

Информационные технологии

УДК 658.512.011

ОПТИМИЗАЦИЯ КОНТРОЛЯ ДОСТОВЕРНОСТИ ИНФОРМАЦИИ В ЭЛЕКТРОННЫХ ДОКУМЕНТАХ НА ОСНОВЕ ГИБРИДНОЙ МОДЕЛИ МЯГКИХ ВЫЧИСЛЕНИЙ

Тишликов С.А., Ахатов А.Р., Рахмонкулов Ф.
Самаркандский государственный университет,
140104, г. Самарканд, Университетский бульвар, 15.

Разработаны основные подходы и методы проектирования и реализации систем интеллектуального анализа данных для повышения достоверности передачи и обработки электронных документов на основе синтеза моделей и алгоритмов нейро-нечеткой сети, сети Петри и эволюционного моделирования. Предложенные методики позволяют получить эффективный инструментарий оптимизации определения и настройки параметров гибридной модели при контроле и коррекции искажений в электронных документах.

Ил. 2. Библиогр. 5 назв.

Ключевые слова: электронный документ, нестационарный объект, интеллектуальный анализ данных, база данных, база знаний, нейро-нечеткая сеть, нечеткие правила, генетические алгоритмы, идентификация, аппроксимация, контроль, достоверность информации.

Тишликов Султон Абдураимович, старший преподаватель кафедры прикладной математики и информатики ГулГУ, тел.: +998933973840, e-mail: tsa_sultonbek@bk.ru

Ахатов Акмал Рустамович, доктор технических наук, профессор кафедры информационных технологий СамГУ, тел.: +998902716418, e-mail: akmalar@rambler.ru.

Рахмонкулов Феруз Пардабоевич, магистрант информационных технологий СамГУ.

ЎМШОҚ ҲИСОБЛАШЛАР ГИБРИД МОДЕЛИ АСОСИДА ЭЛЕКТРОН ҲУЖЖАТ МАЪЛУМОТЛАРИ ИШОНЧЛИГИНИ НАЗОРАТ ҚИЛИШНИ МАҚБУЛЛАШТИРИШ

Тишликов С.А., Ахатов А.Р. Рахмонкулов Ф.
Самарканд давлат университети,
140104, Самарканд ш., Университет хиёбони, 15.

Нейро-нотиниқ тармоқ, Петри тармоғи ҳамда эволюцион моделлаштириш модел ва алгоритмлари синтези асосида электрон ҳужжатни узатиш ва қайта ишлаш ишончлигини ошириш учун маълумотларни тафаккурли таҳлил қилиш тизимини лойиҳалаш ва жорийлаштириш ёндошув ва усуллари ишлаб чиқилган. Таклиф этилган услубият электрон ҳужжатда хатоларни назорат ва таҳрир қилиш гибрид модели параметрини аниқлаш ва мослаштиришни мақбуллаштирувчи самарали воситаларни олиш имкониятини беради.

Ил. 2. Библиогр. 5 номда.

Калитли сўзлар: электрон ҳужжат, ностационар объект, маълумотларни тафаккурли таҳлил қилиш, маълумотлар базаси, билимлар базаси, нейро-нотиниқ тармоқ, нотиниқ қоидалар, генетик алгоритм, идентификация, назорат қилиш, маълумот ишончлиги.

Тишликов Султон Абдураимович, ГулДУ амалий математика ва информатика кафедраси катта ўқитувчиси, тел.: +998933973840, e-mail: tsa_sultonbek@bk.ru

Ахатов Акмал Рустамович, техника фанлари доктори, СамДУ ахборотлаштириш технологиялари кафедраси профессори, тел.: +998902716418, e-mail: akmalar@rambler.ru.

Рахмонкулов Феруз Пардабоевич, СамДУ ахборотлаштириш технологиялари кафедраси магистранти.

