

ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА И ПРИКЛАДНАЯ МАТЕМАТИКА

УДК 681.325.5

В.Н. Локтюхин, А.В. Антоненко

СИНТЕЗ СТРУКТУРЫ ПРЕОБРАЗОВАТЕЛЯ ФОРМЫ ИНФОРМАЦИИ В ВИДЕ АНАЛОГО-ЦИФРОВОЙ НЕЙРОСЕТИ

Рассматриваются этапы и содержание синтеза структурных моделей преобразователя формы информации (ПФИ), представленных в виде аналого-цифровой нейросети. Приводятся примеры структур одно- и многотактных ИНС-преобразователей, даются оценки затрат на их возможную физическую реализацию.

Ключевые слова: искусственная нейронная сеть, преобразователь формы информации, нейросекция, структурно-алгоритмическая организация.

Введение. Разработка моделей структур нейросетевых преобразователей формы информации (ПФИ), ориентированных на функциональное преобразование аналоговых величин, представленных в виде значений частоты, периода, временного интервала, амплитуды сигнала, в их цифровой эквивалент (и обратно), является одной из важных составляющих создания нейрочипов для имплантантов сенсорных систем человека, нейропроцессоров ввода-вывода аналоговых сигналов, интеллектуальных датчиков и др.

Нейросетевые ПФИ относятся к классу аналого-цифровых (АЦ) искусственных нейронных сетей (ИНС), физическая реализация которых осуществляется с применением аппаратных средств [1]. При разработке преобразователя необходимо на начальных этапах его синтеза сформировать такую структурную модель нейросетевого (ИНС-) преобразователя, которая обеспечивала бы минимизацию затрат на его предполагаемую реализацию.

Необходимость решения указанной проблемы обусловлена наличием большого многообразия в принципах построения входящих в состав разрабатываемого устройства нейронов-преобразователей (НП) и форм представления информации на их входах-выходах. Под НП понимается математический нейрон с гибридной формой представления информации и возможностью ее преобразования данным устройством совместно с выполнением несложной математической операции [2]. Указан-

ное выше обстоятельство приводит к тому, что, например, функциональное (нелинейное) преобразование «аналог – код», обеспечиваемое реализацией нейросетевой аппроксимирующей зависимости, может породить большое число структур ИНС-преобразователя для ее воспроизведения. Вместе с тем, известные методы построения структурных моделей ИНС [1] для решения широкого класса задач базируются в основном на программных методах их реализации и не учитывают многообразия аппаратного построения НП.

В связи с этим целью работы является рассмотрение вопросов синтеза структуры ИНС-преобразователя как АЦ-нейросети и оценки возможных затрат на его дальнейшую реализацию в виде логической (структурной) схемы, например, на базе программируемых логических интегральных схем (ПЛИС) с использованием языка описания аппаратуры.

Этапы синтеза структурной модели ИНС-преобразователя как АЦ-нейросети. Процедура синтеза структуры преобразователя с использованием крупнозернистого набора нейросетевых операций для описания его модели содержит 5 этапов, основное содержание которых приводится далее.

Этап 1. Выбирается подход к получению структуры ИНС-преобразователя на основе классификации нейросетей по способу решения задач преобразования, в которую входят: формируемые сети, сети с формируемой матрицей свя-

зей, обучаемые сети и комбинированные (смешанные) [3].

Этап 2. Формируется нейросетевая модель синтезируемого устройства в виде исходного математического представления ИНС, выбираемого разработчиком для построения устройства. Так, например, для ИНС с обучением математическая модель структуры ИНС-преобразователя «аналог – код» $x \rightarrow y^*$ или «код – аналог» $x^* \rightarrow y$ может быть представлена в виде двухслойной нейронной сети прямого распространения с линейной функцией активации выходного слоя. Сеть оперирует переменными в заданных диапазонах изменения $x \in [x_{\min}, x_{\max}]$, $y \in [y_{\min}, y_{\max}]$.

Рассматриваемая сеть состоит из двух обрабатывающих слоев математических нейронов ($MH_j^{(1)}$ и $MH_1^{(2)}$) и описывается выражением:

$$y = F_{MH_1}^{(2)} \left(\sum_{j=1}^M \left(w_{MH_j,1}^{(2)} \cdot F_{MH_j}^{(1)} \left(G_{MH_j}^{(1)} \right) \right) \right), \quad (1)$$

где $w_{MH_j,1}^{(2)}$ – весовые коэффициенты нейрона $MH_1^{(2)}$ 2-го выходного слоя сети, $j = \overline{[1, M]}$, M – количество нейронов в 1-ом слое; $F_{MH_j}^{(1)}$ – функция активации j -го нейрона $MH_j^{(1)}$ первого слоя; $F_{MH_1}^{(2)}$ – активационная функция нейрона $MH_1^{(2)}$ выходного слоя сети; $G_{MH_j}^{(1)} = w_{MH_1,j}^{(1)} \cdot x + w_{MH_2,j}^{(1)} \cdot x_0$ – взвешенная сумма сигналов x и x_0 , поступающих с нейронов $MH_1^{(0)}$ и $MH_2^{(0)}$ 0-го (распределительного) слоя сети на входы нейронов $MH_j^{(1)}$ первого слоя; $w_{MH_1,j}^{(1)} = 1$ и $w_{MH_2,j}^{(1)}$ – веса входов нейронов $MH_j^{(1)}$ первого слоя, $j = \overline{[1, M]}$; x – входная величина; x_0 – опорная величина (этапон преобразования).

В частности, при решении задач функционального преобразования информации в качестве функции активации $F_{MH_j}^{(1)}$ целесообразно применение упрощенной сигмоиды

$$F_{MH_j}^{(1)}(G_{MH_j}^{(1)}) = \begin{cases} \frac{G_{MH_j}^{(1)}}{b_j + G_{MH_j}^{(1)}}, & \text{если } G_{MH_j}^{(1)} > 0; \\ 0, & \text{иначе,} \end{cases}$$

где b_j – коэффициент, задающий индивидуальный вид функции $F_{MH_j}^{(1)}$ [4].

Этап 3. Осуществляется выбор метода структурно-алгоритмической организации ИНС-преобразователя.

При выборе структурно-алгоритмической организации устройств, реализующих функцию преобразования импульсно-аналоговых сигналов в код (или обратно) на базе аппаратных средств, целесообразно выделить два основных режима (или принципа) работы (действия) сети нейронов: 1) непрерывного и 2) циклического принципов действия. В соответствии с этими режимами различают: 1) одноктактные ИНС-преобразователи (непрерывного принципа действия); 2) многотактные ИНС-преобразователи (циклического принципа действия).

В одноктактных устройствах заданная функция ПФИ, представленная, например, в виде аппроксимирующей зависимости $y = \varphi(x)$, реализуется в один прием (такт), когда выходной сигнал $y(t_{\text{вых}})$ формируется на выходе устройства начиная с момента $t = t_{\text{ex}}$ изменения входного сигнала $x(t_{\text{ex}})$, т.е. $t_{\text{вых}} = t_{\text{ex}}$. При этом нейросетевой алгоритм вычисления зависимости $y = \varphi(x)$ выражается в организации структуры ИНС-преобразователя в пространстве размещения нейронов, связанных между собой синаптическими связями в соответствии с исходной записью нейросетевого представления аппроксимирующей функции.

В многотактных устройствах для выполнения функции преобразования $y = \varphi(x)$ многократно используется минимум операторов преобразования, образующих операционное устройство ПФИ в виде одной нейросекции, работающей по определенной программе (такт за тактом) под управлением схемы управления сетью (СУС).

Этап 4. Формируется структура ПФИ в виде АЦ-нейросети на основе нейронов-преобразователей.

Этап 5. Производится оценка возможных аппаратных затрат на реализацию устройства.

При необходимости возможен переход к одному из предыдущих этапов с целью их повторения для изменения некоторых исходных предпосылок и получения в итоге структуры ИНС-преобразователя, удовлетворяющей поставленной задаче его разработки.

Формирование модели преобразователя в виде АЦ-нейросети (этап 4). Учитывая многообразие форм представления физических переменных в проектируемом ПФИ, необходимо для его представления в виде АЦ-нейросети перейти от математических переменных, которыми оперируют при описании исходной сети (см. этап 2), к сигналам, имеющим физическую форму представления.

Далее приведены варианты структурных моделей ИНС-преобразователей непрерывного и

циклического принципов действия с учетом форм представления сигнала в виде кода, частоты и временных интервалов (периодов).

Структура ИНС-преобразователя «код – частота» непрерывного принципа действия. Память АЦ-сети для рассматриваемого варианта разбивается по слоям ИНС-преобразователя (рисунок 1), в которых также выделяют одинаковые по структуре M нейронных секций. Использование в них паде-нейронов $НП_j^{(1)}$ «код – частота» ($j = \overline{[1, M]}$) является удобной формой представления j -го нейрона с «упрощенной» сигмоидой при его синтезе и последующей аппаратной реализации. Выходной нейрон-преобразователь $НП_1^{(2)}$ j -й нейросекции осуществляет операцию суммирования частот $f_{wyj}^{(1)}$, которая реализуется посредством сумматора частот.

В итоге ИНС-преобразователь (см. рисунок 1) реализует аппроксимирующую функцию в виде суммы простых дробей

$$f_y = \sum_{i=1}^m \frac{N_{w_{2,1}}^{(2)} \cdot \left(\frac{f_{w_{1,1}}^{(1)} \cdot \frac{N_x}{2^k} - f_0 \cdot \frac{N_{w_{2,i}}^{(1)}}{2^k} \right)}{2^k \cdot \left(N_{b_i} + (N_x - N_{w_{2,i}}^{(1)}) \right)}$$

с осуществлением операции преобразования «код – частота».

Структура многотактного ИНС-преобразователя «код – временной интервал». С целью сокращения аппаратных затрат преобразователь содержит только одну нейросекцию «код – период» ($N_x \rightarrow T_{yj}$) (рисунок 2), которая состоит из двух слоев и работает в циклическом режиме. Первый слой содержит $НП_1^{(1)}$, реализующий дробно-линейную функцию вида

$$T_{yj} = (N_x - N_{a0j}) / (f_0 N_x / 2^k + f_0 N_{b0j} / 2^k), \quad (2)$$

что позволяет создать ИНС-преобразователь (рисунок 2) «код – период» ($N \rightarrow T$), базирующийся на представлении нейросетевой функции в виде суммы простых дробей. Второй слой является суммирующим. Он формирует на основе $НП_1^{(2)}$ временной интервал T_y путем последовательного суммирования интервалов T_{yj}

($j = \overline{[1, m]}$, m – количество тактов преобразования), образуемых в j -х тактах. В итоге за m тактов устройство реализует функциональную зависимость в виде суммы простых дробей:

$$T_y = \sum_{j=1}^m \frac{N_x - N_{a0j}}{f_0 N_x / 2^k + f_0 N_{b0j} / 2^k}. \quad (3)$$

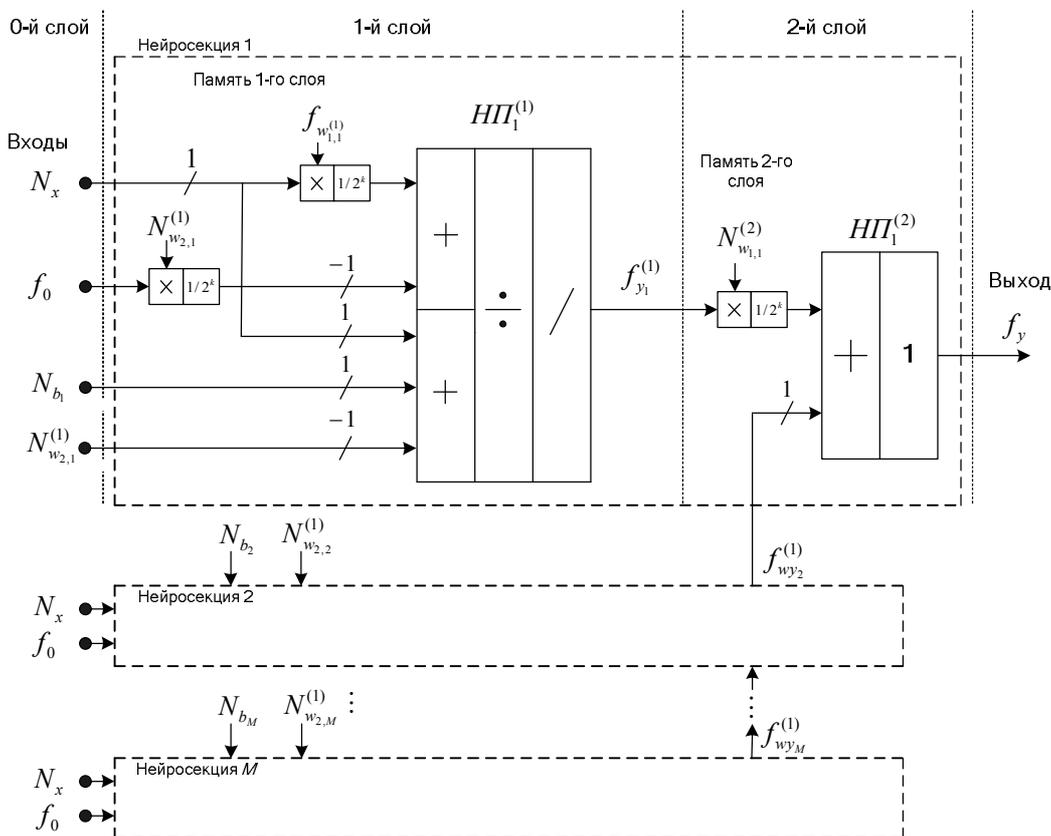


Рисунок 1 – Структура однотоктного ИНС-преобразователя «код – частота»

Время преобразования T_{np} преобразователя (рисунок 2) определяется текущим значением формируемого периода T_y .

Алгоритмическая организация многотактного преобразователя описывается содержательным (рисунок 3, а) и закодированным (рисунок 3, б) графами.

Содержательным является граф [5], операторные и условные вершины которого содержат полное описание операций и булевых выражений. В закодированном графе операции представляются наименованиями управляющих сигналов q_1 , а логические условия – наименованиями осведомитель-

ных сигналов q_{y_j} . Таким образом, согласно содержательному графу (рисунок 3, а), схема управления сетью (СУС) осуществляет начальную загрузку кода N_{a0j} в счетчик, находящийся в составе $НП_1^{(1)}$ нейросекции «код – период» (рисунок 2), – тем самым СУС запускает j -й такт работы нейросекции. Также СУС производит счет тактов j работы нейросекции, $j = [\overline{1, m}]$, m – количество тактов преобразования. Начальная загрузка кода N_{a0j} , согласно закодированному графу (рисунок 3, б), осуществляется сигналом q_1 .

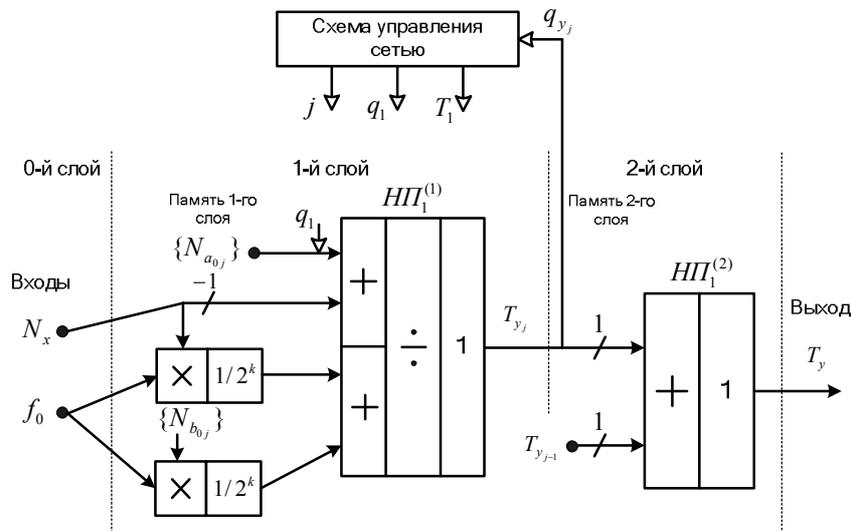


Рисунок 2 – Структура ИНС-преобразователя «код – период»

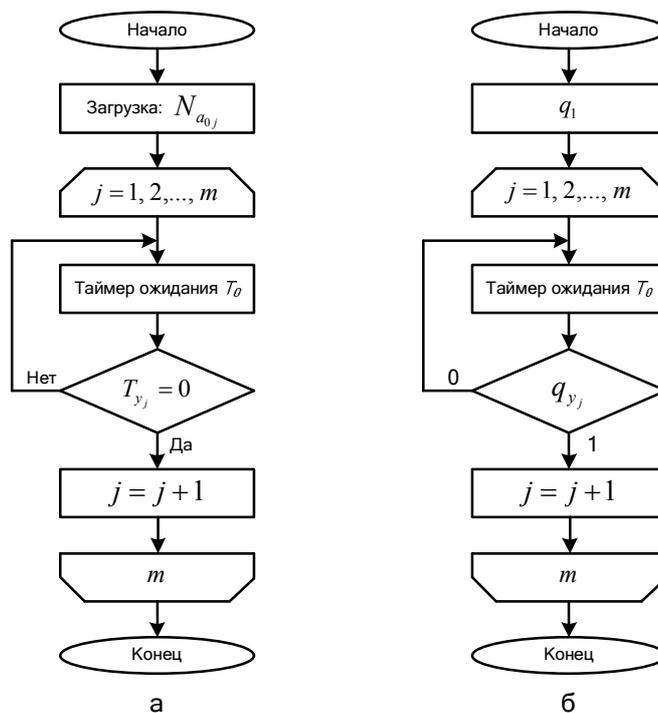


Рисунок 3 – Содержательный (а) и закодированный (б) графы работы многотактного ИНС-преобразователя «код – период»

Оценка аппаратных затрат на реализацию ИНС-преобразователя (этап 5). При синтезе ИНС-преобразователя следует учитывать ряд параметров (характеристик) его структуры, влияющих на предполагаемые аппаратные затраты $C^{(l)} = \sum_{i=1}^{\mu_l} c_i^{(l)}$ на реализацию устройства [6]. В их качестве могут выступать:

1) число $n = \sum_{i=1}^L \mu_i$ нейронов-преобразователей $НП_i^{(l)}$ ($i = [1, \mu_l]$) L -слойной сети без учета 0-го слоя, так как в случае выполнения им только распределительных функций принимается, что условная цена 0-го слоя $c^{(0)} = 0$; μ_l – количество нейронов в l -м слое;

2) количество $V = \sum_{l=1}^L V^{(l)}$ ненулевых аппроксимирующих w и согласующих w_c синаптических связей сети нейронов-преобразователей, составляющих ее память, где $V^{(l)}$ – число этих связей, идущих от нейронов $(l-1)$ -го слоя к нейронам l -го слоя (память l -го слоя) с учетом в $V^{(l)}$ количества единичных $V_1^{(l)}$ и неединичных $V_w^{(l)}$ связей: $V^{(l)} = V_1^{(l)} + V_w^{(l)}$; $V^{(l)} \leq (M^{(l)} \times K^{(l)})$, где $M^{(l)}$ и $K^{(l)} = \mu_l$ – число входов и выходов l -го слоя сети;

3) аппаратные затраты на построение СУС для многотактных ИНС-преобразователей, выраженные с помощью условной цены c_y .

