На правах рукописи

Щерба Мария Витальевна

ОБНАРУЖЕНИЕ НИЗКОАКТИВНЫХ РАСПРЕДЕЛЕННЫХ АТАК ТИПА «ОТКАЗ В ОБСЛУЖИВАНИИ» В КОМПЬЮТЕРНЫХ СЕТЯХ

05.13.19 – Методы и системы защиты информации, информационная безопасность

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата технических наук

Омск 2012

Работа выполнена в Омском государственном техническом университете.

Научный руководитель: доктор технических наук, профессор
Файзуллин Рашит Тагирович

Официальные оппоненты: доктор технических наук, профессор каф. «Вычислительные системы и сети» Московского государственного института электроники и математики
Саксонов Евгений Александрович

кандидат технических наук, доцент каф. «Информационная безопасность» Сибирской государственной автомобильно-дорожной академии
Семенова Ирина Ивановна

Ведущая организация: ФГБУ ВПО «Омский государственный университет путей сообщения»

Защита состоится «22» мая 2012г. в 16-00 часов на заседании диссертационного совета Д 212.133.03 при Московском государственном институте электроники и математики (МИЭМ) по адресу: 109028, Москва, Б. Трехсвятительский пер., 3.

С диссертацией можно ознакомиться в библиотеке Московского государственного института электроники и математики (МИЭМ) по адресу: 109028, Москва, Б. Трехсвятительский пер., 3.

Автореферат разослан «____» _апреля_ 2012 г.

Ученый секретарь
диссертационного совета
Д 212.133.03,
д.т.н., доцент Ю. Л. Леохин

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность работы

Одной из основных тенденций последних лет в сфере компьютерных преступлений является рост количества и сложности атак на доступность информации (ресурсов автоматизированной системы), как один из трех основных критериев (наряду с конфиденциальностью и целостностью) информационной безопасности объекта. Данные атаки образуют класс атак типа «отказ в обслуживании» (DoS-атаки). В этот класс попадают атаки на компьютерную систему, цель которых - довести систему до такого состояния, в котором её легитимные пользователи не смогут получить доступ к предоставляемым системой ресурсам (серверам, сервисам), либо этот доступ будет затруднён. Если атака выполняется одновременно с большого числа компьютеров, имеет место DDoS-атака (Distributed Denial of Service, распределённая атака типа «отказ в обслуживании»). За последние несколько лет количество таких атак выросло многократно и на сегодня данный класс атак имеет максимальную долю от общего числа атак. Так, в 2010 году была зафиксирована DDoS-атака с мощностью потока более 100 Гбит/сек., которая является самой мощной атакой за все время наблюдений. А в 2011 году был поставлен абсолютный рекорд по суммарному объему DDoS-трафика, который превысил трафик за все вместе взятые года предыдущих исследований.

В целях минимизации последствий DDoS-атак, их обнаружение и классификация является крайне важной и вместе с тем сложной задачей. Данная проблема широко рассматривается в работах П.Д. Зегжды, Б.Н. Оныкия, А.А. Молдовяна, А.В. Лукацкого, И. Яблонко, K.J. Houle, C. Patrikakis и других исследователей. Основной способ распознавания DDoS-атаки заключается в обнаружении аномалий в структуре трафика. Традиционные механизмы обеспечения безопасности - межсетевые экраны и системы обнаружения вторжений – не являются эффективными средствами для обнаружения DDoS-атак и защиты от них, особенно атак трафиком большого объема. Фундаментальной предпосылкой для обнаружения атак является построение контрольных характеристик трафика при работе сети в штатных условиях с последующим поиском аномалий в структуре трафика (отклонения от контрольных характеристик). Аномалия сетевого трафика – это событие или условие в сети, характеризуемое статистическим отклонением от стандартной структуры трафика, полученной на основе ранее собранных профилей и контрольных характеристик. Любое отличие в структуре трафика, превышающее определенное пороговое значение, вызывает срабатывание сигнала тревоги.

Вместе с тем, существующие методы обнаружения DDoS-атак, позволяющие эффективно распознавать DDoS-атаки транспортного уровня (SYN-флуд, UDP-флуд и другие), малоэффективны для обнаружения

низкоактивных DDoS-атак прикладного уровня («медленный» HTTP GET флуд и «медленный» HTTP POST флуд). Подробное описание этой проблемы приводится в работе W.O. Chee и T. Brennan. Указанный класс DDoS-атак возник сравнительно недавно и на сегодняшний день представляет основную угрозу доступности информации в распределенных компьютерных сетях. Данные атаки приводят к потерям запросов и ответов, т.е. фактическому отказу веб-серверов на основе Microsoft IIS, Apache и других систем. Кроме того, атака может быть адаптирована для воздействия на SMTP и даже DNS-серверы.

Таким образом, разработка системы обнаружения низкоактивных распределенных атак типа «отказ в обслуживании» является актуальной и практически важной задачей.

Цель работы

В связи с этим цель диссертационной работы заключается в разработке и практической реализации методики обнаружения распределенных атак типа «отказ в обслуживании» транспортного и прикладного уровней в компьютерных сетях. Для достижения поставленной цели были определены и решены следующие задачи:

1. Анализ проблемы обнаружения распределенных атак типа «отказ в обслуживании» и классификация существующих DDoS-атак, методов и средств их обнаружения.
2. Построение математической модели атак типа «отказ в обслуживании» в сетях массового обслуживания и метода их обнаружения.
3. Разработка архитектуры и реализация программно-аналитического комплекса для имитационного моделирования и расчета статистических характеристик сетей массового обслуживания.
4. Разработка архитектуры и реализация программно-аналитического комплекса, предназначенного для обнаружения низкоактивных атак типа «отказ в обслуживании» в распределенных компьютерных сетях на основе разработанной методики.

Объектами исследования являются распределенные компьютерные сети, процессы передачи информации и конкретные реализации атак типа «отказ в обслуживании» на ресурсы информационной системы.

Предметами исследования выступают модели и методы моделирования вычислительных сетей сетями массового обслуживания, а также методы обнаружения распределенных атак типа «отказ в обслуживании».

Методы исследований

В диссертационной работе используются методы математического моделирования, теории вероятностей и математической статистики, теории

систем и сетей массового обслуживания. Полученные теоретические результаты подтверждены экспериментальными исследованиями, выполненными с применением среды программирования Microsoft Visual C++ и библиотек OpenMP, Boost, WinPcap.

Достоверность

Достоверность результатов работы обеспечивается корректной постановкой задач, строгостью применения математических моделей, непротиворечивостью полученных результатов, а также практическим применением разработанных методов.

Научная новизна

В диссертационной работе получены следующие научные результаты:

1. Доказана сходимость разработанной итерационной процедуры аппроксимации сети массового обслуживания с конечными очередями сетью Джексона для вычисления интенсивностей входящих в узлы потоков заявок (для различных топологий исходной сети).
2. Предложена методика построения цепи Маркова на основе характеристик сети массового обслуживания для оценки вероятности потери произвольной заявки в сети.
3. Исследована и экспериментально обоснована адекватность предложенных математических моделей низкоактивных атак типа «отказ в обслуживании» в сетях массового обслуживания.
4. Разработана и исследована методика обнаружения низкоактивных атак типа «отказ в обслуживании» в компьютерных сетях на основе оценки вероятности потерь заявок.

Практическая значимость

Практическая значимость результатов подтверждена внедрением разработанного программно-аналитического комплекса в систему защиты распределенной информационно-телекоммуникационной сети Администрации города Омска. Результаты диссертационного исследования отмечены дипломом III-степени на XIX Международной студенческой конференции-школе-семинаре «Новые информационные технологии» (2011г.). Основные результаты работы внедрены и используются при преподавании дисциплин кафедры «Комплексная защита информации» в Омском государственном техническом университете. Представленные в диссертации модели и методы могут быть использованы в качестве базы для дальнейших исследований.

