На правах рукописи

Бекенева Яна Андреевна

ФОРМАЛИЗАЦИЯ ПРОЦЕССОВ ОБРАБОТКИ И ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ИНФОРМАЦИИ ОТ РАЗНОРОДНЫХ ИСТОЧНИКОВ В СИСТЕМАХ РАСПРЕДЕЛЕННОГО МОНИТОРИНГА

Специальность 05.13.01 - Системный анализ, управление и обработка информации (технические системы)

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Санкт-Петербург – 2019

Работа выполнена на кафедре вычислительной техники федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)» (СПбГЭТУ «ЛЭТИ»).

Научный руководитель:
кандидат технических наук **Шоров Андрей Владимирович,** федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)» (СПбГЭТУ «ЛЭТИ»), кафедра вычислительной техники, ведущий научный сотрудник

Официальные оппоненты:
доктор технических наук, профессор **Саенко Игорь Борисович**, ФГБУН Санкт-Петербургский институт информатики и автоматизации Российской академии наук (СПИИРАН), лаборатория проблем компьютерной безопасности, ведущий научный сотрудник

кандидат технических наук, доцент **Коршунов Игорь Львович**, ФГБОУ ВО Санкт-Петербургский государственный экономический университет, кафедра информационных систем и технологий, заведующий кафедрой

Ведущая организация: АО "Концерн «Океанприбор», г. Санкт-Петербург

Защита состоится «09» декабря 2019 года в 14.00 на заседании диссертационного совета Д 212.238.07, созданном при Санкт-Петербургском государственном электротехническом университете «ЛЭТИ» им. В.И. Ульянова (Ленина) по адресу: 197376, Санкт-Петербург, ул. Профессора Попова, 5.

С диссертацией можно ознакомиться в библиотеке ФГАОУ ВО «Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)» и на сайте университета www.eltech.ru в разделе «Подготовки кадров высшей квалификации» - «Объявление о защитах»

Автореферат разослан «08» октября 2019 года.

Ученый секретарь
диссертационного совета Д 212.238.07
к.т.н., доцент /В. В.Цехановский/

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Современные системы мониторинга распределенных объектов имеют сложную иерархическую структуру и часто состоят из большого количества разнородных устройств. Разные типы устройств предназначены для определения разных параметров, связанных с выполнением различных процессов. Каждый процесс может быть представлен в виде последовательности событий. В рамках данной работы под событием понимается некоторое действие, совершенное объектом наблюдения в определенном месте в определенное время. Под объектом наблюдения понимается сущность, непосредственно участвующая в реализации процесса (человек, транспортное средство). В таких системах мониторинга регистрация некоторого события осуществляется несколькими устройствами: видеокамерами, датчиками движения, системами контроля доступа, межсетевыми экранами, измерительными системами и т.п. При этом возникают проблемы, связанные как с объединением информации из разных, слабосвязанных между собой систем, где одно событие может быть зарегистрировано разными устройствами в разные моменты времени. В некоторых случаях отсутствует единая идентификация контролируемых объектов. Также возникают сложности с прагматическим анализом данных, который должен проводиться на большом объеме информации, поступающей из разных источников, и выдавать результат за минимальное время. Таким образом, требуется разработка новых методов и моделей, обеспечивающих агрегирование разнородных данных с целью дальнейшего анализа.

Для анализа информации требуется не только её сбор в едином хранилище, но и её корреляция, а также представление в виде, приемлемом для ее анализа известными методами. Поступающие от разнородных источников данные имеют не только разный формат, но и зачастую требуют предварительной подготовки для проведения их анализа. В связи с этим можно выделить ряд задач, которые в настоящее время требуют решения:

1. Объединение разнородной информации, описывающей одно событие.
2. Устранение повторяющейся информации.
3. Устранение неопределенностей, связанных с идентификацией движущихся объектов и отсутствием временной синхронизации при генерации данных разнородными источниками.

В связи с этим, актуальными являются задачи интеграции разнородных данных для описания исследуемых событий, процессов и их анализа средствами интеллектуального анализа с целью построения профилей типового поведения движущихся объектов, а также выявления отклонений в ходе текущих процессов.

Целью работы является разработка средств обработки информации, поступающей от разнородных источников распределенной системы мониторинга для формирования единой последовательности событий и ее

интеллектуального анализа. Для достижения заявленной цели в работе решаются следующие **задачи**:

- анализ существующих средств обработки разнородной информации;
- разработка формальной модели последовательностей событий в системах распределенного мониторинга;
- разработка метода формирования последовательностей событий на основе информации от разнородных источников системы распределенного мониторинга;
- разработка методики решения задач интеллектуального анализа на основании информации от распределенных источников для поддержки принятия решений;
- программная реализация разработанного метода формирования последовательностей событий;
- экспериментальное исследование разработанных средств обработки информации.

Объектом исследования является система распределенного мониторинга.

Предметом исследования является обработка информации из разнородных источников систем распределенного мониторинга.

Методы исследований. Для достижения поставленных задач использовались методы системного анализа, сравнения и аналогий, классификации, кластеризации, секвенциального анализа. При программной реализации разработанного метода применялись методы объектно-ориентированного программирования.

Основные положения, выносимые на защиту:

1. Формальная модель последовательностей событий в системах распределенного мониторинга.
2. Метод формирования последовательностей событий на основе информации от разнородных источников.
3. Методика решения задач интеллектуального анализа на основании информации от разнородных источников распределенной системы мониторинга.

Научная новизна:

1. Предложена формальная модель последовательностей событий в системах распределенного мониторинга, в отличие от существующих, учитывающая неопределенности в информации, поступающей от разнородных источников.
2. Разработан метод формирования последовательностей событий на основе информации от разнородных источников, обеспечивающий устранение неопределенностей, связанных с идентификацией объектов и отсутствием единого времени регистрации события.