При генерировании в ходе синтеза новой структуры преобразователя, особенно на стадии обучения (этап 2), эти параметры выступают в качестве первичных критериев, определяющих ее поиск в направлении минимизации возможных затрат C на его реализацию.

Для конкретной структуры ИНС-преобразователя элемент $c_i^{(l)} = c_{i,l}$ отражает аппаратные затраты на реализацию нейрона-преобразователя $НП_i^{(l)}$. Затраты $c_{i,l}$ предлагается рассчитывать как:

$$c_{i,l} = c_{i,l}^{(W)} + c_{i,l}^{(F)},$$

где $c_{i,l}^{(W)}$ – затраты (или цена) на формирование всех синаптических связей $w_{j,i}^{(l)} \cdot y_j^{(l-1)}$, идущих от $НП_j^{(l-1)}$ ($j = 1, 2, \dots, \mu_{l-1}$) предыдущего слоя к $НП_i^{(l)}$; $c_{i,l}^{(F)}$ – затраты на реализацию активационной функции нейрона-преобразователя $НП_i^{(l)}$.

Величина условных затрат $c_{i,l}^{(W)}$ зависит от формы представления компонентов произведений (синаптических связей) $w_{j,i}^{(l)} \cdot y_j^{(l-1)}$, а также от конкретных значений весов $w_{j,i}^{(l)}$, задаваемых при обучении сети, и $w_{c_{j,i}}^{(l)}$, согласующих синаптических связей, причем затраты на связи с весами $w = 0$ и $w = |1|$ принимаются равными нулю. С учетом этого, если затраты $c_{j,i}^{(W)}$ на реализацию любой синаптической связи с $w_{j,i}^{(l)} \neq |1|$ и $w_{c_{j,i}}^{(l)} \neq |1|$ равны некоторой постоянной величине $c_{(l)}^{(W)}$, то суммарные затраты $c_{\Sigma(l)}^{(W)}$ на построение всех связей между $(l-1)$ -м и l -м слоями сети целесообразно определять как:

$$c_{\Sigma(l)}^{(W)} = V^{(l)} \cdot c_{(l)}^{(W)}.$$

На величину затрат $c_{i,l}^{(F)}$ влияют форма представления, число входных сигналов, количество и величины параметров активационной функции $F^{(l)}$, а также ее вид. Причем частотно- и время-импульсный способы представления сигналов предопределяют возможность реализации $F^{(l)}$ с достаточно малыми затратами. Поскольку затраты $c_{i,l}^{(F)}$ на реализацию функций активации НП в l -м слое ИНС-преобразователя могут быть различны, то суммарные затраты $c_{\Sigma(l)}^{(F)}$ на реализацию всех μ_l НП в этом слое наводятся как: $c_{\Sigma(l)}^{(F)} = \sum_{i=1}^{\mu_l} c_{i,l}^{(F)}$. Цена $c_{i,l}^{(F)}$ уменьшается с сокращением числа составляющих аргумента $G_i^{(l)}$ функции активации НП с одновременным увеличением на его входе количества связей с единичными весами $w_{j,i}^{(l)} = |1|$. С учетом найденных величин $c_{\Sigma(l)}^{(W)}$ и $c_{\Sigma(l)}^{(F)}$ затраты на реализацию сети можно определять как:

$$C = \sum_{l=1}^L \left((V^{(l)} - V_1^{(l)}) \cdot c_{(l)}^{(W)} + \sum_{i=1}^{\mu_l} c_{i,l}^{(F)} \right) + c_y. \quad (4)$$

Зависимость (4) может выступать в качестве целевой функции $\min \leftarrow C$ синтеза ИНС-преобразователя на этапе обучения сети. Это связано с тем, что ее аргументами являются элементы (вид и число нейронов и синаптических связей между ними), определяющие искомую структуру преобразователя.

Для однократных и многотактных ИНС-преобразователей значения параметров, отражаю-

щих возможные затраты на их реализацию для одной и той же функции преобразования, приведены в таблице.

Таким образом, при синтезе ИНС-преобразователя как L -слойной аппроксимирующей сети необходимо уменьшать значения параметров $V^{(l)}$ и μ_l , составляющих содержание V и n соответственно и максимизировать при этом характеристику $V_1^{(l)}$ – как составляющую $V^{(l)} = V_1^{(l)} + V_w^{(l)}$.

Оценка эффективности применения того или иного способа организации нейросетевой структуры (однотактной или многотактной) может быть произведена на основе функций выигрыша λ_n , λ_V и λ_{V-V_1} нейросетевых структур [4]. Это – функции выигрыша следующего вида:

$$\lambda_n = n_{OT} / n_{MT},$$

$$\lambda_V = V_{OT} / V_{MT},$$

$$\lambda_{V-V_1} = (V - V_1)_{OT} / (V - V_1)_{MT},$$

где n_{OT} и n_{MT} – число нейронов; V_{OT} и V_{MT} – число ненулевых синаптических связей; $(V - V_1)_{OT}$ и $(V - V_1)_{MT}$ – число неединичных (существующих) синаптических связей однотактной (OT) и многотактной (MT) сетей соответственно.

Оценка возможных аппаратных затрат на реализацию ИНС-преобразователя

Параметр сети	Значения параметров сети					
	Однотактный ПФИ			Многотактный ПФИ		
	Слой 1	Слой 2	СУС	Слой 1	Слой 2	СУС
Число n нейронов сети без учета 0	M	M	0	1	1	c_y
Число V ненулевых синаптических связей	$5 \cdot M$	$2 \cdot M$		4	2	
Число V_1 единичных синаптических связей из V	$3 \cdot M$	M		2	2	

Библиографический список

1. Галушкин А.И. Нейронные сети: основы теории. М.: Горячая линия–Телеком, 2010. 496 с.
2. Локтюхин В.Н., Челебаев С.В., Антоненко А.В. Методика построения функциональных моделей нейронных преобразователей // Вестник РГРТУ, Рязань, 2012. Часть 1. № 4. С. 26-32.
3. Комарцова Л.Г., Максимов А.В. Нейрокомпьютеры: учеб. пособие для вузов. – 2-е изд., перераб. и доп. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2004. – 400 с.

С учетом полученных соотношений при условии, что сравниваемые структуры удовлетворяют решению задачи преобразования с заданной погрешностью, следует отдать предпочтение структурной модели преобразователя с циклическим принципом действия (см. рисунок 2).

Заключение. 1. Предложены этапы синтеза структурной модели функционального преобразователя формы информации как аналого-цифровой нейросети.

2. Рассмотрены два подхода к структурно-алгоритмической организации работы ИНС-преобразователя на основе непрерывного и циклического принципов действия, отличием которых является распределение выполнения нейроопераций в пространстве и во времени.

3. Предложена методика оценки и сравнительного анализа аппаратных затрат на реализацию синтезируемой структуры ИНС-преобразователя.

4. В качестве примера рассмотрены структуры преобразователей «код – частота» и «код – период», произведена оценка аппаратных затрат. Показано что структура многотактного преобразователя существенно выигрывает по сравнению с однотактным устройством по аппаратным затратам.

Работа выполнена при финансовой поддержке Российского Фонда фундаментальных исследований РАН.

4. Локтюхин В.Н., Челебаев С.В., Антоненко А.В. Синтез структурной модели ИНС-преобразователя с использованием нейронов с «упрощенными» сигмоидальными функциями // Вестник Саратовского государственного технического университета. 2012. № 2 (65). Выпуск 1. С. 111-117.

5. Майоров С.А., Новиков Г.И. Принципы организации цифровых машин. Л.: «Машиностроение» (Ленингр. отд-ние), 1974. 432 с.

6. Локтюхин В.Н., Челебаев С.В. Принципы применения технологии искусственных нейронных сетей для проектирования преобразователей частотно-временных параметров сигналов в код // Датчики и системы. 2006. № 2. – С. 43-56.

УДК 621.398:651; 62-519:651

А.В. Кузнецов, О.О. Басов, П.К. Толмачев

**ОЦЕНКА ДОСТОВЕРНОСТИ ИНФОРМАЦИИ,
СОДЕРЖАЩЕЙСЯ В ДОКУМЕНТАХ
НА БУМАЖНЫХ НОСИТЕЛЯХ**

Рассмотрена возможность оценки достоверности информации, содержащейся в документах на бумажных носителях без использования специальных технических средств и программного обеспечения. Предложены математическая модель, критерий и алгоритм оценки достоверности информации, содержащейся в документах на бумажных носителях, обеспечивающие вероятность правильной оценки 97 % при времени выполнения процедуры до 1 мин.

Ключевые слова: достоверность информации, документ на бумажном носителе, морфологическая обработка, распознавание образов.

Введение. Информационное обеспечение деятельности аппарата управления, его документирование, хранение и использование ранее созданных документов в организации реализуется посредством документооборота. При этом основным инструментом в документообороте является документ и содержащаяся в нем информация.

Развитие электронных технологий послужило толчком к появлению электронного документооборота, однако документы на бумажных носителях (ДБН) по-прежнему представляют большую ценность, а «электронные архивы» образуют дублирующую систему. И если современные способы защиты электронных документов близки к совершенству, то достоверность информации, содержащейся в ДБН, под которой понимается степень надежности информации, находится на низком уровне по причине возросшего влияния человеческого фактора. Нарушение достоверности информации в таких документах наиболее вероятно в процессе доставки оригинала текстового документа после прохождения им всех согласований, подписаний и утверждений ответственным лицам. Существующие на сегодняшний день методы полиграфической защиты ДНБ (Bugter F, Kezer K., Архипов О.П., Бородина Л.Н., Зыкова З.П., Богданова В.Н., Барсукова В.С., Иванова М.А.) в основном обеспечивают требуемую достоверность информации, но в большинстве случаев не оправдывают себя по причине дороговизны и узкой специализации реализующих их средств [1-4].

Целесообразным подходом в возникшей ситуации является использование для оценки достоверности информации, содержащейся в ДБН, традиционных средств офисной техники (ПЭВМ,

лазерного принтера и планшетного сканера), а целью работы является разработка соответствующего алгоритма оценки, реализующего установление схожести документа при его получении для ознакомления ответственным исполнителем с цифровой копией его оригинала, согласованного, подписанного и утвержденного ответственными лицами оригинала, единожды занесенного в электронную базу данных.

Постановка задачи. При такой интерпретации задача оценки достоверности информации может трактоваться как одна из задач распознавания образов. Процесс распознавания в этом случае состоит в том, что на основании сопоставления апостериорной информации относительно каждого поступившего на вход системы объекта (отсканированного бумажного документа) с априорным описанием единственного класса, соответствующего оцифрованному оригиналу, принимается решение о принадлежности этого объекта к указанному классу. Формальная постановка задачи распознавания в рассматриваемом случае выглядит следующим образом.

Пусть задано множество объектов $\Omega = \{\omega_1, \dots, \omega_m\}$, представляющих собой цифровые изображения, полученные в результате многократного сканирования оригинала текстового документа и его копий, содержащих различные варианты частичной подделки (частичная или полная замена слов, строк, абзацев и страниц текста), пусть также определено множество возможных решений $L = \{l_1, \dots, l_k\}$, которые могут быть приняты системой, где решения $l \in L$ определяют степень отличия копии документа от его оригинала. Например, в простейшем случае можно использовать $L = \{l_1, l_2\}$, где решение l_1

указывает на полное соответствие копии документа его оригиналу (вероятность схожести $P_{\text{схож}} = 1$), а l_2 – на любые случаи подделки ($P_{\text{схож}} < 1$), или каждым решением $l \in L$ определять степень отличия (на основе вероятностной меры) копии документа от его оригинала.

Введем в рассмотрение множество возможных вариантов разбиения объектов на классы $A = \{A_1, \dots, A_r\}$. Будем полагать, что если выбран вариант разбиения A_α , $\alpha = 1, \dots, r$, то множество Ω подразделяется на m_α классов.

Пусть первоначальная информация позволяет построить априорный вектор признаков $x_a = [x_1, \dots, x_N]$. Информация относительно множества решений $L = \{l_1, \dots, l_k\}$ позволяет произвести исходное разбиение множества объектов на классы, то есть составить априорный алфавит классов.

Обозначим как $P_{\text{пр}}[\Omega_i^{A_i} / x_a]$ оценку апостериорной вероятности правильного решения задачи оценки достоверности информации, содержащейся в ДБН, усредненную по всем возможным значениям признаков априорного словаря, описываемого вектором x_a . Эта оценка может быть получена проведением статистических испытаний (метод Монте-Карло).

В условиях ограничений на вычислительные ресурсы, когда сформировать вектор признаков $x_a = [x_1, \dots, x_N]$ в полном объеме не представляется возможным (или нецелесообразно), приходится его сокращать по сравнению с априорным, то есть переходить от априорного словаря признаков к рабочему.

Рассмотрим вектор $\Lambda = [\lambda_1, \dots, \lambda_N]$, компоненты которого принимают значения 1 или 0 в зависимости от того, используется ли данный признак априорного словаря в рабочем словаре или нет. Кроме того, введем обозначение для рабочего словаря x_p (рабочий словарь представляет собой подмножество множества признаков априорного словаря).

Обозначим T_j время, требуемое для определения x_j -го признака, $x_j \in x_p$, а T_0 – общую величину временных ресурсов, ассигнованных на осуществление всех процедур выделения признаков. Если $\sum_{j=1}^N T_j = T_0$, то в качестве рабочего словаря системы оценки целостности может быть использован априорный словарь. В противном случае затраты на измерение признаков

рабочего словаря определяются величиной $C = \sum_{j=1}^N \lambda_j T_j \leq T_0$.

Обозначим $G(\Omega_i^{A_i})$ выигрыш, связанный с реализацией возможных решений при распознавании объекта ω , отнесенного к классу $\Omega_i^{A_i}$ в варианте классификации A_i . Тогда математическое ожидание выигрыша от выбора варианта A_i при использовании априорного словаря признаков

$$R = \sum_{i=1}^{m_1} P_{\text{пр}}[\Omega_i^{A_1} / x_a] G(\Omega_i^{A_1}).$$

Величину R уместно рассматривать в качестве критерия эффективности системы оценки достоверности информации, содержащейся в ДБН. В условиях ограничений, определяемых величиной T_0 , возникает следующая экстремальная задача: необходимо в пределах T_0 найти такой вариант разбиения объектов на классы и такое пространство признаков, при которых обеспечивается максимальное значение критерия эффективности системы R . Другими словами, необходимо определить $A_\alpha = A^0$ (оптимальный алфавит классов) из множества $A = \{A_1, \dots, A_r\}$ и вектор $\lambda = \lambda^0$ (оптимальный рабочий словарь признаков), которые при наилучшем решающем правиле доставляют экстремальное (максимальное) значение величины R при соблюдении ограничений на величину $T \leq T_0$:

$$\max_{A_\alpha, \lambda} R = \max_{A_\alpha, \lambda} \sum_{i=1}^{m_\alpha} P_{\text{пр}}[\Omega_i^{A_\alpha} / x_a] G(\Omega_i^{A_\alpha})$$

с учетом $T = \sum_{j=1}^N \lambda_j T_j \leq T_0$.

Математическая модель ДБН. Для решения поставленной задачи разработана структурная модель текстового документа, включающая в себя:

1) исходное множество объектов

$$\Omega = \{\omega_1, \dots, \omega_r\};$$

2) множество возможных решений

$$L = \{l_1, \dots, l_k\};$$

3) априорный вектор признаков

$$x_a = [x_1, \dots, x_N];$$

4) меру близости объектов;

5) значения выигрышей, получаемых от принятия конкретных решений из множества $L = \{l_1, \dots, l_k\}$;

6) величину временных ресурсов T_0 , ассигнованных на осуществление процедур выделения признаков.