Апробация работы

Результаты работы прошли апробацию в виде выступлений на научных конференциях и семинарах:

1. Региональная молодежная научно-техническая конференция «Омское время – взгляд в будущее» (2010, г. Омск).
2. Международный информационный конгресс «МИК-2010» (2010, г. Омск).
3. III Всероссийская молодёжная научно-техническая конференция «Россия молодая: передовые технологии – в промышленность!» (2010, г. Омск).
4. III Международная научно-практическая конференция «Перспективы развития информационных технологии» (2011, г. Новосибирск).
5. XIX Международная студенческая конференция-школа-семинар «Новые информационные технологии» (2011, г. Судак, Украина).

Публикации

Результаты диссертации отражены в 9 публикациях, в том числе 2 публикации в изданиях, рекомендованных ВАК для публикации основных научных результатов диссертации.

Структура и объём работы

Диссертационная работа состоит из введения, трех глав, заключения, списка литературы и двух приложений. Общий объем работы составляет 123 страницы, в том числе 20 рисунков и 5 таблиц.

Личный вклад

Все исследования, изложенные в диссертационной работе, проведены автором в процессе научной деятельности. Все результаты, выносимые на защиту, получены автором лично. Из совместных публикаций включен лишь тот материал, который непосредственно принадлежит диссертанту, заимствованный материал обозначен в работе ссылками.

Основные результаты, выносимые на защиту

1. Итерационная процедура аппроксимации сети массового обслуживания с конечными очередями сетью Джексона для вычисления интенсивностей входящих в узлы потоков заявок.
2. Методика обнаружения низкоактивных распределенных атак типа «отказ в обслуживании» в компьютерных сетях на основе оценки вероятности потерь заявок.
3. Программно-аналитический комплекс, предназначенный для:
   - имитационного моделирования атак типа «отказ в обслуживании» и расчета статистических характеристик сетей массового обслуживания под воздействием атак;
   - обнаружения низкоактивных распределенных атак типа «отказ в обслуживании» в компьютерных сетях на основе оценки вероятности потерь заявок в сети.

СОДЕРЖАНИЕ РАБОТЫ

Во введении обосновывается актуальность темы, определяется цель и задачи исследования, излагается научная новизна и практическая значимость работы.

В первой главе представлен аналитический обзор проблемы обнаружения распределенных атак типа «отказ в обслуживании», приводится классификация существующих DDoS-атак, методов и средств их обнаружения, а также сформулирована постановка задачи исследований.

Для выполнения DDoS-атаки используются так называемые «ботнеты» или сети «ботов», т.е. компьютеров, зараженных определенной троянской программой, которыми нападающая сторона может управлять удаленно. В ходе подготовки к DDoS-атаке нападающий взламывает ряд хостов и устанавливает на них демона. Затем на такой демон посылают запрос на генерацию лавинообразного трафика пакетов того или иного вида в адрес целевого хоста. Обработка такого масштабного потока данных приводит к исчерпанию ресурсов хоста или маршрутизатора, на которые направлена атака, и приводит к недоступности сервиса или услуги, которую они предоставляли.

В работе было рассмотрено несколько различных подходов к обнаружению распределенных атак типа «отказ в обслуживании». Большинство существующих методов и систем обнаружения DDoS-атак позволяет эффективно распознавать и бороться с лавинообразными DDoS-атаками сетевого и транспортного уровня, направленными на заполнение пропускной способности каналов (Smurf, UDP-флуд и др.) и превышение нормальной загрузки отдельных узлов сети (SYN-флуд, Teardrop, Ping of death и др.).

Вместе с тем, существующие методы и средства обнаружения DDoS-атак малоэффективны для обнаружения классических DDoS-атак, протекающих в низкоактивном режиме (Low Rate DoS), а также современных DDoS-атак прикладного уровня, направленных против конкретных сетевых сервисов и служб. Отличить трафик, генерируемый в ходе данных атак, от легального прикладного трафика достаточно сложно, что делает затруднительным применение сигнатурного метода обнаружения атак. Кроме того, низкоактивные DDoS-атаки практически не приводят к образованию статистических аномалий, т.к. каналы передачи данных практически не перегружаются.

Данные факты обуславливают необходимость разработки специализированной методики обнаружения низкоактивных распределенных атак прикладного уровня типа «отказ в обслуживании» в компьютерных сетях. Разрабатываемая методика обнаружения строится на основе оценки вероятности потерь пакетов или запросов прикладного уровня в компьютерной сети, при условии ее функционирования в стационарном режиме. Достаточно часто для расчёта параметров потоков данных, а также для оценки

вероятностей потерь пакетов в вычислительных сетях применяют математические модели в виде сетей массового обслуживания (СеМО). Использование СеМО в качестве модели вычислительной сети действительно дает возможность проводить анализ работы сети со сложной структурой и разнообразными сетевыми сервисами.

В качестве отправной точки исследований было установлено соответствие между объектами реальных компьютерных сетей и систем и рассматриваемой математической моделью. В математической модели под узлом подразумевается некоторая сущность, которая способна обрабатывать и передавать дальше запросы (заявки). Указанная сущность на обработку каждой заявки тратит некоторые временные ресурсы. В каждом узле находится ровно одно устройство для обработки заявок. Кроме этого, узел содержит очередь ограниченной длины для поступающих заявок. Узлы также имеют связи с другими узлами. Формально можно считать, что каждый узел связан с каждым, учитывая, что вероятности использования этих связей в некоторых случаях равны нулю.

Каждому узлу математической модели соответствует исполняемый в операционной системе процесс (сетевая служба, сервис). Процесс имеет ограниченные временные ресурсы по процессорному времени для обработки заявок (ограниченное значение тактовой частоты процессора или ограничения, налагаемые планировщиком ОС). Процесс также имеет в своем распоряжении ограниченные ресурсы для промежуточного хранения поступающих заявок (оперативная память, число открытых сетевых соединений).

Под заявкой подразумевается некоторая сущность, которая может перемещаться по сети от одного узла к другому, возможно изменяя свое состояние. Путь заявки по узлам полностью определяется самими узлами. В математической модели все заявки равнозначны. В реальной системе такой заявке соответствует (с некоторым приближением) запрос/ответ прикладного уровня или пакет транспортного уровня соответствующего сетевого протокола.

Потерей заявки (из-за ограниченности ресурсов узлов) является как «физическая» потеря пакета из-за переполнения отведенной под хранение пакетов памяти, так и преждевременное завершение сеанса работы с протоколом из-за истечения тайм-аута для обработки. Очередь образуется вследствие последовательного характера обработки заявок узлом сети (процессом). Исполняемый процесс последовательно извлекает из области памяти (очереди) заявки и тратит некоторые временные ресурсы на их обработку. Область памяти, выделяемая узлу (процессу) ограничена, отсюда следуют ограниченность очереди и последовательность обработки. Конечный размер очереди обоснован как аппаратными ограничениями узлов, так тайм-аутами, предусмотренными сетевыми протоколами при обработке пакетов. Далее в терминах предложенной математической модели были описаны классифицированные ранее атаки на «отказ в обслуживании».

В случае атаки UDP-флуд нарушителем генерируется большое количество произвольных UDP пакетов на узлы атакуемой компьютерной сети. Заявкой является UDP-пакет. Атакуемые узлы помещают входящие заявки в очередь для их последующей обработки. Ввиду существующих ограничений по системным ресурсам (памяти), отведенным на хранение поступающих заявок, и максимальному количеству одновременных подключений, возможно переполнение очереди. В этом случае поступающим заявкам (в том числе от легальных пользователей) не удается встать в очередь и происходит потеря пакетов, т.е. возникает частичный или полный отказ в обслуживании.

