ПРЕДИСЛОВИЕ

Важнейшим атрибутом нашего времени является глобальная информационная интеграция, основанная на построении компьютерных сетей масштаба предприятия и их объединении посредством сети Интернет. Сложность логической и физической организации современных сетей приводит к объективным трудностям при решении вопросов управления и защиты сетей. В процессе эксплуатации компьютерных сетей администраторам приходится решать две главные задачи:

- диагностировать работу сети и подключенных к ней серверов, рабочих станций и соответствующего программного обеспечения;
- защищать информационные ресурсы сети от несанкционированной деятельности хакеров, воздействий вирусов, сетевых червей и т. п., обеспечивая их конфиденциальность, целостность и доступность.

При решении задач, связанных с диагностикой и защитой сетевых ресурсов, центральным вопросом является оперативное обнаружение состояний сети, приводящих к потере полной или частичной ее работоспособности, уничтожению, искажению или утечке информации, являющихся следствием отказов, сбоев случайного характера или результатом получения злоумышленником несанкционированного доступа к сетевым ресурсам, проникновения сетевых червей, вирусов и других угроз информационной безопасности. Раннее обнаружение таких состояний позволяет своевременно устранить их причину, а также предотвратить возможные катастрофические последствия.

Для обнаружения угроз используется большой спектр специализированных средств. Так, при решении проблем диагностики сетей применяются средства систем управления, анализаторы сетевых протоколов, системы нагрузочного тестирования, системы сетевого мониторинга и др. Проблемы защиты информационных ресурсов сетей решаются с помощью межсетевых экранов (firewall), антивирусов, систем обнаружения атак (СОВ) (Intrusion Detection System, IDS), систем контроля целостности, криптографических средств защиты. Характерными особенностями использования этих систем является либо их периодическое и кратковременное применение для решения определенной проблемы, либо постоянное использование, но со статическими настройками. В результате методы анализа, используемые в современных системах, направлены на обнаружение известных и точно описанных типов воздействий, но зачастую оказываются не в состоянии обнаружить их модификации или новые типы, что делает их использование малоэффективным.

На сегодняшний день актуальной задачей является поиск эффективных методов выявления недопустимых событий (аномалий) в работе сети, являющихся следствием технических сбоев или несанкционированных воздействий. Основным требованием предъявляемым к этим методам является возможность обнаружения произвольных типов аномалий, в том числе новых, а также воздействий, распределенных во времени.

Первые работы, посвященные данной проблеме, были опубликованы в 90-х годах прошлого столетия. В настоящий момент исследования в этой области интенсивно продолжаются. Общий подход, лежащий в основе этих исследований, заключается в поиске методов анализа, позволяющих выявлять аномальные состояния информационных ресурсов в виде отклонений от обычного («нормального») состояния. Эти отклонения могут являться результатами сбоев в работе аппаратного и программного обеспечения, а также следствиями сетевых атак хакеров. Такой подход теоретически позволит обнаруживать как известные, так и новые типы проблем. От эффективности и точности аппарата, определяющего «нормальное» состояние и фиксирующего отклонение, зависит в целом эффективность решения вопросов диагностики и защиты сетевых ресурсов. Особую важность на текущий момент представляет проблема обнаружения аномальных состояний в работе сети, имеющих распределенный во времени характер.

В *первой главе* рассмотрены наиболее распространенные типы сетевых аномалий и их классификация. На основе рассмотренной классификации методов и систем обнаружения анализируются существующие системы и инструменты обнаружения сетевых аномалий, включая поведенческие методы, методы машинного обучения, методы вычислительного интеллекта, методы основанные на знаниях. Анализируются существенные аспекты обнаружения сетевых аномалий, включая обоснование требуемого объема обучающих выборок, выбор метрик, используемых для оценки методов и систем обнаружения сетевых аномалий и др.

В *главе 2* анализируются проблемы контроля и анализа сетевого трафика с помощью сетевых анализаторов, включая задачи и средства анализа сетевого трафика. Излагаются особенности сбора данных с помощью протоколов NetFlow, SNMP, технологии и подходы сетевого мониторинга, включая программный сниффер Tcpdump, технологию DPI. Анализируются известные наборы обучающих и тестовых данных, используемые в задачах обнаружения сетевых атак, включая синтетические, эталонные, реальные наборы данных. Рассмотрены системы и инструменты, используемые для выбора и обработки характеристик трафика, включая Rattle, Weka, MOA, Python, Orange, RapidMiner, Scikit-learn и др.

Глава 3 посвящена анализу статистических характеристик сетевых аномалий, вызванных атаками. Рассмотрены примеры аномалий трафика, вызванные типовыми сетевыми атаками в локальной сети: ICMP flooding, Flash crowd, Smurf, Fraggle, Synflooding, Udp-storm и др. Анализируются возможности классификации аномальных вторжений статистическими методами, а

также особенности фрактальных свойств телекоммуникационного трафика, которые могут быть положены в основу алгоритмов обнаружения сетевых аномалий.