В качестве признаков априорного словаря формализованы следующие структурные компоненты текстового документа:

1) количество строк: $N_{\text{строк}}$;

2) номера неполных строк:

$$N_{\text{непол. строк}} = [n_{\text{непол. строк } 1}, \dots, n_{\text{непол. строк } i}],$$

где $1 \leq n_{\text{непол. строк}} \leq N_{\text{строк}}$ – компонента вектора, соответствующая порядковому номеру неполной строки (как правило, первая и последняя строка абзаца);

3) количество слов в каждой строке:

$$N_{\text{слов}} = [n_{\text{слов } 1}, \dots, n_{\text{слов } N_{\text{строк}}}],$$

где $n_{\text{слов } i}$ – компонента вектора, равная числу слов в i -й строке;

4) расположение коротких слов:

$$P_{\text{корот. слов}} = (p_{ij})_{i=1, j=1}^{N_{\text{строк}}, m}, \text{ где элемент}$$

матрицы

$$p_{ij} = \begin{cases} 1, & \text{если } j\text{-е слово в } i\text{-й строке короткое;} \\ 0, & \text{в противном случае;} \end{cases}$$

$i = 1, \dots, N_{\text{строк}}, j = 1, \dots, m$, m – максимальное число слов в строке анализируемого текста;

5) площадь слов: $S_{\text{слов}} = (s_{\text{слов } ij})_{i=1, j=1}^{N_{\text{строк}}, m}$,

где элемент матрицы s_{ij} – площадь (число пикселей области) j -го ($j = 1, \dots, m$) слова в i -й ($i = 1, \dots, N_{\text{строк}}$) строке (строки с числом слов меньшим m дополнялись справа нулями);

6) относительное расстояние между словами

$$L_{\text{слов}} = (l_{\text{слов } ij})_{i=1, j=1}^{N_{\text{строк}}, m}, \text{ где элемент мат-}$$

рицы $l_{\text{слов } ij} = \sqrt{(x_{ij} - x_{i1})^2 + (y_{ij} - y_{i1})^2}$; x_{ij} и y_{ij} – горизонтальная и вертикальная координаты центра масс j -го ($j = 1, \dots, m$) слова в i -й ($i = 1, \dots, N_{\text{строк}}$) строке соответственно;

7) количество отверстий в буквах (по строкам) $N_{\text{отв}} = [n_{\text{отв } 1}, \dots, n_{\text{отв } N_{\text{строк}}}]$, где $n_{\text{отв } i}$ –

компонента вектора, равная числу отверстий в буквах слов i -й строки;

8) относительное расстояние между отвер-

стиями в буквах $L_{\text{отв}} = (l_{\text{отв } ij})_{i=1, j=1}^{N_{\text{строк}}, o}$, где

$$l_{\text{отв } ij} = \sqrt{(x_{\text{отв } ij} - x_{\text{отв } i1})^2 + (y_{\text{отв } ij} - y_{\text{отв } i1})^2},$$

$x_{\text{отв } ij}$ и $y_{\text{отв } ij}$ – горизонтальная и вертикальная координаты центра масс j -го ($j = 1, \dots, o$) отвер-

стия в i -й ($i = 1, \dots, N_{\text{строк}}$) строке; o – максимальное число отверстий в строке анализируемого текста;

9) количество вертикальных линий в буквах $N_{\text{верт}} = [n_{\text{верт } 1}, \dots, n_{\text{верт } N_{\text{строк}}}]$, где $n_{\text{верт } i}$ –

компонента вектора, равная числу вертикальных линий в словах i -й строки;

10) относительное расстояние между вертикальными линиями $L_{\text{верт}} = (l_{\text{верт } ij})_{i=1, j=1}^{N_{\text{строк}}, v}$,

где

$$l_{\text{верт } ij} = \sqrt{(x_{\text{верт } ij} - x_{\text{верт } i1})^2 + (y_{\text{верт } ij} - y_{\text{верт } i1})^2},$$

$x_{\text{верт } ij}$ и $y_{\text{верт } ij}$ – горизонтальная и вертикальная координаты центра масс j -й ($j = 1, \dots, v$) вертикальной линии в i -й ($i = 1, \dots, N_{\text{строк}}$) строке; v – максимальное число вертикальных линий в строке анализируемого текста.

Для рассмотренных признаков в среде MATLAB реализованы соответствующие алгоритмы их оценивания, основанные на морфологической обработке оцифрованного текстового документа, в частности, на операциях дилатации и эрозии.

В качестве меры близости признаков объектов выбрана евклидова метрика:

$$\begin{aligned} d^2(\omega_{pk}, \omega_{ql}) &= \sum_{j=1}^{10} (x_j^{(p,k)} - x_j^{(q,l)})^2 = \\ &= (N_{\text{строк}}^{(p,k)} - N_{\text{строк}}^{(q,l)})^2 + \\ &+ \sum_{j=1}^t (n_{\text{непол.строк } j}^{(p,k)} - n_{\text{непол.строк } j}^{(q,l)})^2 + \\ &+ \sum_{j=1}^{N_{\text{строк}}} (n_{\text{слов } j}^{(p,k)} - n_{\text{слов } j}^{(q,l)})^2 + \\ &+ \sum_{j=1}^{N_{\text{строк}}} \sum_{i=1}^m (s_{\text{слов } j, i}^{(p,k)} - s_{\text{слов } j, i}^{(q,l)})^2 + \\ &+ \sum_{j=1}^{N_{\text{строк}}} \sum_{i=1}^m (p_{\text{корот.слов } j, i}^{(p,k)} - p_{\text{корот.слов } j, i}^{(q,l)})^2 + \\ &+ \sum_{j=1}^{N_{\text{строк}}} \sum_{i=1}^m (l_{\text{слов } j, i}^{(p,k)} - l_{\text{слов } j, i}^{(q,l)})^2 + \\ &+ \sum_{j=1}^{N_{\text{строк}}} (n_{\text{отв } j}^{(p,k)} - n_{\text{отв } j}^{(q,l)})^2 + \\ &+ \sum_{j=1}^{N_{\text{строк}}} \sum_{i=1}^o (l_{\text{отв } j, i}^{(p,k)} - l_{\text{отв } j, i}^{(q,l)})^2 + \end{aligned}$$

$$\begin{aligned}
& + \sum_{j=1}^{N_{\text{строк}}} \left(n_{\text{верт},j}^{(p,k)} - n_{\text{верт},j}^{(q,l)} \right)^2 + \\
& + \sum_{j=1}^{N_{\text{строк}}} \sum_{i=1}^o \left(l_{\text{верт},j,i}^{(p,k)} - l_{\text{верт},j,i}^{(q,l)} \right)^2 = d_1^2(\omega_{pk}, \omega_{ql}) + \\
& + d_2^2(\omega_{pk}, \omega_{ql}) + d_3^2(\omega_{pk}, \omega_{ql}) + d_4^2(\omega_{pk}, \omega_{ql}) + \\
& + d_5^2(\omega_{pk}, \omega_{ql}) + d_6^2(\omega_{pk}, \omega_{ql}) + d_7^2(\omega_{pk}, \omega_{ql}) + \\
& + d_8^2(\omega_{pk}, \omega_{ql}) + d_9^2(\omega_{pk}, \omega_{ql}) + d_{10}^2(\omega_{pk}, \omega_{ql}),
\end{aligned}$$

где $p=1, 2, \dots, r$; $q=1, 2, \dots, r$; $k=1, 2, \dots, k_p$;

$l=1, 2, \dots, k_q$; $x_j^{(pk)}$ – значение j -го признака k -го

объекта p -го класса, т.е. объекта ω_{pk} ; $x_j^{(ql)}$ – зна-

чение j -го признака l -го объекта q -го класса, т.е.

ω_{ql} ; $d_j^2(\omega_{pk}, \omega_{ql})$, $j=1, \dots, 10$ – эвклидова мера

близости параметров $N_{\text{строк}}$, $N_{\text{неполн. строк}}$,

$N_{\text{слов}}$, $P_{\text{корот. слов}}$, $S_{\text{слов}}$, $L_{\text{слов}}$, $N_{\text{отв}}$, $L_{\text{отв}}$,

$N_{\text{верт}}$, $L_{\text{верт}}$ k -го объекта p -го класса и l -го объ-

екта q -го класса соответственно.

С учетом различного вклада отдельных параметров в формирование общей меры близости совокупность признаков объектов, используемых в рабочем словаре, описана N -мерным вектором $\Lambda = [\lambda_1, \dots, \lambda_N]$, компоненты которого определяют вес соответствующего признака. С учетом Λ квадрат расстояния между объектами составил:

$$\begin{aligned}
d^2(\omega_{pk}, \omega_{ql}) &= \lambda_1 d_1^2(\omega_{pk}, \omega_{ql}) + \lambda_2 d_2^2(\omega_{pk}, \omega_{ql}) + \\
& + \lambda_3 d_3^2(\omega_{pk}, \omega_{ql}) + \lambda_4 d_4^2(\omega_{pk}, \omega_{ql}) + \lambda_5 d_5^2(\omega_{pk}, \omega_{ql}) + \\
& + \lambda_6 d_6^2(\omega_{pk}, \omega_{ql}) + \lambda_7 d_7^2(\omega_{pk}, \omega_{ql}) + \lambda_8 d_8^2(\omega_{pk}, \omega_{ql}) + \\
& + \lambda_9 d_9^2(\omega_{pk}, \omega_{ql}) + \lambda_{10} d_{10}^2(\omega_{pk}, \omega_{ql}).
\end{aligned} \quad (1)$$

Для определения значений выигрышей, получаемых от принятия конкретных решений из множества $L = \{l_1, \dots, l_k\}$ и определения коэффициентов λ_j , $j=1, \dots, 10$, в работе использовался множественный регрессионный анализ.

Важным условием его применения является независимость и нормальность распределения независимых величин. Для оценки указанных свойств производилось 100-кратное сканирование оригинала текстового документа с разрешением 150 dpi и определение параметров априорного словаря. Анализ полученных результатов показал, что:

1) параметры $d_j^2(\omega, \omega_1)$, $j=1, \dots, 3$ имеют нулевую дисперсию;

2) выбросы, присутствовавшие на гистограммах распределения оставшихся параметров, свойственны изображениям, имеющим значительный поворот ($\geq 2^\circ$) относительно оригинала;

3) распределения мер близости $d_j^2(\omega, \omega_1)$, $j=4, \dots, 10$, исключая выбросы, можно аппроксимировать нормальным законом.

Подбор распределений осуществлен с помощью средства DistributionFittingTool пакета Statistics программы технических расчетов MATLAB. Для проверки гипотезы о нормальности распределения мер близости параметров априорного словаря признаков использовался критерий согласия Колмогорова – Смирнова. При этом для каждого распределения нулевая гипотеза состояла в том, что распределение генеральной совокупности не противоречит стандартному нормальному закону, а альтернативная – в том, что распределение генеральной совокупности противоречит стандартному нормальному закону. Для всех параметров на критическом уровне значимости $p_{кр} = 0,05$ подтвердилась нулевая гипотеза.

Критерии оценки достоверности информации. Выражение (1) позволяет определить меру близости двух текстовых документов, представленных в пространстве признаков рабочего словаря, но не дает оценок достоверности информации, содержащейся в них. Поэтому в работе на основе структурных мер информации, учитывающих только дискретное строение информационного комплекса, в частности, количество содержащихся в нем информационных элементов, предложено оценивать достоверность с использованием коэффициента схожести:

$$K_{\text{схож}} = \frac{N_{\Sigma \text{копия}} - N_{\Sigma \text{изм. копия}}}{N_{\Sigma \text{оригинал}}}, \quad (2)$$

где $N_{\Sigma \text{оригинал}}$ – информационная емкость (общее число символов) текстового документа, занесенного в базу данных организации после прохождения им всех согласований, подписаний и утверждений ответственными лицами (оригинала); $N_{\Sigma \text{копия}}$ – информационная емкость текстового документа, полученного и отсканированного исполнителем (копия); $N_{\Sigma \text{изм. копия}}$ – число информационных элементов (символов) в копии, отличающихся от информационных элементов оригинала.

Кроме того, следует учесть, что оригинал бумажного документа сканируется один раз, поэтому решаемую задачу следует рассматривать как задачу распознавания для случая единственности эталона. Тогда в исходном множестве Ω

определим объект ω_{11} , соответствующий оригиналу документа, и подмножество $\Omega_1 = \{\omega_{11}, \omega_{12}, \dots, \omega_{1k_1}\}$, включающее объекты, соответствующие копиям документа, не содержащим подделки. Для указанного подмножества в рамках решаемой задачи коэффициент схожести $K_{\text{схож}} = 1$.

Для копий документов в решаемой задаче целесообразно для всех объектов, не принадлежащих подмножеству Ω_1 , определять меру близости (1) как расстояние между распознаваемым объектом ω и эталонным объектом ω_{11} , т.е. $d^2(\omega, \omega_{11})$, и установить связь между полученной мерой и коэффициентом схожести следующим образом:

$$K_{\text{схож}} \xrightarrow{d^2(\omega, \omega_{11}) \rightarrow 0} 1.$$

Тогда зависимость выбранной меры близости текстовых документов (1) и критерия достоверности (2) можно определить как

$$K_{\text{схож}} = \lambda_1 d_1^2(\omega, \omega_{11}) + \lambda_2 d_2^2(\omega, \omega_{11}) + \lambda_3 d_3^2(\omega, \omega_{11}) + \lambda_4 d_4^2(\omega, \omega_{11}) + \lambda_5 d_5^2(\omega, \omega_{11}) + \lambda_6 d_6^2(\omega, \omega_{11}) + \lambda_7 d_7^2(\omega, \omega_{11}) + \lambda_8 d_8^2(\omega, \omega_{11}) + \lambda_9 d_9^2(\omega, \omega_{11}) + \lambda_{10} d_{10}^2(\omega, \omega_{11}). \quad (3)$$

Таблица 1 – Значения коэффициентов регрессии

λ_j	Коэффициенты регрессии	Ошибки оценивания	Значения уровней значимости по критерию Стьюдента
λ_0	1,02428583470	0,018664	0
λ_1	0,00656096614	0,003331	0,0492
λ_2	0,00034963509	0,0000247	0
λ_3	0,00086461683	0,00013	0
λ_4	-0,0068071581	0,0014	0
λ_5	-0,0000000005	0,0000000002	0,0099
λ_6	0,00000000545	0,0000000007	0
λ_7	-0,0005825931	0,000022	0
λ_8	0,00000000095	0,0000000004	0,0099
λ_9	0,0000194619	0,000006	0,0018
λ_{10}	0,00000000038	0,0000000002	0,1014

Для определения коэффициентов регрессии в (3) использовался множественный регрессионный анализ признаков 2200 оцифрованных изоб-

ражений оригинала документа и его 30 копий, содержащих подделки (таблица 1). В результате на уровне значимости $p = 0,0000001$ (по критерию Фишера) получен коэффициент множественной корреляции $R_{\text{регр}} = 0,9868$ и признаны значимыми признаки λ_j для $j = 1, \dots, 9$ (исключен признак $L_{\text{верг}}$).

На основании чего критерий оценки достоверности информации, содержащейся в ДБН, представлен в следующем виде:

$$K_{\text{схож}} = \lambda_0 + \sum_{j=1}^9 \lambda_j d_j^2(\omega, \omega_{11}). \quad (4)$$

Полученный критерий (4) позволяет определить конкретные решения из множества возможных $L = \{l_1, \dots, l_k\}$ [5-17].

Алгоритм оценки достоверности информации, содержащейся в ДБН. Разработанная модель позволила представить алгоритм оценки достоверности информации, содержащейся в ДБН, следующим образом (рисунок 1). Для оценки его точности по результатам статистических испытаний использовалась средняя абсолютная ошибка. Полученные результаты (для 3 оригиналов и 12 документов, содержащих частичную подделку) свидетельствуют о высокой точности разработанного алгоритма. Однако разброс оцениваемых параметров для копий документа, не содержащих подделки, указал на необходимость представления множества возможных решений в виде $L = \{l_1, l_2\}$, где решение l_1 ($K_{\text{схож}} \geq 0,97$) указывает на соответствие копии документа его оригиналу с вероятностью правильной оценки 97 % при вероятности ошибок второго рода 3 % (значения получены экспериментально), а l_2 ($K_{\text{схож}} < 0,97$) – на любые случаи подделки, требующие принятия дополнительных мер по обеспечению достоверности.

Произведено сравнение программной реализации [18] разработанного алгоритма с существующими средствами анализа документов (совокупностью программ ABBYY FineReader и Microsoft WORD). При проведении экспериментов использовались следующие показатели: время сканирования одной страницы – 25 с; время сохранения и распознавания одной страницы – 2,5 с (ПЭВМ Pentium 4). Полученные результаты (таблица 2) показали превосходство разработанного прототипа по критерию времени оценивания достоверности (оперативности) и позволили установить требования к величине временных ресурсов в виде $T_0 = 1$ мин безотносительно требований к ПЭВМ.

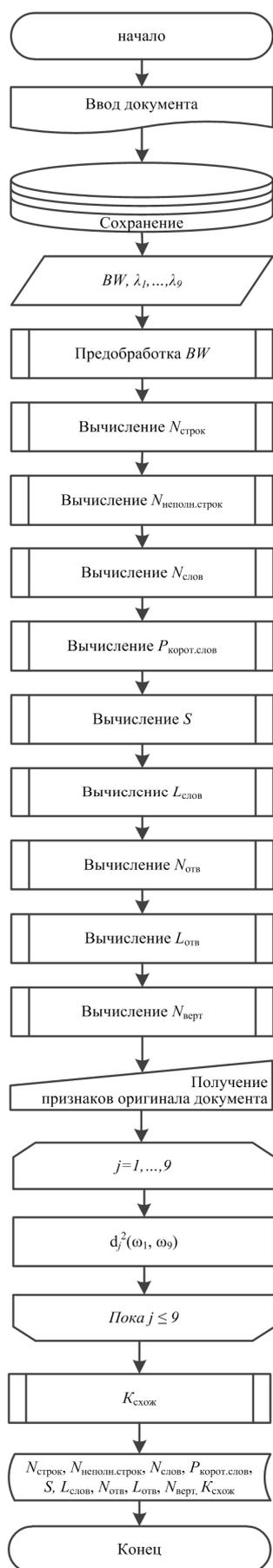


Рисунок 1 – Алгоритм оценки достоверности информации, содержащейся в ДБН

При этом следует учесть, что сравнение документов с помощью Microsoft WORD осуществляется в ручном режиме, а для качественного распознавания текста с помощью ABBYY FineReader необходимо разрешение сканирования 300 dpi, что увеличивает общее время оценки.