Практическая ценность работы:

1. Предложена методика решения задач интеллектуального анализа на основании информации от разнородных источников распределенной системы мониторинга с целью поддержки принятия решений.

2. Программная реализация метода формирования последовательностей событий на основе информации от разнородных источников, позволяющая автоматизировать обработку информации о событии.

Реализация и внедрение результатов работы. Результаты исследования были использованы в работах, выполняемых в АО «НИЦ СПб ЭТУ», а также использованы при проведении практических занятий по дисциплине «Технология анализа и извлечения знаний» для студентов направления «Информатика и вычислительная техника» и чтении лекций и проведении практических занятий по дисциплине «Интеллектуальный анализ данных» для бакалавров направления «Информационные системы и технологии», что подтверждено актами о внедрении.

Апробация работы. Основные положения и результаты диссертационной работы докладывались и обсуждались на международной конференции по передовым проводным и беспроводным сетям и системам нового поколения NEW2AN, 2016 г., международных конференциях по мягким вычислениям и измерениям SCM'2017, SCM'2018, Санкт-Петербург, 2017-2018 гг, международных конференциях по управлению качеством, транспортной и информационной безопасности и информационным технологиям IT&QM&IS 2017, 2018 гг, международной конференции по управлению в технических системах CTS 2017, международной конференции по человеческому фактору в сложных технических системах ERGO 2018, международной конференции по виброинженерии (JVE) 2017, конференциях профессорско-преподавательского состава СПбГЭТУ «ЛЭТИ», Санкт-Петербург, 2014-2019 гг, международном научном симпозиуме "INTELS-2018", Санкт-Петербург, 2018 г, международной конференции молодых ученых ElConRus 2015, 2018, 2019 гг, международной конференции по интернету вещей, умных пространств ruSMART 2019.

Обоснованность и достоверность представленных в диссертационной работе научных положений обеспечивается проведением анализа состояния исследований в данной области, подтверждается согласованностью теоретических результатов с практическими, полученными при компьютерной реализации, а также апробацией основных теоретических положений в печатных трудах и докладах на научных конференциях. Достоверность результатов диссертационной работы подтверждается корректностью применяемого математического аппарата, строгими доказательствами предложенных утверждений, результатами эксперимента.

Публикации. Основные теоретические и практические результаты диссертации опубликованы в 34 научных работах, среди которых: 10 статей – в изданиях, рекомендованных в действующем перечне ВАК, 21 работ – в материалах и трудах международных и всероссийских научно-технических конференций и 3 свидетельства о государственной регистрации программ для ЭВМ.

Личный вклад соискателя состоит в непосредственном участии в получении исходных данных и научных экспериментах, разработке формальной модели и метода формирования последовательности событий,

методики решения задач интеллектуального анализа, подготовке ключевой части публикаций по выполненной работе и представлению результатов работы на конференциях различного уровня, в том числе международных.

Структура и объем диссертации. Диссертационная работа состоит из введения, четырех глав, заключения, одного приложения, списка литературы (105 наименований). Общий объем работы составляет 154 страницы машинописного текста, который включает 28 рисунков, 12 таблиц, 2 приложения.

Соответствие паспорту специальности. Данное диссертационное исследование выполнено в соответствии с паспортом специальности 05.13.01 «Системный анализ, управление и обработка информации (технические системы)», а именно соответствует следующим областям (номера соответствуют пунктам в паспорте специальности): п. 2 – Формализация и постановка задач системного анализа, оптимизации, управления, принятия решений и обработки информации; п. 12 – Визуализация, трансформация и анализ информации на основе компьютерных методов обработки информации; п. 13 – Методы получения, анализа и обработки экспертной информации.

ОСНОВНОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во **введении** дано обоснование актуальности темы диссертационного исследования, сформулированы цели и задачи работы, ее научная новизна и практическая значимость, представлены положения, выносимые на защиту.

В **первой главе** описана постановка решаемой задачи, рассмотрены методы интеграции разнородных данных, подходы к устранению неопределенностей, в том числе связанных с временными характеристиками.

В ходе обзора были проанализированы существующие системы мониторинга, варианты структуры их организации, процедуры регистрации событий, хранения и обработки данных.

Современные системы мониторинга включают в себя различные устройства, предназначенные для определения параметров, связанных с реализацией процессов.

Все данные, получаемые от устройств мониторинга, направляются на центральный узел сбора данных. Такие данные поступают в разном формате, который зависит от типа средства мониторинга, зарегистрировавшего событие. Как правило, на центральном узле осуществляются некоторые первичные преобразования данных с целью их приведения к виду, удобному для представления и хранения, однако их формат по-прежнему остается разным и зависящим от типа источника. Кроме того, могут возникать неопределенности, связанные с отсутствием единого времени регистрации события разными устройствами контроля и генерации информации о событии.

При подготовке данных для применения к ним средств интеллектуального анализа данных необходимо решить следующие задачи:

1. Объединить информацию от разнородных источников.
2. Интегрировать информацию, относящуюся к одному событию.
3. Устранить повторяющуюся информацию от разнородных источников.

4. Устранить неопределенности, связанные с различиями в идентификации объектов для разных устройств контроля.

5. Устранить неопределенности, связанные с неодновременностью регистрации события разными устройствами контроля.