Для атаки SYN-флуд нарушителем генерируется большое количество TCP пакетов с установленным флагом SYN на узлы атакуемой компьютерной сети. В качестве заявок в данном случае выступают TCP-пакеты с установленными управляющими флагами. Атакуемый узел принимает запрос на подключение, ожидая подтверждения от атакующего узла. В рамках рассматриваемой модели это соответствует случаю, когда пакет после обработки снова становится в очередь узла. Атакующий хост не отправляет пакет с установленным флагом ACK для подтверждения соединения, инициируя вместо этого новое соединение. Очередь узла переполняется заявками ввиду большого количества полуоткрытых соединений, т.е. происходит потеря заявок и возникает отказ в обслуживании легальным пользователям.

Атаки типа «отказ в обслуживании» прикладного уровня, например HTTP-флуд, построены на том, что нарушитель генерирует HTTP GET (либо HTTP POST) запросы на атакуемый узел компьютерной сети. Инициализируется соединение, на сервере возникает новый экземпляр потока в рамках процесса атакуемого сервиса. Открытым соединениям по тому или иному правилу планировщиком ОС выделяется процессорное время для обработки. В таких условиях (возникающих из-за ограниченности ресурсов) время между сеансами работы с заявкой может превысить предусмотренный протоколом тайм-аут, и соединение будет сброшено, т.е. может произойти отказ в обслуживании легальным пользователям. В качестве заявок для данной атаки выступают запросы HTTP GET (либо HTTP POST), очередь и время обработки ограничены и являются ненулевыми соответственно из-за характера работы сетевых служб.

Вторая глава диссертационной работы посвящена описанию математической модели разрабатываемой методики обнаружения низкоактивных DDoS-атак. Для распознавания атаки типа «отказ в обслуживании» оценивается вероятность потери произвольной заявки при ее прохождении по сети. Поскольку атаки прикладного уровня на различные сетевые службы происходят независимо, в рамках каждой службы для моделирования узлов СеМО можно использовать одноканальную систему массового обслуживания (СМО) с очередью длины m.

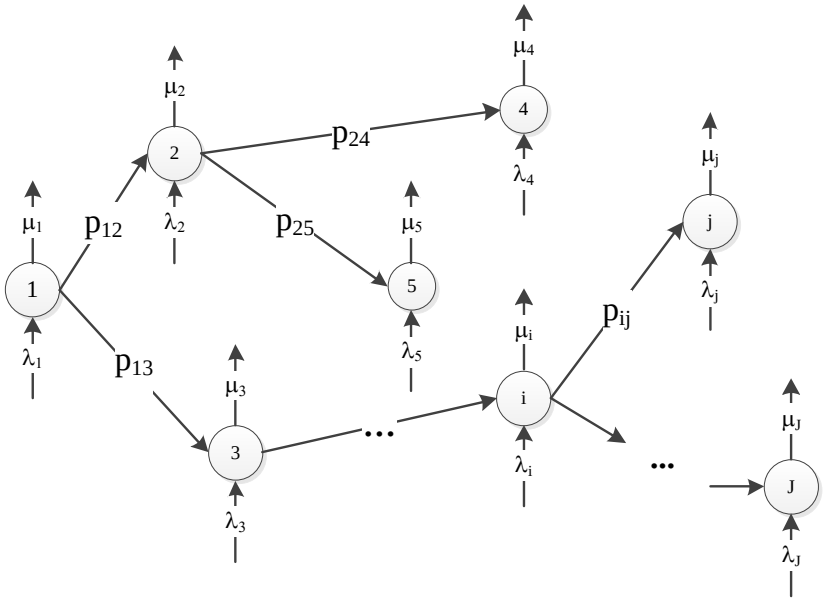

Рис. 1. Исходная сеть массового обслуживания.

Рассмотрим отдельный узел СеМО (рис. 1). Предположим, что вся сеть в целом и выбранный узел в частности функционирует в стационарном режиме, и что этот режим существует. Извне поступает пуассоновский поток заявок с параметром λ. Узел содержит одно устройство для обслуживания заявок, для которого задана интенсивность μ обработки заявок. После обработки заявка покидает узел. Если во время поступления заявки обслуживающее устройство занято обработкой другой заявки, то входящая заявка становится в очередь, максимальная длина очереди m. Если заявка поступает, и очередь полностью заполнена, заявка теряется.

Исходящий из узла поток успешно обработанных заявок является пуассоновским с параметром:

$$\lambda_R = \mu(1-p_0),\ p_0 = \frac{1}{\sum_{k=0}^{m}\left(\frac{\lambda}{\mu}\right)^k}.$$

Поток, управляющий потерями заявок в узле в стационарном режиме, также является пуассоновским с параметром:

$$\lambda_F = p_F\lambda,\ p_F = \frac{\left(\frac{\lambda}{\mu}\right)^m}{\sum_{k=0}^{m}\left(\frac{\lambda}{\mu}\right)^k}.$$

Здесь p_0 - вероятность того, что при функционировании в стационарном режиме очередь узла будет пуста, а p_F - вероятность того, что в стационарном режиме очередь узла будет полна, p_0 и p_F можно найти с помощью второй формулы Эрланга.

Далее рассматривается СеМО без потерь заявок (СеМО Джексона), состоящая из J узлов. Каждый узел содержит одно устройство для обработки заявок, время обработки одной заявки в узле j имеет экспоненциальное распределение с параметром μ_j. Значения μ_j вектора $\vec{\mu}$ положительны.

Очередь для поступающих в узел заявок не ограничена по длине. Также известна субстохастическая матрица маршрутизации P (сумма элементов по строке меньше или равна 1). Её элемент p_{ij} задает вероятность, что заявка, которая успешно завершила обслуживание в узле номер i, отправится в узел j. Матрица P неприводима. С матрицей маршрутизации связан ориентированный граф, в котором из узла i в узел j есть ребро только в случае, когда p_{ij} отлично от нуля (рис. 1).

Величина:

$$p_i^* = 1 - \sum_{k=1}^{J} p_{ik} .$$

задает вероятность, что после обработки в узле заявка покинет сеть. На вход узла сети j поступает пуассоновский поток заявок извне с параметром λ_j. Значения элементов λ_j вектора $\vec{\lambda}$ неотрицательны. Также вводится вектор $\vec{\rho}$.

Величины ρ_i - интенсивности суммарных входящих в узлы потоков. В общем случае суммарный поток не обязан быть пуассоновским, но при сделанных предположениях о сети он таковым является. Если матрица $(I - P)$ обратима (это ограничение эквивалентно требованию $\|P\| < 1$), то:

$$\vec{\rho} = \vec{\lambda}(I - P)^{-1} \qquad (1)$$

Далее была предложена *процедура аппроксимации исходной сети массового обслуживания с потерями заявок (с конечными очередями) сетью Джексона (с бесконечными очередями, без потерь заявок) для определения интенсивностей суммарных входящих в узлы исходной сети потоков заявок*.

Рассмотрим узел сети массового обслуживания j, не передающий заявок другим узлам. Этот узел заменяется узлом сети Джексона с аналогичными макро характеристиками в стационарном режиме (рис. 2).