Таблица 2 – Результаты обработки документов с помощью программ ABBYY FineReader и Microsoft WORD и разработанного алгоритма

Условный шифр документа	ABBYY FineReader и Microsoft WORD			Разработанный алгоритм	
	Время распознавания, с	Время сравнения, с	Общее время, с	Время обработки, с	Общее время, с
К.1	20,5	2,5	48	19,12	44,12
П.1.1	22,5	2,5	50	20,37	45,37
П.1.2	18,5	2,5	46	20,14	45,14
П.1.3	20,5	2,5	48	19,83	44,83
К.2	14,5	2,5	42	12,94	37,94
П.2.1	15,5	2,5	43	12,53	37,53
П.2.2	16	2,5	43,5	12,78	37,78
П.2.3	16,5	2,5	44	12,48	37,48
К.3	16	2,5	43,5	13,78	38,78
П.3.1	16	2,5	43,5	13,91	38,91
П.3.2	15,5	2,5	43	13,78	38,78
П.3.3	16	2,5	43,5	13,85	38,85
Среднее время:			44,83		40,46

Выводы. Наиболее уязвимым местом в системе оборота ДБН остается целостность информации (нарушаемая при несанкционированном уничтожении, добавлении лишних элементов и модификации данных), а существующие автоматизированные системы контроля и сопровождения документооборота в достаточной степени не решают вопросы оценки достоверности информации, содержащейся в ДБН. В связи с этим в работе на основе разработанных модели ДБН и критерия оценивания достоверности информации, содержащейся в нем, предложен алгоритм такого оценивания с помощью классических средств офисной техники: ЭВМ и планшетного сканера. При времени оценивания до 1 мин разработанный алгоритм обеспечивает точность 97 %.

Библиографический список

1. Кузнецов А.В. Организация сопровождения жизненного цикла документов / А.В. Кузнецов. // Информационные системы и технологии. – Орел: Изд-во ОрелГТУ, 2011. – № 1/(63) – С. 68 – 72.

2. *Архипов О.П.* Технологии создания уникальных растров / О.П. Архипов, П.О. Архипов, Л.Н. Бородин. – М.: ИПИ РАН, 2006. – 76с.

3. *Щербаков А.* Возможности маркеров подлинности по обеспечению непрерывности защиты электронных и бумажных документов / А. Щербаков, Н. Ключко. [Online] 2007. www.directum-journal.ru.

4. *Ларионов В.Г.* Как защититься от подделки? / В.Г. Ларионов, М.Н. Скрыпникова. // Маркетинг в России и за рубежом, №3, 2001. [В Интернете] http://www.mavfiz.ru.

5. *Кузнецов А.В.* Способ определения схожести содержательной части документов / А.В. Кузнецов // Наука и бизнес: пути развития. Труды конференции. – Тамбов, 2011. – № 1. – С.147-155.

6. *Кузнецов А.В.* Метод установления схожести содержательной части бумажного документа с цифровой копией его оригинала / А.В. Кузнецов, О.О. Басов // Информационные системы и технологии. – Орел: ФГОУ ВПО «Госуниверситет-УНПК», 2011. – Т.3 – С. 67-71.

7. *Кузнецов А.В.* Структурная модель текстового документа / А.В. Кузнецов, О.О. Басов, И.В. Блинов // Новые информационные технологии в научных исследованиях. – Рязань, 2011. – С. 286-287.

8. *Бутаков Е.А.* Обработка изображений на ЭВМ / Е.А. Бутаков, В.И. Островский, И.Л. Фадеев – М.: Радио и связь, 1987. – 240с.: ил.

9. *Анисимов Б.В.* Распознавание и цифровая обработка изображений: учеб. пособие для студентов вузов / Б.В. Анисимов, В.Д. Курганов, В.К. Злобин. – М.: Высш. шк., 1983. – 295 с. ил.

10. *Иглин С.П.* Математические расчеты на базе

MATLAB / С.П. Иглин. – СПб.: БХВ-Петербург, 2005. – 640 с.: ил.

11. *Гонсалес Р.* Цифровая обработка изображений в среде MATLAB: пер. с англ. / Р. Гонсалес, Р. Вудс, С. Эддинс. – Техносфера, 2006. – 616с.

12. *Дуда Р.* Распознавание образов и анализ сцен: пер. с англ. / Р. Дуда, П. Харт. – М.: Мир, 1976. – 512с.

13. *Верхаген К.* Распознавание образов: состояние и перспективы: пер. с англ. / К. Верхаген, Р. Дейн, Ф. Грун и др. – М.: Радио и связь, 1985. – 104 с., ил.

14. *Кобзарь А.И.* Прикладная математическая статистика. Для инженеров и научных сотрудников / А.И. Кобзарь. – М.: ФИЗМАТЛИТ, 2006. – 816с.

15. *Халафян А.А.* Statistica 6. Статистический анализ данных / А.А. Халафян. – М.: ООО «Бином-Пресс», 2007. – 512с.: ил.

16. Патент №2482542 РФ. Способ установления подлинности оригиналов бумажных документов / А.В. Кузнецов, О.О. Басов, А.И. Офицеров; заявитель и патентообладатель Академия Федеральной службы охраны Российской Федерации. – заявка № 2011131428; заявл. 26.07.2011; зарег. 20.05.2013.

17. Патент на полезную модель №112790 РФ. Система считывания изображения / А.В. Кузнецов, О.О. Басов, А.И. Офицеров, И.Ю. Баранов; заявитель и патентообладатель Академия Федеральной службы охраны Российской Федерации. – заявка № 2011133337; заявл. 09.08.2011; зарег. 20.01.2012.

18. Свидетельство о государственной регистрации программы для ЭВМ №2011619222 РФ. Морфологическая обработка текстовых документов / заявитель и патентообладатель А.В. Кузнецов, О.О. Басов. – заявка № 2011617543; заявл. 11.10.2011; зарег. 30.11.2011.

УДК 681.317.75:519.2

Д.А. Перепелкин

АЛГОРИТМ ПАРНЫХ ПЕРЕСТАНОВОК МАРШРУТОВ НА БАЗЕ ПРОТОКОЛА OSPF ПРИ ДИНАМИЧЕСКОМ ПОДКЛЮЧЕНИИ УЗЛОВ И ЛИНИЙ СВЯЗИ КОРПОРАТИВНОЙ СЕТИ

Предложен алгоритм адаптивной ускоренной маршрутизации на базе протокола OSPF при динамическом подключении узлов и линий связи корпоративной сети, повышающий эффективность ее функционирования.

Ключевые слова: адаптивная ускоренная маршрутизация, алгоритмы маршрутизации, алгоритм парных перестановок маршрутов, протокол OSPF, динамические изменения, динамические корпоративные сети.

Введение. В настоящее время при проектировании корпоративных сетей большое значение приобретают вопросы эффективности применяемых процессов маршрутизации. Задача маршрутизации в корпоративных сетях решается при условии, что оптимальный маршрут, обеспечивающий передачу пакета за минимальное

время, зависит от топологии сети, пропускной способности и нагрузки на линии связи. Топология сети изменяется в результате подключения узлов и линий связи телекоммуникационной системы.

Применение новых перспективных подходов для решения задачи маршрутизации позволяет

повысить эффективность функционирования корпоративной сети за счет уменьшения трудоемкости построения оптимальных маршрутов.

Цель работы – разработка нового эффективного алгоритма поиска оптимальных маршрутов на базе протокола OSPF при динамическом подключении узлов и линий связи корпоративной сети, повышающего эффективность ее функционирования.

Теоретическая часть. Для повышения эффективности функционирования корпоративных сетей наиболее важной задачей является выбор алгоритма маршрутизации, который будет обеспечивать поиск оптимальных маршрутов с учетом различных свойств той или иной корпоративной сети. В настоящее время широкое применение получили алгоритмы адаптивной маршрутизации. Эти алгоритмы обеспечивают автоматическое обновление таблиц маршрутизации после изменения конфигурации сети.

Протокол OSPF (Open Shortest Path First) при решении задачи маршрутизации базируется на алгоритме состояния каналов. Характеристики и параметры качества обслуживания протокола OSPF подробно рассматриваются в работах [1-3]. Выбор оптимального маршрута в протоколе OSPF определяется по алгоритму Дейкстры. Трудоемкость построения таблиц маршрутизации с использованием данного алгоритма составляет порядка $O(N^2)$, где N – число маршрутизаторов в корпоративной сети.

В работах [4-6] предложен алгоритм парных переходов, позволяющий за счет сбора дополнительной информации учесть возможные изменения параметров линий связи корпоративной сети и не производить полный пересчет маршрутных таблиц. Это позволило снизить трудоемкость расчета таблиц маршрутизации до величины порядка $O(k \cdot N)$, где k – число фактически выполненных парных переходов.

В работах [1-2, 7-9] предложены эффективные алгоритмы адаптивной ускоренной маршрутизации при динамическом подключении узлов и линий связи корпоративной сети, которые также позволили снизить трудоемкость построения оптимальных маршрутов передачи данных до величины $O(k \cdot N)$.

Недостатком предложенных алгоритмов является то, что при динамических изменениях параметров линий связи и в структуре корпоративной сети после каждого парного перехода необходимо рассчитывать дополнительную информацию для того, чтобы определить оптимальный маршрут до других узлов в сети.

В работах [2-3, 10-13] предложен алгоритм парных перестановок маршрутов при динамиче-

ских изменениях параметров линий связи корпоративной сети, что позволило снизить трудоемкость построения оптимальных маршрутов передачи данных до величины $O(N)$. Однако данный алгоритм не учитывает возможность динамического подключения узлов и линий связи корпоративной сети. В связи с этим трудоемкость построения оптимальных маршрутов передачи данных оказывается выше существующих алгоритмов.

Разработка новых более эффективных алгоритмов адаптивной ускоренной маршрутизации позволяет уменьшить трудоемкость построения таблиц маршрутизации до величины $O(N)$ в корпоративных сетях, использующих в своей работе протокол OSPF.

Разработка алгоритма. Для повышения эффективности функционирования корпоративных сетей на базе протокола OSPF предлагается алгоритм парных перестановок маршрутов, который позволяет уменьшить трудоемкость построения таблиц маршрутизации до величины $O(N)$ при динамическом подключении узлов и линий связи корпоративной сети.

В общем случае для решения данной задачи применяется графовая модель корпоративной сети, в которой множество вершин графа соответствует множеству узлов связи или маршрутизаторов в сети, а множество ребер соответствует возможным каналам связи между этими узлами. Каждое ребро, соответствующее каналу связи, имеет свой вес. На практике весу ребра могут соответствовать стоимость аренды канала связи, затраты на оплату единицы трафика, передаваемого по каналу связи, соответствующему данному ребру, либо более сложная функция, учитывающая большее число параметров корпоративной сети.

Представим корпоративную сеть в виде неориентированного взвешенного связного графа $G = (V, E, W)$, где V – множество вершин (узлов связи), $|V| = N$, E – множество ребер (каналов или линий связи), $|E| = M$, W – множество весов ребер (стоимость каналов или линий связи).

Пусть на графе G в некоторый момент времени уже решена задача поиска оптимальных маршрутов до всех узлов множества $V_s = \mathcal{V} \setminus \{v_s\}$ из начального узла v_s , т.е. построено дерево оптимальных маршрутов с корнем в узле v_s . Обозначим это дерево как T_g . На рисунке 1 жирными линиями обозначено построенное дерево оптимальных маршрутов.

Рассмотрим множество каналов E графа G . По признаку вхождения каналов в дерево T_g можно разделить исходное множество E на два подмножества: $E_T \in T_g$ и $E_R \notin T_g$, $E_T \cup E_R = E$.

Множество каналов дерева E_T – множество каналов дерева T_g для графа G . Для заданного графа G согласно свойству дерева мощность множества E_T будет равняться мощности множества V минус единица $|E_T|=|V| - 1$.

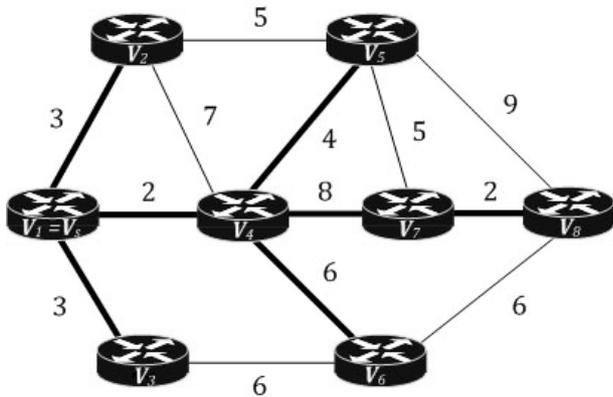


Рисунок 1 – Граф G корпоративной сети

Множество каналов замены E_R для дерева – множество каналов графа G , не вошедших в дерево T_g . При соответствующих условиях некоторый канал $e_{i,j} \in E_R$, инцидентный узлам v_i и v_j , может перейти в множество каналов дерева E_T , заменив собой некоторый канал $e_{k,p} \in E_T$. При этом инцидентность канала $e_{k,p}$ узлу v_i или v_j является обязательным условием. В свою очередь, канал $e_{i,j}$ перейдет в множество E_R .

Будем называть такие переходы парными переходами и обозначать $e_{i,j} - e_{k,p}$.

В множестве E_R можно выделить два подмножества.

Множество каналов замены E_S для дерева – это такое подмножество множества E_R , элементы-каналы которого участвуют, по крайней мере, в одном отношении парного перехода.

Множество непарных каналов E_P – это такое подмножество множества E_R , элементы-каналы которого не участвуют ни в одном отношении из множества R .

В общем случае множество E_P может быть пустым $|E_P| = 0$. Множество E_S будет пустым только при условии, что исходный связный граф G является деревом, и задача поиска оптимальных маршрутов в этом случае лишена смысла.

Для каждого канала связи $e_{i,j} \in E$ на шкале значений весов определены точка вхождения в дерево $w_{i,j}^t$ и точка вхождения в множество замены $w_{i,j}^s$, причем $w_{i,j}^t \leq w_{i,j}^s$, под которыми понимается максимально возможный вес канала $e_{i,j}$ при его вхождении в множество каналов дерева $E_T \in T_g$ и в множество каналов замены для дерева $E_S \notin T_g$ соответственно.

Обозначим $w_{i,j}$ – вес канала, соединяющего узлы v_i и v_j . Узел v_i располагается ниже по иерархии в дереве оптимальных маршрутов отно-

сительно v_j . Множество E_T – множество каналов, каждый элемент которого входит, по крайней мере, в один оптимальный маршрут из начального узла связи, E_R – множество остальных каналов. $E_R \cup E_T = E$, $E_R \cap E_T = \emptyset$. Обозначим V_T – множество узлов, до которых найден оптимальный маршрут из начального узла связи, V_R – множество остальных узлов. $V_R \cup V_T = V$, $V_R \cap V_T = \emptyset$.

Будем называть V_k – маршрутом R_k или совокупностью подмножества $V^{(k)} \subseteq V$ узлов, через которые проходит оптимальный маршрут до узла v_k из исходного узла v_s , и подмножества $E^{(k)} \subseteq E$ каналов, составляющих этот маршрут.

Назовем V_k – деревом T_k или совокупностью подмножества $V_T^{(k)} \subseteq V$, состоящего из всех узлов, оптимальные маршруты до которых из исходного узла содержат узел v_k и подмножества $E_T^{(k)} \subseteq E$ каналов, составляющих эти маршруты после v_k при движении от узла v_s .

Обозначим множество маршрутов до узла v_i из исходного узла v_s через Π_i , где элемент множества $\pi_{i,k} \in \Pi_i$ будет множеством не повторяющихся каналов $e_{i,j} \in E$, образующих вместе маршрут, соединяющий v_s и v_i . Каждому $\pi_{i,k} \in \Pi_i$ поставим в соответствие число, равное сумме весов входящих в него каналов, т. е. длину маршрута $d_{i,k} \in D_i$, где D_i представляет собой множество оценок оптимальных маршрутов до узла v_i из исходного узла v_s . На множестве Π_i задан селектор H , возвращающий оптимальный маршрут из множества Π_i . В том случае, если существует несколько маршрутов в Π_i с минимальной длиной, то выбирается один из них. Оптимальный маршрут до узла v_i будем обозначать $\pi_i = H(\Pi_i)$, оценку его длины – d_i .

Так, для графа G , представленного на рисунке 1, множество каналов дерева составляет $E_T = \{e_{1,2}; e_{1,3}; e_{1,4}; e_{4,5}; e_{4,6}; e_{4,7}; e_{7,8}\}$; множество каналов замены $E_S = \{e_{2,4}; e_{2,5}; e_{3,6}; e_{5,7}; e_{6,8}\}$; множеством непарных каналов будет $E_P = \{e_{5,8}\}$. Если рассмотреть канал связи $e_{4,5}$, то для него точка вхождения в дерево будет составлять 6, а точка вхождения в множество замены – 13. При этом данный канал связи находится в отношении парного перехода с каналом $e_{2,5}$, который, в свою очередь, находится в отношении парного перехода множества замены с каналом $e_{5,7}$. После попадания $e_{4,5}$ в множество непарных каналов эта парная перестановка примет вид: $e_{2,5} - e_{5,7}$.

Рассмотрим работу алгоритма парных перестановок маршрутов на примере графа G корпоративной сети, показанного на рисунке 1, в котором уже решена задача поиска оптимальных маршрутов и построено соответствующее дере-

во. После того как сформирован список парных переходов, дополнительно рассчитываем список оптимальных маршрутов и их маршрутов замены из исходного узла связи до каждого узла сети.

Таким образом, списки оптимальных маршрутов и их маршрутов замены до всех узлов графа корпоративной сети из исходного узла v_s будут сформированы следующим образом.

Узел связи V_2

Маршрут $\pi_2 = \{e_{1,2}\}$ с оценкой $d_2 = 3$ – оптимальный маршрут;

маршрут $\pi_2^{(1)} = \{e_{1,4}; e_{4,2}\}$ с оценкой $d_2^{(1)} = 2 + 7 = 9$ – маршрут замены;

маршрут $\pi_2^{(2)} = \{e_{1,4}; e_{4,5}; e_{5,2}\}$ с оценкой $d_2^{(2)} = 2 + 4 + 5 = 11$ – маршрут замены.

Узел связи V_3

Маршрут $\pi_3 = \{e_{1,3}\}$ с оценкой $d_3 = 3$ – оптимальный маршрут;

маршрут $\pi_3^{(1)} = \{e_{1,4}; e_{4,6}; e_{6,3}\}$ с оценкой $d_3^{(1)} = 2 + 6 + 6 = 14$ – маршрут замены.

Узел связи V_4

Маршрут $\pi_4 = \{e_{1,4}\}$ с оценкой $d_4 = 2$ – оптимальный маршрут;

маршрут $\pi_4^{(1)} = \{e_{1,2}; e_{2,4}\}$ с оценкой $d_4^{(1)} = 3 + 7 = 10$ – маршрут замены;

маршрут $\pi_4^{(2)} = \{e_{1,2}; e_{2,5}; e_{5,4}\}$ с оценкой $d_4^{(2)} = 3 + 5 + 4 = 12$ – маршрут замены;

маршрут $\pi_4^{(3)} = \{e_{1,3}; e_{3,6}; e_{6,4}\}$ с оценкой $d_4^{(3)} = 3 + 6 + 6 = 15$ – маршрут замены.

