

**МОДЕЛИ И МЕТОД АВТОМАТИЗИРОВАННОГО ВЫБОРА МАРШРУТОВ  
ЭЛЕКТРОННЫХ ДОКУМЕНТОВ НА УРОВНЕ ПОДРАЗДЕЛЕНИЙ****И. В. Шевченко, А. А. Скрыль, Д. А. Расторопов, Ю. Н. Свиченский**

Кременчугский национальный университет имени Михаила Остроградского

ул. Первомайская, 20, г. Кременчуг, 39600, Украина. E-mail: ius.shevchenko@gmail.com

Предложен подход к автоматизации процедуры маршрутизации в системе электронного документооборота. Разработаны и формализованы критерии выбора маршрута по тематическому принципу, по принципу интенсивности переписки по аспектам и по частоте регулярных контактов. Это позволяет осуществить многокритериальный выбор промежуточных адресов на маршруте. Разработано модельное представление процесса выбора маршрута документа, включающее онтологию предметной области, формальное описание взаимосвязей ключевых слов документа, аспектов функциональности подразделений и критериев выбора маршрута. Разработан метод автоматизированной параллельной маршрутизации электронных документов, отличающийся тем, что формирование зоны ответственности документа происходит по тематическому критерию и критерию частоты контактов подразделений. Как результат пользователь получает ранжированный список подразделений, потенциально входящих в зону ответственности документа. Это дает возможность уменьшить ошибки пользователя при составлении маршрута и устранить коллизии при прохождении документа. Предложенный метод отличается логической прозрачностью и обеспечивает адаптацию подсистемы маршрутизации к условиям предприятия и простое масштабирование. Разработана методика настройки и обучения подсистемы маршрутизации.

**Ключевые слова:** электронный документооборот, маршрутизация, критерии выбора адресов, модель выбора, метод автоматизированной маршрутизации.

**МОДЕЛІ І МЕТОД АВТОМАТИЗОВАНОГО ВИБОРУ МАРШРУТІВ  
ЕЛЕКТРОННИХ ДОКУМЕНТІВ НА РІВНІ ПІДРОЗДІЛІВ****І. В. Шевченко, О. О. Скриль, Д. А. Расторопов, Ю. М. Свіченський**

Кременчуцький національний університет імені Михайла Остроградського

вул. Першотравнева, 20, м. Кременчук, 39600, Україна. E-mail: ius.shevchenko@gmail.com

Запропоновано підхід до автоматизації процедури маршрутизації в системі електронного документообігу. Розроблено та формалізовано критерії вибору маршруту за тематичним принципом і за частотою регулярних контактів. Це дозволяє здійснити багатокритеріальний вибір проміжних адрес на маршруті. Для вирішення завдання автоматизованої маршрутизації розроблено комплекс моделей, що перетворюють значення первинних параметрів документа в підмножину передбачуваних адрес розсилки. Модельне уявлення процесу вибору маршруту документа включає онтологію предметної області, формальний опис взаємозв'язків ключових слів документа, аспектів функціональності підрозділів і критеріїв вибору маршруту. Розроблено і формально описано тематичний критерій для вибору маршруту та критерій інтенсивності листування, які дозволяють обчислити інтегральне значення ступеня асоціації змісту документа з кожним аспектом предметної області. Ранжування аспектів за ступенем асоціації дозволяє вибрати найбільш важливі аспекти, визначити відповідні повноваження підрозділів та вибрати підмножину підрозділів для паралельної адресації. Розроблено і формально описано критерій регулярних контактів. Критерій призначений для аналізу та відбору підрозділів, у яких є регулярне листування і з підрозділом-джерелом, і з підрозділом-одержувачем. Розроблено метод автоматизованої паралельної маршрутизації електронних документів, що відрізняється тим, що формування зони відповідальності документа відбувається за тематичним критерієм і критерієм частоти контактів підрозділів. Як результат користувач отримує ранжований список підрозділів, які потенційно входять до зони відповідальності документа. Це дає можливість зменшити помилки користувача при складанні маршруту і усунути колізії при проходженні документа. Запропонований метод відрізняється логічною прозорістю і забезпечує адаптацію підсистеми маршрутизації до умов підприємства і просте масштабування. Розроблено методіку настройки і навчання підсистеми маршрутизації.

**Ключові слова:** електронний документообіг, маршрутизація, критерії вибору адрес, модель вибору, метод автоматизованої маршрутизації.

