На правах рукописи

Алшаиа Хайдер Яхья Атоун

МЕТОД И АЛГОРИТМ ОБРАБОТКИ ДАННЫХ НА ОСНОВЕ ИДЕНТИФИКАТОРОВ В СПЕЦИАЛИЗИРОВАННОМ ВЫЧИСЛИТЕЛЬНОМ УСТРОЙСТВЕ

Специальность: 05.13.05 – Элементы и устройства вычислительной техники и систем управления

АВТОРЕФЕРАТ
диссертации на соискание учёной степени
кандидата технических наук

Курск – 2021

Работа выполнена на кафедре «Информационная безопасность» Федерального государственного бюджетного образовательного учреждения высшего образования «Юго-Западный государственный университет».

Научный руководитель: Доктор физико-математических наук, профессор
Добрица Вячеслав Порфирьевич

Официальные оппоненты: **Бехтин Юрий Станиславович**
Доктор технических наук, профессор, Рязанский государственный радиотехнический университет имени В. Ф. Уткина, кафедра «Автоматика и информационные технологии в управлении», профессор

Калмыков Игорь Анатольевич
Доктор технических наук, профессор, Северо-Кавказский федеральный университет, кафедра «Информационная безопасность автоматизированных систем», профессор

Ведущая организация: Федеральное государственное бюджетное образовательное учреждение высшего образования «Орловский государственный университет имени И.С. Тургенева»

Защита диссертации состоится «5» октября 2021 г. в 16:00 на заседании диссертационного совета Д 212.105.02, созданного на базе Юго-Западного государственного университета по адресу: 305040, г. Курск, ул. 50 лет Октября, 94, конференц-зал.

С диссертацией можно ознакомиться в библиотеке Юго-Западного государственного университета и на сайте университета https://swsu.ru/upload/iblock/c0e/wbzx3l9l24ky7b5k4tq90r0ke4s9v5x4/Dissertatsiya-Alshaia-2021.07.27.pdf

Автореферат разослан «___»_____2021 г.

Ученый секретарь
диссертационного совета Титенко Евгений Анатольевич
Д 212.105.02

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Создание объединённых в сложные иерархические комплексы специализированных устройств мониторинга и управления технологическими процессами, оценки и анализа состояний технических систем является одним из перспективных направлений развития средств вычислительной техники и систем управления. Характерными их чертами являются: разброс скоростей передачи данных, динамический характер процессов обмена между адресатами, наличие изменяющихся параметров в передаваемых сообщениях, сложность алгоритмов анализа и разбора их структуры и ограниченный размер самих сообщений (в диапазоне от нескольких байтов до нескольких десятков байтов). Всё это обусловливает необходимость создания специализированных методов, алгоритмов и средств аппаратной обработки таких сообщений, которые обладают преимуществом перед программными способами решения указанных выше задач с точки зрения производительности оконечного вычислительного оборудования и его аппаратной сложности. На основе входящих в состав принимаемых сообщений кодовых последовательностей – идентификаторов – приёмные, коммутирующие и ретранслирующие устройства (например, беспроводные микроконтроллеры, объединённые в распределённую информационно-управляющую систему) формируют непересекающиеся множества сообщений для их последующей ретрансляции, передачи на обработку и т.п. в соответствии с определёнными для каждого такого множества правилами.

Корректность обработки поступающих данных в части выделения из общей информационно-управляющей последовательности наборов сообщений, относящихся к одному адресату, является важным аспектом обеспечения правильности функционирования как вычислительных, информационных и управляющих систем в целом, так и отдельных их компонентов. Ошибки, возникающие в том числе и из-за ограниченности размеров цифровых последовательностей, содержащих в себе информацию об источниках и приемниках сообщений, нарушают работоспособность оконечных устройств и приводят к снижению их производительности из-за необходимости переспросов обработанных с ошибками данных.

В основе создания средств аппаратной обработки данных на основе идентификаторов лежит необходимость обеспечения требуемой достоверности обработки, высокой производительности разрабатываемых устройств и их приемлемой аппаратной сложности, так как они являются частью оконечного оборудования приема и обработки потока сообщений.

Степень разработанности темы исследования. Созданию различных алгоритмов, технических решений и устройств для обработки сообщений посвящены работы Гурина О. Д. Ёкокава Т. Маллади Д. Бухарина В. В. Горохова А. В основе их методов и проектируемых устройств лежит отнесение сообщений множествам на основе результатов обработки идентификаторов, введённых в состав сообщений. Вопросы разработки и

применения процессоров общего назначения и специализированных процессоров, в том числе для задач обработки и анализа идентификаторов сообщений, отражены в трудах В.В. Корнеева, А.В. Киселева, А.В. Шарапова, И.В. Кривченко, В.В. Гребнева, В.В. Жилы, Н.И. Витиски, О.Б. Макаревича. Тем не менее, выделение из сообщений и обработка идентификаторов ограниченной длины как ведущий замысел исследования, влияющий на производительность устройств, не рассматривались в этих трудах. Методам обработки объединённых в группы сообщений ограниченной длины посвящены работы М. Белла (M.Bellare), В. Сталлингса (W. Stallings), (M. Dworkin), Д. Блэка (J. Black), Р. Мишры (R. Mishra), С. Шарма (S. Sharma), Б. Отмана (B. Othman).

Анализ литературы показал, что существующие алгоритмы обработки данных на основе идентификаторов обладают либо недостаточной достоверностью при ограниченном размере идентификатора (алгоритмы, основанные на анализе каждого отдельного сообщения), либо высокой вычислительной сложностью (алгоритмы анализа групп сообщений). Таким образом, актуальной является **научно-техническая задача** разработки метода и алгоритмов определения принадлежности групп сообщений адресатам, ориентированных на высокую скорость выполнения при их аппаратной реализации на микроконтроллерах и ПЛИС.

Объект исследования: элементы и устройства обработки данных на основе идентификаторов.

Предмет исследования: методы и алгоритмы определения принадлежности сообщений независимым адресатам.

Цель работы. Целью диссертационной работы является повышение производительности устройств обработки данных, представленных в виде сообщений ограниченной длины.

Задачи исследований:

1. Проведение аналитического обзора современных методов, алгоритмов и технических решений обработки данных на основе их идентификаторов.
2. Создание метода обработки данных на основе идентификаторов.
3. Разработка алгоритмов обработки данных и отнесения групп сообщений целевому множеству на основе идентификаторов.
4. Разработка вычислительного специализированного устройства обработки данных на основе идентификаторов.
5. Оценка вычислительной сложности алгоритма обработки данных на основе идентификаторов и схемотехнической сложности и производительности специализированного вычислительного устройства, его реализующего.

Новыми научными результатами и положениями, выносимыми на защиту, являются:

1. Метод обработки данных на основе идентификаторов, основанный на формировании групп сообщений и проверки для таких сформированных групп условий принадлежности их конкретному множеству сообщений, отличающийся исключением из анализа по результатам обработки содержимого

поступающего сообщения части множеств, позволяющий снизить аппаратную сложность специализированного вычислительного устройства обработки данных на основе идентификаторов и повысить его производительность.