Узел связи V_5

Маршрут $\pi_5 = \{e_{1,4}; e_{4,5}\}$ с оценкой $d_5 = 2 + 4 = 6$ – оптимальный маршрут;

маршрут $\pi_5^{(1)} = \{e_{1,2}; e_{2,5}\}$ с оценкой $d_5^{(1)} = 3 + 5 = 8$ – маршрут замены;

маршрут $\pi_5^{(2)} = \{e_{1,4}; e_{4,7}; e_{7,5}\}$ с оценкой $d_5^{(2)} = 2 + 8 + 5 = 15$ – маршрут замены;

маршрут $\pi_5^{(3)} = \{e_{1,2}; e_{2,4}; e_{4,5}\}$ с оценкой $d_5^{(3)} = 3 + 7 + 4 = 14$ – маршрут замены.

Узел связи V_6

Маршрут $\pi_6 = \{e_{1,4}; e_{4,6}\}$ с оценкой $d_6 = 2 + 6 = 8$ – оптимальный маршрут;

маршрут $\pi_6^{(1)} = \{e_{1,3}; e_{3,6}\}$ с оценкой $d_6^{(1)} = 3 + 6 = 9$ – маршрут замены.

Узел связи V_7

Маршрут $\pi_7 = \{e_{1,4}; e_{4,7}\}$ с оценкой $d_7 = 2 + 8 = 10$ – оптимальный маршрут;

маршрут $\pi_7^{(1)} = \{e_{1,4}; e_{4,5}; e_{5,7}\}$ с оценкой $d_7^{(1)} = 2 + 4 + 5 = 11$ – маршрут замены;

маршрут $\pi_7^{(2)} = \{e_{1,2}; e_{2,5}; e_{5,7}\}$ с оценкой $d_7^{(2)} = 3 + 5 + 5 = 13$ – маршрут замены.

Узел связи V_8

Маршрут $\pi_8 = \{e_{1,4}; e_{4,7}; e_{7,8}\}$ с оценкой $d_8 = 2 + 8 + 2 = 12$ – оптимальный маршрут;

маршрут $\pi_8^{(1)} = \{e_{1,4}; e_{4,6}; e_{6,8}\}$ с оценкой $d_8^{(1)} = 2 + 6 + 6 = 14$ – маршрут замены;

маршрут $\pi_8^{(2)} = \{e_{1,3}; e_{3,6}; e_{6,8}\}$ с оценкой $d_8^{(2)} = 3 + 6 + 6 = 15$ – маршрут замены;

маршрут $\pi_8^{(3)} = \{e_{1,4}; e_{4,5}; e_{5,7}; e_{7,8}\}$ с оценкой $d_8^{(3)} = 2 + 4 + 5 + 2 = 13$ – маршрут замены.

Для разработки алгоритма парных перестановок маршрутов на базе протокола OSPF при динамическом подключении узлов и линий связи корпоративной сети сформулируем следующие теоремы.

Теорема 1. При подключении канала связи $e_{i,j}$, инцидентного узлам связи V_i и V_j , причем V_i лежит ниже чем V_j по дереву иерархии с весом $w_{i,j}$, то без изменения окажутся оптимальные маршруты и их оценки для узлов множества $V^{(V_i)}$.

Доказательство. Пусть канал связи $e_{i,j}$ входит в маршрут $\pi_{i,k} < > \pi_i$ и $\pi_{j,p} < > \pi_j$. Если данный канал не уменьшает оценок обоих инцидентных ему узлов V_i и V_j , то есть $d_{i,k} \geq d_i$ и $d_{j,p} \geq d_j$, то оптимальные маршруты до всех узлов не изменятся, так как канал $e_{i,j}$ оказывает влияние, прежде всего, на инцидентные ему узлы множества V . Так как существовавшие до изменения маршруты до узлов V_i и V_j имели меньшую оценку, то канал $e_{i,j}$ не включается, и оптимальные маршруты до всех узлов не изменятся. Если уменьшилась оценка какого-либо инцидентного узла, например V_j , то эта оценка $d_{j,p}$ будет оценкой оптимального маршрута до узла V_j и канал $e_{i,j}$ войдет в состав этого маршрута, так как не существует другого оптимального маршрута π_j до узла V_j , кроме маршрута $\pi_{j,p}$, содержащего канал $e_{i,j}$. Этот оптимальный маршрут $\pi_{j,p}$ не будет существовать, если не будут существовать оптимальные маршруты до всех промежуточных узлов $V_k \in V^{(V_j)}$ этого маршрута. Невозможно будет сказать, останутся ли неизменными оптимальные маршруты до остальных узлов графа. Теорема доказана.

Следствие 1. При подключении канала связи $e_{i,j}$, инцидентного узлам связи V_i и V_j , с весом $w_{i,j}$ при не изменении оценок d_i и d_j оптимальные маршруты до всех узлов не изменятся. Если изменится оценка какого-либо узла V_i и V_j , то необходимо определить новые оптимальные маршруты и их маршруты замены для множества узлов, не принадлежащих множеству $V^{(V_i)}$.

Следствие 2. При подключении канала связи $e_{i,j}$, инцидентного узлам связи V_i и V_j , причем V_i лежит ниже чем V_j по дереву иерархии, с весом $w_{i,j}$, то необходимо определить новые маршруты замены (с учетом канала $e_{i,j}$) для множества узлов, не принадлежащих множеству $V^{(V_i)}$.

Данному случаю соответствует графовая модель корпоративной сети, представленная на рисунке 2, при подключении канала связи $e_{6,7}$.

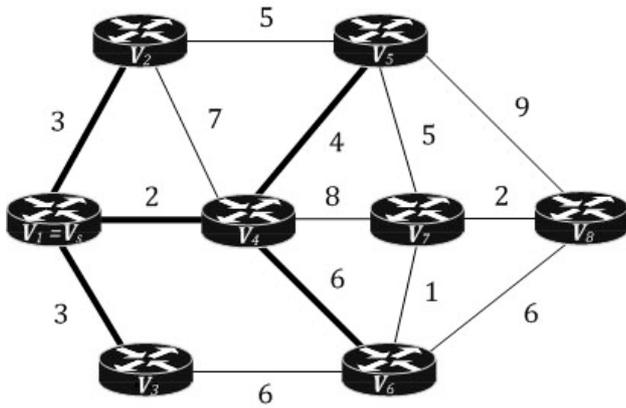


Рисунок 2 – Подключение канала связи $e_{6,7}$

На рисунке 2 жирными линиями обозначено дерево оптимальных маршрутов, которое не требует изменения.

То есть при подключении канала связи $e_{6,7}$ дерево оптимальных маршрутов до узлов V_2, V_3, V_4, V_5 и V_6 согласно теореме 1 не изменится. Для узлов V_7 и V_8 необходимо переформировать список парных переходов, пересчитать точки вхождения в дерево и в множество замены, а также переформировать списки оптимальных маршрутов и их маршруты замены с учетом подключенного канала связи $e_{6,7}$. Так для узла связи V_7 будут добавлены маршруты $\pi_7^{(1)} = \{e_{1,4}; e_{4,6}; e_{6,7}\}$ с оценкой $d_7^{(1)} = 2 + 6 + 1 = 9$ и $\pi_7^{(2)} = \{e_{1,3}; e_{3,6}; e_{6,7}\}$ с оценкой $d_7^{(2)} = 3 + 6 + 1 = 10$. Для узла связи V_8 будут добавлены маршруты $\pi_8^{(1)} = \{e_{1,4}; e_{4,6}; e_{6,7}; e_{7,8}\}$ с оценкой $d_8^{(1)} = 2 + 6 + 1 + 2 = 11$ и $\pi_8^{(2)} = \{e_{1,3}; e_{3,6}; e_{6,7}; e_{7,8}\}$ с оценкой $d_8^{(2)} = 3 + 6 + 1 + 2 = 12$. После того как новые списки маршрутов для узлов V_7 и V_8 сформированы, из них необходимо выбрать маршрут с минимальной оценкой. Оптимальный маршрут до узла связи V_7 составит $\pi_7 = \{e_{1,4}; e_{4,6}; e_{6,7}\}$ с оценкой $d_7 = 9$. Оптимальный маршрут до узла связи V_8 составит $\pi_8 = \{e_{1,4}; e_{4,6}; e_{6,7}; e_{7,8}\}$ с оценкой $d_8 = 11$.

Таким образом, для графовой модели корпоративной сети, представленной на рисунке 2, множество каналов дерева составит $E_T = \{e_{1,2}; e_{1,3}; e_{1,4}; e_{4,5}; e_{4,6}; e_{6,7}; e_{7,8}\}$; множество каналов замены $E_S = \{e_{2,4}; e_{2,5}; e_{3,6}; e_{4,7}; e_{5,7}; e_{6,8}\}$; множество непарных каналов $E_P = \{e_{5,8}\}$.

Теорема 2. При подключении некоторого узла связи V_i для всех узлов, не инцидентных узлу V_i (то есть не имеющих канала связи $e_{i,k}$), оптимальные маршруты и их оценки окажутся без изменения.

Доказательство. Новый узел связи V_i не входит в оптимальный маршрут к некоторому узлу V_j . Следовательно, и в маршруте к этому узлу $\pi_{v,j}$ отсутствует канал связи $e_{i,k}$. Поэтому новый маршрут к узлу V_j , содержащий данный

канал, будет иметь оценку $d_{j,i} > d_j$ и оптимальные маршруты не изменятся для всех узлов, не инцидентных узлу V_i . Теорема доказана.

Следствие 1. При подключении некоторого узла связи V_k , имеющего каналы связи с узлами V_i и V_j , если $d_i < d_j$, то оптимальные маршруты и их маршруты замены до узла V_i не изменятся.

Следствие 2. При подключении некоторого узла связи V_k , имеющего каналы связи с узлами V_i и V_j , если $d_i < d_j$, то необходимо определить новые оптимальные маршруты и их маршруты замены для множества узлов, инцидентных узлу V_k , кроме узла V_i .

Следствие 3. При подключении некоторого узла связи V_k , имеющего каналы связи с узлами V_i и V_j , необходимо определить оптимальные маршруты и их маршруты замены до узла V_k с учетом инцидентных ему каналов $e_{i,k}$ и $e_{k,j}$.

Данному случаю соответствует графовая модель корпоративной сети, представленная на рисунке 3, при подключении узла связи V_9 .

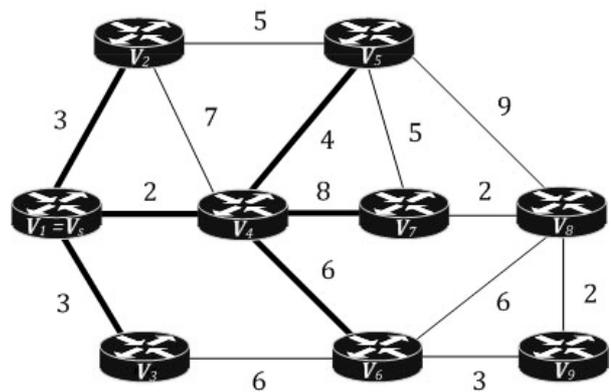


Рисунок 3 – Подключение узла связи V_9

На рисунке 3 жирными линиями обозначено дерево оптимальных маршрутов, которое не требует изменения.

То есть при подключении узла связи V_9 дерево оптимальных маршрутов до вершин V_2, V_3, V_4, V_5, V_6 и V_7 согласно теореме 2 не изменится. Для узлов связи V_8 и V_9 необходимо переформировать список парных переходов, пересчитать точки вхождения в дерево и в множество замены, а также переформировать списки оптимальных маршрутов и их маршруты замены с учетом подключенного узла связи V_9 . Так для узла связи V_8 будут добавлены маршруты $\pi_8^{(1)} = \{e_{1,4}; e_{4,6}; e_{6,9}; e_{9,8}\}$ с оценкой $d_8^{(1)} = 2 + 6 + 3 + 2 = 13$ и $\pi_8^{(2)} = \{e_{1,3}; e_{3,6}; e_{6,9}; e_{9,8}\}$ с оценкой $d_8^{(2)} = 3 + 6 + 3 + 2 = 14$. Для узла связи V_9 будут добавлены маршруты $\pi_9^{(1)} = \{e_{1,4}; e_{4,7}; e_{7,8}; e_{8,9}\}$ с оценкой $d_9^{(1)} = 2 + 8 + 2 + 2 = 14$, $\pi_9^{(2)} = \{e_{1,4}; e_{4,6}; e_{6,9}\}$ с оценкой $d_9^{(2)} = 2 + 6 + 3 = 11$ и $\pi_9^{(3)} = \{e_{1,3}; e_{3,6}; e_{6,9}\}$ с оценкой $d_9^{(3)} = 3 + 6 + 3 = 12$. После того как новые спис-

ки маршрутов для узлов связи V_8 и V_9 сформированы, из них необходимо выбрать маршрут с минимальной оценкой. Оптимальный маршрут до узла V_8 составит $\pi_8 = \{e_{1,4}; e_{4,7}; e_{7,8}\}$ с оценкой $d_8 = 12$. Оптимальный маршрут до узла V_9 составит $\pi_9 = \{e_{1,4}; e_{4,6}; e_{6,9}\}$ с оценкой $d_9 = 11$.

Таким образом, для графовой модели корпоративной сети, представленной на рисунке 3, множество каналов дерева составит $E_T = \{e_{1,2}; e_{1,3}; e_{1,4}; e_{4,5}; e_{4,6}; e_{4,7}; e_{7,8}; e_{6,9}\}$; множество каналов замены $E_S = \{e_{2,4}; e_{2,5}; e_{3,6}; e_{5,7}; e_{8,9}\}$; множество непарных каналов $E_P = \{e_{5,8}; e_{6,8}\}$.

Использование доказанных выше теорем позволяет разработать алгоритм парных перестановок маршрутов на базе протокола OSPF, уменьшающий размерность задачи поиска оптимальных маршрутов при динамическом подключении узлов и линий связи корпоративной сети.

Рассмотрим работу алгоритма парных перестановок маршрутов на базе протокола OSPF в корпоративных сетях. Укрупненно алгоритм имеет следующий вид.

Шаг 1. Первоначальная инициализация исходных данных. Используя пакет HELLO протокола OSPF, определить веса линий связи $w_{i,j}$.

Шаг 2. Построить дерево оптимальных маршрутов корпоративной сети.

Шаг 3. Для узла связи, являющегося листом дерева, произвести поиск всех парных переходов без ограничений. Эти списки для удобства дальнейшей работы привязываются к узлу, инцидентному рассматриваемому каналу связи и расположенному ниже по иерархии.

Шаг 4. Если узел связи не является листом дерева, то вычислить парные переходы для этого узла и выбрать лучшие значения потенциалов парных переходов для потомков узла и собственных парных переходов. Подобную процедуру выполнить для формирования списков парных переходов в случае динамического подключения узлов и линий связи корпоративной сети.

Шаг 5. Для каждого узла сформировать полный список парных переходов. Число элементов в каждом из этих списков не превышает количество узлов графа. Такое решение позволяет отказаться от предварительной сортировки потенциалов или приращений для парных переходов без значительного усложнения алгоритма обработки изменения.

Шаг 6. Для каждого канала связи корпоративной сети определить точку вхождения в дерево оптимальных маршрутов и точку вхождения в множество замены.

Шаг 7. Для каждого узла связи корпоративной сети сформировать полный список возможных маршрутов, проходящий через каналы, со-

стоящие в отношении парного перехода, включая и каналы, входящие в дерево оптимальных маршрутов.

Шаг 8. Определить, есть ли пакеты на передачу:

- а) если да, то перейти к шагу 9;
- б) иначе – к шагу 17.

Шаг 9. Используя поля «Время жизни» и «Контрольная сумма заголовка» протокола IP, определить, требуется ли уничтожить (отбросить) данный пакет:

- а) если да, то перейти к шагу 23;
- б) иначе – к шагу 10.

Шаг 10. Используя поле «Тип сервиса», определить, требуется ли создание виртуального соединения:

- а) если да, то перейти к шагу 11;
- б) иначе – к шагу 13.

Шаг 11. Организовать виртуальное соединение, послав первый пакет «запрос вызова» адресату:

- а) если пакет согласия на соединение от адресата пришел, то перейти к шагу 14;
- б) если адресат отклонил вызов, то перейти к пункту 13.

Шаг 12. а) установить флаг передачи;

- б) послать адресату пакет ликвидации соединения;
- в) получить пакет подтверждения рассоединения от адресата.

Шаг 13. Проверка флага передачи:

- а) если флаг установлен, то перейти к шагу 23;
- б) иначе – к шагу 8.

Шаг 14. Используя поле «Тип сервиса», определить необходимую таблицу маршрутизации с учетом желаемого уровня качества обслуживания:

- а) если таблица 1, то перейти к шагу 15;
- б) иначе – к шагу 16.

Шаг 15. а) передать пакет, используя первую таблицу маршрутизации;

- б) перейти к шагу 12.

Шаг 16. а) передать пакет, используя вторую таблицу маршрутизации;

- б) перейти к шагу 12.

Шаг 17. Анализируя полученную информацией OSPF информацию, определить, произошло ли подключение узлов или линий связи в структуру корпоративной сети:

- а) если да, то перейти к шагу 18;
- б) иначе – к шагу 8.