OPTIMIZATION of INFORMATION AUTHENTICITY CONTROL in ELECTRONIC DOCUMENTS ON THE BASIS OF HYBRID MODEL of SOFT CALCULATIONS

Tishlikov S.A., Akhatov A.R. Rakhmonkulov F.
Samarkand State University,
15, University Blvd. Samarkand, 140104

In the paper authors are developed the basic approaches and methods of designing and realization of data intellectual analysis systems for increase of the electronic documents authenticity during transfer and processing on the basis of synthesis of neuro-fuzzy network, Petri network and evolutionary modeling models and algorithms. The offered techniques allow to receive effective toolkit of optimization for hybrid model parameters definition and adjustment at the control and correction of distortions in the electronic documents.

Il.2. Bibl. 5 names.

Key words: electronic document, non-stationary object, intellectual analysis of data, database, base of knowledge, neuro-fuzzy network, fuzzy rules, genetic algorithms, identification, approximation, control, information authenticity.

Tishlikov Sul-ton Abdurayimovich, the senior teacher of chair of applied mathematics and informatics, GulSU, tel.: +998933973840, e-mail: tsa_sultonbek@bk.ru

Akhatov Akmal Rustamovich, doctor in technical science, professor of chair of information technologies, SamSU, tel.: +998902716418, e-mail: akmalar@mail.ru.

Rakhmonkulov Feruz Pardaboyevich, master of chair of information technologies SamSU.

Актуальность темы. Проектирование практических приложений, связанных с построением систем электронного документооборота (СЭД) на основе методов интеллектуального анализа данных (ИАД), а также решение задач распознавания, обнаружения и исправления ошибок в электронных документах (ЭД) в реальном масштабе времени представляет перспективную и малоизученную тему исследований [1]. Сложный характер динамических процессов, происходящих на этапах передачи и обработки ЭД, неоднозначность, неполнота исходной информации являются аргументами разработки методов и алгоритмов оптимизации повышения достоверности информации на основе моделей нейро-нечетких сетей (ННС) [2]. Система ИАД, построенная на основе синтеза ННС с моделями и средствами нечеткой логики формирует нелинейные принципы нечетких выводов, в результате чего создается эффективный инструментарий для решения прикладных задач различных проблемных областей, в частности обработки данных нестационарных объектов [3].

Источники информации в СЭД представляют разнообразные формы служебных, организационно-распорядительных документов, содержащие цифровую, текстовую информацию, которая для хранения и обработки преобразовывается в формат ЭД. Формальная модель функционирования СЭД включает вектора: входной информации $x = (x_1, x_2, \dots, x_j)$, выходной информации $y = (y_1, y_2, \dots, y_n)$ и вектора, влияющих помех и возмущений $u = (u_1, u_2, \dots, u_i)$. Входные данные могут искажаться из-за ошибок оператора, погрешностей сканирующих и распознающих устройств, влияния помех в каналах связи, а также из-за сбоев и отказов электронных средств передачи и обработки информации [1]. В итоге результаты функционирования системы представляются с некоторой погрешностью $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m)$. В связи с этим актуальными считаются проблемы поиска, распознавания, классификации образов для разработки алгоритмов повышения достоверности информации в СЭД на основе методов ИАД. Проблема повышения достоверности информации, в свою очередь, связана с идентификацией процессов передачи и обработки ЭД на основе использования концепций моделирования нестационарного объекта, которые сводятся к тому, что входы и выход систем представляются нелинейными функциональными зависимостями.

Основные подходы и принципы контроля достоверности информации ЭД на основе ННС. Модели системы ИАД представляются в виде некоторой функции $y = f(x, \varepsilon, u)$. Эффективность моделирования повышается путем определения и настройки параметров модели передачи и обработки информации на основе ННС, способной учитывать динамику входного воздействия u , погрешности ε и использования свойств данных x при недостаточных априорных сведениях и большой неопределенности параметров.

На рис.1 показана схема выполнения формальной модели системы ИАД в виде нелинейной функции зависимостей входов и выход.