**АКТУАЛЬНОСТЬ РАБОТЫ.** Использование систем электронного документооборота (СЭД) позволяет значительно повысить производительность труда делопроизводственного персонала, сокращает время, затрачиваемое на реализацию управленческих операций. На большинстве предприятий существует сложившаяся система документооборота со своими правилами и устоявшимся регламентом. Во многом повышение производительности офисных подразделений зависит от

особенностей документооборота на предприятии, но нельзя исключать и человеческий фактор. Анализ особенностей работы сотрудников с СЭД свидетельствует, что могут возникать проблемы и задержки, связанные с неопределенностью выбора маршрута документа, особенно для молодых сотрудников. Неправильно выбранный маршрут приводит к коллизиям и заметным задержкам в прохождении документа по точкам согласования [1–4]. Следовательно, автоматизация процесса вы-

бора маршрута будет способствовать повышению скорости и надежности работы управленческих подразделений.

В работах [5, 6] рассмотрен подход к созданию моделей композитного документооборота на основе аппарата теории графов. Описаны методы детерминирования множеств для разработанной модели, предложена алгебра документооборота с использованием графов. Результаты работы могут быть использованы для дальнейшего развития теории и практики электронного документооборота и создания на их основе прикладного программного обеспечения. Однако в данных работах не уделяется внимание таким деталям процесса документооборота, как маршрутизация, и связанные с ней проблемы.

В работе [7] поставлен вопрос о чувствительности систем электронного документооборота к коллизиям при прохождении маршрута. Приведена графовая модель и метод анализа чувствительности СЭД, позволяющий анализировать устойчивость систем по отношению к коллизиям. Но в работе нет примеров коллизий и анализа их причин.

Предлагается методика разработки архитектуры СЕД в научных учреждениях [8]. Выявлены основные формы документов в документопотоке научного учреждения. Разработана обобщающая классификация документов в зависимости от их типов и подтипов. Выделены соответствующие ключевые слова для каждой категории документов. Выполнено распределение СЕД на подсистемы, выделены их функции и назначения. На основе построенных архитектур отдельных подсистем СЭД предложена обобщенная архитектура СЭД научных учреждений. В этой работе решены, так сказать глобальные проблемы синтеза СЭД.

В работах [9–11] представлены модели движения документов, ориентированные на применение графов и сетей Петри. В работе [12] описывается программная реализация СЭД транспортного предприятия.

Анализ перечисленных работ показал, что ни в одной из них не уделено внимание надежности работы СЭД при выполнении операции выбора маршрута, от которой во многом зависит весь жизненный цикл документа.

Таким образом, проблема состоит в отсутствии практически электронных систем документооборота, имеющих в своем составе инструментальные средства поддержки процесса маршрутизации с использованием контекста документа и других данных, опираясь на которые можно сформировать необходимое множество подразделений, участвующих в согласовании и доработке документа.

Цель работы – минимизация количества коллизий в процессе обработки документа на маршруте путём разработки и внедрения моделей и метода автоматизированного выбора маршрутов электронных документов на уровне подразделений.

МАТЕРИАЛ И РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ. Как справедливо отмечено в работе [13], в управ-

лении бизнес-процессами в настоящее время большое значение приобретают интеллектуальные процедуры. В данном случае речь пойдет о применении технологий обработки естественного языка для формирования списка подразделений, входящих в зону ответственности документа (ЗОД).

*Модельное представление процесса маршрутизации.* Для решения задачи автоматизированной маршрутизации требуется разработать комплекс моделей, преобразующих значения первичных параметров документа в подмножество предполагаемых адресов рассылки.

Системный подход к разработке комплекса моделей предполагает на первом этапе построение онтологии. Поскольку онтология предметной области «Документооборот предприятия» содержит множество аспектов, связанных с полномочиями и функциями отдельных подразделений, представим онтологию в виде [14]:

$$O = \langle E(AT), ER, AS, EA, F, AR \rangle, \quad (1)$$

где  $E$  – набор сущностей, имеющих значение в разных аспектах ПрО;  $AT$  – множество атрибутов сущностей;  $ER \subseteq E \times E$  – множество отношений сущностей;  $AS$  – множество аспектов предметной области;  $EA \subseteq E \times AS$  – проекция сущностей на аспекты;  $F: E \times ER$  – функции интерпретации отношений и сущностей;  $AR \subseteq AS \times AS$  – пересечение аспектов. Тогда многоаспектное семантическое пространство можно выразить набором

$$SS = \langle O, M \rangle, \quad (2)$$

где  $M$  – набор метрик для вычисления степени близости сущностей и аспектов. Пространство  $SS$  условно разделено на пересекающиеся подпространства, каждое из которых соответствует одному аспекту ПрО.

