

# ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ МЕТОДОВ КОНТРОЛЯ ДОСТОВЕРНОСТИ ОБРАБОТКИ ЭЛЕКТРОННЫХ ДОКУМЕНТОВ НА ОСНОВЕ НЕЧЕТКОЙ ГИПЕРСЕТИ

Ахатов А.Р. Email: Akhatov17101@scientifictext.ru

*Ахатов Акмал Рустамович – доктор технических наук, профессор,  
кафедра информационных технологий,  
Самаркандский государственный университет, г. Самарканд, Республика Узбекистан*

**Аннотация:** в статье исследованы инструменты оптимизации достоверности передачи и обработки электронных документов (ЭД) для расширения состава известных технологий и применения в различных предметных областях. Разработана модель документооборота, ориентированная на применение UML-диаграмм связей ЭД, которая позволяет отражать связи между структурными компонентами организационно-распорядительных документов. Спроектирована семантическая гиперсеть для поиска признаков ЭД на основе синтеза моделей стохастического поиска, нечеткой логики, отражения функциональных связей переменных входов и выходов с помощью функции принадлежности лингвистических термов. Доказано улучшение показателя достоверности информации при реализации разработанных методов по сравнению с существующими технологиями.

**Ключевые слова:** электронный документ, достоверность, эталонная проверка, изображение, сигнальные характеристики, фрактальная обработка, лексикологический синтез, структура документа.

## INCREASING OF EFFECTIVENESS OF METHODS TO CONTROL OF PROCESSED ELECTRONIC DOCUMENTS AUTHENTICITY ON THE BASIS OF FUZZY GYPHERNET

Akhatov A.R.

*Akhatov Akmal Rustamovich – Dr. Sc. in Technics, Professor,  
INFORMATION TECHNOLOGIES DEPARTMENT,  
SAMARKAND STATE UNIVERSITY, SAMARKAND, REPUBLIC OF UZBEKISTAN*

**Abstract:** the article examines tools for optimizing the authenticity of electronic documents (ED) during transmission and processing to expand the known technologies and applications in various subject areas. The workflow model, focused on the use of UML-diagrams of ED relations is developed for allowing reflect of links between the structural components of organizational and administrative documents. The semantic hypernetwork is designed for search ED features on the basis of synthesis of models of stochastic search, fuzzy logic, reflection functional relationships of inputs and outputs variables by functions of linguistic membership terms. Improvement for authenticity of information by implementation of developed methods is proved by comparing with existing technologies.

**Keywords:** electronic document, authenticity, reference check, image, signal characteristics, fractal processing, lexicological synthesis, document structure.

УДК 658.512.011

**Актуальность темы.** Анализ современной научной литературы показал, что несмотря на широкое распространение существующих сервисных технологий, в них недостаточно решены вопросы обеспечения достоверности информации электронных документов (ЭД), которые являются одним из ключевых факторов повышения качества передачи и обработки данных [1]. Отсутствие специализированных программных продуктов, эффективных инструментов, выполняющих функции оптимизации достоверности передачи и обработки ЭД создаёт дополнительные сложности и приводит к снижению качества функционирования организационных систем управления в различных предметных областях [2].

В связи с этим в настоящей работе обосновывается исследование для разработки программно-алгоритмического комплекса, ориентированного на повышение достоверности информации в ЭД на основе поиска, распознавания и классификации, контроля достоверности реквизитов, атрибутов, форматов с целью обнаружения и коррекции искажений [3].

**Основные подходы к построению программно-алгоритмического комплекса повышения достоверности информации ЭД.** С целью создания программно-алгоритмических комплексов повышения достоверности информации разработана модель документооборота, ориентированная на применение UML-диаграмм связей ЭД, которая позволяет отражать связи между структурными

компонентами организационно-распорядительных документов (ОРД), принятых в качестве базовых при построении СЭД. Кроме того, предложенные схемы документооборота способствуют эффективному проектированию инфологических и даталогических моделей, на основе которых строятся методы и алгоритмы поиска, распознавания, контроля достоверности структурных элементов ЭД. Информационная модель документооборота с UML-диаграммой используется для детализации структуры ЭД с целью отражения картины взаимодействия и взаимосвязей элементов для формирования баз данных и знаний (БД и БЗ), изложения объектной и функциональной частей СЭД.