Подробно были рассмотрены методы подготовки данных, отдельно изучены алгоритмы корреляции событий. В ходе исследования было выявлено, что существующие методы установления взаимосвязей между событиями применимы к данным, получаемым от разных источников схожего типа и имеющим схожий формат представления. Однако практически не решается проблема нечеткости и пропуска данных. Обзор современных исследований, связанных с получением данных от источников разного типа, показал, что значительная часть исследований посвящена проблемам хранения таких данных. В некоторых работах решаются задачи использования разнородных данных для последующего моделирования различных процессов, т.е. не ставится задача анализа данных. Задача, связанная с исследованием событий, которые могут описываться одновременно с помощью нескольких записей от разнородных источников, является актуальной, так как существующие методы интеграции и корреляции данных не подходят для её решения. Кроме того, в ходе исследования было обнаружено, что разработанные для решения задач в сфере сетевой безопасности вероятностные методы корреляции обладают способностью обнаруживать новые сценарии атак, поскольку в их основе лежит построение статистической модели функционирования исследуемой сети, и любое отклонение от нее может быть расценено как возможное действие злоумышленника. Однако такие методы в чистом виде предназначены для обработки и анализа данных одного формата, при этом сами данные связаны исключительно с сетевой активностью, поэтому они не решают задачу приведения данных из разнородных источников к одному формату и не могут быть использованы на наборе данных, отличных от сетевых данных.

Сравнительный анализ рассмотренных методов обработки информации от разнородных источников показал, что данные методы решают не все задачи, упомянутые выше. В частности, обходятся или лишь частично решаются задачи, связанные с устранением неопределенностей, не решаются проблемы объединения разнородной информации о событии.

На основании проведенного обзора поставлены цели и задачи, требующие решения.

Во **второй главе** описана формальная модель последовательности событий в системе распределенного мониторинга.

Для понимания процедуры регистрации событий разнородными устройствами и дальнейшей обработки этих данных была формально описана общая структура распределенной системы мониторинга.

Общая формальная структура системы распределенного мониторинга и процедура регистрации событий представлена на рис. 1.

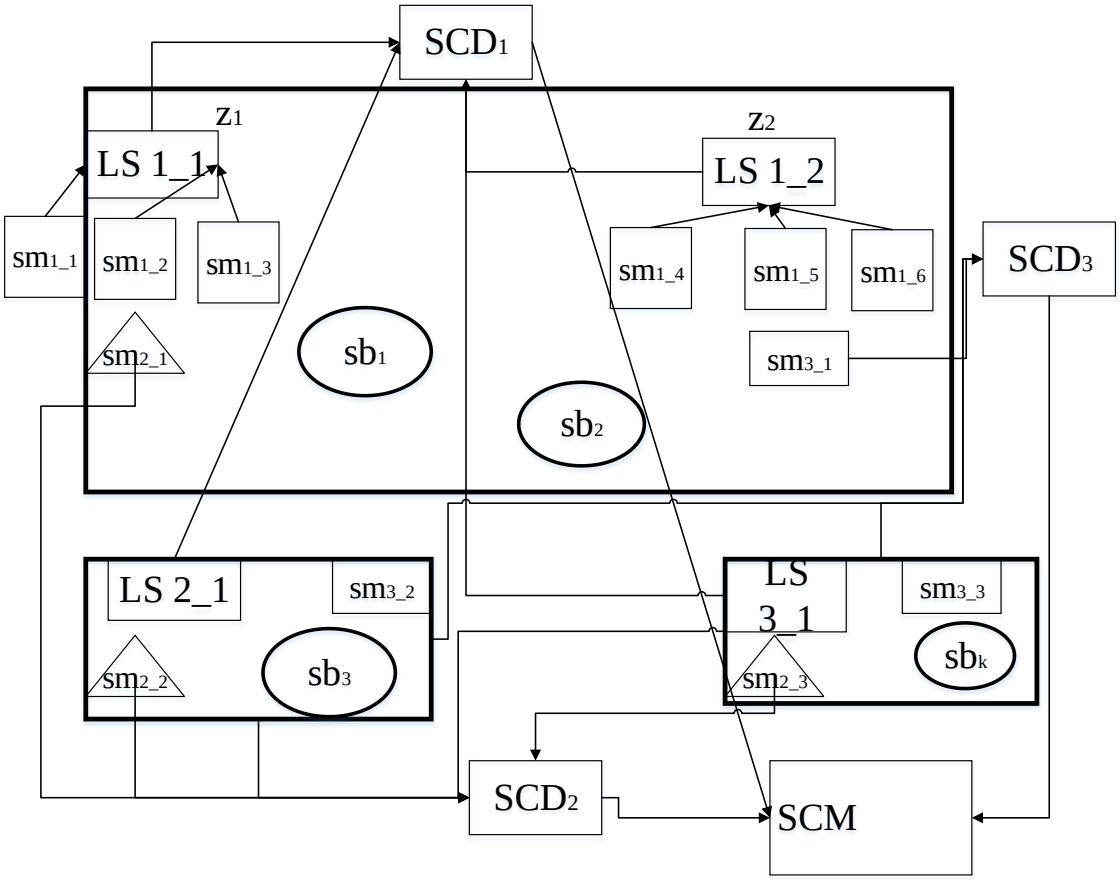

Рисунок 1 – Формальное представление процедуры регистрации событий на производственных объектах

Распределенная система мониторинга включает в себя следующие компоненты:

SB – контролируемые объекты (сущности: сотрудники предприятия, транспортные средства, пользователи сети): $SB = \{sb_1, sb_2, ..., sb_k\}$;

SM – устройства контроля (датчики, измерительные системы, системы контроля доступа, камеры наблюдения, маршрутизаторы): $SM = \{sm_1, sm_2, ..., sm_j\}$;

SCM – центральный узел обработки и анализа;

SCSD – промежуточные системы сбора,

LS – локальные системы сбора.

В границах производственной территории может быть выделено некоторое множество зон, т.е. участков пространства, в пределах которых установлено некоторое множество устройств контроля, формирующие информацию о некоторых событиях, происходящих на данном участке:

$$Z = \{z_1, z_2, ..., z_v\}$$

В рамках данной работы под зоной понимается некоторое пространство, где один контролируемый объект может инициировать одно событие, информация о котором будет получена от расположенных устройств контроля, расположенных на данном участке.

При регистрации события устройство контроля генерирует информацию, которая в общем виде может быть представлена как:

$ent = \{t, sm, add\}$, где

t – время,

sm – идентификатор устройства контроля,

add – дополнительные характеристики, определяемые типом устройства.