Задача автоматизированного выбора маршрута должна решаться, как задача многокритериального выбора подмножества адресов согласования и визирования документа на основе имеющейся информации о содержании документа, источнике документа, адресате и функциональности подразделений. Сформируем критерии отбора адресов согласования и визирования документа.

*Формирование тематического критерия для выбора маршрута.* Пусть  $T$  – текст, выражающий краткое содержание документа, занесенное в регистрационную карточку. В этом тексте имеются слова и словосочетания из множества  $X = \{x_1, \dots, x_i, \dots, x_m\}$ ,  $i=1..m$ , имеющие скрытые отношения с аспектами предметной области. Будем считать, что все эти слова (за исключением стоп-слов) являются выражениями сущностей и их отношений, то есть, ключевыми словами (КС). Общая задача тематической поддержки маршрутизации документа может быть решена при использовании некой функции ассоциации  $\varphi: X \rightarrow SR$ , где  $SR$  – суперпозиция отношений  $EA$ .

Построим матрицу-отношение  $H = \left\| h_{ij} \right\|$  «ключе-

вые слова–аспекты», где  $h_{ij} = 1$ , если  $i$ -е ключевое слово встречается в  $j$ -м аспекте, и  $h_{ij} = 0$ , у противном случае.

Построим структурную матрицу взаимосвязей текста  $T$ , модели выявления ассоциаций, и аспектов ПрО (табл. 1). Модель преобразования  $M$  должна обрабатывать текст  $T$ , и вычислять степень ассоциации текста  $T$  с каждым аспектом ПрО.

Таблица 1 – Структурная матрица взаимосвязей ключевых слов, модели преобразования и аспектов

Входы	Мо- дель	Выходы					
		$F_1$	$F_2$	..	$F_j$	..	$F_n$
$x_1$	$M$	$Y_{11}$	$Y_{12}$	..	$Y_{1j}$	..	$Y_{1n}$
$x_2$		$Y_{21}$	$Y_{22}$	..	$Y_{2j}$	..	$Y_{2n}$
..		..	..	..	..	..	..
$x_i$		$Y_{i1}$	$Y_{i2}$	..	$Y_{ij}$	..	$Y_{in}$
..		..	..	..	..	..	..
$x_m$		$Y_m$	$Y_{m2}$	..	$Y_{mj}$	..	$Y_{mn}$

Выходы системы  $F_j$ ,  $j = \overline{1, n}$  – степени значимости  $j$ -х аспектов, полученные как суперпозиции степеней значимости КС  $x_i$  относительно аспектов  $A_j$ .

Модель  $M$ , описывающая связь КС  $x_i$  и выходной величины  $Y_{ij}$ , представим отображением:

$$M: x_i \rightarrow Y_{ij}. \quad (3)$$

Величину  $Y_{ij}$  можно интерпретировать как весовой коэффициент, учитывающий значимость  $i$ -го слова в  $j$ -м аспекте:

$$Y_{ij} = \left(1 - \frac{n_i}{n}\right) h_{ij}, \quad (4)$$

где  $n_i$  – число аспектов, в которых фигурирует слово  $x_i$ ,  $n$  – общее число аспектов. Значение  $Y_{ij}$ , полученное по выражению (4), является начальным и может быть скорректировано во время настройки системы в пределах  $[0, 1]$ .

Вернемся к функции ассоциации  $\varphi$ , которая, как теперь ясно, должна осуществлять суперпозицию величин  $Y_{ij}$  и играть роль тематического критерия отбора аспектов. Раскроем содержание функции  $\varphi$ :

$$\varphi_j(x) = F_j(Y_{ij}) = \frac{1}{m} \sum_{i=1}^m Y_{ij}. \quad (5)$$

Выражение (5) – тематический критерий для выбора маршрута, который позволяет вычислить интегральное значение степени тематической ассоциации содержания документа с каждым  $j$ -м аспектом предметной области.

Ранжирование аспектов по степени ассоциации позволяет выбрать наиболее важные аспекты, определить соответствующие полномочия подразделений и выбрать подмножество подразделений для параллельной адресации. Критерием адекватности модели является наличие в вершине списка подразделений подразделения-получателя. Из подмножества близких по ассоциативным связям подразделений формируется список на предъявление пользователю в режиме диалога.