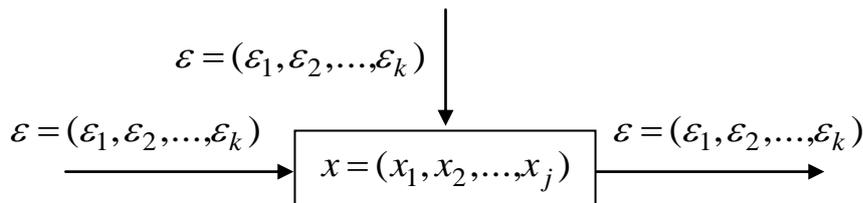


Рис.1. Формирование зависимостей входов и выход системы ИАД.

Конечной задачей ННС системы ИАД является аппроксимация выхода в зависимости от динамики входа для выработки логических выводов и принятия решений. В традиционном подходе для моделирования описаний нестационарного объекта используются линеаризованные дифференциальные (разностные) уравнения с конкретными начальными условиями, в которых не учитывается динамика переходного процесса. Однако, реальная оценка состояний этапов передачи и обработки ЭД требует определения адекватной упрощенной модели нелинейной зависимости $y = \varphi(u)$ при неполно заданной информации с учетом динамики векторов x , y , u , ε . В связи с этим, в предлагаемом подходе настройка параметров модели описаний нестационарного объекта осуществляется на основе реальных выборок, вычислительных схем компонентов ННС, регистрации в системе вектора входов x , выходов y , вектора влияния помех u и статистики искажений информации.

Пусть зарегистрированы фиксированные вектора u и y объекта. Погрешность представления влияния вектора u на выходной вектор y задается в виде гауссовых функций принадлежности (ФП), центр которых определяется значениями в начальном состоянии u_0 и y_0 , а выполнение критерия оценки достоверности информации ЭД задается в виде условия

$$\mu_{notN}(u, y) = 1. \quad (1)$$

Условие (1) выполняется тогда, когда за границу нечеткого лингвистического термина N (normal) выходит какой-либо из компонентов вектора входов u_i или же выходов y_i [2,4].

Область допустимых значений лингвистического термина N определяется на основании статистики искажений информации, специальных требований к отклонениям формата, содержания и структуры ЭД от характеристик модальных примеров, размещенных в БД системы ИАД.

Окончание переходных процессов, связанных с передачей и обработкой ЭД задается в виде условия:

$$\mu_N(u, y) = k, \quad k \in (0,1). \quad (2)$$

Условия (1) и (2) определяют выборку переходного процесса для исследования зависимостей «вход-выход». Нелинейная функция $y = \varphi(u)$ идентифицирует и аппроксимирует зависимости, затем полученные результаты сравниваются с реальным выходом y_r во всех n точках выборки переходного процесса для принятия решений о достоверности ЭД.

Качество моделирования динамических переходных процессов в текущем времени оценивается по критерию относительного среднеквадратического отклонения $\delta_{\text{о.а.е}}$ путем проверки соответствия реального принятого ЭД к максимального значения эталонного примера δ_{max} , что отождествляется с достоверностью передачи и обработки информации, т.е. имеем коэффициент достоверности

$$K_{\bar{A}} = \frac{\delta_{\text{о.а.е}}}{\delta_{\text{max}}},$$

причем в реальных условиях $\delta_{\text{о.а.е}}$ значительно меньше δ_{max} .

Определение набора лингвистических термов, нечеткие правила контроля, проверка соответствия принятого ЭД к эталонному, а также определение зависимости влияния вектора входов u на выход y являются важными задачами формирования базы знаний (БЗ) системы ИАД.

Компоненты вектора $u = (u_1, u_2, \dots, u_i)$ в рассматриваемом контексте представляются следующими тремя нечеткими лингвистическими термами, определяемые относительно начального установившегося режима:

L (low), когда при проверке соответствия принятого ЭД к эталонной ЭД настраиваемый параметр модели снижается;

N (normal), когда отклонение параметра не выходит за определенный диапазон $\pm \Delta$;

H (high), когда значение параметра повышается.