Шаг 18. а) если подключился узел связи V_k , имеющий каналы связи с узлами V_i и V_j , причем $d_i < d_j$ (узел V_i расположен ниже по иерархии в дереве оптимальных маршрутов), то:

- 1) дерево оптимальных маршрутов до узла V_i не изменится;

2) для всех узлов, не инцидентных узлу V_i , дерево оптимальных маршрутов не изменится;

3) потенциал узла V_j (оценка маршрута d_j до узла V_j) с учетом новых каналов $e_{i,k}$ и $e_{k,j}$ не изменился. Первоначальное дерево оптимальных маршрутов не изменится. Из возможных маршрутов с оценками $d_i + d_{i,k}$ и $d_j + d_{k,j}$ выбрать маршрут минимальной длины и включить его в дерево оптимальных маршрутов. Сформировать новый список маршрутов замены до узла V_j . Перейти к пункту 6;

4) потенциал узла V_j (оценка маршрута d_j до узла V_j) с учетом новых каналов $e_{i,k}$ и $e_{k,j}$ с оценками $d_{i,k}$ и $d_{k,j}$ соответственно уменьшится (то есть $d_i + d_{i,k} + d_{k,j} < d_j$). Включить в дерево оптимальных маршрутов маршрут, проходящий через каналы $e_{i,k}$ и $e_{k,j}$ с оценками $d_{i,k}$ и $d_{k,j}$ соответственно. Для всех узлов V_m , инцидентных узлу V_j , определить их новые потенциалы. Если потенциал до узла V_m (оценка оптимального маршрута d_m до узла V_m) уменьшился, то из списка маршрутов замены до узла V_m выбрать маршрут с минимальной оценкой, который будет проходить через канал $e_{j,m}$, и включить его в дерево оптимальных маршрутов, заменив первоначальный оптимальный маршрут до узла V_m . Сформировать новый список маршрутов замены до узла V_j ;

5) в список возможных маршрутов корпоративной сети добавить новый оптимальный маршрут и его маршруты замены до узла V_k с учетом каналов $e_{i,k}$ и $e_{k,j}$;

б) для всех каналов, включенных в дерево оптимальных маршрутов, определить возможные каналы для парных переходов и рассчитать для них точки вхождения в множество каналов замены и в дерево оптимальных маршрутов.

б) если подключился канал связи $e_{i,j}$ с весом $w_{i,j}$, причем узел связи V_i располагается ниже по иерархии в дереве оптимальных маршрутов, чем узел V_j , то:

1) дерево оптимальных маршрутов, список маршрутов замены, а также оценки их длин окажутся без изменения для множества узлов $V^{(V_j)}$;

2) оценка длины оптимального маршрута d_j до узла V_j не изменилась. Дерево оптимальных маршрутов не изменится. Определить, находится ли канал $e_{i,j}$ в отношении парного перехода к какому-либо каналу для узла V_j . Если да, то включить его во множество каналов, состоящих в отношении парного перехода к узлу V_j . Построить новые маршруты замены (с учетом канала $e_{i,j}$) для множества узлов, инцидентных узлу V_j (кроме узла V_i);

3) оценка длины оптимального маршрута d_j до узла V_j уменьшилась. Построить новый маршрут до узла V_j через канал $e_{i,j}$ и включить его в

дерево оптимальных маршрутов. Для всех узлов, кроме V_i , построить новые маршруты, проходящие через канал $e_{i,j}$ и рассчитать новые оценки длин. Для тех узлов, у которых оценки длин маршрутов уменьшились, включить новый оптимальный маршрут, проходящий через канал $e_{i,j}$ в дерево оптимальных маршрутов, заменив в нем первоначальный маршрут с большей оценкой до этого узла. Построить новые маршруты замены (с учетом канала $e_{i,j}$) для множества узлов, инцидентных узлу V_j (кроме узла V_i).

Шаг 19. Используя список парных переходов, определить, требуется ли сделать парный переход:

а) если да, то перейти к шагу 20;

б) иначе – к шагу 21.

Шаг 20. Для каждого узла связи, у которого в списках возможных маршрутов произошли изменения, определить путь минимальной длины и поместить его в дерево оптимальных маршрутов.

Шаг 21. Построить новое дерево оптимальных маршрутов с учетом изменений.

Шаг 22. Сформировать таблицы маршрутизации.

Шаг 23. Проверка окончания работы маршрутизатора:

а) если да, то перейти к шагу 24;

б) иначе – сбросить флаг передачи и перейти к шагу 8.

Шаг 24. Завершение работы маршрутизатора.

Результаты моделирования. Для подтверждения правильности алгоритма парных перестановок маршрутов на базе протокола OSPF при динамическом подключении узлов и линий связи корпоративной сети разработано программное обеспечение моделирования процессов маршрутизации.

При разработке основное внимание уделялось корректности предлагаемого алгоритма и размерности решаемой задачи.

Также при проектировании и создании программных средств рассматривался подход, предложенный в работе [14].

На рисунках 4-6 представлены результаты моделирования алгоритма парных перестановок маршрутов на базе протокола OSPF.

Для каждого испытания на множестве обработанных изменений выбирались минимальное, максимальное и среднее значения размерности задачи, выраженные через количество узлов связи, для которых необходим поиск оптимальных маршрутов. Для каждого эксперимента были найдены значения оценок математического ожидания и среднего квадратичного отклонения числа изменений. Для алгоритма парных пере-

становок маршрутов на базе протокола OSPF определялось число фактически выполненных перестановок маршрутов при динамическом подключении узлов и линий связи корпоративной сети.

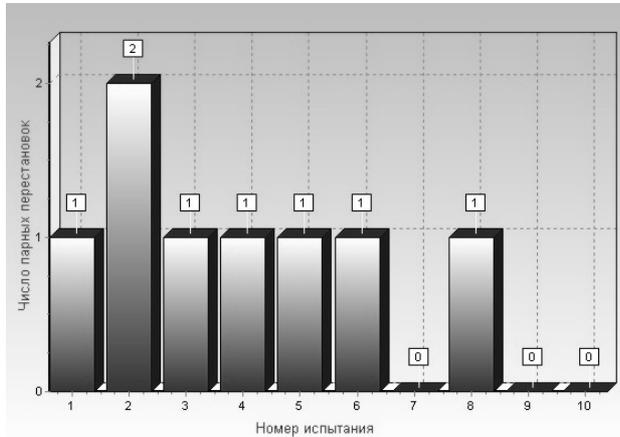


Рисунок 4 – Число изменений корпоративной сети из 10 узлов

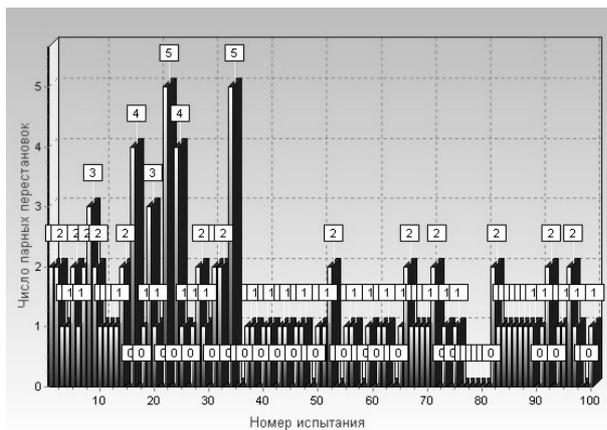


Рисунок 5 – Число изменений корпоративной сети из 100 узлов

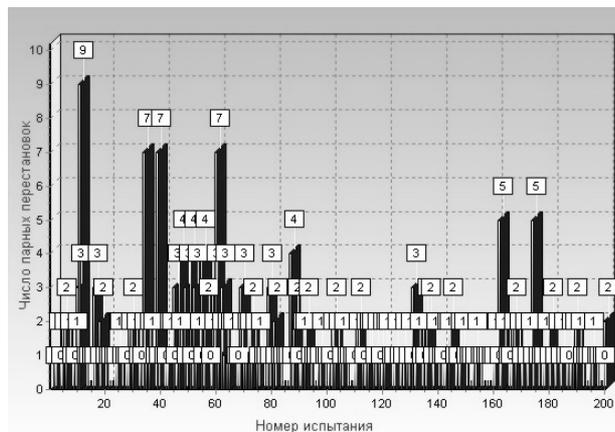


Рисунок 6 – Число изменений корпоративной сети из 200 узлов

Были проведены исследования графовых моделей корпоративных сетей, состоящих из 10, 100 и 200 узлов связи. Исследование разработанного

алгоритма парных перестановок маршрутов на базе протокола OSPF показало, что максимальное значение числа изменений существенно меньше размерности для каждой из рассмотренных моделей сети, а значение оценки математического ожидания не превышает единицы. Более того, обнаружена тенденция уменьшения значения оценки математического ожидания числа изменений дерева оптимальных маршрутов с увеличением количества узлов связи в сети.

В таблице для различного числа узлов связи (N) графа корпоративной сети приведены обобщенные статистические характеристики числа изменений дерева маршрутов: минимальное (Min) и максимальное (Max) значения, значения оценок математического ожидания (МО) и среднего квадратичного отклонения (СКО).

N	Min	Max	МО	СКО
10	0	0,2	0,08	0,0409
100	0	0,05	0,01	0,0085
200	0	0,045	0,0045	0,0061

При динамическом подключении узлов и линий связи корпоративной сети разработанный алгоритм парных перестановок маршрутов на базе протокола OSPF в отличие от алгоритма Дейкстры позволяет производить перестроение таблиц маршрутизации не полностью, а только той ее части, в которой произошли изменения. Для этого необходимо один раз просмотреть списки оптимальных маршрутов и их маршрутов замены только для тех узлов сети, оценки длин которых изменились при динамическом подключении узлов и линий связи корпоративной сети, и на основе предварительно собранной информации о списках маршрутов выполнить изменения для вычисления новых оптимальных маршрутов. При этом трудоемкость изменения дерева оптимальных маршрутов является линейной функцией от числа узлов сети и определяется выражением $O(N)$. Если рассмотреть полученные результаты при моделировании предложенного алгоритма парных перестановок маршрутов и сравнить их с алгоритмом Дейкстры, то видно:

1) для корпоративной сети из 10 узлов связи максимальное число парных перестановок маршрутов составляет два изменения ($= 2$), при этом для алгоритма Дейкстры при каждом подключении узлов и линий связи трудоемкость построения таблицы маршрутизации составляет $O(N^2)$, т.е. порядка $10^2 = 100$ элементарных операций;

2) для корпоративной сети из 100 узлов связи максимальное число парных перестановок маршрутов составляет пять изменений ($= 5$), при этом для алгоритма Дейкстры при каждом подключении узлов и линий связи трудоемкость по-

строения таблицы маршрутизации составляет $O(N^2)$, т.е. порядка $100^2 = 10\,000$ элементарных операций;

3) для корпоративной сети из 200 узлов связи максимальное число парных перестановок маршрутов составляет девять изменений (= 9), при этом для алгоритма Дейкстры при каждом подключении узлов и линий связи трудоемкость построения таблицы маршрутизации составляет $O(N^2)$, т.е. порядка $200^2 = 40\,000$ элементарных операций.

На основе этого можно сделать вывод, что предложенный алгоритм парных перестановок маршрутов на базе протокола OSPF является эффективным при поиске оптимальных маршрутов в условиях динамических изменений в структуре корпоративной сети и нагрузках на линиях связи за счет использования дополнительной информации о возможных маршрутах замены.

Заключение. Разработанный алгоритм парных перестановок маршрутов на базе протокола OSPF позволяет повысить эффективность функционирования корпоративной сети за счет уменьшения трудоемкости построения таблиц маршрутизации до величины порядка $O(N)$ при динамическом подключении узлов и линий связи в ее структуру.

Библиографический список

1. *Перепелкин Д.А.* Алгоритм адаптивной ускоренной маршрутизации на базе протокола OSPF при динамическом добавлении элементов корпоративной сети // Вестник Рязанского государственного радиотехнического университета. – 2010. – № 34. – С. 65-71.
2. *Корячко В.П., Перепелкин Д.А.* Корпоративные сети: технологии, протоколы, алгоритмы. – М.: Горячая линия – Телеком, 2011. – 219 с.
3. *Корячко В.П., Перепелкин Д.А.* Анализ и проектирование маршрутов передачи данных в корпоративных сетях. – М.: Горячая линия – Телеком, 2012. – 235 с.
4. *Уваров Д.В., Перепелкин А.И.* Алгоритм динамической маршрутизации в вычислительной сети // Вестник Рязанской государственной радиотехнической академии. – 2003. – № 12. – С. 77-80.
5. *Уваров Д.В., Перепелкин А.И., Корячко В.П.* Построение дерева кратчайших путей в графе на основе данных о парных переходах // Системы управления и информационные технологии. – 2004. – № 4 (16). – С. 93-96.
6. *Корячко В.П., Перепелкин Д.А.* Построение дерева оптимальных маршрутов корпоративной сети в условиях динамического изменения нагрузки на ее линиях связи // Информационные технологии: межвуз. сб. науч. тр. РГРТУ. – 2011. – С. 7-18.
7. *Перепелкин Д.А., Перепелкин А.И.* Разработка алгоритмов адаптивной маршрутизации в корпоративных вычислительных сетях // Вестник Рязанской государственной радиотехнической академии. – 2006. – № 19. – С. 114-116.
8. *Перепелкин Д.А., Перепелкин А.И.* Разработка алгоритма динамической маршрутизации на базе протокола OSPF в корпоративных вычислительных сетях // Вестник Рязанского государственного радиотехнического университета. – 2009. – № 28. – С. 68-72.
9. *Перепелкин Д.А.* Алгоритм адаптивной ускоренной маршрутизации на базе протокола OSPF при динамическом отказе элементов корпоративной сети // Вестник Рязанского государственного радиотехнического университета. – 2011. – № 37. – С. 53-58.
10. *Корячко В.П., Перепелкин Д.А., Перепелкин А.И.* Алгоритм парных перестановок маршрутов в корпоративных сетях // Системы управления и информационные технологии. – 2010. – Т. 40. № 2. – С. 51-56.
11. *Корячко В.П., Перепелкин Д.А., Перепелкин А.И.* Повышение эффективности функционирования корпоративных сетей при динамических изменениях в их структуре и нагрузках на линии связи // Вестник Рязанского государственного радиотехнического университета. – 2010. – № 33. – С. 49-55.
12. *Перепелкин Д.А., Перепелкин А.И.* Повышение качества функционирования корпоративных сетей на базе протокола OSPF // Качество. Инновации. Образование. – 2010. – № 12. – С. 51-56.
13. *Перепелкин Д.А., Перепелкин А.И.* Алгоритм адаптивной ускоренной маршрутизации в условиях динамически изменяющихся нагрузок на линиях связи корпоративной сети // Информационные технологии. – 2011. – № 3(175). – С. 2-7.
14. *Пруцков А.В., Цыбулько Д.М.* Проблемно-ориентированное объектное программирование // Вестник Рязанского государственного радиотехнического университета. – 2013. – № 45. – С. 57-62.

УДК004.93.1

*Е.М. Дондик, А.Н. Пылькин, О.С. Вавилова***РАСПОЗНАВАНИЕ ИСКАЖЕННЫХ ФОНЕМ МЕТОДОМ
ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ**

Наиболее привычным и естественным для человека является речевое общение. В данной статье представлен метод распознавания искаженных фонем на основе вейвлет-преобразования. Представлены экспериментальные исследования системы, построенной на основе кратко-масштабного вейвлет-анализа.

Ключевые слова: вейвлет-преобразования, кратко-масштабный вейвлет-анализ, определение степени искаженности звуков речи, цифровая обработка речи, акустические свойства звуков речи, распознавание речи, распознавание фонем, обучение глухих произношению звуков речи.

Введение. Для человека естественным и привычным является речевое общение. Не получая звуковой информации человек не может объективно и полно воспринимать окружающую его действительность, возникают серьезные проблемы в общении с другими людьми.

Люди с врожденной глухотой, но с не нарушенным речевым аппаратом, имеют возможность научиться говорить и не являются истинно глухонемыми. Их немота обусловлена тем, что они не слышат звуки, которые произносят, что и является основной трудностью при обучении их речи. Основная задача состоит в том, чтобы дать глухонемым возможность обучения правильно произношению наименьших единиц звукового строя языка – фонем.

В статье приведено описание метода распознавания искаженных фонем с помощью вейвлет-преобразования. В настоящее время около 3 % людей в обществе являются глухонемыми, их немота обусловлена тем, что они не слышат произносимые ими звуки. Эти люди с ограниченными возможностями произношения сталкиваются с серьезными проблемами в общении с другими людьми. Рассмотренный в статье метод может быть основой создания системы, которая поможет глухонемым общаться, понимать и быть понятыми окружающими людьми без знания сложных аспектов дактильной и жестовой речи.

Таким образом, появляется возможность самостоятельного освоения глухонемыми правильного произношения звуков речи в домашней обстановке или в школе для при наличии стандартных средств вычислительной техники.

Математическое представление процесса восприятия речи. Речевой сигнал акустически представляет собой распространяемые в воздушной среде звуковые колебания, и его модель может

быть представлена с помощью преобразования Фурье. Классический анализ Фурье основан на возможности исследования функций во временной ($|t| < \infty$) и частотной ($|v| < \infty$) областях с помощью прямого и обратного преобразований Фурье [1]

$$\hat{f}(v) = \int_{-\infty}^{\infty} f(t)e^{-i2\pi vt} dt \quad (1)$$

$$f(v) = \int_{-\infty}^{\infty} f(t)e^{-i2\pi vt} dt \quad (2)$$

Преобразование Фурье имеет замечательную способность фокусировать в точку распределенную по времени информацию о периодичности функции при переходе из временной области в частотную. Достигается это за счет того, что ядро преобразования Фурье, то есть функция $e^{-i2\pi vt}$, не локализована во времени, но имеет предельную локализацию в частотной области. Это обстоятельство и делает преобразование Фурье прекрасным инструментом для изучения процессов, свойства которых не меняются со временем.

Для обработки иррегулярных функций, характеристики которых меняются во времени, к которым относится речевой сигнал, это обстоятельство делает преобразование Фурье непригодным. Для устранения этого недостатка нужно локализовать преобразование Фурье на промежутках конечной длины, вычисляя оценки спектра мощности не только по всей длине временного ряда, но и по его различным частям [1]. Формализация такого подхода может быть описана, например, с помощью оконного преобразования Гэбора:

$$GT(v, b, a) = \int_{-\infty}^{\infty} f(t) e^{-\frac{(t-b)^2}{2a^2}} e^{-i2\pi vt} dt \quad (3)$$

Сравнивая рассмотренное выражение с (1), мы видим, что введенное под знак интеграла временное окно $e^{-\frac{(t-b)^2}{a}}$ выделяет лишь небольшой отрезок исходного временного ряда с центром в произвольной точке b .