*Формирование критерия интенсивности переписки по аспектам.* Во время обследования предприятия менеджерам предлагается оценить по балльной шкале интенсивность переписки по имеющимся аспектам деятельности (полномочиям).

Используется стандартная процедура преобразования балльных оценок в весовые коэффициенты значимости аспектов по критерию интенсивности переписки. Процедура предусматривает:

1. Вычисление весовых коэффициентов  $\gamma_j$  по формуле:

$$\gamma_j = \frac{I_j}{\sum_{j=1}^n I_j}, \quad (6)$$

где  $I_j$  – оценки интенсивности переписки по  $j$ -му аспекту в баллах.

2. Ранжирование аспектов по убыванию оценок интенсивности  $\gamma_j$ .

Оценки по данному критерию учитываются путем вычисления скалярного произведения с вектором оценок по тематическому критерию, о чем будет сказано ниже, при описании метода маршрутизации.

*Формирование критерия регулярных контактов.* Критерий предназначен для анализа и отбора подразделений, у которых есть регулярная переписка и с подразделением-источником, и с подразделением-получателем.

Пусть имеется распределение  $S_k$  интенсивностей (относительных частот) контактов  $k$ -го подразделения-источника с другими подразделениями в виде гистограммы  $S_k = \{s_{kl}, \dots, s_{kb}, \dots, s_{kp}\}$ ,  $k \neq l$ . Следует установить пороговое значение интенсивности связей  $IT$  и формировать подмножество подразделений, предположительно входящих в зону ответственности по условию

$$p_l \in P_k | s_{kl} \geq IT. \quad (7)$$

Выражение (7) является критерием отбора подразделений по принципу регулярных контактов с источником.

Аналогично происходит отбор подразделений по принципу регулярных контактов с получателем и формируется подмножество  $P_z$  на основании распределения относительных частот  $S_t = \{s_{tz}, \dots, s_{tz}, \dots, s_{tz}\}$ ,  $t \neq z$ :

$$p_z \in P_z | s_{tz} \geq IT, \quad (8)$$

где  $z$  – индекс подразделения, имеющего регулярные контакты с получателем;  $t$  – индекс подразделения-получателя.

Сформируем подмножество  $P_{tz}$  подразделений, входящих в оба подмножества:

$$p_{tz} \in P_{tz} | (p_l \in P_z) \cap (p_z \in P_l). \quad (9)$$

Распределения относительных частот контактов можно интерпретировать как функции принадлежности нечетких подмножеств. Тогда подмножество  $P_{tz}$  формируется как нечеткое пересечение подмножеств  $P_l$  и  $P_z$ .

Таким образом, сформирован критерий (7)–(9)

регулярных контактов для отбора подразделений в зону ответственности исходящего документа.

*Метод автоматизированного выбора маршрута электронных документов.* Общая последовательность действий по выбору подразделений для параллельной маршрутизации с использованием предложенных выше моделей состоит из нескольких этапов. Многие процедуры этапов алгоритмически повторяются, и это обеспечивает универсальность ограниченного набора классов при программировании.

*Этап 1.* Подготовка тезаурусов по аспектам.

От количества аспектов и качества тезаурусов зависит успешность применения тематического критерия выбора маршрута. Аспекты ПрО выбираются на основе функциональности подразделений и предприятий, с которыми ведется достаточно интенсивная переписка. Среди аспектов необходимо выделить предметные, содержащие возможные ключевые слова заданной предметной области (ПрО) или нескольких областей, и утилитарные, содержащие образы действия или оценки в виде лексем типа «ответ на обращение», «служебное несоответствие» и т.п.

*Этап 2.* Предварительная обработка анализируемого текста.

2.1. Преобразование всех символов к верхнему или нижнему регистру.

2.2. Удаление «стоп-слов». Как правило, это слова, не оказывающие влияния на тематику документа, например, артикли, союзы и предлоги.

2.3. Стемминг. Данная процедура заключается в выделении значимой части слова с помощью отсечения суффиксов и окончаний.

*Этап 3.* Определение значимых аспектов. Это делается путем проверки КС на их наличие в аспектных тезаурусах ПрО. Происходит формирование подмножества значимых аспектов по критерию максимального числа совпадений в тезаурусах.

3.1. Соберем КС в матрицу  $X \times F$  (табл. 1), в которой строки – это КС, а столбцы – аспекты. В качестве меры значимости  $j$ -го аспекта необходимо вычислить для каждого аспекта сумму локальных значимостей:

$$F_j = \frac{1}{m} \sum_{i=1}^m Y_{ij} \cdot \quad (10)$$

Результат – вектор  $F$  значимости аспектов по тематическому критерию.