Для большинства устройств $sb \in add$, однако существует ряд устройств, не осуществляющих идентификацию объекта.

В общем случае при анализе и группировке некоторого *d* количества записей, которые лишь частично описывают одно и то же событие, важно следовать трем основным правилам.

1. Временные атрибуты события должны совпадать или разница между ними Δt не должна превышать допустимую задержку τ:

$$\max(t_1 ... t_d) - \min(t_1 ... t_d) = \Delta t \le \tau$$

2. Должны совпадать пространственные атрибуты события:

$$sm_1 \subseteq z_v, ..., sm_d \subseteq z_v$$

3. Субъект должен быть одним и тем же:

$a_1^{sb} = a_1^{sb_k}, ..., a_d^{sb} = a_d^{sb_k}$, где a_i^{sb} – атрибут объекта наблюдения в i-записи.

Однако не всегда можно однозначно соотнести данные от разных источников, в особенности, если в некоторых записях отсутствуют временные параметры или же некоторые атрибуты (в особенности идентификаторы субъектов) были зафиксированы некорректно или отсутствуют у данного типа устройств мониторинга (рис. 2).

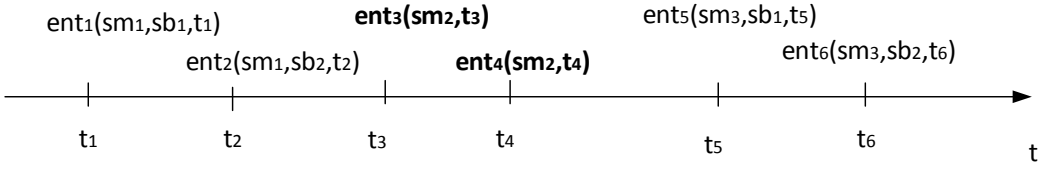

Рисунок 2 – Пример неопределенности отнесения записей к одному из нескольких событий, происходящих одновременно

Предложен метод формирования последовательностей событий, который позволяет сформировать последовательности событий на основе разнородных записей от распределенных источников.

Для более точной интеграции данных предложенный метод использует дополнительные источники данных, такие как различные учетные системы. Разработанный метод позволяет снизить избыточность данных путем объединения одинаковых по смыслу атрибутов и записей, описывающих одно событие, а также решить проблему неопределенностей, связанных с идентификацией объектов наблюдения и регистрацией события разными устройствами в разные моменты времени.

```
read UDS_SM
for number of attributes n do
        if (a_i) describes (par_j) & (a_k) describes (par_j) then
        generate a_i = a_i & a_k; while i=n;
remove useless attributes;
read additional databases;
add information;
        if (sb_k from SB^m is single match to sb_j from SB^q) then
        fill mutual attributes;
        else next step;
sort by zone;
for each zone
sort by time;
combine entries with sb =sb_k and time parameters fulfill condition 1;
if number of entries num = d then
aggregate entries to one event;
else find entry with sb is empty;
find all entries with time fulfill condition 1;
if for all entries sb =sb_k then
                aggregate entries to one event;
else compose possible entries combinations;
for each combination
        for each event in combination
        calculate dist;
calculate avdist;
choose combination with avdist = min;
calculate t_e;
aggregate entries to one event;
form event sequences;
write UDS_corr
```

В **третьей главе** предложена методика решения задач интеллектуального анализа на основании информации от разнородных источников распределенной системы мониторинга с целью поддержки принятия решений.

Методика формирования типовых процессов в общем виде включает в себя все этапы преобразования исходных данных для приведения их к виду, пригодному для анализа, группировки и сортировки преобразованных данных для формирования отдельных процессов и выделения типовых шаблонов процессов.

Условно можно выделить три этапа методики анализа последовательности событий (рис. 3). Первые два этапа направлены на преобразование данных и приведение их к формату, пригодному для применения к ним известных методов анализа данных. Третий этап направлен непосредственно на реализацию анализа данных.