2. Алгоритм формирования цепочек сообщений, основанный на выполнении операций сравнения идентификатора сообщения с предварительно сформированными кодовыми последовательностями, характеризующими предыдущее и последующее сообщения в цепочке сообщений, отличающийся порядком добавления сообщений в формируемые цепочки и проверкой условия возникновения ошибки после каждой сформированной цепочки, позволяющий уменьшить число задействованных при его реализации блоков регистровой памяти специализированного вычислительного устройства обработки данных.

3. Созданные на основе аппарата теории вероятности и теории случайных процессов математические модели определения принадлежности сообщения множеству, отличающиеся представлением записи описателей буферизированных сообщений в регистровой памяти устройства как линейного динамического процесса, позволившие определить целесообразные характеристики синтезируемого специализированного вычислительного устройства обработки данных на основе идентификаторов.

4. Структурно-функциональная организация специализированного вычислительного устройства обработки данных на основе идентификаторов, отличающаяся параллельным выполнением операций предобработки и анализа входящих сообщений в независимых асинхронно работающих блоках, позволяющая повысить производительность устройства обработки данных на основе идентификаторов за счёт уменьшения числа и длительности выполняемых операций при определении принадлежности группы сообщений целевому множеству.

Достоверность результатов диссертационной работы обеспечивается корректным и обоснованным применением методов проектирования цифровых устройств, аппарата математической логики, положений и методов теории вероятностей, теории случайных процессов и математической статистики, а также подтверждается имитационным моделированием с использованием разработанного программного обеспечения.

Практическая ценность результатов исследований:

1. Разработан алгоритм управления статусами множеств сообщений, основанный на отказе от проверки принадлежности сообщений неактивному множеству, отличающийся проверкой порядкового номера сообщения в группе сообщений неактивного множества и его обработке только в случае если номер не превышает определённое для каждого множества значение, меньшее длины формируемой цепочки, что позволяет снизить аппаратную сложность специализированного вычислительного устройства обработки данных на основе идентификаторов за счёт того, что число независимых вычислительных блоков устройства, выполняющих процедуры обработки сообщений, меньше числа анализируемых множеств сообщений.

2. Созданный алгоритм формирования цепочек сообщений, отличающий рекурсивным вызовом процедуры сравнения описателя текущего сообщения, позволяет за счёт изменения порядка добавления сообщения в цепочку и завершения работы после выполнения условия возникновения ошибки сократить в каждом вычислительном блоке специализированного вычислительного устройства обработки данных на основе идентификаторов на 30% число регистров, требуемых для хранения адресов сообщений, за счет досрочного останова формирования цепочки сообщений с общим адресатом.

3. На основе разработанной математической модели записи идентификаторов сообщений во внутренней регистровой памяти специализированного вычислительного устройства обработки данных, отличающейся представлением процедур формирования и анализа логически связанных структур сообщений в виде случайного процесса, определено целесообразное соотношение 2:1 между числом строк и столбцов матричной регистровой памяти, хранящей адреса буферизированных сообщений, что за счёт сокращения числа выполняемых устройством операций сделало линейной вычислительную сложность алгоритма формирования цепочек сообщений в диапазоне вероятности возникновения ошибки обработки от 0 до 0,15.

4. Разработанная структурно-функциональная организация специализированного вычислительного устройства обработки данных на основе идентификаторов, позволяет за счёт снижения числа декодирований сообщений и реализации декодирования в отдельном структурном блоке, повысить число операций обработки сообщений в единицу времени на 15 – 35 %.

Результаты полученных в диссертации теоретических и прикладных и экспериментальных исследований используются в ООО «Щит-СБ» и учебном процессе Юго–Западного государственного университета при обучении студентов по направлениям 10.04.01 «Информационная безопасность» (дисциплины «Математическое моделирование технических объектов и систем управления), 10.03.01 «Информационная безопасность» (дисциплина «Проектирование защищенных автоматизированных систем») и 10.05.02 «Информационная безопасность телекоммуникационных систем» (дисциплина «Проектирование защищённых телекоммуникационных систем»).

Соответствие диссертации паспорту научной специальности

В соответствии с п. 1 формулы научной специальности 05.13.05 – Элементы и устройства вычислительной техники и систем управления в диссертации содержатся результаты разработки специализированного вычислительного устройства для решения задачи обработки данных на основе идентификаторов.

В соответствии с п. 2 формулы научной специальности в диссертации проводились теоретический анализ и экспериментальные исследования достоверности процедур обработки данных на основе идентификаторов, реализуемые специализированными вычислительными устройствами, с целью снижения числа выполняемых вычислительными блоками устройства операций и уменьшению размеров задействованной регистровой памяти.

Методология и методы исследования. Исследования проведены с применением основ цифровой схемотехники и разработки цифровых устройств, теории вероятностей и математической статистики, теории случайных процессов, конструирования средств вычислительной техники, аналитического конструирования. Экспериментальные исследования выполнены с использованием методов математического и имитационного моделирования, технологий объектно-ориентированного программирования.

Апробация работы. Результаты и научные положения диссертационной работы докладывались и обсуждались на следующих всероссийских и международных научных конференциях: Международная научно-техническая конференция «Инфокоммуникации и космические технологии: состояние, проблемы и пути решения» (г. Курск, 2019, 2020, 2021); Всероссийская научно-техническая конференция «Фундаментальные проблемы информационной безопасности в условиях цифровой трансформации» (г. Ставрополь, 2019); Всероссийская научно-техническая конференция для молодых ученых и студентов с международным участием (г. Пенза, 2021); Международная научно-техническая конференция «RusAutoCon» (г. Сочи, 2019, 2020), Международная конференция ISCAU – 2020 (г. Эн-Насирия, Ирак, 2020).

Публикации. По теме диссертационной работы опубликовано 18 научных работ, в том числе: 5 статей в научных рецензируемых изданиях, входящих в перечень ВАК; 5 статей баз данных Web Of Science и Scopus; 9 докладов на международных и всероссийских конференциях; получено одно свидетельство на программный продукт.

Личный вклад автора в получение результатов, изложенных в диссертационной работе. Все результаты диссертационной работы, в том числе постановка задач, разработка и исследование защищаемых метода, моделей и алгоритмов и структурно - функциональные схемы, основные научные результаты, выводы и рекомендации, принадлежат автору лично. В научных работах, выполненных в соавторстве, личный вклад соискателя состоит в следующем: разработаны алгоритмы обработки групп сообщений [1, 2, 11, 15, 17, 18], разработаны принципы организации буферной памяти приёмника данных [3, 6, 7], создан метод определения принадлежности сообщения множеству [4, 5, 9, 10], разработаны модели оценки характеристик метода обработки данных на основе идентификаторов [16].

Структура диссертации. Диссертация общим объемом 139 страниц состоит из введения, четырёх глав и заключения, содержит 111 страниц основного текста, перечень используемой научно-технической литературы из 135 наименований на 17 страницах, приложений на 11 страницах, 32 рисунков и 1 таблицы.

СОДЕРЖАНИЕ РАБОТЫ

Во введении обосновывается актуальность темы диссертации, выделяются и формулируются цели и задачи исследования, показана научная новизна, теоретическая и практическая значимость результатов, описывается

структура диссертационной работы. Изложены основные положения, выносимые на защиту, соответствие содержания диссертации паспорту научной специальности, степень достоверности и апробация результатов.

