Оптимизация документопотоков предполагает периодический анализ текущих количественных и качественных показателей документооборота и бизнес-процессов [3], [6].

Существующие методы анализа документооборота позволяют выявить лишь отдельные характеристики документооборота как такового. При этом не учитываются связи показателей состояния документооборота с параметрами эффективности бизнес-процессов и деятельности организации. Анализ документооборота – задача, которая носит смысловой характер, и успех ее решения определяется прежде всего личными качествами сотрудника,

проводящего анализ. Привлечение сторонних организаций увеличивает время выполнения анализа документооборота, а во многих случаях и проблематично из-за отсутствия в таких фирмах специалистов, имеющих опыт работы в данной предметной области.

Поэтому анализ и оптимизацию документооборота предлагается произвести с применением автоматизированной информационной системы (АИС) [2], в которой используется многоуровневая модель представления знаний, основанная на применении аппарата семантических сетей. Это позволяет прогнозировать не только «узкие места» в движении документов, но и показатели эффективности работы организации, зависящие от особенностей движения документов.

Структура многоуровневой модели представления знаний более подробно рассмотрена в [4]. В ее состав входят следующие компоненты:

— семантическая сеть понятий, содержащая также сведения о характере отношений между понятиями. Применение указанной модели создает основу для накопления и систематизации сведений о предметной области, обеспечивает оперативный обмен информацией между пользователем и системой;

— семантические утверждения в форме предложений естественного языка, описывающих показатели состояния документооборота, а также эффективности бизнес-процессов и деятельности организации, например: «Комплект документов, получаемых из предшествующих процессов, не полон», «Система регистрации документов не оптимальна»;

— семантическая сеть бизнес-процессов, показывающая последовательность выполнения основных и вспомогательных производственных процессов организации, а также взаимосвязи их компонентов;

— семантическая сеть прогнозирования, отражающая причинно-следственные отношения между утверждениями в форме правил (продукций) вида: ЕСЛИ «Объем документооборота организации велик», ТО «Количество подготавливаемых должностным лицом документов велико»; ЕСЛИ «Количество подготавливаемых должностным лицом документов велико», ТО «Число задержек исполнения документов велико».

Такая модель позволяет установить взаимосвязь бизнес-процессов и операций документирования деятельности организации. Использование методов инженерии знаний позволяет отнести разрабатываемую систему к классу экспертных (ЭС). Основными элементами предлагаемой АИС являются: база знаний (БЗ), база данных (БД), подсистема логического вывода, подсистема редактирования знаний, подсистема объяснений [1].

Эффективность ЭС определяется объемом и качеством БЗ, а также тем, насколько возможна адаптация системы к особенностям функционирования организации в рассматриваемой предметной области. Необходимость настройки АИС на особенности

делопроизводства в определенный промежуток времени определяется динамикой, изменчивостью бизнес-процессов, а следовательно, периодической сменой состава, количества и содержания порождаемых бизнес-процессами документов.

Качество настройки ЭС на особенности предметной области и специфику постановки делопроизводства определяется степенью корректности и полноты содержания сведений базы знаний, апеллирование которыми позволяет сформулировать утверждения, характеризующие состояние документооборота «от документов» и «от процессов» [3].

Основное назначение ЭС – анализ документооборота, поэтому утверждения, входящие в семантическую сеть прогнозирования, должны характеризовать особенности документационного обеспечения управления (ДОУ) и бизнес-процессов организации. Определение входных данных и результатов выполнения каждого процесса необходимо для установления связи между бизнес-процессами. Все процессы деятельности организации являются взаимосвязанными. Как правило, они используют выходную информацию предшествующего процесса и сами являются источниками информации для последующего процесса. Несоответствие ожидаемым значениям реальных входных данных процесса, являющихся, в свою очередь, результатом выполнения одного или нескольких предшествующих процессов, может сделать невозможным осуществление этого процесса.

Выявление смысловых зависимостей между отдельными характеристиками процессов позволяет сформулировать причинно-следственные связи между ними в виде правил типа «ЕСЛИ...ТО». Например: ЕСЛИ «Выбор поставщиков продукции неэффективен» ТО «Сырье закуплено несвоевременно»; ЕСЛИ «Сырье закуплено несвоевременно» ТО «Заказ выполняется несвоевременно»; ЕСЛИ «Заказ выполняется несвоевременно» ТО «Низка удовлетворенность партнеров» и т.д.