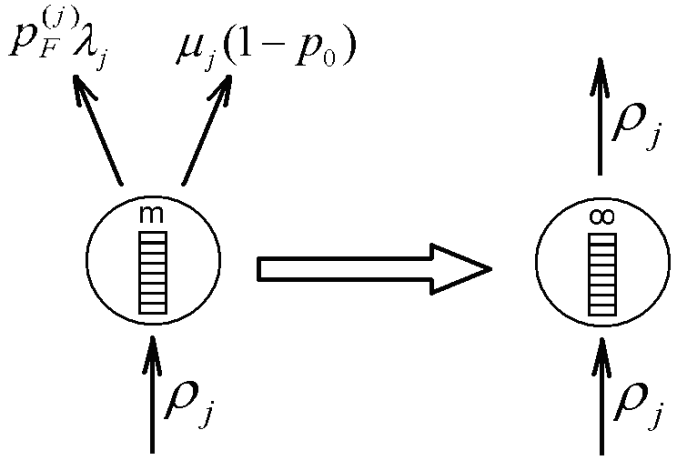

Рис. 2. Замена узла исходной СеМО узлом сети Джексона.

Т.е. принимается предположение, что успешно обработанные заявки объединяются в один поток с потерянными заявками и образуют общий поток успешно обработанных заявок в узле после замены.

Далее рассматривается узел исходной сети, в котором заявки не только завершают работу, но и передаются другим узлам. Вероятности перехода $\tilde{p}_{ji}$ у узла с неограниченной очередью после замены выбираются меньше (учитывается, что часть заявок выходного потока узла после замены на самом деле потеряны и не передаются в другие узлы), а вероятность успешной обработки $\tilde{p}_j^*$ – больше соответствующей вероятности p_j^*. Потенциальный путь поступившей в исходный узел заявки отображен на следующей схеме (рис. 3):

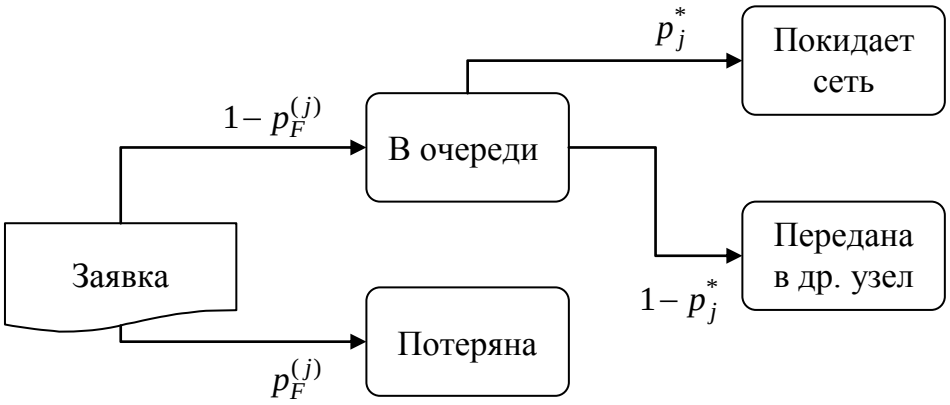

Рис. 3. Потенциальный путь поступившей в исходный узел заявки.

Далее необходимо рассмотреть вероятность $\tilde{p}_j^*$ «успешного» обработки в узле после замены (которое включает в себя и потерю заявки). Событие, отвечающее этой вероятности, состоит из потери заявки при заполненной очереди и успешной обработки и освобождения в случае постановки в очередь:

$$\tilde{p}^*{}_j = p_{j,F} + (1-p_{j,F})p_j^* = p_j^* + (1-p_j^*)p_{j,F}$$

Далее определяются вероятности переходов в другие узлы $\tilde{p}_{ji}$. Их можно рассчитать как:

$$\tilde{p}_{ji} = \beta_j p_{ji},$$

где коэффициент β_j можно найти из условия нормировки:

$$\tilde{p}_{j1} + \tilde{p}_{j2} + \cdots + \tilde{p}_{jJ} + \tilde{p}^* = 1$$

$$\beta_j(p_{j1} + p_{j2} + \cdots + p_{jJ}) + \tilde{p}^* = 1 \quad .$$

$$\beta_j = \frac{1-\tilde{p}^*}{p_{j1} + p_{j2} + \cdots + p_{jJ}} = \frac{1-\tilde{p}^*}{1-p_j^*}$$

Тогда:

$$\beta_j = 1 - p_{j,F}. \qquad (2)$$

После чего пересчитываются параметры стационарного режима, и происходит замена следующего узла.

В работе предложена обобщенная итерационная процедура построения скорректированной матрицы сети. Пусть задана субстохастическая матрица P, интенсивности входных потоков λ_j и обработки заявок в узлах μ_j. В результате итерационной процедуры строится матрица $\widetilde{P} = \vec{\beta} * P$ (данное обозначение соответствует операции, в ходе которой i-ая строка матрицы P умножается на i-ый элемент вектора $\vec{\beta}$). В качестве начального приближения выбирается $\vec{\beta}^{(0)} = (1,1,1,\cdots,1)$.

Шаг итерации:

1. Используя формулу (1) и матрицу $\widetilde{P} = \vec{\beta} * P$, рассчитываются текущие интенсивности входящих потоков в узлы:
$$\vec{\rho}^{(k)} = \vec{\lambda}\left(I - \vec{\beta}^{(k)} * P\right)^{-1}.$$
2. Рассчитываются элементы вектора вероятностей потерь $\vec{p}_F$:
$$p_{j,F}^{(k)} = \frac{\left(\frac{\rho_j^{(k)}}{\mu_j}\right)^{m_j}\left(1 - \frac{\rho_j^{(k)}}{\mu_j}\right)}{1 - \left(\frac{\rho_j^{(k)}}{\mu_j}\right)^{m_j}}.$$
3. Определяется результат итерации по формуле (2):
$$\beta_j^{(k+1)} = 1 - p_{j,F}^{(k)}.$$
4. Переход на следующую итерацию.

Итерации продолжаются, пока $\left\|\vec{\beta}^{(k+1)} - \vec{\beta}^{(k)}\right\| > \varepsilon$, где ε - заданная точность. В работе доказана сходимость предложенной итерационной процедуры.

Пусть теперь известен вектор интенсивностей входящих в узлы исходной сети потоков $\vec{\rho}$ (суммарные потоки извне и из других узлов), полученный с помощью расчета стационарного распределения соответствующей сети Джексона. Также задан вектор интенсивностей входящих потоков заявок извне $\vec{\lambda}$, вектор интенсивностей обработки заявок $\vec{\mu}$, и субстохастическая матрица маршрутизации после коррекции $\widetilde{P}$. В работе предложена *методика построения цепи Маркова с дискретным временем, соответствующая пути произвольной заявки по узлам, для оценки вероятности потерь заявок в сети*.

Для этого вводятся расщепленные состояния этой цепи, т.е. состоянием называется упорядоченная пара чисел $(i,d) = s$. Первое число соответствует номеру узла, в котором находится заявка (меняется в пределах от 1 до J), а второе – количеству занятых мест в очереди узла (меняется в пределах от 1 до

$m_i + 1$). Состояния вида $(i, m_i + 1)$, соответствуют переполненной очереди в узле i. Начальное распределение $\hat{p}$ можно рассчитать как:

$$\hat{p}\{s = (i,d)\} = p_{\lambda_i} \cdot p_{\pi,i,d-1}. \quad (3)$$

Здесь p_{λ_i} - вероятность того, что заявка пришла извне именно в узел i, принимаем ее равной:

$$p_{\lambda_i} = \frac{\lambda_i}{\sum_{j=1}^{J} \lambda_j}. \quad (4)$$

$p_{\pi,i,d}$ – вероятность того, что в очереди узла i находятся d заявок, эту вероятность можно рассчитать, используя формулу (2.6), заменяя λ_i на ρ_i:

$$p_{\pi,i,d} = \frac{\frac{\rho_i^{d-1}}{\mu_i^{d-1}}\left(1 - \frac{\rho_i}{\mu_i}\right)}{1 - \frac{\rho_i^{m_i}}{\mu_i^{m_i}}}. \quad (5)$$

Подставляя (5) и (4) в (3), можно получить выражение для начального распределения:

$$\hat{p}\{s = (i,d)\} = \frac{\lambda_i}{\sum_{j=1}^{J} \lambda_j} \cdot \frac{\frac{\rho_i^{d-1}}{\mu_i^{d-1}}\left(1 - \frac{\rho_i}{\mu_i}\right)}{1 - \frac{\rho_i^{m_i}}{\mu_i^{m_i}}}. \quad (6)$$

Кроме того, вводятся два дополнительных состояния (S) и (F). Первое соответствует успешной обработке заявки, а второе – потере заявки. Начальные вероятности этих состояний равны 0.