В первой главе содержит литературный обзор по теме диссертации. Представлены результаты анализа публикаций, касающихся особенностей формирования и обработки сообщений на основе входящих в их состав идентификаторов в современных информационно-вычислительных системах. Определена целевая характеристика устройств обработки данных – производительность – количество формируемых множеств сообщений в единицу времени:

$$E_{\text{eq}} = \left(T_{\text{dec}} \cdot N_{\text{dec}} \left(|U| \right) + T_{\text{comp}} \cdot N_{\text{comp}} \left(|U| \right) \right)^{-1} \quad (1)$$

где: T_{dec} – длительность операции декодирования сообщения, T_{comp} – длительность операций формирования и сравнения идентификаторов, $N_{\text{dec}}(|U|)$ – число операций декодирования сообщений, являющееся функцией мощности множества сообщений U, поступающих в устройство в единицу времени, $N_{\text{comp}}(|U|)$ – число операций сравнения идентификаторов (функция от $|U|$).

Если определить число сообщений как произведение числа формируемых множеств K на число сообщений M, относящихся к каждому множеству, которые поступают в устройство в единицу времени, то для известных методов, ориентированных на обработку единичных сообщений, производительность определится как:

$$E_{\text{eq}}^{1} = \left(T_{\text{dec}} \cdot K \cdot M + T_{\text{comp}} \cdot K \cdot M \right)^{-1}. \quad (2)$$

Для методов, ориентированных на обработку групп сообщений:

$$E_{\text{eq}}^{2} = \left(\left(T_{\text{dec}} \cdot K + T_{\text{comp}} \cdot N_{\text{comp}} \left(K \cdot M \right) \right) \right)^{-1}, \quad (3)$$

Отсюда сформулированы основные подходы к повышению производительности (увеличению соотношения $E_{\text{eq}}^{2} / E_{\text{eq}}^{1}$) приёмников, обрабатывающих группы сообщений: уменьшение времени T_{comp} выполнения операции формирования и сравнения идентификаторов, использование параллельно работающих модулей формирования множеств сообщений, использование алгоритмов сравнения идентификаторов, сложность которых $O(K \cdot M)$ меньше сложности известных решений.

Вторая глава диссертации посвящена созданию метода и алгоритмов обработки данных на основе идентификаторов ограниченной длины. Сообщения, относящиеся к множеству сообщений адресата, кодируются на основании некоторого идентификатора данного множества Ω^{A}. Каждое сообщение S_i содержит в себе поле идентификатора S_i^{C}, содержащее проверочные данные, по которым осуществляется проверка корректности обработки сообщения, а также определяется принадлежность сообщения конкретному множеству. Его содержимое есть функция следующих данных: Ω^{A} – идентификатора множества сообщений, содержимого информационной части S_i^{inf} предыдущего сообщения и некоторого идентификатора μ_i

сообщения, позволяющего его выделить среди всех сообщений целевого множества (определить его номер i^{p} в анализируемой группе сообщений):

$$S_i^{\mathrm{C}} = \left\{ F_1\left(\Omega^A, S_{i-1}^{\inf}, i^{\mathrm{p}}\right) \middle| \mu_i \right\} = \left\{ F_1\left(\Omega^A, S_{i-1}^{\inf}, i^{\mathrm{p}}\right) \middle| F_2\left(\Omega^A, i^{\mathrm{p}}\right) \right\} \quad (4)$$

где F_1 – функция формирования приёмником имитовставки из слов $S_{i-1}^{\inf}$ и Ω^A (необратимое преобразование); F_2 – выполненное приёмником кодирование числа i^{p} на основе слова Ω^A (обратимое преобразование), | – операция конкатенации двух слов.

Полученное сообщение будет идентифицировано как i^{p}–е сообщение источника, если будет истинным следующее логическое выражение:

$$\left[F_2^{-1}\left(\Omega^A, \mu_i\right) = i^{\mathrm{p}} \right] \wedge \left[S_i^{\mathrm{C}} / \mu_i = F_1\left(\Omega^A, S_{i-1}^{\inf}, i^{\mathrm{p}}\right) \right] \quad (5)$$

где F_2^{-1} – выполняемое приёмником декодирование слова μ_i по идентификатору множества Ω^A, i^{p} – число полученных сообщений обрабатываемой группы, S_i^{C} / μ_i – результат отделения от слова S_i^{C} слова μ_i. Все полученные сообщения, для которых было выполнено условие (5), буферизируются в основной буферной памяти (ОБП) устройства, затем из них формируется последовательность (цепочка) сообщений фиксированной длины M, для каждого элемента которой верно условие (5). По факту формирования цепочки принимается решение о принадлежности всех её M сообщений множеству с идентификатором Ω^{A}.

В основе повышения производительности приёмников (см. формулу (3)) лежит сокращение временных затрат при проверке условия (5). Для этого, вместо многократного чтения из ОБП содержимого информационной части $S_i^{\inf}$ сообщений и выполнения операций F_2^{-1} и F_1, с помощью быстродействующей регистровой памяти адресов сообщении (ПАС) устройства организуется хранение следующих слов:

– адреса сообщения в ОБП;

– результата вычисления выражения $F_1\left(\Omega^A, S_{i^{\mathrm{p}}}^{\inf}, i^{\mathrm{p}}+1\right)$ для данного сообщения для сравнения его за один такт (без предобработки и дополнительных преобразований) с результатом $S_{i^{\mathrm{p}}+1}^{\mathrm{C}} - \mu_{i^{\mathrm{p}}+1}$, а также для сравнения с содержимым поля S^{C} сообщения, у которого $F_2^{-1}(\Omega^A, \mu_i) = i^{\mathrm{p}}+1$;

– слова S_i^{C} / μ_i , для сравнения его за один такт с содержимым поля S^{C} сообщения, у которого $F_2^{-1}(\Omega^A, \mu_i) = i^{\mathrm{p}}-1$.

Для уменьшения вычислительных затрат при обработке сообщений, создан алгоритм управления статусами обрабатываемых специализированным вычислительным устройством множеств сообщений (рис. 1).

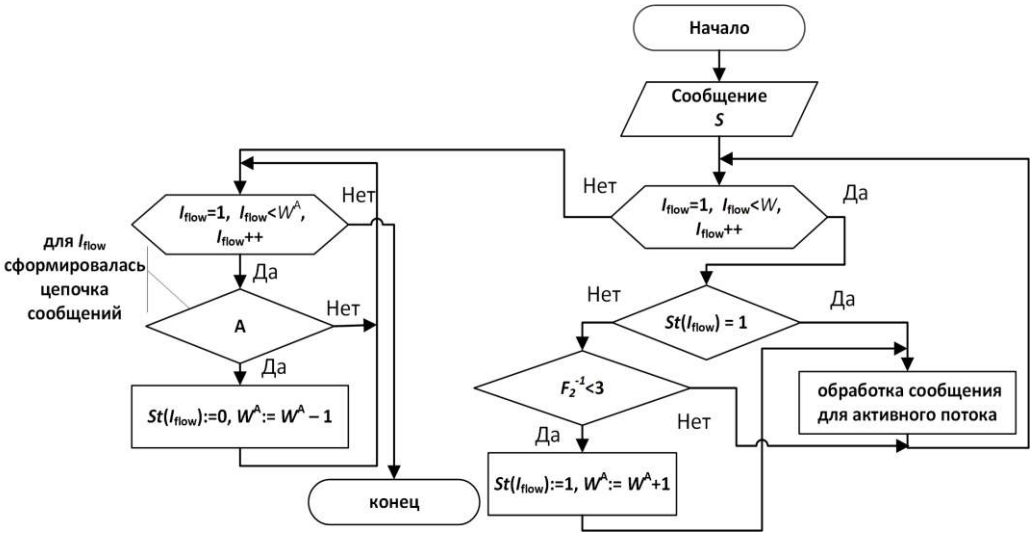

Рисунок 1. Алгоритм управления статусами множеств сообщений.