Выбор этих термов для настройки модели ННС в переходном режиме обосновывается тем, что значение параметров могут увеличиваться, уменьшаться или остаться неизменными в диапазоне $u_{i0} \pm \Delta$ относительно значения начального установившегося режима. БЗ имеет полный набор нечетких правил контроля поведения вектора выходов y при всех возможных изменениях вектора u .

Если полученная идентификационная и аппроксимирующая модель переходных процессов не обеспечивает требуемое соответствие принятого ЭД к эталонной ЭД, то количество лингвистических термов и количество нечетких правил контроля вектора входов может быть увеличено. Этим достигается последовательная настройка параметров модели переходного процесса ко всем возможным изменениям входных параметров на основе ННС.

Переходный процесс, образуемый передачей и обработкой ЭД является вероятностным при произвольных начальных значениях векторов x и u . Ниже для расширения положений нечеткого описания вероятностных процессов, определения предпосылок нечетких правил, аппроксимирующей модели предлагается использовать метод символьной нечеткой переменной.

Отличительные особенности предложенного подхода заключаются в том, что теперь границы лингвистических термов заранее не задаются жестко численными значениями, а подстраиваются к значениям установившегося режима процесса путем задания допустимого отклонения $\pm dx$ для каждого терма. Требуемое условие обеспечения достоверности передачи ЭД отыскиваются по числу переданных примеров (ЭД) N , отклонением dx принятого ЭД от эталонной ЭД, областью существования ФП лингвистических переменных на основе изменения от 0 до 1.

Вывод каждого правила, формирующийся как линейная комбинация параметров предпосылок и свободной составляющей, представлен в табл.1:

Таблица 1

Формирование правил нечеткой аппроксимации

u_1	u_2	...	u_i	u_n
MF_{11}	MF_{12}	...	MF_{1i}	Y_{n1}
MF_{21}	MF_{22}	...	MF_{2i}	Y_{n2}
...
MF_{j1}	MF_{j2}	...	MF_{ji}	Y_{nj}

На начальном этапе функционирования системы ИАД на основе ННС для ФП вектора входов u выбираются гладкие, дифференцируемые функции гауссового типа [4]. Аппроксимирующая модель выхода системы ИАД определяется i – компонентами входного вектора u и j – режимом изменения нестационарного

Слой 6. Единственный неадаптивный узел этого слоя вычисляет полное выходное значение по оценке достоверности передачи ЭД на основе аппроксимирующей модели переходного процесса в рассматриваемой точке (ЭД) как сумму всех нормированных выводов каждого правила.

Отметим, что реализация изложенной методики построения системы ИАД, несмотря на эффективность при проектировании многих практических приложений, связана с большими итерационными процессами и трудоемким поиском глобального оптимума и перебором параметров при обучении НС, что требует проектирования более упрощенных алгоритмов обучения ННС, выполняемых с меньшей вычислительной сложностью и материальными затратами.

В связи с этим, в настоящей работе решена также новая задача, связанная с оптимизацией определения и настройки модели переходного процесса передачи и обработки ЭД (количество параметров, структура и время) путем синтеза ННС с генетическим алгоритмом (ГА) [4]. Однако, оптимизация структуры модели переходного процесса требует учета специфики предметной области для реализации системы ИАД.

Одним из подходящих подходов для предметной области исследуемой проблемы является синтез эволюционного и стохастического моделирования по механизмам выполнения ГА и сети Петри (СП) [5]. Причем, СП свойственно моделирование некоторого действия по преобразованию входных данных в выходные данные, а ГА свойственно определение оптимальных параметров синтезированных моделей на основе настройки с помощью генетических операторов.