Далее необходимо определить вероятности переходов (матрица $\hat{P}$). Прежде всего, надо отметить, что состояния (S) и (F) не сообщаются. Цепь, попав в одно из этих состояний, уже из него не выходит. Вероятность перехода из состояния (i,d) в состояние (j,w) можно рассчитать следующим образом:

$$p\{(i,d) \to (j,w)\} = \tilde{p}_{ij}\, p_{\pi,j,w-1},$$

$$p\{(i,d) \to (j,w)\} = \tilde{p}_{ij} \frac{\frac{\rho_j^{w-1}}{\mu_j^{w-1}}\left(1 - \frac{\rho_j}{\mu_j}\right)}{1 - \frac{\rho_j^{m_j}}{\mu_j^{m_j}}}.$$

Вероятность успешной обработки заявки в узле равна:

$$p\{(i,d)\to(S)\}=\tilde{p}_i^{\,*}.$$

Если заявка находится в переполненной очереди, вероятность потери (перехода в состояние) (F) будет равна 1:

$$p\{(i,m_i+1)\to(F)\}=1.$$

Все остальные вероятности равны 0.

Если возвести матрицу $\hat{P}$ в степень k и умножить справа начальное распределение (6) на результат возведения в степень:

$$\hat{p}^{(k)}=\hat{p}\left(\hat{P}\right)^k,$$

то член вектора $\hat{p}^{(k)}\{(F)\}$, соответствующий состоянию (F), будет являться оценкой вероятности потери заявки при стационарном режиме работы сети.

В случае, когда сеть имеет произвольную структуру, установка параметра k происходит с помощью дополнительной итерационной процедуры. На первом шаге процедуры устанавливается некоторое начальное значение k, строится цепь Маркова и рассчитывается распределение ее состояний после k переходов заявки по узлам. По результатам расчёта вычисляется $\hat{p}_N^{(k)}$ - вероятность того, что заявка всё еще находится в сети, т.е. не потеряна и не обработана полностью:

$$\hat{p}_N^{(k)}=1-\hat{p}^{(k)}\{(F)\}-\hat{p}^{(k)}\{(S)\}.$$

Если в результате вычислений $\hat{p}_N^{(k)}$ превышает некоторую наперед заданную точность, k увеличивается, и расчет повторяется до тех пор, пока не будет достигнута заданная точность указанной вероятности.

В третьей главе работы представлена программная реализация системы и содержатся описания экспериментов. В ходе первого этапа исследований был разработан модуль программно-аналитического комплекса для моделирования рассматриваемой сети массового обслуживания с целью определения вероятности потерь заявок (рис. 4). На вход системы поступает субстохастическая матрица, задающая топологию сети, интенсивности входных потоков и интенсивности обработки заявок в узлах.

Моделирующий модуль программного комплекса содержит объекты двух видов - источники заявок и узлы сети. Каждый источник заявок связан с одним из узлов сети массового обслуживания. Источник заявок генерирует новые объекты заявок и добавляет их в очередь соответствующего узла. Начиная с момента запуска, узел проверяет наличие в очереди заявок. Если в очереди присутствует хотя бы одна заявка, она извлекается, после чего узел ожидает тайм-аут, имеющий экспоненциальное распределение, и далее случайным образом определяет заявку либо успешно обработанной, либо предпринимается попытка передать ее в очередь одного из узлов. Заявка теряется, если при попытке добавления ее в очередь узла эта очередь полностью заполнена.

В процессе обработки заявки сохраняется вся информация о ее маршруте по сети массового обслуживания, что позволяет по завершению процесса моделирования точно рассчитать наблюдаемую частоту потери заявок. Функционирование объектов модели СМО реализовано в параллельном режиме на основе библиотеки OpenMP, генерация псевдослучайных чисел осуществляется при помощи библиотеки Boost.

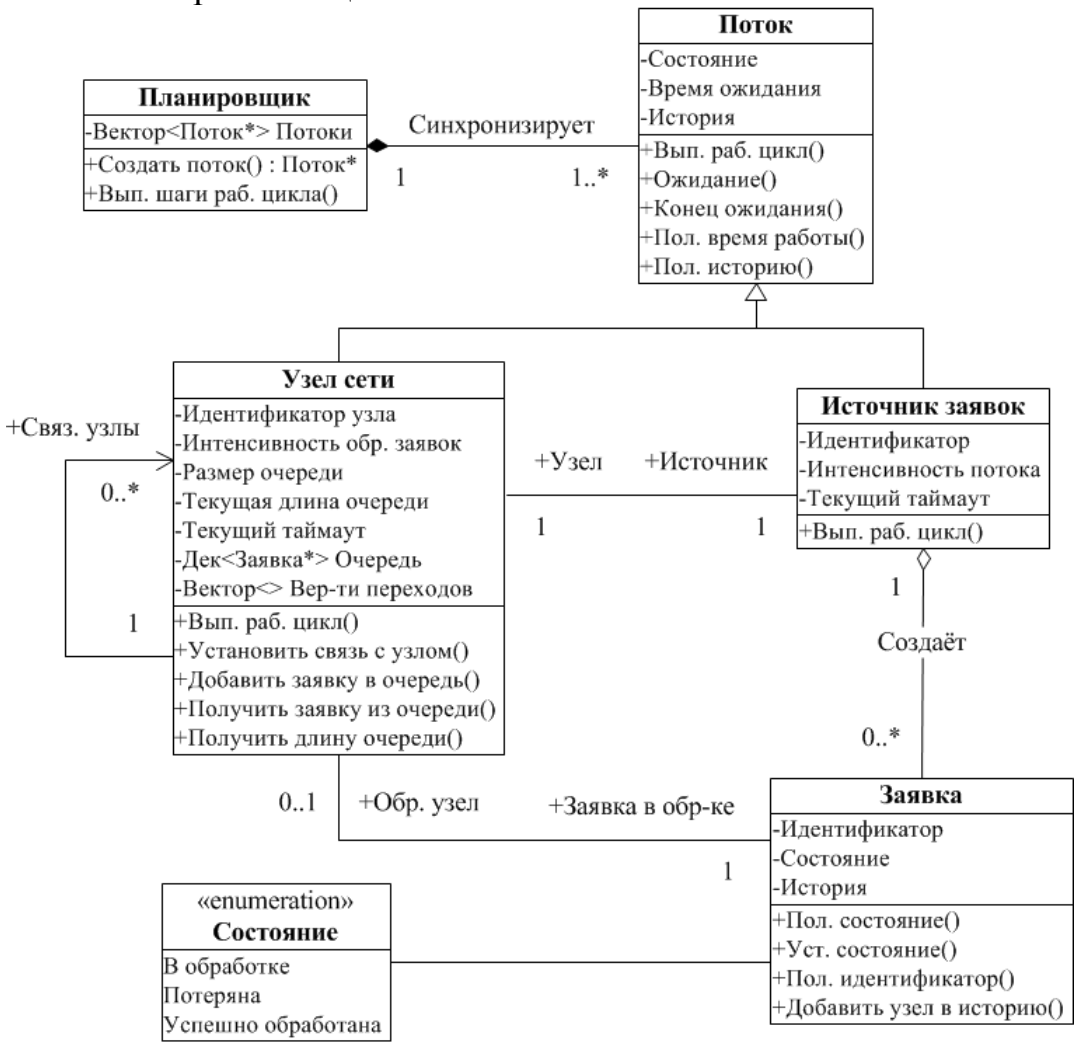

Рис. 4. UML-диаграмма классов моделирующего модуля.