3.2. Нормируем степени значимости аспектов по тематическому критерию:

$$F_j = \frac{F_j}{\sum_{j=1}^n F_j} \cdot \quad (11)$$

3.3. Вычисляем степени значимости аспектов с учетом критерия интенсивности переписки:

$$F_j^L = F_j \gamma_j \cdot \quad (12)$$

3.4. Отбираем в рабочее подмножество значимые аспекты, используя пороговое преобразование

$$H = \begin{cases} 1 & \text{if } F \geq T \\ 0 & \text{otherwise} \end{cases}, \quad (13)$$

где  $F$  – оценка значимости аспекта;  $T$  – заданный порог.

3.5. Ранжируем аспекты по степени значимости и выделяем подразделения, соответствующие выбранным аспектам/полномочиям. Обозначим зону ответственности документа (ЗОД) – как подмножество подразделений, участвующих в обработке документа –  $ZR$  (zone of responsibility). Построим матрицу проекции множества выявленных из контекста документа аспектов/полномочий (set of established powers)  $SP$  на множество подразделений  $SD$  (projection of authority over subdivisions):  $|g_{jk}| = G = SP \times SD$ .

Обозначим показатель значимости подразделений  $z_k$  – его можно интерпретировать, как весовой коэффициент, учитывающий значимость  $k$ -го подразделения в зоне ответственности:

$$z_k = \frac{1}{m} \sum_{j=1}^m g_{jk} \cdot \quad (14)$$

Нормируем оценки значимости подразделений по формуле, аналогичной (11), и при помощи порогового преобразования (13) формируем множество  $S$  подразделений, значимых по тематическому критерию и критерию интенсивности переписки.

Результат – множество  $S$  подразделений, имеющих значимые ассоциации, с присвоенными весовыми коэффициентами  $z_k$ .

Отметим, что полученные оценки значимости могут быть скорректированы при настройке и обучении системы маршрутизации.

*Этап 4.* Формирование подмножества регулярных контактов.

4.1. Анализ и отбор подразделений, у которых есть регулярные контакты с отправителем. На этом этапе используется критерий (7), позволяющий выделить подмножество  $C_s$  подразделений, для которых частота контактов с отправителем (по сути – вероятность вхождения в ЗОД) наиболее велика. При помощи порогового преобразования (13) отбираем подразделения в подмножество  $C_r$ , вычисляем локальную значимость подразделений по критерию частоты контактов  $f_{cs}$  и нормируем коэффициенты значимости по формуле, аналогичной (11).

4.2. Анализ и отбор подразделений, у которых есть регулярные связи с получателем и формирование подмножества  $C_r$ . Используется тот же механизм, что и в п.4.1 и критерий (8). Вычисляем локальную значимость подразделений по критерию частоты контактов  $f_{cr}$  и нормируем коэффициенты значимости по формуле аналогичной (11).

4.3. Отбор подразделений входящих в оба подмножества. Используя коэффициенты значимости  $f_{cs}$  и  $f_{cr}$ , вычисляем пересечение  $C_s \times C_r$  по условию (8). Для этого вычисляются совместные коэффициенты значимости контактов:

$$b_k = f_{crk} f_{csk} \cdot \quad (15)$$

При помощи порогового преобразования (13) отбираем подразделения в подмножество  $C$  значимых по критерию частоты контактов. Нормируем коэффициенты значимости  $b_k$  по формуле, аналогичной (11).

Результат – множество  $C$  подразделений, имеющих значимые регулярные письменные контакты, как с отправителем, так и с получателем с присвоенными весовыми коэффициентами  $b_k$ .

*Этап 5.* Объединение множеств подразделений, полученных по локальным критериям.

5.1. Два множества подразделений  $S$  и  $C$  частично пересекаются:  $S_0 = S \cap C \neq \emptyset$ . Стратегия объединения указанных множеств в общее множество  $ZR$  следующая:

5.1.1. Вычисляем пересечение множеств  $S_0 = S \cap C$ . Выбираем из множества  $S_0$  элементы подмножества  $C_0$ :  $C_0 \subseteq S_0 = \{c \mid c \in S_0\}$ , то есть в подмножество  $C_0$  входят только те элементы множества  $C$ , которые учтены в множестве  $S$ .

5.1.2. Для элементов подмножества  $C_0$  корректируем коэффициенты значимости:

$$z_k^c = z_k^c + ab_k^c, \quad (16)$$

где  $\alpha < 1$  – константа, значение которой подбирается экспериментально.