Представление стохастической модели на основе СП. Запишем формализованную модель функционирования системы ИАД на основе СП для контроля достоверности ЭД в следующем виде [5]:

$$S = \left\langle In, Out, \{S_k\}_{k=1}^K, \{f_a\}_{a=1}^A, \{F_b\}_{b=1}^B \right\rangle,$$

где S – проектируемый комплекс повышения достоверности ЭД на основе системы ИАД;

In – множество входных данных;

Out – множество выходных данных;

S_k – k -ый модуль комплекса S ;

f_a – функция соответствия зависимостей «входы-выход» системы: $f_a : In \rightarrow Out$;

F_b – бинарное отношение на множестве $\{S_k\}_{k=1}^K : F_b \subset \{S_k\}_{k=1}^K$.

Согласно модели СП для заданной функции f_{a_0} требуется подобрать бинарное отношение F_{b_0} такое, чтобы множество модулей комплекса $\{S_k\}_{k=1}^K$ обеспечивало обработку входных данных, т.е. требуется получение выхода в соответствии с некоторой функцией f_{a_0} . Согласно S входы системы отображаются множеством позиций $P_{in} = \{p_m^{in}\}_{m=1}^M$, где M – количество входов, а ее выходы – множеством позиций $P_{out} = \{p_n^{out}\}_{n=1}^N$, где N – количество выходов.

Входы модуля S_k – это множество $P_k^{in} = \{p_{k,m}^{in}\}_{m=1}^{M(k)}$, где $M(k)$ – количество входов модуля S_k , а ее выходы – множество $P_k^{out} = \{p_{k,n}^{out}\}_{n=1}^{N(k)}$, где $N(k)$ – количество выходов модуля S_k ; P^{in} , P^{out} – обозначение множества позиций. Входы модуля $\{S_k\}_{k=1}^K$ связаны как между собой, так и с входами и выходами комплекса S . Эти связи отображаются множеством переходов $T = \{t_q\}_{q=1}^Q$, где Q – количество переходов.

Каждый переход t_q характеризуется входным и выходным позициями. Входами перехода t_q могут быть как любые входы комплекса S , так и любые выходы модуля $\{S_k\}_{k=1}^K$. Выходами перехода t_q могут быть как любые выходы комплекса S , так и любые входы модуля $\{S_k\}_{k=1}^K$. Обозначим входы перехода t_q через In_q , а выходы – через Out_q , которые представляются в виде

$$In_q \subset P_{in} \cup \left(\bigcup_{k=1}^K P_k^{out} \right) \text{ и } Out_q \subset P_{out} \cup \left(\bigcup_{k=1}^K P_k^{in} \right).$$

Для каждого входа S существует переход t_q , соединенный с этим входом, а для каждого выхода модуля S_k существует переход t_q , соединенный с этим выходом, которые представляются в виде

$$\bigcup_{q=1}^Q In_q = P_{in} \cup \left(\bigcup_{k=1}^K P_k^{out} \right).$$

Аналогично, для каждого выхода S существует переход t_q , соединенный с этим выходом и для каждого входа каждого модуля S_k существует переход t_q , соединенный с этим входом, которые записываются в виде

$$\bigcup_{q=1}^Q Out_q = P_{out} \bigcup \left(\bigcup_{k=1}^K P_k^{in} \right).$$

Множество модулей $\{S_k\}_{k=1}^K$ вместе с подмножеством множества переходов $T = \{t_q\}_{q=1}^Q$ полностью определяют текущую структуру комплекса S . Настройка структуры S определяется изменением подмножества задействованных в данный момент переходов.

Следовательно, согласно предложенному подходу нестационарный переходной процесс передачи ЭД идентифицируется на основе ННС, изменение состояния моделируется с помощью СП, а настройка параметров такой модели производится с помощью ГА.