На основании всего сказанного выше можно описать основные этапы разрабатываемого метода обработки данных на основе идентификаторов.

1. Каждому множеству сообщений присваивают статус «анализируемое» или «неактивное». Сообщения проверяются на принадлежность только к множествам со статусом «анализируемое».

2. Для каждого поступающего в устройство сообщения производится операция его декодирования с использованием множества идентификаторов.

3. Результат проверки содержимого сообщения по условию (5) относит его к определённой цепочке каждого множества со статусом «анализируемое».

4. Результат проверки декодированного номера сообщения переводит множество из состояния «неактивное» в состояние «анализируемое».

5. Если для некоторого множества из состояния «анализируемое» получено граничное сообщение и сформирована единственная цепочка сообщений, то вся такая цепочка относится к данному множеству, а множеству присваивают статус «неактивное», что исключает его из анализа при предобработке данных и позволяет снизить число операций декодирования поступающих сообщений.

Новизна метода обработки данных заключается в управлении набором множеств, на принадлежность к которым проверяется каждое поступающее сообщение, что уменьшает аппаратную сложность операционной части устройства за счёт сокращения числа параллельно работающих блоков, выполняющих процедуры анализа принадлежности групп сообщений множествам, и обеспечивает повышение его производительности.

Процедура формирования цепочек сообщений включает отнесение каждого обрабатываемого сообщения к одной из групп $w_1 - w_M$ сообщений с одинаковым порядковым номером i в цепочке и размещающихся в соответствующем столбце ПАС. После чего цепочки могут формироваться двумя способами:

– параллельным формированием всех цепочек путём выполнения итераций по добавлению во все цепочки элементов из групп $w_1 - w_M$;

– формированием каждой цепочки до конца и лишь затем переходом к формированию следующей.

Соответственно, разработаны два алгоритма формирования цепочек сообщений: итерационный, на основе перебора всех сообщений из ПАС, и рекурсивный алгоритма, в котором содержимое столбцов ПАС передаётся в качестве параметров в соответствующую процедуру. На начальном этапе $N_{\text{chain}} = 1,\ i_{\text{chain}} = 1,\ i_{\text{ind}} = 1,\ k = 1\ u_1 = \{s_{\text{start}}\}$, где s_{start} – последнее сообщение предыдущей цепочки множеств (рис. 2).

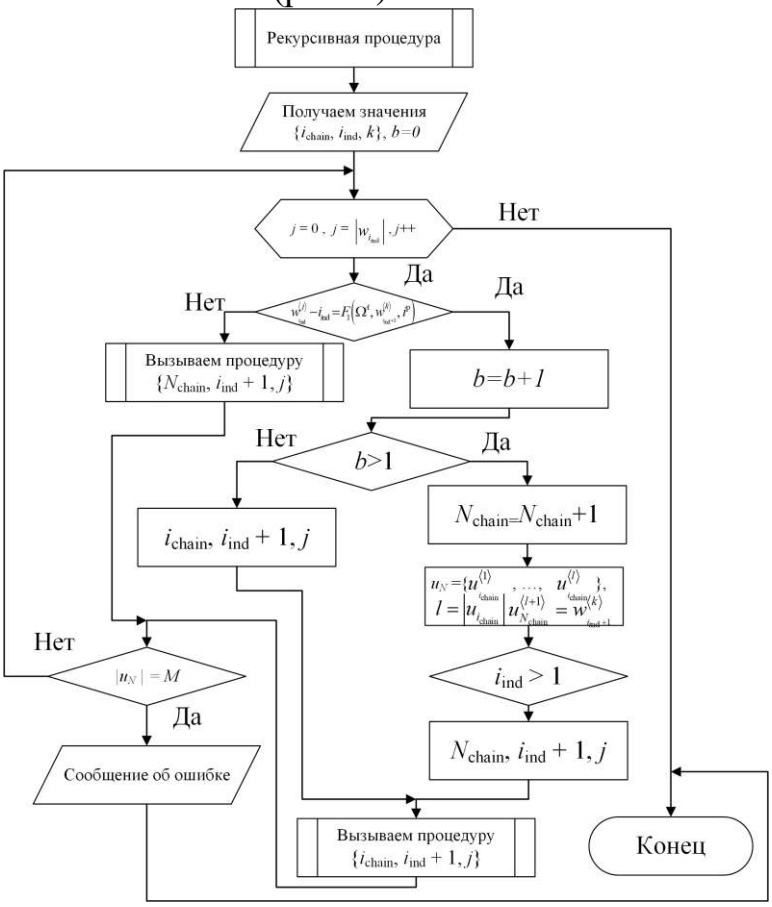

Рисунок 2. Блок-схема рекурсивного алгоритма формирования цепочек сообщений.

Рекурсивный алгоритм отличается последовательным формированием цепочек сообщений и проверкой условия возникновения ошибки после каждой сформированной цепочки и позволяет при обнаружении более чем одной цепочки длиной M завершить работу, прекратив заполнение регистровой памяти устройства описателями сформированных цепочек. Это позволяет снизить требуемый объём такой памяти.

Третья глава посвящена оценке с помощью математического и имитационного моделирования вычислительной сложности разработанных алгоритмов и ресурсных затрат для их реализации. Исследуемыми параметрами были: размер матрицы регистров ПАС N и M (M определяется как максимальная длина формируемой приёмником цепочки сообщений), разрядность h поля идентификатора сообщения и K – число формируемых множеств сообщений. В основе математической модели процедуры обработки

сообщений было положено представление получения сообщений приёмником в виде пуассоновского процесса. Получена рекуррентная формула для среднего числа сформированных цепочек сообщений при длине цепочки *r*:

$$p\left(n_{\text{ch}}^{(1,r)},r\right)=\sum_{i=0}^{N^r}\left(p\left(n_{\text{ch}}^{(1,r-1)}\right)\left(\sum_{k=1}^{N}\frac{K^k\times e^{-K}}{k!}C_{N+1}^{k}\left(2^{-h}\right)^k\left(1-2^{-h}\right)^{N-k}\right)\right). \quad (6)$$

При $r=M$ получим функцию плотности вероятности числа сформированных цепочек. Итоговая формула для вероятности возникновения ошибки $p_{\text{fl}}(M)$ отнесения цепочки сообщений к множеству получена, исходя из вероятности выполнения условия (5) для хотя бы одной из $n_{\text{ch}}^{(1,M)}-1$ цепочек. Она является вероятностью совместного наступления независимых событий:

$$p_{\text{fl}}\left(M\right)=\sum_{n_{\text{ch}}^{(1,M)}=1}^{\infty}p\left(n_{\text{ch}}^{(1,M)}\right)\left[1-\left(1-2^{-h}\right)^{n_{\text{ch}}^{(1,M)}}\right]. \quad (7)$$

Вероятность $p_{\text{owfl}}(M)$ ошибки нехватки размеров ПАС, при которой число сообщений с одинаковым значением μ_i превышает число *N* строк в матрице регистров ПАС, определится как:

$$p_{\text{owfl}}\left(M,N\right)=\left\{1-\left[1-\sum_{n_{\text{ch}}^{(1,M)}=N}^{\infty}p\left(n_{\text{ind}}^{m},M\right)\right]^M\right\}. \quad (8)$$

Общая же вероятность возникновения ошибки $p_{\text{err}}(M, N)$ определится как вероятность наступления хотя бы одного из описанных выше событий.

$$p_{\text{err}}\left(M,N\right)=1-\left(1-p_{\text{fl}}\left(M\right)\right)\left(1-p_{\text{owfl}}\left(M,N\right)\right). \quad (9)$$

Полученное соотношение позволяет определить зависимость между количеством столбцов *M* и строк *N* в ПАС при требуемых значениях вероятности ошибки p_{err} (рис. 3).