Окно преобразования Гэбора имеет постоянную ширину, которая определяется параметром a . Эффективная ширина окна определяет длину интервала ΔT , который вносит главный вклад в значение интеграла в выражении (3). Длина ΔT спектральной линии $\Delta \nu$ определяет меру частотного разрешения. Обе эти характеристики связаны между собой соотношением $\Delta \nu$ пропорционально $1/\Delta T$.

Естественное для анализа нерегулярных сигналов стремление повысить временное разрешение всегда приводит к уменьшению разрешающей способности в области частот. При использовании преобразования Гэбора возникает проблема выбора ширины окна во временной области. Слишком широкое окно может обеспечить разумное представление низкочастотных компонентов ряда, но его ширина будет избыточной для гармоник с высокой частотой, поскольку все интересные нерегулярности в высокочастотной области спектра сгладятся. Слишком узкое окно даст возможность изучить вариации во времени высокочастотных компонентов, но оно не будет адекватным для низкочастотных гармоник.

Если сделать оконную функцию зависящей от частоты так, чтобы для низких частот окно становилось шире, а для высоких – уже, то оконное преобразование Фурье переходит в новый класс преобразований, который и получил название вейвлет-преобразования [2].

Вейвлет-преобразование одномерного сигнала (формула 4) – это его представление в виде обобщенного ряда или интеграла Фурье по системе базисных функций, сконструированных из материнского вейвлета $\psi(t)$, обладающего определенными свойствами за счет операций сдвига во времени b и изменения временного масштаба a . Множитель $1/\sqrt{a}$ обеспечивает независимость нормы этих функций от масштабирующего числа a

$$\psi_{ab}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (4)$$

Для заданных значений параметров a и b функция $\psi_{ab}(t)$ – вейвлет, порождаемый материнским вейвлетом [2].

Вейвлеты локализованы как во временной, так и в частотной областях.

В отличие от традиционного преобразования Фурье вейвлет-преобразование обеспечивает двумерное представление исследуемого сигнала

в частотной области в плоскости частота – положение. Аналогом частоты при этом является масштаб аргумента базисной функции – времени, а положение характеризуется ее сдвигом. Это позволяет разделять крупные и мелкие особенности сигналов, одновременно локализуя их на временной шкале. Другими словами, вейвлет-анализ можно охарактеризовать как спектральный анализ локальных возмущений.

Спектральное представление вейвлетов аналогично заданию окна в оконном преобразовании Фурье [3]. Главное отличие состоит в том, что свойства окна, его ширина и перемещение по частоте присущи самим вейвлетам. Это служит предпосылкой их адаптации к сигналам, представляемым совокупностью вейвлетов. Поэтому с помощью вейвлетов можно осуществить анализ и синтез локальной особенности речевого сигнала.

Вейвлет-преобразование не просто «режет» исследуемый объект на куски, а выделяет из него компоненты разных масштабов, и каждая компонента анализируется с той степенью детальности, которая соответствует его масштабу. Следовательно, вейвлет-преобразование можно использовать для обработки нестационарных сигналов, которым и является речевой сигнал.

Структура органов восприятия эталонной речи. Внутреннее ухо человека представляет собой сложный лабиринт соединенных друг с другом каналов, которые расположены в височной кости и заполнены специальной жидкостью. Часть этого лабиринта образует вестибулярную систему, а одна из костных структур, имеющая форму спирали, относится к слуховой системе и называется улиткой.

В улитке происходит преобразование механических процессов в нервные сигналы, которые затем передаются в мозг и анализируются (рисунок 1).

При передаче звуковых колебаний основание стремечка в такт с изменением звукового давления вдавливают мембрану овального окна в жидкость вестибулярного канала, создавая в нем избыточное давление, которое практически мгновенно распространяется по всей длине улитки. При этом с информационной точки зрения со звуком происходят определенные преобразования, которые сводятся в основном к усилению. Таким образом, большинство элементов в улитке имеют определенный коэффициент усиления k . Увеличение давления приводит в движение податливые структуры среднего канала – Рейснерову мембрану, жидкость среднего канала и базилярную мембрану, что, в свою очередь, вызывает изменение давления жидкости в тимпанальном канале и смещение мембраны круглого окна. Реакция базилярной мембраны на

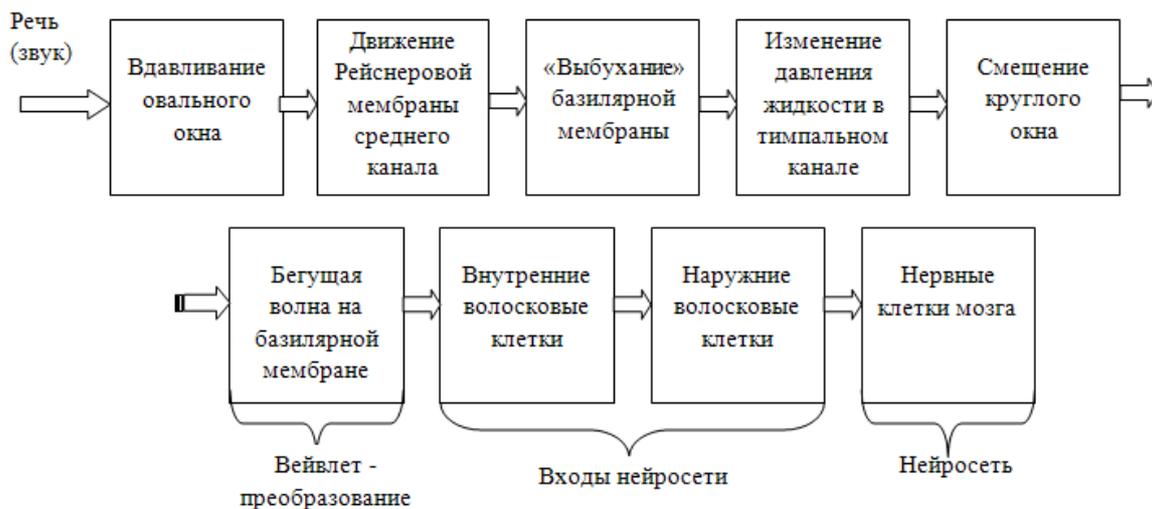


Рисунок 1 – Схема преобразования информационных процессов в нервные сигналы

отдельный импульс увеличения давления в жидкости улитки не является мгновенной и может быть описана следующим образом. Сначала мембрана «выбухает» в сторону тимпанального канала у своего основания, затем в силу эластичности возвращается в положение равновесия. Возникшее выбухание перемещается вдоль длины мембраны к вершине за время, равное нескольким мс. Таким образом, на мембране возникает бегущая волна смещения ее локальных участков относительно положения, которое они занимают вне звука.

Реакция базилярной мембраны на частоту звуковых колебаний является следствием ее упругих свойств и анатомического строения. Разные участки мембраны отличаются шириной и жесткостью. Ширина мембраны увеличивается по направлению от основания к вершине примерно в десять раз, а упругость постепенно уменьшается: у основания мембрана более чем в 100 раз жестче, чем у вершины. Приблизительно половина мембраны, считая от ее вершины, реагирует на колебания с частотами от 0 до 2000 Гц, в то время как оставшаяся часть (у основания) реагирует на остальные частоты звукового диапазона. Такое распределение означает, что различия между звуками в области низких частот обнаруживаются слуховой системой гораздо лучше, чем различия в высоких частотах. Можно сказать, что реакция мембраны на частоту сигнала приблизительно следует логарифмической шкале.

На колебательный процесс, возникающий в мембране, влияет интенсивность звука. Чем сильнее звук, тем больше размах колебаний всех участков мембраны, в том числе и участка с максимальным смещением. Размер зоны возбуждения вблизи точки максимального смещения определяется как частотой звука, так и его ампли-

тудой. Связь местонахождения максимального возбуждения базилярной мембраны с частотой звука свидетельствует о частотно-избирательной или фильтрующей функции этого слухового органа в физиологии слуха; связь между местом максимального смещения базилярной мембраны и частотой возбуждающего сигнала называют тонотопической организацией. Этот же термин распространяется и на более высокие уровни слуховой системы для обозначения пространственной упорядоченности нервных элементов в соответствии с их частотной избирательностью [4].

Таким образом, во внутреннем ухе человека осуществляется перекодирование частоты и интенсивности звукового сигнала в активность, локализованную на определенном участке базилярной мембраны. Различные участки мембраны являются своеобразными биомеханическими фильтрами, а мембрана в целом может рассматриваться как набор фильтров, упорядоченных по частоте и покрывающих в совокупности всю область частот, доступных восприятию человека. В техническом спектральном анализе фильтр является устройством, которое, обрабатывая сигнал, пропускает колебания только определенных частот. Результирующее преобразование сигнала во внутреннем ухе человека с точностью до константы совпадает с вейвлет-преобразованием сигнала. Следовательно, восприятие человеческим ухом звуковых колебаний, и в том числе человеческой речи, с математической точки зрения соответствует вейвлет-разложению. Поэтому естественно для распознавания речи использовать вейвлет-разложение.

Вейвлет-представление речевого сигнала.

Представим речевой сигнал как функцию f из бесконечномерного Гильбертова пространства $L^2(R)$. Эту функцию f разложим на некотором

заданном уровне разрешения j_n , (где $n=1,2,3\dots$) в ряд вида, представленного формулой

$$f(t) = \sum_k s_{j_n,k} \phi_{j_n,k} + \sum_{j \geq j_n,k} d_{j_n,k} \psi_{j_n,k}, \quad (5)$$

где $\phi_{j_n,k}$ и $\psi_{j_n,k}$ – масштабированные и смещенные версии скейлинг-функции (масштабной функции) ϕ и «материнского вейвлета» ψ ; $s_{j,k}$ – коэффициенты аппроксимации; $d_{j,k}$ – детализирующие коэффициенты.

Вейвлет-коэффициенты аппроксимации соответствуют передаточной характеристике фильтра низких частот, а коэффициенты детализации соответствуют высокочастотному фильтру.

Вейвлет-коэффициенты $s_{j,k}$ и $d_{j,k}$ можно вычислить по формулам

$$s_{j,k} = \int f(x) \phi_{j,k}(x) dx; \quad (6)$$

$$d_{j,k} = \int f(x) \psi_{j,k}(x) dx. \quad (7)$$

Первая сумма в (5) со скейлинг-функциями $\phi_{j,k}$ содержит средние значения f по диадным интервалам $[\kappa 2^j, (\kappa+1) 2^j]$. Усреднение проводится с весовыми функциями $\phi_{j,k}(x)$. Второе слагаемое содержит все флуктуации f на данном интервале. Эти флуктуации проистекают из всех меньших интервалов, заключенных внутри данного и соответствующих большим значениям параметра масштабирования f . Можно сказать, что это слагаемое как бы фокусирует наше внимание на все более тонких деталях изучаемого сигнала. Это последовательное изучение свойств сигнала происходит через равные промежутки, как следует из постоянства масштабирующего множителя, причем детали размером 2^j располагаются в точках $k 2^j$. На низшем уровне разрешения j_0 , наиболее широких интервалах, первая сумма содержит всего одно слагаемое, дающее общее усредненное взвешенное значение сигнала $\langle f \rangle = s_{j_0, k_0}$, где k_0 обозначает центр гистограммы. Вторая сумма в (5) показывает флуктуации на всех без исключения уровнях. На следующем уровне с более мелким разбиением $j_1 > j_0$ в первой сумме содержатся два слагаемых, отвечающих за средние значения f в полуинтервалах с центрами, расположенными в k_1, k_2 . Число слагаемых во второй сумме уменьшается на единицу – исчезает слагаемое, которое раньше показывало величину флуктуаций на половинном масштабе. Общее число слагаемых в разложении остается неизменным. Число слагаемых в каждой сумме зависит от выбранного уровня разрешения. Меняя на единицу этот уровень, мы передвигаем некоторые слагаемые в другую сумму, причем каждое из этих представлений является истинным представлением исходной гистограммы на разных уровнях разрешения.

Формально эту процедуру можно провести и в обратном направлении, переходя к все более широким интервалам усреднения, то есть к уровням разрешения $j < j_0$. с задания f , можно пытаться продолжить это с усреднением. Тогда среднее значение уменьшается, и в L^2 -смысле можно пренебречь первой суммой в формуле (5), потому что ее L^2 -норма стремится к нулю. В примере с гистограммой оно падает, как $|\langle f \rangle| \propto N^{-1}$ и $|\langle f \rangle|^2 \propto N^{-2}$, тогда как область интегрирования пропорциональна N , то есть

$$\|\langle f \rangle\|^2 \propto N^{-1} \rightarrow 0 \text{ при } N \rightarrow \infty.$$

При решении данной задачи рассматривается только вторая сумма в (5) и результат называют вейвлет-разложением f [5].

Пример вейвлет-преобразования речевого сигнала фонемы «к» представлен на рисунке 2.

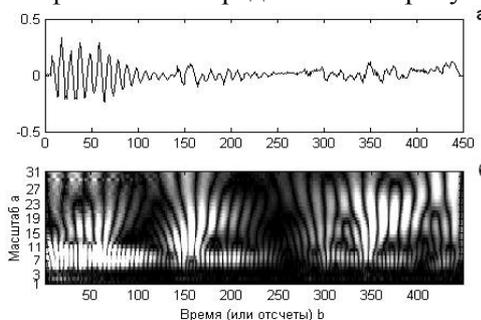


Рисунок 2 – Пример вейвлет-преобразования речевого сигнала фонемы «к»;
а) анализируемый сигнал,
б) картина вейвлет-коэффициентов

Для рассмотренной задачи целесообразно использовать многомасштабный или кратномасштабный вейвлет-анализ. Идея метода состоит в представлении сигнала последовательностью образов с разной степенью детализации, что позволяет выявлять локальные особенности сигнала и классифицировать их по интенсивности.

Экспериментальные исследования. В процессе исследования фонем были проведены эксперименты в системе Matlab, в результате которых были выделены характерные признаки каждой из фонем, представленных с помощью вейвлет-преобразования. В таблице приведены векторы из восьми коэффициентов, которые представляют собой средние значения энергий детализирующих коэффициентов, рассчитанные на каждом уровне вейвлет-разложения. Жирным шрифтом выделены характерные особенности каждой из фонем.

Из таблицы видно, что вейвлет-преобразование методом многомасштабного вейвлет-анализа позволяет выявить характерные признаки каждой фонемы подобно тому, как это происходит на базилярной мембране внутреннего уха человека. Затем векторы подаются на вход искусст-

венной нейронной сети и распознаются подобно тому, как это происходит в человеческом мозге [6]. Следовательно, предложенная система работает по алгоритму человеческого уха – идеальной системы распознавания фонем, созданной самой природой.

Таблица – Векторы средних значений энергий детализирующих коэффициентов, рассчитанные на каждом уровне вейвлет-разложения для фонем ‘а’, ‘и’, ‘о’

a1	и2	o1
1.0e+003	1.0e+003	1.0e+003
0.0020	0.0142	0.0001
0.0117	0.0575	0.0006
0.1657	0.0099	0.0041
5.1521	0.0231	0.2986
0.7928	0.1897	7.1588
0.3462	2.8531	0.2790
0.0895	0.0623	0.0159
0.0388	0.0081	0.0133

На рисунке 3 показаны результаты распознавания фонем системой с помощью оконного преобразования Фурье.

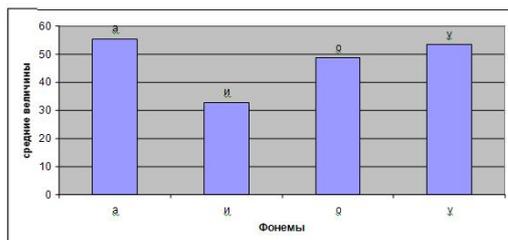


Рисунок 3 – График средних арифметических невзвешенных величин результатов распознавания фонем ‘а’, ‘и’, ‘о’, ‘у’ нескольких людей одного пола системой, использующей оконное преобразование Фурье, обученной на фонемах одного из них

Из рисунка 3 следует, что предполагаемая модель и ее исполнение показывают хорошие результаты правильности распознавания фонем, но все средние величины результатов распознавания меньше шестидесяти. Это обусловлено тем, что метод Фурье преобразования не решает проблему дикторозависимости системы. Следовательно, все системы, построенные на преобразовании Фурье, являются дикторозависимыми, то есть предназначены для работы только с одним пользователем, на фонемах которого обучалась данная система. Однако в нашем случае это невозможно, так как человек, чьи фонемы должна распознавать система, не может правильно произнести их.

На рисунке 4 приведены результаты распознавания фонем системой с помощью вейвлет-преобразований.

Из рисунка 4 видно, что система, основанная на методе вейвлет-преобразования, близка к дикторонезависимым, что для решения нашей задачи является необходимым. Система обучается на векторах фонем правильного произношения, а затем распознает искаженное произношение этой фонемы.

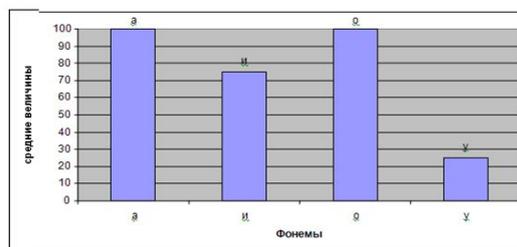


Рисунок 4 – График средних арифметических невзвешенных величин результатов распознавания фонем ‘а’, ‘и’, ‘о’, ‘у’ нескольких людей одного пола системой, использующей вейвлет-преобразование, обученной на фонемах одного из них

Заключение. По результатам проделанной работы можно сделать вывод, что вейвлет-анализ как средство многомасштабного анализа может быть успешно применен для исследования особенностей речевого сигнала. Исходя из результатов эксперимента, показано, что при использовании вейвлет-преобразования для решения задач распознавания речи процент распознавания фонем гораздо выше, чем при использовании преобразования Фурье. Это объясняется тем, что вейвлет-преобразование гораздо ближе к естественному восприятию звука человеческим ухом.