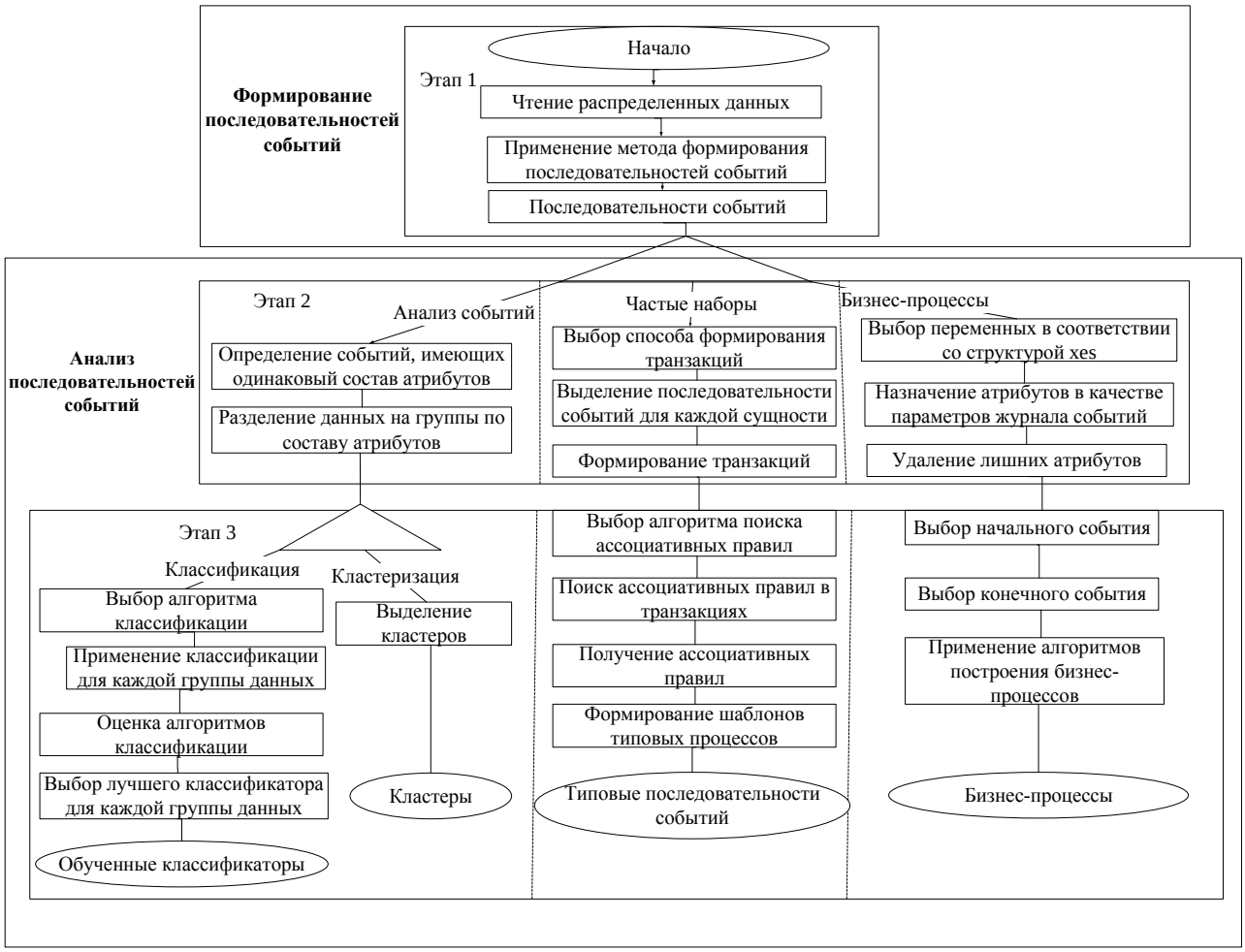

Рисунок 3 – Методика решения задач последовательности событий

Первый этап включает в себя общие преобразования данных и применяется в случае наличия разнородных средств в составе системы мониторинга. В таком случае информация о зарегистрированных событиях имеет разный формат и требует приведения к единому виду. На первом этапе выполняется преобразование исходных данных, в результате которого должен быть получен набор данных, в котором каждому зафиксированному параметру соответствует единственный атрибут. Затем выполняется объединение записей об одном событии, если одно событие может быть зафиксировано более чем одним средством мониторинга. Результатом является получение набора данных, в котором каждому отдельно взятому событию соответствует единственная запись, полученная на основе информации от разнородных устройств контроля. Далее события должны быть упорядочены по времени для формирования последовательностей событий.

Второй этап методики анализа последовательностей событий включает в себя преобразования, необходимые для приведения данных, полученных на первом этапе, к формату, пригодному для применения к ним выбранного метода анализа. В рамках методики рассматривается три задачи интеллектуального анализа данных: анализ отдельно взятых событий с целью выявления аномалий, построение типовых последовательностей событий с помощью частых наборов и построение бизнес-процессов. Для перехода ко

второму этапу анализа требуется выбрать желаемый тип анализа событий. Исследователь может реализовать как один желаемый тип анализа, так и все три, выполнить их как последовательно, так и одновременно.

Для выполнения классификации или кластеризации событий исходные данные будут автоматически поделены на группы в зависимости от состава атрибутов записей. Полученные группы данных будут использованы на третьем этапе.

Для формирования ассоциативных правил на основе транзакций имеющиеся записи упорядочиваются по времени, т.е. получается последовательность различных событий, относящихся к различным процессам. Сначала выбирается способ, который будет использоваться формирования транзакций, затем осуществляется сам процесс формирования. В результате будет получен набор транзакций, соответствующий заданному условию формирования. Транзакции представляют собой наборы последовательно зафиксированных событий, относящихся к одному процессу. Полученные транзакции будут использованы для анализа и выявления типовых шаблонов.

Для построения моделей бизнес процессов следует сначала сформировать журналы событий, с которым способны работать алгоритмы интеллектуального анализа процессов. Для этого выбираются атрибуты, соответствующие атрибутам формата .xes. Если для анализа требуется учесть большее количество атрибутов, следует осуществить конкатенацию значимых полей, связанных друг с другом по смыслу и цели анализа, и объединить их в наиболее подходящий атрибут.

Третий этап включает в себя выполнение выбранного типа анализа данных. При осуществлении классификации событий для каждой группы данных, полученной на втором этапе, определяется наилучший алгоритм классификации. Затем каждый из выбранных классификаторов обучается на полной выборке данных и после обучения может быть использован для классификации новых данных.

При выявлении типовых шаблонов в последовательностях событий сформированные на предыдущем этапе транзакции подаются на вход выбранного алгоритма поиска частых наборов. В результате применения таких алгоритмов получаются наборы часто встречающихся элементов (событий) для различных процессов. Полученные наборы определяют последовательности ключевых событий, входящих в процессы различных типов. Такие наборы могут быть использованы для выявления аномального поведения или отклонений в ходе выполнения процесса.

При построении моделей бизнес процессов строятся модели, отвечающие целям анализа. Таких моделей может быть как одна, так и несколько, если необходимо проанализировать разные процессы или разные характеристики одного процесса. Построенные модели могут использоваться для оценки хода реального процесса.

В **четвертой главе** рассмотрена программная реализация метода формирования последовательности событий на основе информации от

разнородных источников. Произведена экспериментальная оценка методики решения задач интеллектуального анализа применительно к реальным данным от разнородных источников мониторинга, включающая в себя три серии экспериментов. В рамках каждого экспериментального исследования были последовательно проведены оценки отдельно взятых событий методами классификации или кластеризации, построены ассоциативные правила с помощью алгоритмов поиска частых наборов, а также сформированы журналы событий и построены модели бизнес-процессов.