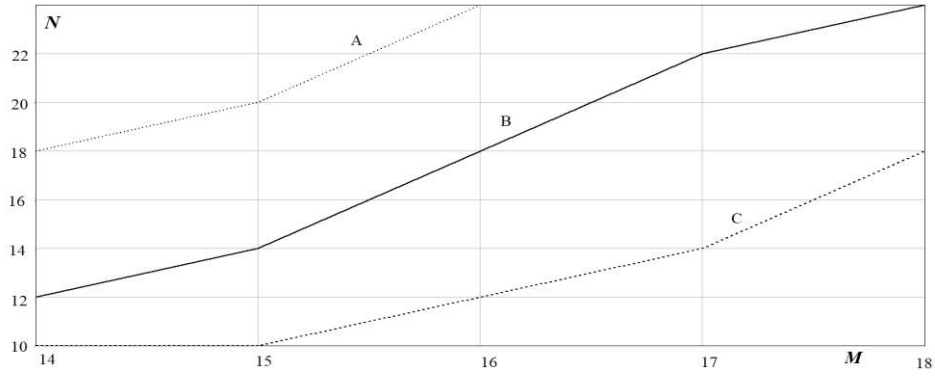

Рисунок 3. Зависимость между количеством столбцов *M* и строк *N* в матрице ПАС A) $p_{\text{err}} = 0.25$ B) $p_{\text{err}} = 0.15$ C) $p_{\text{err}} = 0.1$.

Исходя из данных графиков, можно утверждать, что целесообразное соотношение между количеством строк *N* и столбцов *M* регистровой памяти ПАС лежит в диапазоне 1.5 … 2.0. Большее соотношение нецелесообразно,

так как влечёт лишь увеличение объёма ПАС без уменьшения вероятности возникновения ошибок при обработке сообщений и их идентификаторов.

Затраты внутренней регистровой памяти устройства возникают, во-первых, из-за необходимости хранить адреса сообщений в ПАС, а во-вторых, из-за необходимости для каждой сформированной цепочки хранить номера регистров ПАС, её составляющих. Размеры ПАС выбираются исходя из требуемой протоколом передачи данных достоверности (p_{err}) и полученных выше соотношений. На основе формулы (6) определяется число регистров, требуемых для хранения сформированных цепочек. Проведённые серии математических экспериментов позволили определить целесообразный диапазон для данной характеристики как 20 … 100 (рис. 4).

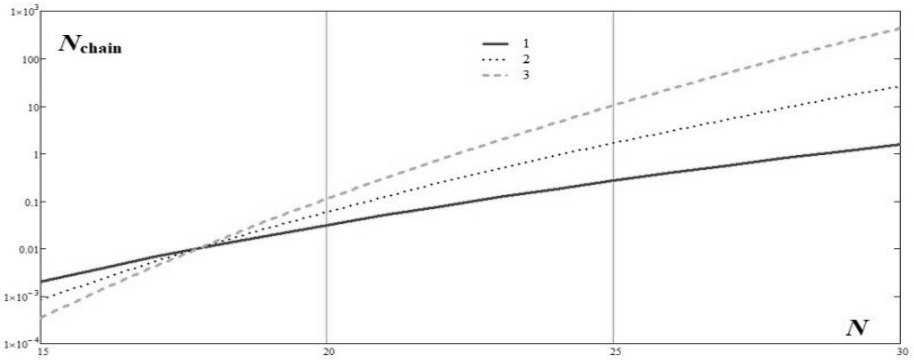

Рисунок 4. Зависимость числа сформированных цепочек сообщений N_{chain} от числа строк ПАС N и K=30. 1) M = 10, 2) M = 15, 3) M = 20.

Для оценки влияния типа алгоритма формирования цепочек сообщений на число необходимых регистров внутренней памяти устройства и число операций, выполняемых таким устройством, оба алгоритма были реализованы с помощью имитационной модели. Получена зависимость N_{chain} среднего числа формируемых устройством цепочек для каждого алгоритма от числа анализируемых сообщений $|U|$ и размера поля служебных данных h (рис. 5). Установлено, что тип алгоритма незначительно влияет на вычислительную сложность процедуры формирования цепочки сообщений: выигрыш в пользу рекурсивного составляет менее 10%.

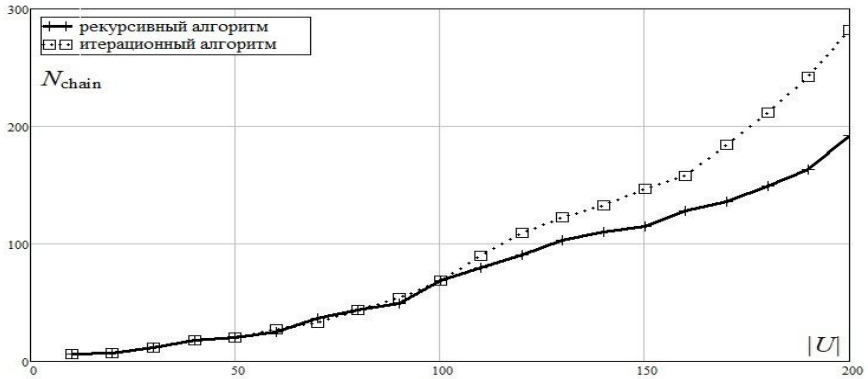

Рисунок 5. Зависимость N_{chain} среднего числа цепочек от числа анализируемых сообщений $|U|$ при h = 5.

Числа цепочек, формируемых при использовании рекурсивного алгоритма, снижается на величину до 30% (при вероятности возникновения ошибок p_{err}>0,15) от числа цепочек, формируемых при использовании итерационного, что определяет выбор рекурсивного алгоритма в качестве основного алгоритма обработки данных на основе идентификаторов.

Создана математическая модель оценки числа операций проверки условия (5) при использовании рекурсивного алгоритма обработки данных и получена зависимость отношения *T/M* числа сравнений к длине цепочки сообщений от числа формируемых множеств сообщений *K* . Установлено, что данная зависимость является линейной: $T/M = 1{,}5\ldots2.0 \times K$.