Оптимизация достоверности ЭД на основе синтеза ННС, СП и ГА. Обобщенная гибридная модель переходного процесса передачи и обработки ЭД настраивается путем задания операторов ГА. Генотип G , представляющий последовательность передаваемых ЭД имеет следующий вид:

$$G = (g_1, \dots, g_q, \dots, g_Q, h_1, \dots, h_k, \dots, h_K),$$

где $g_q \in \{0,1\}$ и $h_k \in \{0,1, \dots, r, \dots, R(k)\}$.

Если $g_q = 0$, то переход t_q отсутствует в модели комплекса S , иначе, когда $g_q = 1$ переход t_q присутствует в модели комплекса S . Если $h_k = 0$, то в модели модуля S_k отсутствуют переход. Если $h_k = r$ ($r \in \{1, \dots, R(k)\}$), то в модели модуля S_k представлен переход $t_{k,r}$. Число C всех гипотетически возможных моделей комплекса S определяется, как

$$C = 2^Q \prod_{k=1}^K (R(k) + 1).$$

Отсюда видно, что даже при небольшом количестве связей количество модулей комплекса S будет очень велико. Так например, при $Q = 10$, $K = 10$ и $R(1) = R(2) = \dots = R(10) = 10$ получаем $C = 2^{10} \times 11^{10} = 22^{10} = 26559922791424$, что обосновывает решения задачи оптимизации определения и настройки параметров модели переходного процесса на основе ГА.

Теперь из имеющегося множества модулей $\{S_k\}_{k=1}^K$ и переходов $T = \{t_q\}_{q=1}^Q$ требуется построить такой комплекс S , который на входной вектор реагировал бы соответствующим выходным результатом для принятия решений по обеспечению достоверности передачи ЭД. При этом, чем значение выходного вектора (принятого ЭД) будет ближе к требуемому (эталонной ЭД), тем адекватнее будет модель переходного процесса передачи информации. Для оценки этой близости, т.е. соответствия переданного ЭД к эталонной ЭД введем следующие обозначения:

вектор $V = (v_1, \dots, v_n, \dots, v_N)$, где v_n – количество меток, которые должны находиться в позиции P_n^{out} комплекса S , а $v_n \in \{0,1,2,3, \dots\}$;

вектор $W = (w_1, \dots, w_n, \dots, w_N)$, который получается в результате работы комплекса S , где w_n – количество меток, которые реально находятся в позиции P_n^{out} комплекса S , а $w_n \in \{0,1,2,3, \dots\}$.

Расстояние между признаками требуемого и реального векторов оценивается

$$\rho(V, W) = \sum_{n=1}^N |v_n - w_n|.$$

При реализации ГА рекомендуется применить

$$\rho(V, W) = \left(\sum_{n=1}^N |v_n - w_n|^p \right)^{1/p},$$

которая для $p = 1$, $p = 2$ и $p \rightarrow \infty$ выполняет рассмотренные выше формулы.

Изменение параметра p с помощью ГА приводит к изменению эффективности системы ИАД. Чем меньше p , тем ближе принятая ЭД соответствуют к требуемой, т.е. достоверной модели. При $p = 0$ переданная ЭД в системе полностью соответствует заданному требованию и считается достоверным.

Теперь изложим выполнение ГА для оптимизации определения и настройки параметров модели переходного процесса передачи ЭД на основе генетических операторов.

Оператор отбора предназначен для скрещивания тех генотипов, которые ближе всего к эталонной ЭД. Принцип контроля достоверности передачи ЭД заключается в упорядочивании всех принятых генотипов по качеству (от лучшего к худшему) и скрещивании. Разрешается лучшим генотипам участвовать в нескольких скрещиваниях. Для простоты рассмотрим два генотипа

$$G_1 = (g_1^1, \dots, g_q^1, g_{q+1}^1, \dots, g_Q^1, h_1^1, \dots, h_k^1, h_{k+1}^1, \dots, h_K^1) \text{ и } G_2 = (g_1^2, \dots, g_q^2, g_{q+1}^2, \dots, g_Q^2, h_1^2, \dots, h_k^2, h_{k+1}^2, \dots, h_K^2).$$

Случайным образом выбирается два числа: q из диапазона $\{1,2,\dots,Q\}$ и k из диапазона $\{1,2,\dots,K\}$. А затем меняются соответствующие участки генотипов.