5.1.3. Ранжирование элементов множества  $S$  и предъявление списка подразделений, потенциально участвующих в согласовании и обсуждении документа, пользователю.

*Методика настройки и обучения системы маршрутизации.* Для настройки системы необходимо иметь заполненную базу данных по предметной области, сформированные тезаурусы по аспектам, список подразделений с привязкой к аспектам/полномочиям.

1. Формирование примеров для обучения системы. Каждый пример содержит данные из регистрационной карточки документа – краткое содержание и конечный адрес документа.

2. Системе предъявляются примеры из обучающего множества, и для каждого примера она формирует ранжированный список подразделений с определенной долей вероятности попадающих в ЗОД. Список разделен на три категории: первая категория – подразделения с высокой вероятностью нахождения в ЗОД; вторая категория – подразделения со средней вероятностью нахождения в ЗОД; третья категория – подразделения с низкой вероятностью нахождения в ЗОД.

3. Экспертная оценка правильности выбора подразделений и фиксация ошибок в категоризации подразделений. Величина ошибки рассчитывается как расстояние Левентштейна, так как требуется коррекция списка путем перестановок кодов подразделений. Полная ошибка на обучающем множестве рассчитывается как усредненное количество частных ошибок на отдельных примерах.

4. Корректировка значений величин  $z_k$  по формулам

$$z_k^d = z_k - \beta, \quad (17)$$

$$z_k^u = z_k + \beta, \quad (18)$$

где  $z_k^d$  – значимость подразделения, которое нужно переместить в нижнюю категорию;  $z_k^u$  – значимость

подразделения, которое нужно переместить в верхнюю категорию;  $\beta < 1$  – константа, значение которой подбирается экспериментально.

*Эксперименты и предварительные результаты.*

Для проверки работоспособности моделей и метода была разработана база данных и программа, состоящая из нескольких модулей. Тезаурусы заполнялись вручную, на основе системного анализа деятельности подразделений городского исполнительного комитета. В качестве экспертов при составлении тезаурусов привлекались работники исполнительного комитета. Настройка подсистемы автоматизированной маршрутизации осуществлялась по приведенной выше методике.

В базу данных были занесены подразделения, краткое описание их функциональности, заполнены ключевыми словами тезаурусы. Кроме того, использовалась таблица типовых поручений по обработке документов (ознакомление, согласование, выполнение).

Краткие аннотации документов были извлечены из регистрационных карточек существующей системы электронного документооборота и занесены в текстовый файл, откуда программа могла их выбирать случайным образом. Из регистрационных карточек и книг регистрации входящих и исходящих документов были взяты также конечные адреса документов и частоты контактов подразделений.

Программа выдавала ранжированные списки подразделений зоны ответственности документа, разделяя их на три категории, как описано в п.п. 2 методики настройки.

Правильность выполнения задач маршрутизации проверяли эксперты. Фиксировались ошибки в ранжированных списках. При обнаружении ошибки выполнялась корректировка значений коэффициентов согласно п.п. 3 и 4 методики настройки.

Динамика уменьшения относительного числа ошибок на общем массиве задач (24 задачи) с каждой итерацией показана на рис. 1.

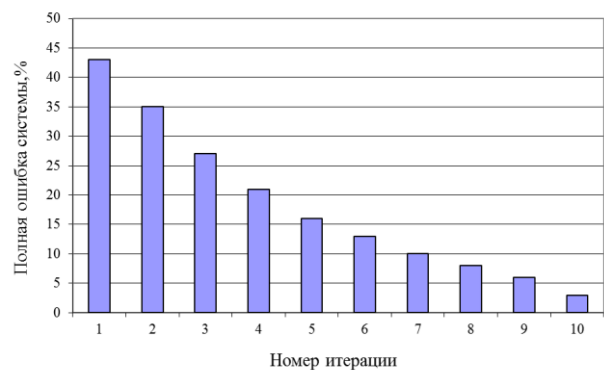


Рисунок 1 – Динамика уменьшения ошибки при формировании зоны ответственности документа

Видно, что система при обучении уменьшает долю ошибок практически по экспоненте, что доказывает правильность предлагаемого подхода и метода.

Предложенный метод отличается простотой и логической прозрачностью и обеспечивает адаптацию подсистемы маршрутизации к условиям пред-

прияття путем дообучення, и ее масштабирование.

**ВЫВОДЫ.** Предложен подход к автоматизации процедуры выбора промежуточных адресов согласования для подсистемы маршрутизации системы электронного документооборота.