Первая серия экспериментов посвящена подготовке и анализу данных, описывающих перемещения грузовых транспортных средств на территории предприятия. Исследуемые данные содержат информацию о следующих событиях: перемещения железнодорожных составов в зонах железнодорожных переездов и станций, манипуляции, связанные со взвешиванием грузов, перевозимых составами;перемещения грузовых транспортных средств в различных зонах (въезд, выезд, движение в сторону) а также, действиях, связанных с их погрузкой и разгрузкой (взвешивание); неисправности средств наблюдения (камер); различного вида нарушения регламента. Данные представляют собой записи, сгенерированные такие средствами мониторинга как камеры фото- и видеофиксации, системы контроля доступа, системы измерения веса. Данные от средств мониторинга хранятся в виде двух таблиц, одна из которых содержит 6 506 825 записей, другая содержит 15 522 224 записи.

Обработка данных производилась в среде RapidMiner (рис. 4). Некоторые преобразования данных производились с помощью стандартных блоков среды, для реализации отдельных функций был написан скрипт.

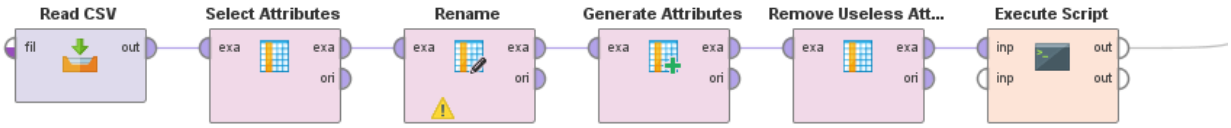

Рисунок 4 – Обработка данных в среде RapidMiner

Диаграмма классов плагина, реализующего разработанный метод, представлена на рис. 5.

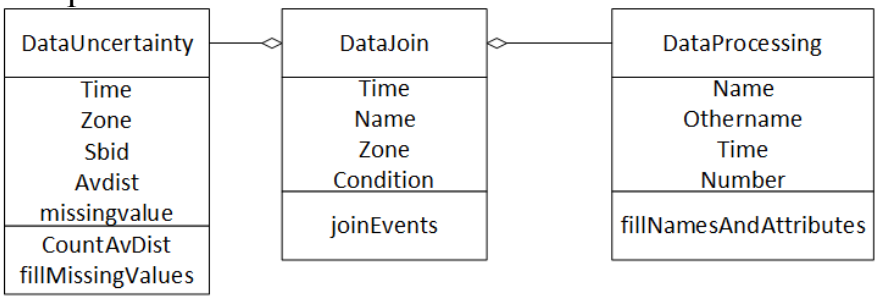

Рисунок 5 – Диаграмма классов плагина

В результате произведенных преобразований и формирования единого события на основе данных от разнородных источников удалось сократить размерность таблицы в 20 раз и на 17% снизить количество пропущенных значений.

В результате группировки данных были выделены группы данных с одинаковым составом атрибутов. Для некоторых из них были протестированы различные алгоритмы классификации, произведена их оценка, выбран лучший классификатор для каждой группы. Для поиска частых наборов был выбран алгоритм FPG, так как он считается одним из самых эффективных алгоритмов поиска частых наборов. В результате был получен набор часто повторяющихся событий в рамках посещения различных зон.

Полученные данные могут быть использованы при анализе реального хода процесса. В случаях, когда текущий ход процесса отклоняется от выявленных типовых последовательностей, это может свидетельствовать о возможном нарушении регламента.

Заключительный этап анализа представлял собой формирование журналов событий и построение модели типового бизнес-процесса (рис. 6).

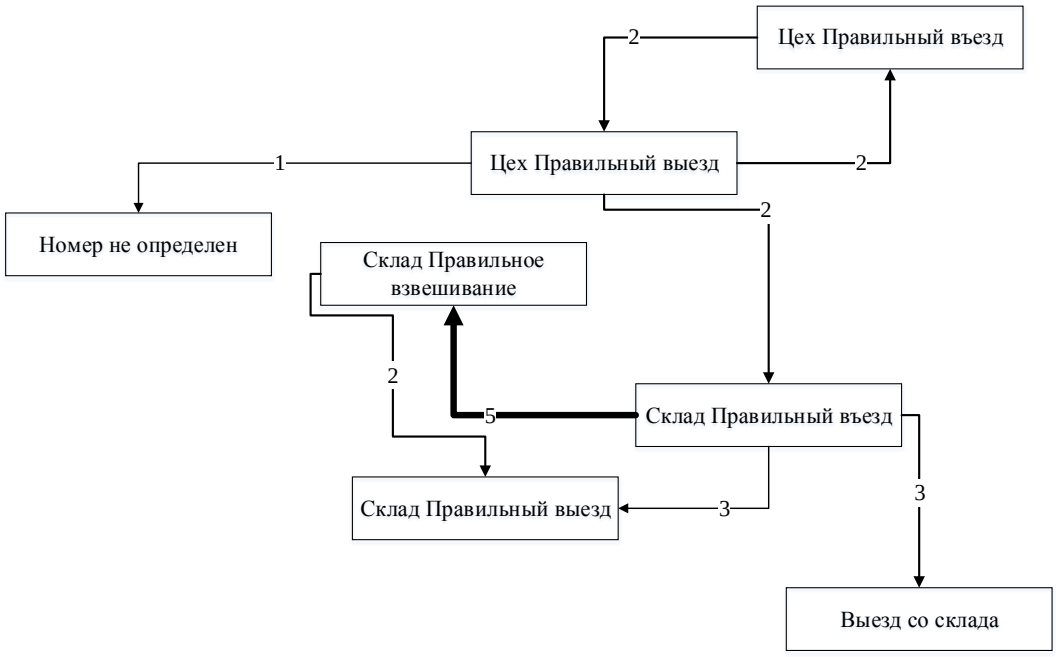

Рисунок 6 – Модель бизнес-процесса для перемещений транспортного средства на территории предприятия