Из родителей G_1 и G_2 получаются потомки G_3 и G_4 , наследующие свойства родителей:

$$G_3 = (g_1^1, \dots, g_q^1, g_{q+1}^2, \dots, g_Q^2, h_1^1, \dots, h_k^1, h_{k+1}^2, \dots, h_K^2) \text{ и } G_4 = (g_1^2, \dots, g_q^2, g_{q+1}^1, \dots, g_Q^1, h_1^2, \dots, h_k^2, h_{k+1}^1, \dots, h_K^1).$$

Можно ограничиться и выбором какого-то одного числа (q или k), а также выбором для каждого диапазона $\{1,2,\dots,Q\}$ и $\{1,2,\dots,K\}$ по две точки и меняться «серединками».

Оператор мутации. Рассматривается один генотип $G = (g_1, \dots, g_q, \dots, g_Q, h_1, \dots, h_k, \dots, h_K)$. Случайно выбирается два числа: q из диапазона $\{1,2,\dots,Q\}$ и k из диапазона $\{1,2,\dots,K\}$. Затем меняются значения g_q и h_k . Если $g_q = 0$, то меняем его на единицу: $g_q = 1$. И наоборот: если $g_q = 1$, то станет $g_q = 0$.

Если $h_k = r$, то меняем его на любое другое из диапазона $\{0,1,\dots,r-1,r+1,\dots,R(k)\}$. Когда $r = 0$, то диапазон замены будет $\{1,\dots,R(k)\}$ и когда $r = R(k)$, то диапазон $\{0,1,\dots,R(k)-1\}$.

Оператор редукции удаляет генотипы, несоответствующие эталонной модели ЭД. Для этого, после скрещивания определяется качество всех потомков, т.е. оценивается соответствие принятого ЭД к эталонной ЭД. Затем потомков и родителей объединяют в одно множество генотипов и упорядочивают их по качеству от лучшего к худшему. Генотипы (ЭД), попавшие в худшую половину удаляются.

Таким образом выполнение алгоритма синтеза СП и ГА в одном цикле контроля достоверности информации сводится к следующим процедурам:

- начальная популяция генотипов G_1, G_2, \dots, G_{2D} размещается в позиции СП;
- оператор селекции *sel* осуществляет отбор генотипов для скрещивания;
- операторы кроссоверы $cross_1, \dots, cross_d, \dots, cross_D$ осуществляют скрещивание генотипов, которые затем подвергаются мутации оператором *mut*;
- цикл работы завершает оператор *red*, удаляющий слабые генотипы.

В случае неблагоприятных исходов работы системы ИАД процесс контроля достоверности информации повторяется. Остановка циклического процесса выполнения ГА может определяться путем задания нижнего предела качества ρ всей популяции или нижнего предела качества ρ части популяции.

Таким образом, в результате теоретических исследований предложены методики реализации алгоритмов: контроля достоверности передачи и обработки ЭД на основе положений ННС; оптимизации достоверности передачи и обработки ЭД на основе упрощенной модели СП и формирования генотипов;

контроля достоверности с функциями обнаружения и исправления искажений при передаче и обработке ЭД на основе генетических операторов скрещивания, мутации и редукции.

Система ИАД для контроля достоверности информации с функциями обнаружения ошибок передачи и обработки ЭД на основе гибридной модели синтеза ННС, СП и ГА реализована в компьютере с операционной системой Windows XP Professional и среде MATLAB 6.1. Для тестирования системы сформированы выборки ЭД размером от 1200 до 8000 символов. В результате функционирования системы ИАД определено, что способность обнаружения искажений передачи и обработки ЭД повышается на два порядка по критерию относительной среднеквадратической ошибки. В заключении отметим, что для реализации системы ИАД с функциями исправления ошибок передачи и обработки ЭД с целью обеспечения достоверности информационных ресурсов СЭД необходимо проведение исследованию по расширению и развитию предложенных методик.