Разработаны и формализованы критерии выбора маршрута по тематическому принципу и по частоте регулярных контактов.

Разработано модельное представление процесса выбора маршрута документа, включающее онтологию проблемной области, формальное описание взаимосвязей ключевых слов документа, аспектов функциональности подразделений и критериев выбора маршрута.

Разработан метод автоматизированной параллельной маршрутизации электронных документов, отличающийся тем, что формирование зоны ответственности документа происходит по тематическому критерию и критерию частоты контактов подразделений, что дает возможность уменьшить ошибки пользователя при составлении маршрута и устранить коллизии при прохождении документа.

Разработана методика настройки и обучения подсистемы маршрутизации.

Предложенный метод отличается логической прозрачностью и обеспечивает адаптацию подсистемы маршрутизации к условиям предприятия и простое масштабирование.

#### ЛИТЕРАТУРА

1. Куперштейн В.И. Современные информационные технологии в делопроизводстве и управлении. СПб.: БХВ, 1999. 248 с.
2. Кипенко А.А. Аналитика системы электронного документооборота. *Вестник Новосибирского государственного университета. Серия: Информационные технологии.* 2010. Т. 8. Вып. 2. С. 69–77.
3. Моделі оперативного моніторингу бізнес-процесів у підрозділах підприємств / Шевченко І.В., Скриль О.О., Захарченко Я.Р., Самойленко Ю.М. *Вісник Кременчуцького національного університету імені Михайла Остроградського.* 2016. Вип. 6/2016 (101). С. 86–93.
4. Поліновський В.В., Огурцов М.І. Впровадження системи електронного документообігу в на-

уковій організації. *Вісник Хмельницького національного університету.* 2010. № 4. С. 60–69.

5. Круковский М.Ю. Методология построения композитных систем документооборота. *Математичні машини і системи.* 2004. № 1. С. 101–114.

6. Круковский М.Ю. Графовая модель композитного документооборота. *Математичні машини і системи.* 2005. № 1. С. 120–136.

7. Круковский М.Ю. Метод анализа чувствительности систем электронного документооборота. *Система підтримки прийняття рішень. Теорія і практика (СППР 2010): зб. тез доповідей.* Київ, 2010. С. 127–128.

8. Поліновський В.В., Огурцов М.І. Архітектура та методи систем електронного документообігу наукової установи. *Інженерія програмного забезпечення.* 2012. № 1 (9). С. 66–72.

9. Малахов Е.В., Филатова Т.В., Ахремчук С.Г. Блочная модель маршрутизации документа. *Труды Одесского политехнического университета.* 2008. Вып. 2 (30). С. 125–129.

10. Van der Aalst W. The Application of Petri Nets to Workflow Management. *The J. of Circuits, Systems and Computers.* 1998. № 8. P. 21–66.

11. Завозкин С.Ю. Прототип системы движения документов в системе электронного документооборота КемГУ. *Сб. труд. молодых ученых Кемеровского государственного университета, посвященный 60-летию Кемеровской области.* Кемерово: Полиграф, 2003. С. 67–72.

12. Свиридова О.В., Макушкина Л.А. Программная реализация математической модели документооборота в транспортном отделе предприятия. *Московское научное обозрение.* 2012. № 6. С. 22–25.

13. Авраменко В.П., Калистратова С.Ю. Интеллектуальные процедуры управления бизнес-процессами предприятий. *Бионика интеллекта.* 2008. № 2 (69). С. 73–76.

14. Тертышный В.А., Шевченко И.В. Модель и метод многоаспектного поиска фактографических данных для поддержки принятия решений. *Вісник Кременчуцького національного університету імені Михайла Остроградського.* 2016. Вип. 5/2016 (100). С. 19–25.

#### MODELS AND METHOD OF AUTOMATED CHOICE OF ROUTES ELECTRONIC DOCUMENTS AT THE DEPARTMENTAL LEVEL

**I. Shevchenko, A. Skryl, D. Rastoropov, Yu. Svichensky**

Kremenchuk Mykhailo Ostrohradskyi National University

vul. Pervomayskaya, 20, Kremenchuk, 39600, Ukraine. E-mail: ius.shevchenko@gmail.com