Поскольку результатом осуществления бизнес-процесса являются не только реализованные действия, но и их документированное отражение, то выходными данными многих процесса являются документы, создаваемые в результате выполнения каждого действия бизнес-процесса. Поэтому задачей специалистов службы ДОУ является определение показателей качества выполнения отдельных операций с документами. Критерием принятия решения могут служить не только правильность оформления документа, но и порядок его прохождения. Для анализа операций с документами и маршрута их движения документы необходимо рассматривать с точки зрения принадлежности к одному из трех видов документопотоков – входящих, исходящих, внутренних. Возможные нарушения показателей качества документооборота являются основанием для формулирования семантических утверждений вида «Маршрут прохождения документа не

закреплен», «Количество ошибок в оформлении документов велико», «Унифицированные формы документов отсутствуют» и т.д.

Для формирования корректного и эпистемологически полного набора семантических утверждений предлагается использовать алгоритм, применение которого позволит учесть в ЭС особенности предметной области. Алгоритм содержит следующие шаги:

1. Обследование организации и выявление всех бизнес-процессов, реализуемых в организации. Поскольку бизнес-процессы, как правило, выполняются в несколько этапов, в рамках каждого процесса необходимо выделить отдельные действия.
2. Формирование для каждого бизнес-процесса следующих сведений: входные и выходные данные; документы, создаваемые в результате осуществления каждого действия бизнес-процесса; показатели эффективности, позволяющие сделать вывод о рациональности и правильности выполнения процесса.
3. Формулирование семантических утверждений – предложений естественного языка, в которых фиксируется некоторое нарушение какого-либо показателя эффективности или риск выполнения процесса, например: «Своевременность закупки сырья нарушена», «Учет изготовленной продукции произведен с ошибкой», «Последовательность действий процесса нарушена», «Низка удовлетворенность потребителя» и т.д.
4. Выявление возможных проблем в работе с документами и формулирование соответствующих семантических утверждений.
5. Детальный анализ деятельности каждого должностного лица организации. При этом исследованию подлежат не только бизнес-процессы, в которых задействовано должностное лицо, но и его функции по обработке и подготовке документов. Результатом данного этапа является выявление семантических утверждений вида «Количество согласуемых должностным лицом документов велико», «Количество подписываемых должностным лицом документов велико», «Количество утверждаемых должностным лицом документов велико». Необходимость их включения в семантическую сеть прогнозирования объясняется тем, что в случае перегрузки должностного лица работой с документами под угрозой оказываются бизнес-процессы, в которых задействован сотрудник.
6. Экспертная оценка взаимовлияния семантических утверждений, позволяющая оценить состояние документооборота «от документов» и «от процессов». Например, ЕСЛИ «Маршрут прохождения документа не закреплен в нормативных документах», ТО «Последовательность подготовки документа нарушена»; ЕСЛИ «Последовательность подготовки документа нарушена», ТО «Документ подготовлен несвоевременно».

Результатом применения описанного алгоритма является формирование сети прогнозирования, позволяющей определить реальную картину взаимовлияния показателей

состояния документооборота и эффективности бизнес-процессов и отражающей особенности функционирования организации и постановки ДООУ.

Как известно, наиболее важным элементом экспертной системы [4] является механизм логического вывода, который предназначен для получения новых фактов на основе сопоставления исходных данных из рабочей памяти и знаний из базы знаний.

Каждое семантическое утверждение сети продукций характеризуется некоторой вероятностью его появления, поэтому для настройки подсистемы логического вывода и организации логического вывода в ЭС предлагается использовать известную теорему Байеса [7]:

$$p(c_k / c_n) = \frac{p(c_n / c_k) \cdot p(c_k)}{p(c_n)}, \quad (1),$$

где $p(c_k)$ – вероятность выполнения следствия c_k правила $H(c_k / c_n)$ (априорная вероятность);

$p(c_n)$ – вероятность события c_n , состоящего в выполнении условия c_n правила $H(c_k / c_n)$;

$p(c_n / c_k)$ – условная вероятность события c_n при гипотезе c_k (апостериорная вероятность).