Отдельно оценивалась теоретическая вероятность потери заявки, с помощью описанной выше итерационной процедуры и построения цепи Маркова, соответствующей пути заявки в системе. В работе представлены результаты серии экспериментов по моделированию потерь заявок в сетях массового обслуживания, функционирующих в стационарном режиме.

По итогам реализации программно-вычислительного комплекса для имитационного моделирования и расчёта характеристик сетей массового обслуживания была произведена серия вычислительных экспериментов для различных логических сетевых топологий. Ниже приводится описание эксперимента для топологии «произвольный ориентированный граф», исходные параметры СеМО, граф (рис. 5), определяемый матрицей переходов,

рассчитанные теоретические оценки параметров СеМО, а также практические результаты моделирования. В рамках каждого эксперимента производилось 10 сеансов имитационного моделирования сетевого взаимодействия.

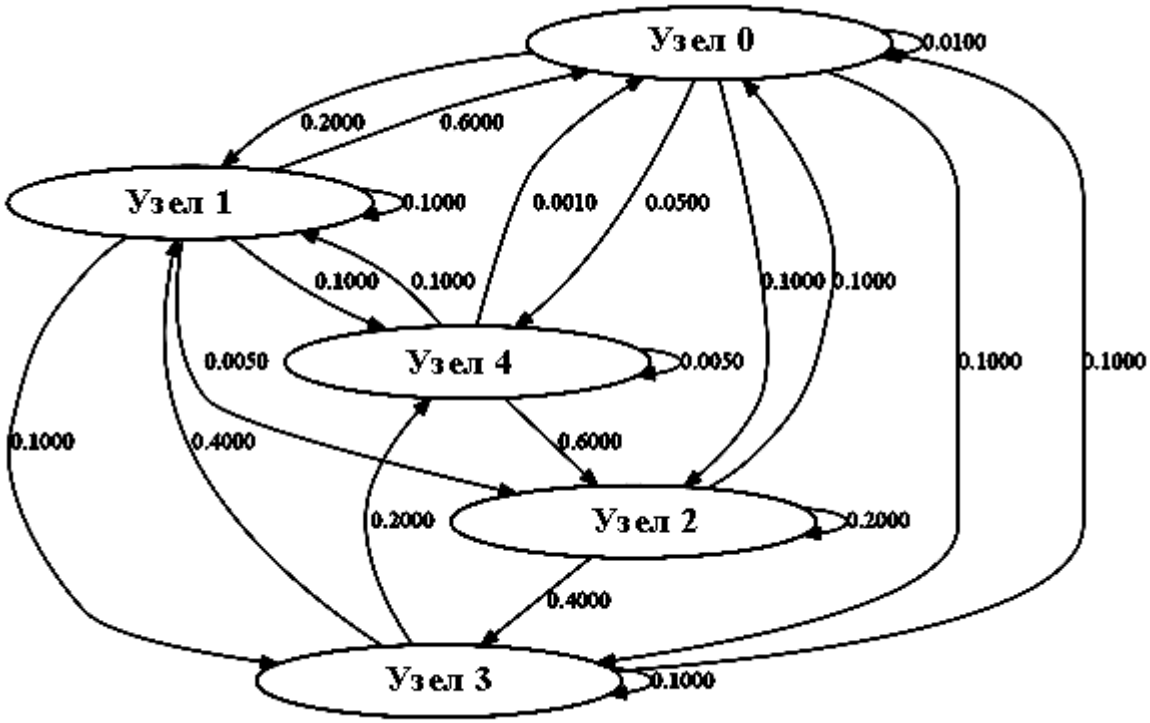

Рис. 5. Моделируемая сеть типа «произвольный ориентированный граф».

Матрица P и вектора $\vec{m}, \vec{\lambda}, \vec{\mu}$, соответствующие этой сети, имеют вид:

$$P = \begin{pmatrix} 0.01 & 0.2 & 0.1 & 0.1 & 0.05 \\ 0.6 & 0.1 & 0.005 & 0.1 & 0.1 \\ 0.1 & 0 & 0.2 & 0.4 & 0 \\ 0.1 & 0.4 & 0 & 0.1 & 0.2 \\ 0.001 & 0.1 & 0.6 & 0 & 0.005 \end{pmatrix}, \vec{m} = \begin{pmatrix} 5 \\ 4 \\ 8 \\ 9 \\ 5 \end{pmatrix}; \quad \vec{\lambda} = \begin{pmatrix} 2.8 \\ 0.9 \\ 2.1 \\ 1.0 \\ 0.5 \end{pmatrix}; \quad \vec{\mu} = \begin{pmatrix} 8.1 \\ 10.1 \\ 8.1 \\ 7.3 \\ 9.3 \end{pmatrix}.$$

В табл. 1 приведены статистические параметры полученной выборки. По результатам наблюдений можно отметить, что средняя ошибка рассчитанной оценки вероятности потерь заявок в СеМО не превышает среднеквадратичное отклонение наблюдаемой частоты потерь в серии экспериментов (табл. 1).

Количество сеансов моделирования	q	10
Ср. знач. наблюдаемой частоты потерь заявок	$p_{ср.}$	9,42E-2
Среднеквадратичное отклонение частоты потерь	$\sigma(p)$	1,08E-3
Рассчитанная оценка вероятности потерь заявок	$p_{рассч.}$	9,38E-2
Средняя ошибка	$\Delta_{ср.}(p)$	4,0E-4

Таблица 1. Результаты имитационного моделирования заданной сети.

На рис. 6 представлен график зависимости теоретической оценки вероятности потери произвольной заявки в СеМО от количества шагов предложенной итерационной процедуры. Кроме того, на графике отмечена наблюдаемая частота потерь заявок в моделируемой сети. Представленный график позволяет сделать вывод о необходимом количестве итераций для обеспечения заданной точности.

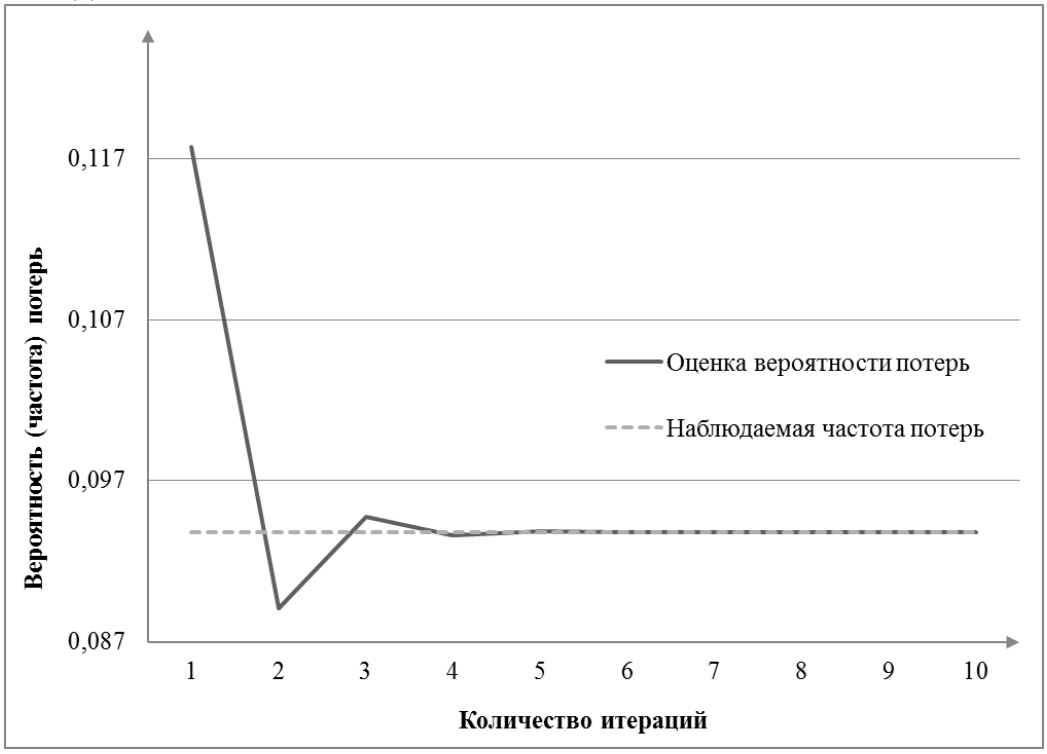

Рис. 6. Зависимость теоретической оценки вероятности потери заявки от количества шагов итерационной процедуры