**Purpose.** An approach to automating the routing procedure in the electronic document management system is proposed. The criteria for choosing a route based on the thematic principle and on the frequency of regular contacts were developed and formalized. This allows a multicriteria selection of intermediate addresses on the route. To solve the problem of automated routing, a set of models was developed that converted the values of the primary parameters of the document into a subset of prospective mailing addresses. **Results.** A model representation of the document route selection process, includes a domain ontology, a formal description of the relationship of the document keywords, aspects of the unit functionality and route selection criteria. The thematic criterion for choosing a route and the criterion of the intensity of correspondence, which allow us to calculate the integral value of the degree of association of the contents of the document with each aspect of the subject area, is developed and formally described. Ranking the aspects by the degree of association allows you to select the most important aspects, determine the appropriate authority units and select a

subset of subunits for parallel addressing. The criterion for the adequacy of the model is the presence at the top of the list of subdivisions of the recipient subdivision. The criterion of regular contacts is developed and formally described. The criterion is intended for the analysis and selection of subdivisions that have regular correspondence with both the source subdivision and the recipient subdivision. A method for automated parallel routing of electronic documents has been developed, characterized in that the formation of the responsibility area of the document takes place according to the topic criterion and the frequency of contacts of the units. **Practical value.** As a result, the user receives a ranked list of units potentially entering the responsibility area of the document. This makes it possible to reduce user errors when compiling the route and to eliminate collisions during the passage of the document. The proposed method is characterized by logical transparency and ensures the adaptation of the routing subsystem to enterprise conditions and simple scaling. The technique of tuning and learning the routing subsystem has been developed. References 14, figure 1, table 1.

**Key words:** electronic document management, routing, criteria for selecting addresses, model of choice, automated routing method.

#### REFERENCES

1. Kupershtein, V.I. (1999), *Sovremennyye informatsionnyie tehnologii v deloproizvodstve i upravlenii* [Modern information technologies in office management and management], BHV, St. Petersburg, Russia.
2. Kipenko, A.A. (2010), "Analytics of electronic document management system", *Transactions of Novosibirsk State University. Series: Information technology*, vol. 8, iss. 2, pp. 69–77.
3. Shevchenko, I.V., Skril, O.O., Zaharchenko, Ya.R., Samoylenko, Yu.M. (2016), "Models of operational monitoring of business processes in business units", *Transactions of Kremenchuk Mikhaylo Ostrohradskyi National University*, iss. 6, no. 101, pp. 86–93.
4. Polinovskiy, V.V., Ogurtsov, M.I. (2010), "Implementation of the system of electronic document circulation in a scientific organization", *Transactions of Khmelnytsk National University*, no. 4, pp. 60–69.
5. Krukovskiy, M.Yu. (2004), "Methodology for building composite systems of document circulation", *Mathematical Machines and Systems*, no. 1, pp. 101–114.
6. Krukovskiy, M.Yu. (2005), "Graphic model of composite document circulation", *Mathematical Machines and Systems*, no. 1, pp. 120–136.
7. Krukovskiy, M.Yu. (2010), "Method of analysis of sensitivity of electronic document management systems", *Decision support systems. Theory and practice (DSS 2010): collection of abstracts*, pp. 127–128.
8. Polinovskiy, V.V., Ogurtsov, M.I. (2012), "Archi-tecture and methods of systems of electronic document circulation of a scientific institution", *Software Engineering*, vol. 1, no. 9, pp. 66–72.
9. Malahov, E.V., Filatova, T.V., Ahremchuk, S.G. (2008), "Block routing document model", *Proceedings of Odessa Polytechnic University*, iss. 2, no. 30, pp. 125–129.
10. Van der Aalst, W. (1998), "The Application of Petri Nets to Workflow Management", *The J. of Circuits, Systems and Computers*, no. 8, pp. 21–66.
11. Zavozkin, S.Yu. (2003), "Prototype of the document movement system in the electronic document management system of KemsU", *Collection of works of young scientists of Kemerovo State University, dedicated to the 60th anniversary of Kemerovo region*, pp. 67–72.
12. Sviridova, O.V., Makushkina, L.A. (2012), "Software implementation of the mathematical model of document circulation in the transport department of the enterprise", *Moscow Scientific Review*, no. 6, pp. 22–25.
13. Avramenko, V.P., Kalistratova, S.Yu. (2008), "Intellectual procedures for managing business processes of enterprises", *Bionics of Intellect*, no. 2 (69), pp. 73–76.
14. Tertyishnyiy, V.A., Shevchenko, I.V. (2016), "The model and method of multidimensional search for factor data for decision support", *Transactions of Kremenchuk Mikhaylo Ostrohradskyi National University*, iss. 5, no. 100, pp. 19–25.

Стаття надійшла 09.04.2018.