Четвёртая глава посвящена разработке специализированного вычислительного устройства обработки данных на основе идентификаторов (СВУОДОИ). Структурная схема СВУОДОИ приведена на рис. 6, где сплошной линией показаны направления передачи данных, пунктирной – управляющих и информационных сигналов. Блоки устройства обмениваются данными по двум внутренним шинам: данных сообщения и адреса сообщения в ОБП.

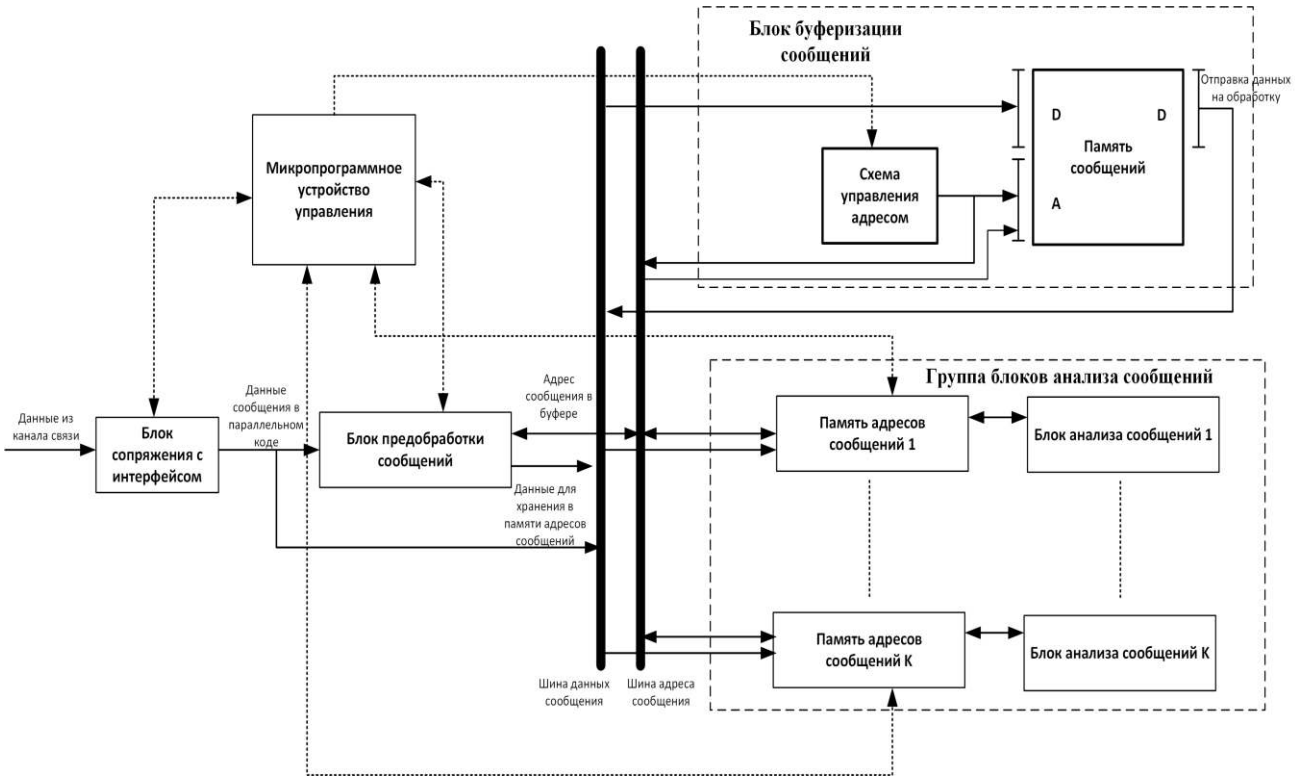

Рисунок 6. Структурная схема специализированного вычислительного устройства обработки данных на основе идентификаторов (СВУОДОИ).

Микропрограммное устройство управления (МПУУ) управляет взаимодействием всех блоков устройства через внутренние шины, вместе с блоком предобработки сообщений осуществляет управление работой блоков анализа сообщений, закрепляя за ними функции определения принадлежности сообщений конкретному множеству. Также МПУУ обеспечивает синхронизацию работы независимых модулей СВУОДОИ.

Блок сопряжения с интерфейсом обеспечивает преобразование сигнала из канала связи в параллельный код сообщения, пригодный для обработки узлами СВУОДОИ. Блок буферизации сообщений состоит из ОЗУ ОБП и схемы управления адресами, которая обеспечивает формирование адреса для очередного записываемого в буфер сообщения, а также поддержание битовой карты занятых областей ОБП.

Несколько независимых блоков анализа сообщений, подключаемых к общей шине, обеспечивают формирование содержимого регистровой памяти адресов сообщений (ПАС) для нескольких множеств. После формирования цепочки сообщений, их адреса передаются на шину адреса, и производится их обработка в соответствии с определёнными для множества правилами.

Функциональная схема одного блока памяти адресов сообщений, работающего в связке с блоком анализа сообщений, приведена на рисунке 7. Она представляет собой матрицу регистров *M* на *N* , в которые записывается адрес сообщения в ОБП (с шины адреса), а также слова $F_1\left(\Omega^A, S_{i^{\mathrm{p}}}^{\inf}, i^{\mathrm{p}}+1\right)$ и $S_i^{\mathrm{C}} - \mu_i$ (с шины данных). Выбор регистра хранения происходит по сигналу от дешифраторов DC 1 – DC *M*, на которые поступает двоичный номер регистра либо из блока анализа сообщений (при анализе сообщений), либо со считающих регистров CRg 1 – CRg *M* (при заполнении ПАС).

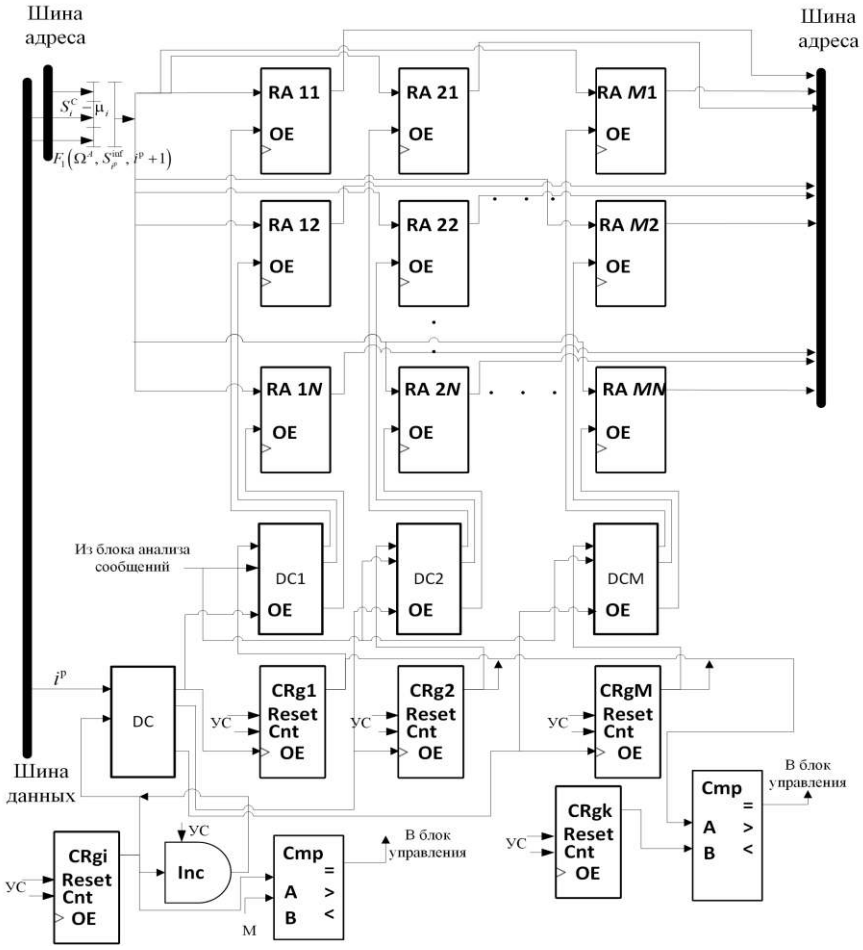

Рисунок 7. Функциональная схема памяти адресов сообщений.