Методы и алгоритмы обработки данных, реализуемые в комплексе повышения достоверности информации ЭД предназначены для решения следующих задач [1]:

- построение вычислительных схем упрощенной идентификации, аппроксимации переходных процессов сигнальных характеристик кадров ЭД;
- оптимизация распределения и закрепления потоков в соответствии с заданиями обработки ЭД, возможностью вычислительных ресурсов, нагрузкой составных программных модулей, приоритетов, ограничений на длительность обработки и исполнения ОРД;
- разработка методов обнаружения и исправления искаженной информации, обеспечения достоверности, целостности и сохранности ЭД;
- разработка алгоритмов оперативного анализа и обработки ЭД, учёта ОРД, внесения необходимых коррекций в составе документов;
- моделирование и идентификация инфологической последовательности задач обработки данных, формирование БД и БЗ;
- разработка моделей описания связей между входами и выходами каждого программного модуля СЭД, проектирование семантической гиперсети для поиска, анализа, обработки ЭД;
- разработка схем интеграции в виде семантической гиперсети на основе синтеза методов оптимизации обработки ЭД.

Программный комплекс включает следующие типовые модули:

- интегратор - средства настроек, объединений, импорта, экспорта данных;
- семантический синтезатор, анализатор текста и средства генерации ЭД;
- адаптер, построенный на основе методов и алгоритмов ИАД;
- интерфейс загрузки данных, библиотек ключевых слов и словарей;
- конвертор - преобразователь данных; распределитель движения ОРД по задачам;
- встроенный функционал CASE-технологий, в которые встроены группы таблиц БД, БЗ, конвертирующиеся в формате SQL.

При запуске СЭД входными являются разрозненные версии ЭД, библиотека классов и программ ЭД, задания на обработку ЭД, конечный формат представления документа-шаблона [2].

Входные ЭД представляются в виде записей, составляющие которых разбиваются на заданное количество полей.

**Гиперсемантическая сеть поиска информативных признаков ЭД.** Предложен проект модуля четырехуровневой семантической гиперсети и оптимизации функционирования СЭД, который выполняет функции в следующих блоках [3]:

- выполнение функционала семантического анализа как поступающего, так хранящегося в БД;
- формирование базовых словарных библиотек, ключевых слов заданий, сообщений, замечаний, рекомендаций;
- формирования ответвлений заданного тематического направления СЭД, сохранение ЭД на момент обработки, формирование тематических заголовков и их состава;
- синтез, анализ, генерацию, трансляцию текстов.

Семантическая гиперсеть в среде СЭД выполняет функции поиска необходимого элемента ЭД, анализа связей между ними. Сеть построена на основе синтеза моделей стохастического поиска, нечеткой логики, методов формирования и правил использования БЗ нечетких правил, а также отражения функциональных связей переменных входов и выходов с помощью функции принадлежности лингвистических термов.

Для оптимизации обучения семантической гиперсети сформированы рациональные обучающие наборы, производится редукция избыточных связей между составными компонентами ЭД и неэффективными нечеткими правилами в БЗ.

Алгоритмы, выдающие результаты работы НСГС синтезируют в себе процедуры настройки параметров. Поиск структурных элементов ЭД осуществляется на основе реализации соответствующего модуля НСГС, в котором картины полученных связей объектов-узлов представляются в виде нечеткого графа.

Графовая модель в свою очередь используется для оптимизации настройки связей полей, поиска связей, слов в библиотеках, определения тематических направлений с последующим формированием групп тем и их состава с последующими дополнениями.

Алгоритмы обеспечения достоверности, целостности и сохранности информации основываются на сравнение сходимости поискового компонента ЭД с эталонными формами из БД, необходимые свойства и характеристики которых помещены в БЗ.

Для выполнения функции поиска, анализа, обеспечения достоверности, целостности и сохранности информации отражаются картины связей между компонентами ЭД на следующих уровнях:

- на первом отражаются картины связей между ЭД;
- на втором отражаются связи между атрибутами ЭД;
- на третьем отражаются связи между признаками ЭД;
- на четвертом осуществляется увязка правил контроля достоверности, целостности и сохранности информации;
- на пятом отражаются связи входов и выходов НСГС;
- на шестом располагается выходной слой сети для принятия решений о достоверности информации ЭД.

**Заключение.** Представленные результаты составляют методическую основу реализации программно-алгоритмического комплекса обеспечения целостности, сохранности и повышения достоверности передачи и обработки информации ЭД на основе НСГС. Установлено, что примененные средства описания ИМ обладают такими преимуществами, как компактность и наглядность, удобство для перевода в компьютерный код, возможность трансформации и развития модели. Сфера применения предложенной методики не ограничивается конкретным примером задачи обеспечения достоверности передачи и обработки ЭД.

#### *Список литературы / References*

1. Коновалов М.Г. Методы адаптивной обработки информации и их приложения. М.: ИПИ РАН, 2007. 212 с.
2. Джордж Ф.Л. Искусственный интеллект, стратегии и методы решения сложных проблем, Четвёртое издание. «Нью-Мексиковский университет». «Вильямс». М-СПБ-Киев, 2005 г. 863 с.
3. Жуманов И.И., Ахатов А.Р. Нечеткая семантическая гиперсеть контроля достоверности информации в системах электронного документооборота // 4-th AICT. Section 2. IEEE. Tashkent, 2010. P. 21-25.