Разработанная методика позволяет получать адекватную оценку частоты потери заявок в сети в случае, если сеть массового обслуживания работает в стационарном режиме. Во время возникновения DDoS-атаки один или несколько узлов СеМО выходят из стационарного режима на некоторое время, после чего устанавливается стационарный режим с другими параметрами. На время перехода между режимами рассматриваемая методика неприменима. Таким образом, оценка времени перехода между режимами имеет определяющее значение. Очевидно, что время перехода сильно зависит от топологии сети и параметров узлов.

В работе был произведен вычислительный эксперимент для оценки указанного параметра. В течение сеанса моделирования СеМО, спустя некоторое время после запуска один из узлов производил атаку типа «отказ в обслуживании» на заполнение пропускной способности канала. Общий интервал времени моделирования был разбит на 30 равных интервалов, для каждого из которых оценивалась теоретическая вероятность потери заявок и наблюдаемая частота потерь заявок. При максимальном размере заявки (пакета) 8760 байт (размер окна, максимальное количество данных, которое может быть

отправлено по протоколу TCP/IP без подтверждения) и пропускной способности каналов 100 Мбит/сек, время перехода между стационарными режимами варьируется от 10 минут (при средней загруженности сети на уровне 30%) до 35 минут (при средней загруженности сети на уровне 10%).

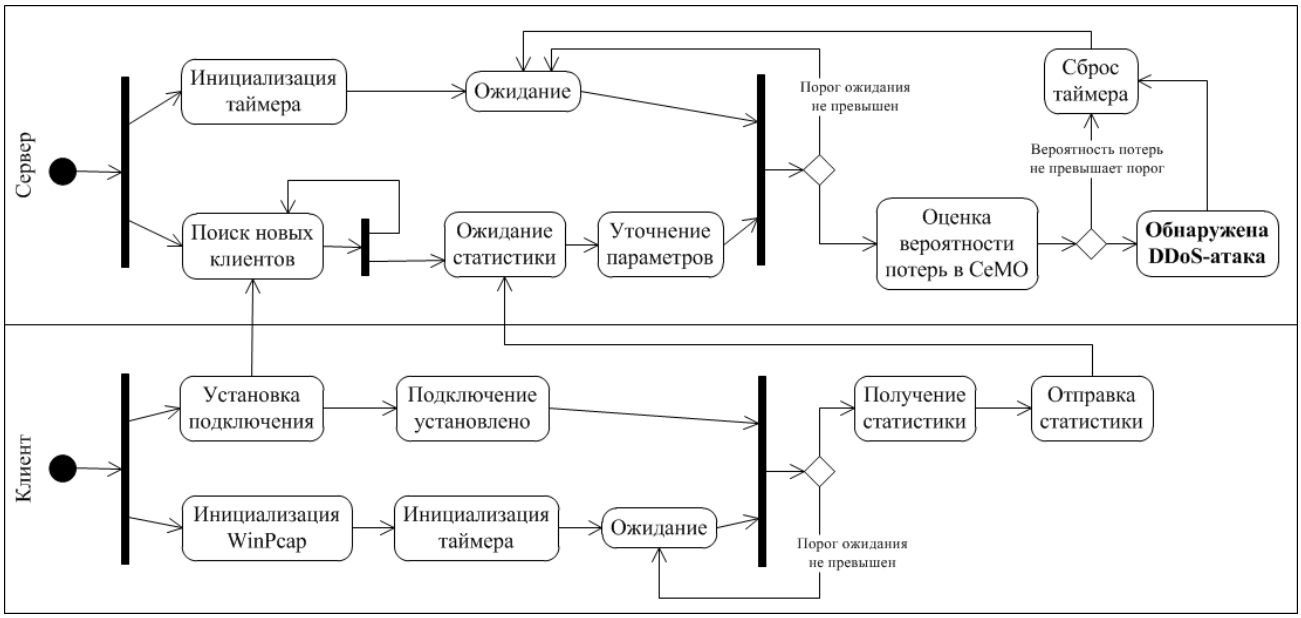

Рис. 7. UML-диаграмма деятельности комплекса обнаружения низкоактивных атак типа «отказ в обслуживании».

В ходе второго этапа исследований на основе представленной методики был разработан программно-аналитический комплекс, предназначенный для обнаружения низкоактивных атак типа «отказ в обслуживании» в распределенных компьютерных сетях (рис. 7).

Клиентский модуль программно-аналитического комплекса при запуске выполняет следующие действия:

1. Получает список доступных в системе сетевых адаптеров, и информацию о них.
2. На каждом из адаптеров запускается процесс захвата пакетов. Для этого фильтр пакетов конфигурируется таким образом, чтобы перехватывать фреймы Ethernet и протокола TCP с портами, соответствующими распространенным сетевым службам (NetBIOS, HTTP, FTP, SMTP, и т. д.)
3. Начинается отсчет времени. Когда время, прошедшее с начала захвата, превышает заданное значение $t_{\max}$, захват трафика останавливается. Подсчитывается объем исходящего и входящего трафика, переданного другим узлам распределенной сети по контролируемым портам. Соответствующая информация передается на серверный модуль. Далее, начиная со второго шага, процесс повторяется. Следует отметить, что значение временного периода $t_{\max}$, за который

оценивается статистика работы сетевых приложений узла, может быть различным для разных узлов сети.

Серверная часть программно-аналитического комплекса при запуске выполняет следующие действия:

1. Получает список доступных в системе сетевых адаптеров, и информацию о них.
2. На каждом из адаптеров запускается процесс захвата пакетов. Для этого фильтр пакетов конфигурируется таким образом, чтобы перехватывать фреймы Ethernet, а вернее пакеты протокола TCP с портами, соответствующими распространенным сетевым службам (NetBIOS, HTTP, FTP, SMTP, и т. д.)
3. Открывается порт для входящих соединений от клиентских модулей программно-аналитического комплекса.
4. В случае входящего соединения от клиента производятся следующие действия:
    a. Принимается информация о получателе трафика (пара адрес: порт) и количестве трафика, а также времени, за который эта информация была собрана.
    b. Для каждой пары адрес: порт в сообщении от клиента уточняется полученное ранее среднее значение объема трафика, переданного текущим сокетом другим узлам по формуле:

$$Av_{k_1,k_2,t+t_{\max}} = \frac{Av_{k_1,k_2,t} \cdot t + V_{k_1k_2,t_{\max}}}{t + t_{\max}}$$

Здесь:

$Av_{k_1,k_2,t+t_{\max}}$ – среднее значение объем трафика от сокета k_1 сокету k_2 к моменту времени $t + t_{\max}$

t – время, прошедшее с начала наблюдений до предыдущего получения информации от клиентской части

$t_{\max}$ – время, за которое клиентская часть собирает данные о трафике сокета

$Av_{k_1,k_2,t}$ – средний объем трафика от сокета k_1 сокету k_2 за время, прошедшее с начала наблюдений до предыдущего получения информации от клиентской части

$V_{k_1k_2,t_{\max}}$ – объем трафика от сокета k_1 сокету k_2, прошедший за время $t_{\max}$ (последний период наблюдения клиентом)

    c. используя информацию о среднем значении объёма трафика сокета k_1 сокету k_2 и суммарному входящему трафику в сокет k_1, оцениваются параметры субстохастической матрицы, задающей топологию сети.