Разработанная методика также была применена к задаче исследования перемещений сотрудников на территории офисного здания и к задаче исследования сетевого поведения пользователей. Для данных о перемещениях сотрудников вместо классификации была выполнена кластеризация, которая позволила выявить некоторые возможные аномалии. Построенные для этих данных ассоциативные правила также выявили некоторые аномальные перемещения сотрудников (рис. 7). Также были построены модели бизнес процессов, которые позволяют оценить маршруты перемещений сотрудников и длительность посещений ими различных зон.

lazada001	Engineering	2-1	6	0.7	23.6	2-1 _ 6 _ 0.7 _ 23.6	
lazada001	Engineering	2-4	6	0.7	0.7	2-4 _ 6 _ 0.7 _ 0.7	Правило2-4 _ 6 _ 0.7 _ 0.7 -> 1-4 _ 6 _ 0.7 _ 0
lazada001	Engineering	1-4	6	0.7	0	1-4 _ 6 _ 0.7 _ 0	

Рисунок 7 – Пример выполнения ассоциативных правил для событий

При анализе сетевого трафика выявить шаблоны легитимного трафика и некоторые отклонения от него, указывающие на проведение внешних атак на сервер и участие нескольких узлов сети в проведении атаки на внешний узел.

Проведенные эксперименты подтвердили работоспособность методики и ее применимость к данным о различных процессах.

ЗАКЛЮЧЕНИЕ

В соответствии с целью и задачами диссертации получены основные результаты работы, заключающиеся в следующем:

1. Проведен анализ существующих средств обработки разнородной информации.
2. Разработана формальная модель последовательности событий, зарегистрированных в системе распределенного мониторинга.
3. Разработан метод формирования последовательностей событий на основе информации от разнородных источников системы распределенного мониторинга.
4. Предложена методика решения задач интеллектуального анализа на основании информации от распределенных источников с целью поддержки принятия решений.
5. Разработана программная реализация метода формирования последовательностей событий.
6. Проведено экспериментальное исследование разработанных средств обработки информации.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ ПО ТЕМЕ ДИССЕРТАЦИИ

Статьи, опубликованные в изданиях, включенных в перечень ВАК:

1. Бекенева, Я.А. Преобразование данных от разнородных систем мониторинга /Я. А. Бекенева // Программные продукты и системы. - 2019. - Т. 32. - № 2. - С. 197–206.
2. Классификация событий по составу независимых атрибутов / Я. А. Бекенева [и др.] // Известия СПбГЭТУ «ЛЭТИ». – 2018. – №9. – С. 22-27.
3. Обзор алгоритмов корреляции событий безопасности для обеспечения безопасности облачных вычислительных сред / Е. С. Новикова [и др.] // Информационно-управляющие системы. – 2017. - Т. 90. - №5. - с. 95-104.
4. Новикова, Е. С. Исследование методов корреляции событий безопасности для обеспечения безопасности облачных вычислительных сред / Новикова Е. С., Бекенева Я. А., Шоров А. В. // Известия СПбГЭТУ «ЛЭТИ». - 2017. - №7. - с. 23-30.
5. DRDoS-атаки и механизмы защиты от них / Я. А. Бекенева [и др.] // Известия СПбГЭТУ «ЛЭТИ». – 2017. - №1. - с. 3-7.
6. Бекенева, Я. А. Анализ актуальных типов DDoS-атак и методов защиты от них / Я. А. Бекенева // Известия СПбГЭТУ «ЛЭТИ». – 2016. - №1. - с. 7-14.
7. Система имитационного моделирования для разработки и тестирования методов защиты от ddos-атак с возможностью подключения реальных узлов / К.А. Борисенко [и др.] // Безопасность информационных технологий. – 2015. -

№4. - с. 6-17.

8. Система имитационного моделирования для разработки и тестирования методов защиты от DDoS-атак с возможностью подключения реальных узлов / К. А. Борисенко [и др.]// Известия СПбГЭТУ «ЛЭТИ». – 2015. - №6. - с. 22-29.

9. Моделирование DDoS-атак и механизмов защиты от них / Я. А. Бекенева [и др.] // Известия СПбГЭТУ «ЛЭТИ». – 2015. - №3. - с. 32-40.

10. Бекенева, Я. А. Формальные модели систем распределенных вычислений и типы сетевых атак на них / Я. А. Бекенева, А. В. Дорохов // Известия СПбГЭТУ «ЛЭТИ». – 2014. - №7. – с. 26-29.

Публикации, входящие в международные базы цитирования Scopus и WoS:

11. Violation Detection in Heterogeneous Events Streams / Y. A. Bekeneva [et al.] //Procedia Computer Science. – 2019. – Vol. 150. – P. 381-388.

12. Novikova, E. The Location-Centric Approach to Employee's Interaction Pattern Detection / E. Novikova, Y. Bekeneva, A. Shorov //2019 27th Euromicro International Conference on Parallel, Distributed and Network-Based Processing (PDP), Pavia, 13-15 feb. 2019. – P. 373-378.

13. Bekeneva, Y. A. Event Pattern Constructing Technique for Prediction of Violations in Technological Processes / Y. A. Bekeneva [et al.] //2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), St-Petersburg, 28-31 jan. 2019. – P. 182-186.

14. Novikova, E., The Motif-Based Approach to the Analysis of the Employee Trajectories within Organization / E. Novikova, Y. Bekeneva, A. Shorov //Security and Communication Networks. – 2018. – Vol. 2018.

15. Novikova, E. S. Approach for the analysis of the contacts of the critical infrastructure employees / E. S. Novikova [et al.] // Young Researchers in Electrical and Electronic Engineering (EIConRus), St-Petersbutg, 28 jan. – 1 feb. 2018. - P. 347-350.