Глава 4 посвящена статистическим методам обнаружения аномалий. Рассмотрен широкий спектр известных статистических методов обнаружения аномалий и проведена их классификация. Введены в рассмотрение критерии аномального поведения трафика. Анализируются параметрические методы регистрации аномальных изменений трафика; статистические алгоритмы и методы идентификации аномалий включая алгоритмы обобщенных экстремальных отклонений; алгоритм обнаружения выбросов на основе вычисления моментов статистических распределений; алгоритм Шапиро–Вилка, различные алгоритмы обнаружения выбросов, а также матричные методы обнаружения выбросов.

Рассмотрены методы, критерии и алгоритмы математической статистики включая методы описательной статистики, критерии Кохрана–Кокса, Беренса–Фишера, Фишера, а также композицию статистических критериев, используемых при работе алгоритмов обнаружения аномалий в режиме on-line. Приводятся численные результаты обнаружения аномалий статистическими методами в режиме on-line.

Анализируются особенности обнаружения аномалий с использованием информационных критериев, а также алгоримов обнаружения с использованием кумулятивных сумм, методом разладки Бродского–Дарховского и подобных им. Описываются особенности поиска и оценки аномалий сетевого трафика на основе циклического ананализа.

Глава 5 посвящена обнаружению аномалий методами кратномасштабного анализа (КМА). Излагаются основные положения КМА, включая разложение и реконструкцию; дискретное вейвлет-преобразование (ДВП) с максимальным перекрытием, пакетные вейвлеты и др.

Рассмотрены вопросы обнаружения аномалий методами ДВП в режиме off-line и on-line, а также особенности реализации предварительной фильтрации трафика методами трешолдинга. Исследуется оценка достоверности обнаружения аномалий методами кратномасштабного анализа. Приводится сравнительный анализ результатов обнаружения аномалий при использовании различных систем вейвлетов

Глава 6 посвящена обнаружению аномалий методами моно и мультифрактального анализа. Излагаются основные положения теории фракталов и мультифракталов. Анализируются особенности имитационного моделирования мультифрактальных характеристик телекоммуникационного трафика в условиях DoS-атак. Исследуются вопросы обнаружения аномалий методом мультифрактального кратномасштабного анализа в реальном времени, а также путем оценки скачка фрактальной размерности в режиме on-line. Анализируются структура и особенности реализации программного обеспечения методов оценки фрактальной размерности. Приводятся при-

меры обнаружения аномалий сетевого трафика реализующие оценку фрактальной размерности.

В *главе 7* рассматриваются вопросы обнаружения и локализации аномалий в крупномасштабных сетях. Излагаются основные положения сетевой томографии базирующиеся на понятии матрицы трафика (TM). Рассмотрены особенности моделирования и оценки матрицы трафика на основе информационно-теоретического подхода. Важное место в главе занимают вопросы оценки TM методом максимального правдоподобия. Анализируются вопросы рекуррентной оценки TM.

Анализируются методы обнаружения и локализации аномалий объема. Рассматривается обнаружение аномалий с использованием фильтра Калмана, а также обнаружение и локализация аномалий методом главных компонент. Анализируется структура системы мониторинга аномалий объема в больших распределенных системах. Приводятся результаты имитационного моделирования

Глава 8 посвящена вопросам обнаружения и прогнозирования аномалий трафика в потоковых данных. Приведен анализ существующих методов прогнозирования включая линейные авторегрессионные (AR) модели; авторегрессионные модели скользящего среднего (ARMA); авторегрессионные интегральные модели скользящего среднего (ARIMA); фрактальные авторегрессионные интегральные модели скользящего среднего (FARIMA) и др.

Рассмотрены особенности обнаружения аномалий на основе прогнозирования профиля нормального функционирования компьютерной системы.

Основное внимание в главе уделено вопросам потокового обнаружения аномалий. Обнаружение аномалий в дискретных последовательностей является сложной задачей, поскольку предполагает использование для обнаружения аномалий последовательный (потоковый) характер поступления данных. Анализируются особенности потоков данных, которые рассматриваются как дополнительные ограничения. Показано, что, как правило, формулировки задачи обнаружения аномалий в дискретных последовательностях сетевых технологий принципиально различны и требуют эксклюзивных методов их решения. Рассматриваются различные формулировки проблемы обнаружения аномалий в последовательностях и возможные сценарии их реализации методами основанными на ядре, скользящем окне, на моделях Маркова различной сложности. Анализируются особенности обнаружения аномалий в потоковых данных на основе скрытых марковских моделей.