Литература

1. Тишликов С.А. Модели и алгоритмы проверки достоверности элементов текста в системе контроля орфографических ошибок // «Проблемы информатики и энергетики», Издательство «Фан» АН РУз, - Ташкент, 2012 - №4-5. - с.91-96
2. Ахатов А.Р. Зарипова Г.И. Методы и алгоритмы обучения нейро-нечеткой системы контроля достоверности передачи и обработки данных нестационарных объектов // Журнал «Вестник ТУИТ», Ташкентский университет информационных технологий. - Ташкент, 2013. - № 3/2013. - с. 101-110
3. Jang J.S.R. ANFIS: Adaptive network based fuzzy inference systems // IEEE Trans. on Systems, Man, and Cybernetics.- May 1993.- 23(03).- P. 665-685.
4. Ахатов А.Р., Исроилов Н.О. Применимость нечетких генетических алгоритмов в системе анализа и обработки данных нестационарных объектов // «Актуальные проблемы прикладной математики и информационных технологий – Ал-Хоразмий 2014». – Самарканд, 2014. - с.67-68.
5. Петросов Д.А. Представление генетических алгоритмов Сетями Петри в задачах проектирования компьютерной техники / Д.А. Петросов, В.Г. Лобода, Д.Б. Ельчанинов // Материалы научно-практической конференции «Информационные технологии в науку и образование». – Х. : ХНУРЭ, 2005. – С. 48–51.

Отзыв

на статью С.А. Тишликова, А.Р. Ахатова, Ф.Рахмонкулова «Оптимизация контроля достоверности информации в электронных документах на основе гибридной модели мягких вычислений»

Сложный характер динамических процессов, происходящих на этапах передачи и обработки электронных документов (ЭД), неоднозначность, неполнота исходной информации являются аргументами разработки методов и алгоритмов оптимизации повышения достоверности информации на основе моделей нейро-нечетких сетей (ННС). Проектирование практических приложений, связанных с построением систем электронного документооборота (СЭД) на основе методов интеллектуального анализа данных (ИАД) и аппарата мягких вычислений, в частности синтеза ННС с моделями и средствами нечеткой логики формирует нелинейные принципы нечетких выводов, в результате чего создается эффективный инструментарий для решения прикладных задач различных проблемных областей.

В рецензируемой работе предложен новый подход построения системы передачи и обработки текстовой информации на основе принципов многоступенчатого формального и логического контроля правильности заполнения полей баз данных. Предложены математические модели для разработки алгоритмов кодирования и преобразования полей баз данных.

При реализации методики построения системы ИАД для оптимизации больших итерационных процессов и поиска глобального оптимума обучения ННС разработаны упрощенные алгоритмы, выполняемые с меньшей вычислительной сложностью и материальными затратами.

В работе решена задача определения и настройки модели переходного процесса передачи и обработки ЭД (количество параметров, структура и время) путем синтеза ННС с генетическим алгоритмом (ГА). Осуществление оптимизации структуры модели переходного процесса проведено с учетом специфики предметной области. Проведен синтез эволюционного и стохастического моделирования по механизмам выполнения ГА и сети Петри (СП). Используются особенности ГА и СП при моделировании преобразования входных данных в выходные данные и определении оптимальных параметров синтезированных моделей.

Представленная на рецензию статья авторов С.А.Тишликова, А.Р.Ахатова и Ф.Рахмонкулова посвящена решению актуальной научно-технической задачи, выполнена на высоком теоретическом уровне, соответствует предъявляемым требованиям и может быть рекомендована к открытой публикации.

Профессор кафедры «Информационные технологии» СамГУ, д.т.н.

И.И.Жуманов