Библиографический список

1. *Витязев В.В.* Вейвлет-анализ временных рядов: учеб. пособие. – СПб.: Изд-во С.-Петербург. ун-та, 2001. – 58 с.
2. *Яковлев А.Н.* Введение в вейвлет-преобразование: учеб. пособие. – Новосибирск: Изд-во НГТУ, 2003. – 104 с.
3. *Дондик Е.М., Пылькин А.Н., Скоробогатова Н.Е.* Информационная система распознавания образов переменной структуры // Вестник Рязанского государственного радиотехнического университета. – № 3 (выпуск 37), 2011. С. 17-23.
4. *Фланаган Д.Л.* Анализ, синтез и восприятие речи: пер. с англ. / под редакцией А.А. Пироговой – М.: Связь 1968. – 392 с.
5. *Вавилова О.С., Пылькин А.Н., Дондик Е.М.* Автоматизированный анализ искажений произносимых фонем // Оптимизация процесса образования и воспитания детей с ограниченными возможностями. Монография. Пенза, 2013. – С.101-111.
6. *Пылькин А.Н., Тишкин Р.В.* Методы и алгоритмы сегментации изображений. – М.: Горячая линия – Телеком, 2010. – 93с.

УДК 621.391

И.В. Баландин

МОДИФИКАЦИЯ ГИБРИДНОЙ МОДЕЛИ РЕЧЕВОСПРИЯТИЯ ДЛЯ АЛГОРИТМА ПОИСКА КЛЮЧЕВЫХ СЛОВ В ПОТОКЕ СЛИТНОЙ РЕЧИ

Предлагается модификация гибридной модели речевосприятия с помощью процедуры моделирования акустического окружения ключевого слова с учетом контекста для обеспечения работы дикторонезависимого алгоритма поиска ключевых слов (АПКС) в потоке слитной речи. Показано, что предложенная модификация обеспечивает уменьшение вероятности ошибки обнаружения ключевых слов до 0,1 и снижение вычислительных затрат до 6 раз по сравнению с моделями, использующими процедуру поиска границ слова.

Ключевые слова: распознавание речи, скрытые марковские модели, гибридные модели, речевой сигнал.

Введение. В настоящее время при решении задачи распознавания речи лучшие результаты показывают системы на основе гибридной модели речевосприятия [1], в которой изменчивость речевого сигнала (РС) в частотной области учитывается при акустическом моделировании искусственной нейронной сетью (ИНС), а моделирование временных зависимостей осуществляется скрытой марковской моделью (СММ). Для построения АПКС в потоке слитной речи [2] они не эффективны, так как требуют наличия фоновой паузы для определения границ слова [3] и недостаточно устойчивы к воздействию мешающих факторов.

В слитной речи, вследствие непрерывного движения органов речеобразования, фоновые паузы между словами могут отсутствовать. Следовательно, в АПКС необходимо включать процедуру поиска границ слова. Для этого, как правило, используется метод скользящего окна [4, 5], в котором расчет меры сходства для каждой модели производится после каждого поступившего вектора параметров сегмента речевого сигнала, соответствующего символу наблюдения СММ. Перебором всех возможных пар начала и конца вхождения ключевого слова определяется наиболее вероятная последовательность состояний каждого ключевого слова. Необходимость перебора создает значительную вычислительную сложность, что затрудняет применение системы на практике.

Другой подход к определению границ слов состоит во включении в словарь «обобщенных» слов [5, 6], представляющих собой сегменты сигнала незнакомого слова или неречевого акустического события. Подготовка таких моделей является сложным процессом, нацеленным на оп-

ределенный набор команд, что не позволяет динамически изменять словарь ключевых слов с сохранением показателей качества обнаружения.

Цель работы – модифицировать гибридную модель речевосприятия для обеспечения работы АПКС в потоке слитной речи и повышения устойчивости к воздействию мешающих факторов.

Разработка модифицированной модели. Для обеспечения работы АПКС в потоке слитной речи предлагается модифицировать гибридную модель восприятия речи, включив в СММ ключевого слова состояние-заполнитель для явного моделирования акустического окружения и введя контекст в ИНС.

Рассмотрим предлагаемую модифицированную гибридную модель, для этого введем параметры СММ [3]:

– множество состояний модели:

$$\{s_1, s_2, \dots, s_U\},$$

где U – число возможных состояний модели;

– множество различных символов наблюдения, порождаемых моделью:

$$\{y_1, y_2, \dots, y_V\},$$

где V – число символов наблюдения модели;

– распределение вероятностей переходов между состояниями:

$$a_{ij} = P(q_t = s_j | q_{t-1} = s_i), \quad i, j = \overline{1, U},$$

где q_t , $t = \overline{1, T}$, – состояние модели в дискретный момент времени t , T – количество наблюдений в последовательности;

– распределения вероятностей появления символов наблюдения y_k в состоянии s_j или эмиссионных вероятностей:

$$b_j(y_k) = P(h_t = y_k | q_t = s_j), \quad j = \overline{1, U}, \quad k = \overline{1, V},$$

где $h_t, t = \overline{1, T}$, — символ наблюдения, поступивший в момент времени t .

Совокупность параметров $a_{ij}, b_j(y_k)$, $i, j = \overline{1, U}, k = \overline{1, V}$, соответствующих z -му ключевому слову, обозначим λ_z .

Для явного моделирования акустического окружения в СММ ключевого слова вводится состояние-заполнитель s_0 . Символы наблюдения, соответствующие акустическому окружению, не являются частью ключевого слова, поэтому эмиссионные вероятности задаются из условия:

$$\begin{cases} b_j(y_k) = 1, j = 0, k > V; \\ b_j(y_k) = 0, j \neq 0, k > V. \end{cases} \quad (1)$$

Вероятность правдоподобия последовательности наблюдений H модели ключевого слова λ_z оценивается алгоритмом прямого-обратного хода [3]:

$$P(H | \lambda_z) = \sum_{i=1}^U \alpha_T(i), \quad (2)$$

где $\alpha_t(i), t = \overline{0, T}$, — прямая переменная, вычисляемая рекуррентно по выражению:

$$\alpha_{t+1}(j) = b_j(h_{t+1}) \sum_{i=0}^U \alpha_t(i) a_{ij}, t = \overline{0, T-1}, j = \overline{0, U}. \quad (3)$$

Начальные значения прямой переменной задаются из условия:

$$\begin{cases} \alpha_0(j) = 1, j = 0; \\ \alpha_0(j) = 0, j \neq 0. \end{cases} \quad (4)$$

На рисунке 1 представлена СММ ключевого слова «клад» с состоянием-заполнителем s_0 , которому соответствуют символы наблюдения y_8, y_9 , моделирующие акустическое окружение.

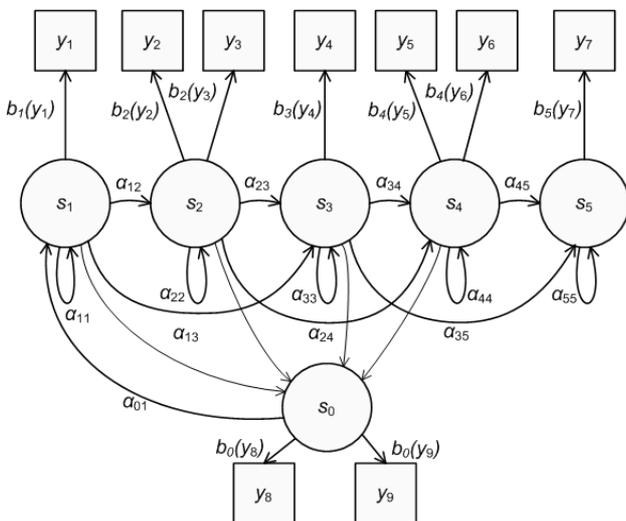


Рисунок 1 – СММ ключевого слова с состоянием-заполнителем

В момент времени $t = 0$ СММ находится в состоянии s_0 . При поступлении на вход модели последовательности символов наблюдения h_t , соответствующих ключевому слову, модель последовательно переходит в состояние s_j , что приводит к обнаружению ключевого слова.

При поступлении на вход модели символа наблюдения h_t , не соответствующего ключевому слову, значение прямой переменной $\alpha_t(j)$, $j = \overline{1, U}$, в соответствии с выражением (3) и условием (1) станет равным нулю, и модель перейдет в исходное состояние s_0 .

Для уменьшения вероятности ложного обнаружения необходимо исключить предположение СММ о статистической независимости последовательных наблюдений [1]:

$$P(h_t | H_1^{t-1}, Q_1^{t-1}) = P(h_t | q_t), \quad (5)$$

где $H_1^{t-1} = [h_1, h_2, \dots, h_{t-1}]$ – частичная последовательность наблюдений от момента времени $t = 1$ до момента времени $t - 1$; $Q_1^{t-1} = [q_1, q_2, \dots, q_{t-1}]$ – частичная последовательность состояний модели от момента времени $t = 1$ до момента времени $t - 1$.

Так как в гибридной модели значение на выходе ИНС является оценкой эмиссионной вероятности

$$s_j^{MI}(y_k) = P(s_j | h_t = y_k), k = \overline{1, V}, j = \overline{1, U},$$

предлагается для повышения точности оценки вероятности последовательности наблюдений на вход ИНС акустического моделирования совместно с вектором акустических параметров символа наблюдения ввести контекст, в качестве которого используется прямая переменная, определяемая выражениями (3), (4):

$$u_t = [\alpha_{t-1}(1), \dots, \alpha_{t-1}(U)]. \quad (6)$$

В этом случае значение на выходе ИНС представляет собой оценку условной вероятности СММ состояния s_j при наличии на входе вектора наблюдения h_t и вектора контекста u_t [6]:

$$s_j^{MI}(y_k) = P(s_j | h_t = y_k, u_t), k = \overline{1, V}, j = \overline{1, U}.$$

Выполнив преобразование, с учетом выражений (3), (6) получим:

$$s_j^{MI}(y_k) = P(s_j | h_t = y_k, H_1^{t-1}), k = \overline{1, V}, j = \overline{1, U}. \quad (7)$$

С учетом вышеизложенного выражение (5) запишется в виде:

$$P(h_t | Q_1^T, H_1^T) = P(h_t | q_t, H_1^{t-1}),$$

где $H_1^T = [h_1, h_2, \dots, h_T]$ — последовательность наблюдений от момента времени $t = 1$ до момен-

та времени $t=T$; $Q_1^T = [q_1, q_2, \dots, q_T]$ – последовательность состояний модели от момента времени $t=1$ до момента времени $t=T$.

Таким образом, применение глобальной рекуррентной связи позволяет частично преодолеть недостаток СММ, связанный с предположением о статистической независимости наблюдений.

Структурная схема поиска ключевых слов на основе модифицированной гибридной модели речевосприятия представлена на рисунке 2.

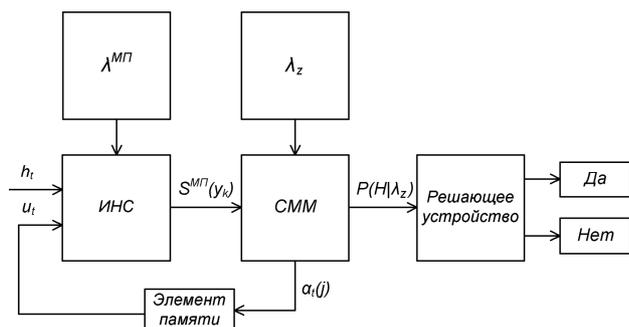


Рисунок 2 – Структурная схема поиска ключевых слов на основе модифицированной гибридной модели

На вход системы поступает очередной вектор акустических параметров, соответствующих символу наблюдения h_t , который совместно с текущим вектором контекста u_t формирует входной вектор ИНС. Весовые коэффициенты λ^{MPP} нейронной сети содержат параметры распределения эмиссионных вероятностей $b_j(y_k)$. Значение на выходе ИНС $S^{MPP}(y_k)$ является оценкой условной вероятности состояний $s_j, j = \overline{1, U}$ (7), которая подается на вход СММ. Алгоритмом прямого-обратного хода (2), (3) оценивается вероятность правдоподобия $P(H|\lambda_z)$ последовательности векторов наблюдений H заданной модели λ_z . Вектор прямых переменных $\alpha_i(j)$, задержанный элементом памяти, используется для формирования вектора контекста на следующем временном шаге u_{t+1} . Решающее устройство в соответствии с заданным порогом принимает решение о наличии в РС ключевого слова.

Результаты экспериментальных исследований. Для оценки эффективности и сравнения с известными алгоритмами предложенной модификации гибридной модели речевосприятия, учитывающей акустический контекст, была построена модель поиска ключевых слов.

Для экспериментального исследования был собран фонетический материал в виде наборов

записей 20 дикторов разного пола без заметных дефектов артикуляции в возрасте от 18 до 35 лет из различных районов страны. Запись осуществлялась дважды в комнате с отсутствием акустических отражений, в разное время суток, при нормальном темпе произношения, обеспечивающем разборчивость речи не ниже пяти баллов в соответствии с ГОСТ Р50840-95.

Ввод и дискретизация РС осуществлялась с помощью ЭВМ при частоте дискретизации 48 кГц и 16-битным представлением отсчетов. Методами цифровой фильтрации выделялась полоса частот 0,3-3,4 кГц, и частота дискретизации понижалась до 8 кГц. Зашумление в виде равномерного в полосе частот РС гауссовского шума моделировалось программно.

Обучение алгоритма производилось на обучающей выборке в виде набора ключевых слов, автоматически сегментированных и размеченных вручную. Вектор параметров символов наблюдения h_t формировался с помощью непрерывного вейвлет-преобразования.

Для каждого из 20 ключевых слов были созданы отдельные СММ и ИНС.

После предварительной оценки параметров СММ обучающая выборка была расширена набором из акустически сбалансированных фраз по ГОСТ Р50840-95, что позволило увеличить акустический контекст для обучения ИНС и моделирования постороннего РС.

При сравнительных исследованиях рассматривались алгоритмы поиска ключевых слов:

- без учета акустического контекста на основе известной СММ с процедурой поиска границ слова алгоритмом Витерби [3];
- с моделями «обобщенных» слов [6];
- на основе предложенной модифицированной гибридной модели речевосприятия.

Результаты оценки временных затрат на поиск ключевого слова представлены в таблице.

Алгоритм	Время поиска, мс
На основе модифицированной гибридной модели	2,5
С поиском алгоритмом Витерби	15
С моделями «обобщенных» слов	10

Для оценки качества поиска ключевых слов использовался критерий Неймана – Пирсона. На рисунке 3 представлены результаты исследования влияния введения контекста в гибридную модель на работу алгоритма поиска ключевого слова в виде зависимости вероятности правильного обнаружения P_t ключевого слова (кривые 1, 2, 3, 4) и ложного обнаружения P_f (кривые 5, 6, 7, 8) от отношения сигнал-шум (ОСШ) q для

гибридной модели без учета контекста (кривые 1, 2, 5, 6) предложенным методом моделирования акустического окружения с учетом контекста (кривые 3, 4, 7, 8).

Получены зависимости для ключевого слова «клад» (кривые 1, 3, 5, 7) и усредненные по 20 ключевым словам (кривые 2, 4, 6, 8).

Анализ результатов эксперимента показал, что предлагаемая модификация гибридной модели речевосприятия с учетом акустического контекста обеспечивает обнаружение ключевых слов в потоке слитной речи без использования процедуры поиска границ слова, что позволяет, в среднем по набору из 20 ключевых слов, в 6 раз снизить вычислительные затраты.

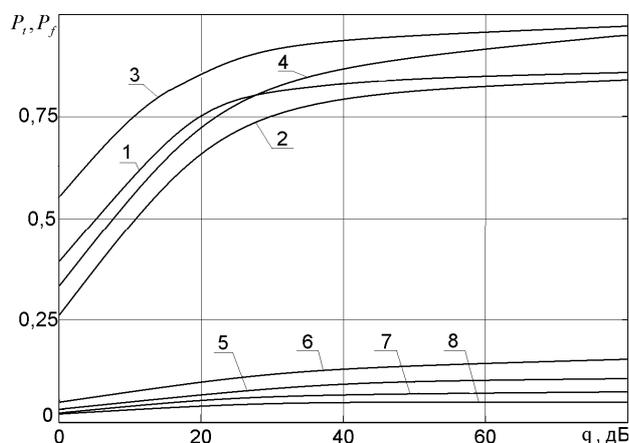


Рисунок 3 – Зависимости вероятности правильного и ложного обнаружения ключевого слова от ОСШ

Вероятность правильного обнаружения по сравнению с известными моделями возросла на 0,1, а вероятность ложного обнаружения снизилась на 0,05. Анализ усредненных по 20 ключе-

вым словам зависимостей показал, что средняя вероятность правильного обнаружения составила 0,95, а вероятность ложного обнаружения – 0,05.

Заключение. В статье предлагается модификация гибридной модели речевосприятия, включающая явное моделирование акустического окружения состоянием-заполнителем с использованием контекста, позволившая исключить из АПКС процедуру поиска границ ключевого слова, что в 6 раз уменьшило время работы алгоритма, и достигнуть вероятности правильного обнаружения ключевого слова 0,95 при вероятности ложного обнаружения 0,05.

Библиографический список

1. Bourlard H., Morgan N. Connectionist Speech Recognition. A Hybrid Approach // The Kluwer International Series in Engineering and Computer Science. – Vol. 247. Kluwer Academic Publishers, Boston, 1994. – 266 p.
2. Дмитриев В.Т., Баландин И.В. Дикторонезависимая система автоматического поиска ключевых слов в потоке слитной речи, устойчивая к акустическим шумам // Вестник РГРТУ. – 2008. – № 2 (24). – С. 15-19.
3. Рабинер, Л.Р. Скрытые марковские модели и их применение в избранных приложениях при распознавании речи // ТИИЭР. – 1989. – Т. 77. – № 2. – С. 86-98.
4. Bridle J. An efficient elastic template method for detecting given words in running speech // British Acoustical Society Meeting, Apr. – 1973.
5. Гребнов С.В. Аналитический обзор методов распознавания речи в системах голосового управления // Вестник ИГЭУ. – Иваново, 2009. – Вып. 3.– С. 83-85.
6. Viterbi A.J., Error bounds for convolutional codes and an asymptotically optimal decoding algorithm. IEEE Trans. Informat. Theory, – Apr. 1967. – Vol. IT-13. – P. 260-269.