Исходными данными для метода Байеса служат значения априорных вероятностей $p(c_k)$ для каждого утверждения c_k . Они задаются экспертом при создании семантической сети на основании предварительного обследования организации и являются основой для вычисления апостериорных вероятностей, соответствующих измененной вероятности c_k . Если E' – событие, в котором реально зафиксированы все факты выполнения и нарушения утверждений сети в некоторый момент времени работы АИС, то апостериорная вероятность $p(c_k / E')$ есть текущая условная вероятность c_k , если произошло событие E' .

Предлагаемый алгоритм процедуры настройки системы логического вывода на особенности текущей ситуации в организации с использованием метода Байеса включает следующие этапы.

1. Определение активного фокуса семантической сети. Под активным фокусом сети понимается множество семантических утверждений, выбранных пользователем в соответствии с его предпочтениями, или утверждений, в которых фактически зафиксировано некоторое нарушение в работе с документами или в бизнес-процессах. Каждому утверждению семантической сети соответствует вероятность его выполнения.

2. Задание экспертом априорной вероятности $p(c_k)$ для утверждения c_k и априорной условной вероятности $p(c_k/c_n)$, которая показывает выполнения утверждения c_k при выполнении утверждения c_n .

3. Задание пользователем для интересующего его утверждения в диалоговом режиме, значение меры определенности $C(c_n/c'_n)$, которая представляет собой апостериорную вероятность, нормализованную по отношению к априорной вероятности и изменяющуюся в некотором интервале. При создании базы знаний эксперт задает Δ_m – максимально возможное значение исследуемого параметра, при котором вероятность появления соответствующего нарушения равна единице. Пользователь для конкретной ситуации задает реальное значение Δ_r оценочного параметра, при котором можно говорить о реализации нарушения, а также идеальное значение Δ_i , при котором невозможно говорить о каком-либо нарушении исследуемого параметра. Мера определенности вычисляется по формуле:

$$C(c_n/c'_n) = \frac{10 * |\Delta_i - \Delta_r|}{|\Delta_i - \Delta_m|} - 5 \quad . \quad (2)$$

Использование меры определенности $C(c_n/c'_n)$ позволяет снизить требования к математической подготовке пользователя, выполняющего анализ документооборота и являющегося специалистом в области ДООУ, а не математики.

4. Определение вероятности выполнения утверждения $p(c_n/c'_n)$ в реальных условиях, заданных пользователем, по формуле:

$$p(c_n/c'_n) = \begin{cases} p(\overline{c_n}) + \frac{C(c_n/c'_n) * (1 - p(c_n))}{5}, & C(c_n/c'_n) > 0 \\ p(c_n) + \frac{C(c_n/c'_n) * p(c_n)}{5}, & C(c_n/c'_n) < 0 \end{cases} \quad . \quad (3)$$

5. Определение условной вероятности выполнения утверждения c_k в реальных условиях $p(c_k/c'_n)$ по формуле:

$$p(c_k/c'_n) = \begin{cases} p(c_k/c_n) + \frac{p(c_k) - p(c_k/\overline{c_n})}{p(c_n)} * p(c_n/c'_n), & 0 \leq p(c_n/c'_n) \leq p(c_n) \\ p(c_n) + \frac{p(c_k/c_n) - p(c_k)}{1 - p(c_k)} * [p(c_n/c'_n) - p(c_n)], & p(c_n) \leq p(c_n/c'_n) \leq 1 \end{cases} \quad . \quad (4)$$

6. Определение вероятности невыполнения условия правила $p(\overline{c_n}/\overline{c'_n})$ по формуле:

$$p(\bar{c}_n / \bar{c}'_n) = \begin{cases} p(\bar{c}_n) + \frac{C(\bar{c}_n / \bar{c}'_n) * (1 - p(\bar{c}_n))}{5}, & p(\bar{c}_n / \bar{c}'_n) > p(\bar{c}_n) \\ p(\bar{c}_n) + \frac{C(\bar{c}_n / \bar{c}'_n) * p(\bar{c}_n)}{5}, & p(\bar{c}_n / \bar{c}'_n) \leq p(\bar{c}_n) \end{cases} . \quad (5)$$