16. Kholod, I. I. Intellectual model for violations detection in the business process / I. I. Kholod [et al.] // Young Researchers in Electrical and Electronic Engineering (EIConRus), St-Petersbutg, 28 jan. – 1 feb. 2018. - P 313-317.

17. Method for Transformation of Data from Heterogeneous Monitoring Devices for Violations Detection / Ya. Bekeneva [et al.] // Soft Computing and Measurements (SCM), 2017 XXI IEEE International Conference on, St-Petersburg, 23-25 May 2018. - P.753-756.

18. Novikova, E. S. Visualization Technique for Representing Structure of the Cluster of Objects / E. S. Novikova, Y. A. Bekeneva, A. V. Shorov //2018 Third International Conference on Human Factors in Complex Technical Systems and Environments (ERGO) s and Environments (ERGO), St-Petersburg, 4-7 jul. 2018. – P. 97-100.

19. Bekeneva, Y.A. Towards Simulation of the Processes Related to Transport Movement within Industrial Objects / Y.A. Bekeneva [et al.] //2018 IEEE International Conference" Quality Management, Transport and Information Security, Information Technologies"(IT&QM&IS), St-Petersbutg, 24-28 sept. 2018. – P. 304-

307.

20. Bekeneva, Y. Principles of designing distributed data flow control systems with high resistance to malicious actions / Ya. Bekeneva, E. Novikova, A. Shorov // Control in Technical Systems (CTS), 2017 IEEE II International Conference, St-Petersburg, 25-27 oct. 2018. - P. 353-355.

21. Bekeneva, Y. Simulation of DRDoS-attacks and protection systems against them / Y. Bekeneva, A. Shorov // Soft Computing and Measurements (SCM), 2017 XX IEEE International Conference on, St-Petersburg, 24-26 May 2017. – P. 165 – 167.

22. Bekeneva, Y. Development of protection mechanisms against DRDoS-attacks and combined DRDoS-attacks / Y. Bekeneva, A. Shorov // Vibroengineering PROCEDIA vol. 12 26-th International Conference on VIBROENGINEERING, St-Petersburg, 29-30 jun. 2017. - P. 178-183.

23. Bekeneva, Y. Investigation of protection mechanisms against DRDoS attacks using a simulation approach / Y. Bekeneva, N. Shipilov, A. Shorov // 16th International Conference on Next Generation Wired/Wireless Advanced Networks and Systems New2AN, St-Petersburg, 26-28 sept. 2016. – P. 316-325.

24. Investigation of DDoS Attacks by Hybrid Simulation / Y. Bekeneva [et al.] // In Information and Communication Technology. – Berlin: Springer. – 2015. - P. 179-189).

25. Bekeneva, Y. Simulation of DDoS-attacks and protection mechanisms against them / Y. Bekeneva [et al.] // Proceedings of the 2015 IEEE North West Russia Section Young Researches in Electrical and Electronic Engineering Conference, St-Petersburg, 2-4 feb. 2015. - P. 50-56.

Другие статьи и материалы конференций:

26. Метод преобразования данных от разнородных средств контроля для выявления нарушений / Я. А. Бекенева [и др.] // XXI международная конференция по мягким вычислениям и измерениям, г. Санкт-Петербург, 23-25 мая 2018 – Т. 2. – С. 34-37.

27. Новикова, Е.С. Подход к формированию мотивов перемещений сотрудников внутри организации отраслей / Е.С. Новикова, Я.А. Бекенева // Региональная информатика и информационная безопасность. Сборник трудов. - 2017. Вып. 3. – С.86-89.

28. Бекенева, Я.А. Имитационное моделирование DRDOS-атак и систем защиты от них /Я. А. Бекенева, А. В. Шоров // XX международная конференция по мягким вычислениям и измерениям, г. Санкт-Петербург, 24-26 мая 2017. – Т. 1. – С. 392-396.

29. Бекенева, Я. А. Исследование DRDoS-атак и методов защиты от них методами имитационного моделирования /Я. А. Бекенева //Наука настоящего и будущего. – 2017. – Т. 1. – С. 28-29.

30. Борисенко, К. А.. Система моделирования для разработки и тестирования техник безопасности сетевой инфраструктуры от DDoS-атак /К. А. Борисенко, Я. А. Бекенева, А. В. Шоров // Информационная безопасность регионов России,

Санкт-Петербург, 28-30 октяб. 2015. – С. 136-137.

31. Бекенева, Я.А. Анализ методов обнаружения DDoS-атак на системы облачных вычислений / Я. А. Бекенева, А. В. Шоров // Ежемесячный научный журнал "Prospero". – 2014. - № 2. – С. 60-62.

Свидетельства о государственной регистрации программы для ЭВМ

32. Бекенева, Я. А. Программа формирования событий на основе данных от разнородных источников: Свидетельство о государственной регистрации программы для ЭВМ №2019660468 / Я. А. Бекенева, К. С. Кучинский, А. В. Шоров; правообладатель ФГБОУ СПбГЭТУ «ЛЭТИ». – Зарегистрировано в Реестре программ для ЭВМ 6 августа 2019 г.

33. Программа поиска шаблонов нештатных ситуаций в последовательности разнотипных событий: Свидетельство о государственной регистрации программы для ЭВМ №2018619555 /Д. Шевелев, И. И. Холод, Я. А. Бекенева, А. В Шоров; правообладатель ФГБОУ СПбГЭТУ «ЛЭТИ». – Зарегистрировано в Реестре программ для ЭВМ 8 августа 2018.

34. Бекенева, Я. А. Система для имитационного моделирования для разработки и тестирования методов защиты от DDoS-атак, основанных на отражении и усилении трафика: Свидетельство о государственной регистрации программы для ЭВМ № 2016660779 /Я. А. Бекенева, Н. Н. Шипилов; правообладатель ФГБОУ СПбГЭТУ «ЛЭТИ». – Зарегистрировано в Реестре программ для ЭВМ 21 сентября 2016.