Для оценки характеристик СВУОДОИ и сравнения их с аналогами, оно было реализовано в системе синтеза и анализа схем Vivado фирмы Xilinx как проект программируемой логической интегральной схемы (ПЛИС). Реализовывалось устройство в виде двух независимых модулей. Первый модуль состоит из блока предобработки сообщений, блока буферизации сообщений и МПУУ. Второй модуль включает в себя блок ПАС, блок анализа сообщений и их независимый блок управления. В одном СВУОДОИ через общую шину объединяются один первый модуль и несколько вторых.

Реализация первого модуля показала, что его схемотехническая сложность составила 3456 логических блоков (LUT Elements) и 114 задействованных блоков ввода – вывода (IOBs). При этом максимальное время распространения сигнала между ножками составило 95 нс. Это время, определяющее длительность такта работы соответствующих частей СВУОДОИ, необходимого для выполнения операции декодирования содержимого поля служебных данных. Сложность второго модуля 4125 логических блоков и 52 задействованных блока ввода – вывода при минимальной длительности такта работы в 16 нс. Более высокая скорость работы схем второго модуля объясняется невысокой сложностью выполняемых им операций сравнения и записи данных в регистры.

Исходя из полученных данных, пригодной для реализации СВУОДОИ является микросхема xc7s75fgga484-1 фирмы Xilinx. На основании полученных временных характеристик работы различных частей устройства получены временные характеристики работы самого СВУОДОИ, которые сравнивались с аналогичными характеристиками для известных решений. В соответствии с полученным средним числом тактов работы каждого модуля и длительностью соответствующих тактов были получены зависимости между производительностью разрабатываемого СВУОДОИ E_{eq}^2 и производительностью известных решений E_{eq}^1 (рис. 8).

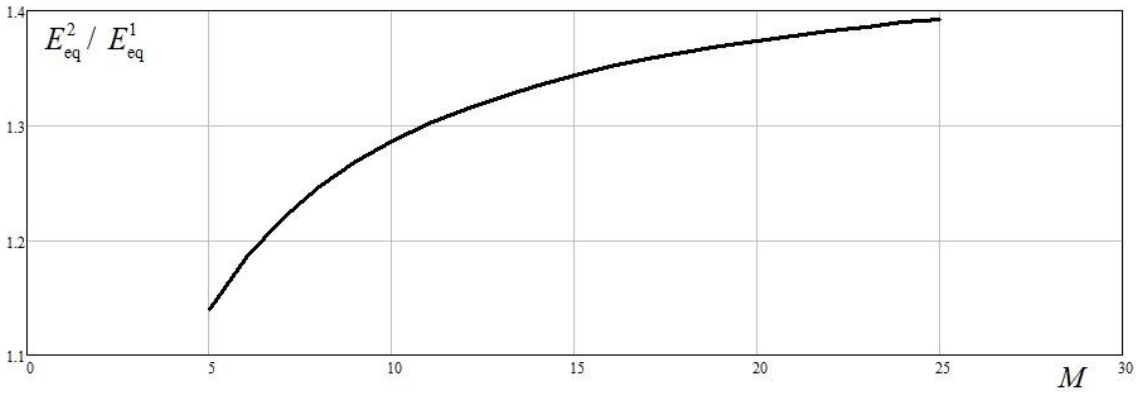

Рисунок 8 – График зависимости отношения производительности разработанного СВУОДОИ к производительности устройства, обрабатывающих отдельные сообщений, от длины цепочек сообщений M.

Реализация метода обработки данных на основе идентификаторов ограниченного размера в разработанном СВУОДОИ, позволяет повысить

число операций обработки сообщений и выработки решений о принадлежности их множествам на 15 – 35 %. Указанное повышение производительности обеспечивается ускорением выполнения операций сравнения содержимого служебных полей и снижением числа трудоёмких операций декодирования данных сообщения.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

Диссертация посвящена решению актуальной научно–технической задачи разработки метода и алгоритмов определения принадлежности групп сообщений адресатам, ориентированных на высокую скорость выполнения при их аппаратной реализации на микроконтроллерах и ПЛИС. Получены следующие результаты:

1. Проведен сравнительный обзор известных методов, алгоритмов и технических средств обработки данных и определения принадлежности сообщений конкретному множеству. Показано, что из-за ограниченной пропускной способности каналов связи в некоторых классах современных информационно-вычислительных систем накладываются существенные ограничения на размер дополнительных служебных полей сообщений. Это требует при обработке данных использования специализированных методов и средств обработки групп сообщений, повышающих достоверность такой обработки.

2. Разработан метод обработки данных на основе идентификаторов ограниченной длины, предполагающий формирование из сообщений логически связанных структур на основе содержимого нескольких битов их служебных полей. Для повышения скорости обработки сообщений метод предусматривается сокращение числа множеств, к которым может быть отнесено каждое поступающее сообщение. При этом процедуры обработки данных реализуются в независимых вычислительных блоках устройства, и число таких блоков меньше числа формируемых устройством множеств сообщений.

3. Разработан алгоритм обработки сообщений на основе идентификаторов и отнесения их к определённым множествам, отличающийся проверкой условия возникновения ошибки после каждой сформированной цепочки сообщений. Это позволяет при формировании цепочек сообщений уменьшить объём задействованной памяти устройства и реализовать память в виде матрицы регистров, повысив тем самым скорость выполнения процедур обработки сообщений.

4. Создана математическая модель размещения описателей сообщений во внутренней регистровой памяти специализированного вычислительного устройства обработки данных на основе идентификаторов ограниченной длины, основанная на теории вероятностей и теории случайных процессов, отличающаяся представлением процедур формирования и анализа логически связанных структур сообщений в виде случайного процесса, позволившая оценить требуемый объём регистровой памяти одного вычислительного блока для реализации метода обработки данных. Соотношение между числом строк и числом столбцов матрицы регистров определено как 2:1, что обеспечи-

вает линейную зависимость вычислительной сложности алгоритма от числа формируемых множеств сообщений при вероятности ошибки определения принадлежности сообщения множеству, не превышающей 0,15.

5. На основании созданной имитационной модели реализации метода обработки данных на основе идентификаторов произведена оценка влияния типа алгоритма формирования цепочки сообщений на требуемый объём регистровой памяти устройства. Это позволило за счёт использования рекурсивного алгоритма формирования цепочек сократить на 30% число регистров в каждом вычислительном блоке устройства.