7. Определение значения вероятности выполнения утверждения c_k при реальном невыполнении утверждения c_n $p(c_k / \bar{c}'_n)$ по формуле:

$$p(c_k / \bar{c}'_n) = \begin{cases} p(\bar{c}_n) + \frac{p(c_k / \bar{c}_n) - p(c_k)}{1 - p(c_k)} * [p(\bar{c}_n / \bar{c}'_n) - p(\bar{c}_n)], & p(\bar{c}_n) \leq p(\bar{c}_n / \bar{c}'_n) \leq 1 \\ p(c_k / \bar{c}'_n) + \frac{p(c_k) - p(c_k / \bar{c}_n)}{p(c_n)} * p(\bar{c}_n / \bar{c}'_n), & 0 \leq p(\bar{c}_n / \bar{c}'_n) \leq p(\bar{c}_n) \end{cases} . \quad (6)$$

8. Определение характеристик продукций, постоянных в условиях конкретной задачи:

а) апостериорное значение предпочтительности гипотезы c_k по формуле:

$$O(c_k / c'_n) = \frac{p(c_k / c'_n)}{1 - p(c_k / c'_n)}; \quad (7)$$

б) априорное значение предпочтительности гипотезы c_k по формуле:

$$O(c_k) = \frac{p(c_k)}{1 - p(c_k)}; \quad (8)$$

с) априорное значение $O(c_k / \bar{c}'_n)$

$$O(c_k / \bar{c}'_n) = \frac{p(c_k / \bar{c}'_n)}{1 - p(c_k / \bar{c}'_n)}; \quad (9)$$

д) константа LS , выражающая достаточность правила, которая в условиях одной задачи определяется свойствами c_n и c_k :

$$LS = \frac{O(c_k / c'_n)}{O(c_k)}; \quad (10)$$

е) константа LN , которая соответствует мере необходимости правила:

$$LN = \frac{O(c_k / \bar{c}'_n)}{O(c_n)}. \quad (11)$$

Действия, составляющие рассмотренный алгоритм, выполняются однократно для каждого семантического утверждения сети прогнозирования, если пользователь находит это нужным. В противном случае вероятностные параметры определяются по умолчанию.

Описанный алгоритм успешно апробирован [5].

Рассмотренный алгоритм формирования сети прогнозирования предметной области и процедура настройки подсистемы логического вывода являются инвариантными к

предметной области и могут быть применены для широкого спектра задач в области разработки систем управления документооборотом.

Список литературы

1. Уотермен Д. Руководство по экспертным системам. – М.: Мир, 1989. – 174 с.
2. Усманова И.В. Модель представления знаний о маршрутах проведения анализа и оптимизации документооборота организации // Известия ПГПУ им. В.Г. Белинского. – 2012. – № 30. – С. 296-301.
3. Усманова И.В., Баканова Н.Б. Автоматизация процедуры оценки состояния документооборота организации // Materiały VII Międzynarodowej naukowo-praktycznej konferencji «Perspektywiczne opracowania są nauką i technikami – 2011» Vol. 57. Techniczne nauki.: Przemysł. Nauka i studia - 104 str. – P. 73-78.
4. Усманова И.В., Забродина Л.В. Взаимодействие интеллектуальных компонентов информационного обеспечения автоматизированной системы анализа и оптимизации документооборота организации // Известия ПГПУ им. В.Г. Белинского. – 2012. – № 30. – С. 310-318.
5. Усманова И.В., Соколова О.Г., Коровина Л.В. Адаптация подсистемы логического вывода интеллектуальной системы анализа документооборота к особенностям организации // Materiály IX mezinárodní vědecko - praktická konference «Dny vědy – 2013». – Díl 33. Moderní informační technologie. Fyzika: Praha. Publishing House «Education and Science» – P. 59-64.
6. Фионова Л.Р. Факторы, влияющие на структуру документооборота // Делопроизводство. – 2012. – № 2. – С. 51-55.
7. Duda R.O., Hart P.E., Nilsson N.J. Subjective Methods for Rule-Based Inference Systems // Proc. of National Computer Conf., 45, AFIPS Press, Oct. 1976. – Artificial Intelligence Center, SRI Int., Menlo Park. – 1976. – P. 1075-1082.

Рецензенты:

Фионова Л.Р., д.т.н., профессор, декан факультета вычислительной техники, заведующий кафедрой «Информационное обеспечение управления и производства» ФГБОУ ВПО «Пензенский государственный университет», г. Пенза.

Косников Ю.Н., д.т.н., профессор, заведующий кафедрой «Информационно-вычислительные системы» ФГБОУ ВПО «Пензенский государственный университет», г. Пенза.