6. Реализована структурно-функциональная схема специализированного вычислительного устройства данных на основе идентификаторов ограниченной длины, отличающаяся выполнением длительных по времени операций декодирования сообщений и коротких операций сравнения содержимого дополнительных полей в независимых параллельно работающих модулях. Такое распараллеливание повышает производительность устройства, выраженную в числе операций обработки сообщений в единицу времени, на 15 – 35 % в сравнении с аналогами, реализующими обработку на основе сравнений, выполняемых непосредственно после декодирования сообщений.

Список научных работ, опубликованных по теме диссертации

Статьи в научных рецензируемых журналах из перечня ВАК:

1. Алшаиа Х.Я. Рекурсивный алгоритм формирования структурированных множеств информационных блоков для повышения скорости выполнения процедур определения их источника [Текст] / М.О. Таныгин, Х.Я.А. Алшаиа, В.П. Добрица, О.Г. Добросердов // Известия Юго-Западного государственного университета. 2021. – № 2. – С. 51-64.

2. Алшаиа Х.Я. Сложность алгоритма определения источника данных [Текст] / Таныгин М.О., Алшаиа Х.Я., Митрофанов А.В.// Труды МАИ. 2021. Выпуск № 117. DOI: https://doi.org/10.34759/trd-2021-117-12.

3. Алшаиа Х.Я. Оценка влияния организации буферной памяти на скорость выполнения процедур определения источника сообщений [Текст] / Таныгин М.О., Алшаиа Х.Я., Добрица В. П, // Труды МАИ. 2020. Выпуск № 114. DOI: https://doi.org/10.34759/trd-2020-114-15.

4. Алшаиа Х.Я. Установление доверительного канала обмена данными между источником и приёмником информации с помощью модифицированного метода одноразовых паролей [Текст] / Таныгин М.О., Алшаиа Х.Я., Алтухова В.А. // Известия Юго-Западного государственного университета. Серия: Управление, вычислительная техника, информатика. Медицинское приборостроение. 2018. – № 4(29). – С. 63-71.

5. Алшаиа Х.Я. Об одном методе контроля целостности передаваемой поблоково информации [Текст] / Таныгин М.О., Алшаиа Х.Я., Алтухова В.А., Марухленко А.Л.// Телекоммуникации. 2019. № 3. С. 12-21.

Публикации в изданиях, индексируемых в международных наукометрических базах Scopus или в Web of Sciense:

6. Alshaeaa H.Y., Tanygin M. O., Dobritsa V. P., " Study of the influence of the unauthorized blocks number on the cost of the speed and memory RAM during the analysis data process", IOP Conference Series: Materials Science and Engineering, Volume 928, 2nd International Scientific Conference of Al-Ayen University (ISCAU-2020). DOI: https://doi.org/10.1088/1757-899X/928/3/032020

7. Alshaeaa H.Y., Tanygin M. O., Dobritsa V. P., "Study of the influence of the unauthorized blocks number on the collisions probability", in International Russian Automation Conference, RusAutoCon 6-12 September 2020, Sochi, Russian Federation. DOI: https://doi.org/10.1007/978-3-030-71119-1_12.

8. Alshaeaa H.Y., Tanygin M. O., Kuleshova E. A., " A Method of The Trans-mitted Blocks Information Integrity Control," Radio Electronics, Computer Science, Control. 2020. № 1, P. 182-189.

9. Alshaeaa H.Y., Tanygin M. O., Efremov M. A., "Analysis of the Secure Data Transmission System Parameters," in International Russian Automation Conference, RusAutoCon 8-14 September 2019, Sochi, Russian Federation. DOI: https://doi.org/10.1007/978-3-030-39225-3_74

10. Alshaeaa H.Y., Tanygin M. O., Altukhova V.,"Establishing Trusted Channel for Data Exchange between Source and Receiver by Modified One-time Pass-word Method," in International Russian Automation Conference, RusAutoCon. DOI: https://doi.org/10.1109/RUSAUTOCON.2019.8867590

Свидетельство на программный продукт

11. Свидетельство о государственной регистрации программы для ЭВМ № 2021614840 Российская Федерация. Программа для формирования структурированных множеств информационных блоков для определения источника сообщений: № 2021613982: заявл. 26.03.2021: опубл. 30.03.2021 / М. О. Таныгин, Х. Я. А. Алашаиа, А. В. Митрофанов.

Научные труды в других изданиях

12. Алшаиа Х.Я. Принципы организации буферной памяти специализированного приёмника, определяющего источник поступающих данных // XVI Международная научно-техническая конференция «Распознавание – 2021. Курск, 2021. С. 44-46.

13. Алшаиа Х.Я. Формальное описание модели предобработки блока данных для систем с ограниченным размером дополнительных служебных полей // Инфокоммуникации и космические технологии: состояние, проблемы и пути решения: сборник научных статей по материалам V Всероссийской НПК. в Ч. 2. Курск, 2021. С. 131-134.

14. Алшаиа Х.Я. Построение модели для сокращения быстродействующих аппаратных затрат на анализ служебных полей блоков информаций / Алшаиа Х.Я.// Инновации технических решений в машиностроении и транспорте: сборник научных статей по материалам VII Всероссийской научно-техническая конференция для молодых ученых и студентов с международным участием / отв. ред. В.В. Салмина. Пенза, 2021. С. 25-29

15. Алшаиа Х.Я. Формальное описание модели взаимодействия устройств в условиях ограничения размера полей метаданных/ Таныгин М.О., Алшаиа Х.Я., Добрица В. П. // Инфокоммуникации и космические технологии: состояние, проблемы и пути решения: сборник научных статей по материалам III Всероссийской НПК. Курск, 2020. С. 7-10.

16. Алшаиа Х.Я. Использование метаданных для исправления ошибок аутентификации при сетевом взаимодействии / Таныгин М.О., Алшаиа Х.Я., Хемраев Д., // Фундаментальные проблемы информационной безопасности в условиях цифровой трансформации: сборник научных статей по материалам XXIII пленума ФУМО ВО ИБ и Всероссийской научной конференции " (ИНФОБЕЗОПАСНОСТЬ -2019). Ставрополь, 2019. С. 14-18.

17. Алшаиа Х.Я. Алгоритм необратимых преобразования для системы контроля целостности цепочек пакетов в сетях с низкой пропускной способностью / Таныгин М.О., Алшаиа Х.Я., Берлизева В.А. // Инфокоммуникации и космические технологии: состояние, проблемы и пути решения: сборник научных статей по материалам III Всероссийской НПК. Курск, 2019. С. 165-169.

18. Алшаиа Х.Я. Алгоритм обратимых преобразования для контроля аутентичности пакетов в сетях с низкой пропускной способностью / Таныгин М.О., Алшаиа Х.Я., Берлизева В.А. // Инфокоммуникации и космические технологии: состояние, проблемы и пути решения: сборник научных статей по материалам III Всероссийской НПК. Курск, 2019. С. 169-173.

Соискатель Алшаиа Хайдер Яхья Атуон

Подписано в печать___.___.2021г. Формат 60×84 1/16.
Печ.л.1,0. Тираж 100 экз. Заказ_____.
Юго-Западный государственный университет.
Издательско-полиграфический центр
Юго-Западного государственного университета.
305040, г. Курск, ул. 50 лет Октября, 94