5. Когда истекло время t_{max} и захват трафика останавливается, подводятся итоги по объему трафика, отправленного текущим узлом другим узлам и характеристикам анализируемой сети массового обслуживания:
    a. Используя субстохастическую матрицу, строится модель анализируемой сети массового обслуживания, и оцениваются параметры стационарного режима.
    b. Используя параметры стационарного режима, строится цепь Маркова, соответствующая пути заявки (пакета) по СеМО.
    c. Рассчитывается эволюция цепи Маркова и оценивается вероятность потерь заявок.
    d. Далее, начиная с четвертого шага, процесс повторяется.

Если в некоторый момент времени оценочная вероятность потерь превысила некоторое, наперёд заданное администратором значение порога, система принимает вывод о реализации в распределенной компьютерной сети атаки типа «отказ в обслуживании».

Оценка эффективности разработанной методики обнаружения DDoS-атак в компьютерных сетях и её сравнительный анализ с другими подходами, методами и системами представляет сложную задачу. Основные трудности оценки и анализа заключаются в следующем:

- Эффективность обнаружения атак типа «отказ в обслуживании» напрямую зависит от параметров работы сети в штатном режиме (загрузка сети, среднее значение потерь пакетов/запросов), при этом воспроизведение параметров работы для двух различных экспериментов не всегда является возможным.
- Разрабатываемые методы и системы обладают рядом настраиваемых индивидуальных параметров (точность вычислений, значение порога принятия решения об обнаружении атаки), существенно влияющих на эффективность обнаружения атак и количество ложных срабатываний.
- Кроме того, для каждой разновидности DDoS-атаки существует множество различных модификаций и параметров (интенсивность атаки, уникальные идентификаторы), которые также непосредственно влияют на эффективность обнаружения атаки.
- Все вышеперечисленные факторы обуславливают отсутствие единого стандарта экспериментальных условий для оценки эффективности систем и методов обнаружения атак типа «отказ в обслуживании».

В диссертационной работе для экспериментальной оценки эффективности разрабатываемой системы был использован подход, предложенный в работе исследователей F.-Y. Leu и I-L. Lin, посвященной разработке системы обнаружения DDoS-атак на основе статистического критерия χ^2.

В ходе поставленного эксперимента ботнет, образованный группой из 4 узлов нарушителей в одной подсети, производит атаки типа «отказ в

обслуживании» (SYN-флуд, ICMP-флуд, UDP-флуд) на узел-жертву в другой подсети (рис. 8). В задачи эксперимента входила оценка точности обнаружения атак типа «отказ в обслуживании» с помощью разработанной методики.

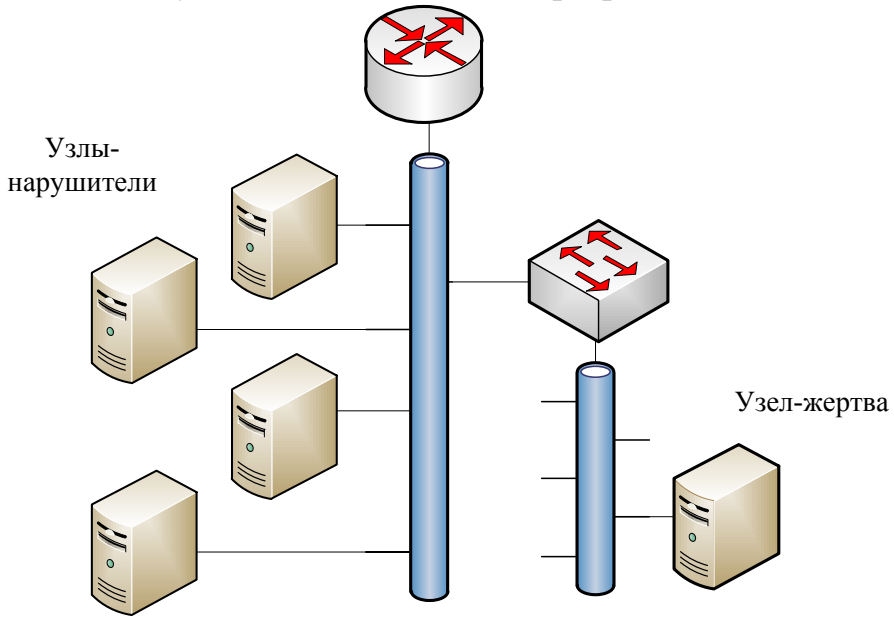

Рис. 8. Топология сети для проведения экспериментальных исследований.

Для решения поставленной задачи ровно 1000 раз в течение заданного временного интервала (10 сек.) происходило накопление передаваемых пакетов в штатном режиме работы сети, после чего ровно 1000 раз происходило накопление передаваемых пакетов на узле под воздействием различных атак типа «отказ в обслуживании». Интенсивность атак варьировалась случайным образом в интервале от 500 пакетов/сек. до 15000 пакетов/сек. По завершении каждого временного интервала на основе накопленных пакетов разработанный программно-аналитический комплекс производил построение моделирующей CeMO и детектирование атаки типа «отказ в обслуживании».

В результате обозначенного эксперимента для разработанной методики обнаружения DDoS-атак были получены следующие результаты для ошибок первого рода (количество ложных срабатываний, 11,6%) и ошибок второго рода (количество необнаруженных атак, 3,4%).

Система	Ошибка первого рода, %	Ошибка второго рода, %
Kaspersky Anti-Hacker 1.8.180	0	10,6
Snort	4,8	10,5
AIDS	3,5	7,24

Табл. 2. Значения ошибок обнаружения DDoS-атак первого рода и второго рода для различных систем.

Данные результаты позволяют сравнить разработанную методику с исследованными ранее в указанной работе системами обнаружения (табл. 2).

Таким образом, разработанная методика демонстрирует более высокий процент обнаружения атак типа «отказ в обслуживании» по сравнению с уже исследованными системами, что представляет наибольшую практическую важность. Вместе с тем процент ложных срабатываний для разрабатываемой методики также максимален, что в определенной степени является её недостатком.

ОСНОВНЫЕ РЕЗУЛЬТАТЫ РАБОТЫ

В результате проведенных исследований получены следующие основные научные и практические результаты:

1. Произведен анализ и классификация существующих атак типа «отказ в обслуживании», обобщены и систематизированы основные подходы к обнаружению атак данного типа в распределенных компьютерных сетях. Указаны их ограничения.
2. Разработана итерационная процедура аппроксимации сети массового обслуживания с конечными очередями сетью Джексона для вычисления интенсивностей входящих в узлы потоков заявок (для различных топологий исходной сети).
3. Предложена методика построения цепи Маркова на основе характеристик сети массового обслуживания для оценки вероятности потерь заявок в сети.
4. Разработана методика обнаружения низкоактивных атак типа «отказ в обслуживании» на основе их моделирования в сетях массового обслуживания и оценки вероятности потерь в стационарном режиме функционирования сети.
5. Разработана архитектура и реализован программно-аналитический комплекс, предназначенный для:
   - имитационного моделирования атак типа «отказ в обслуживании» и расчета статистических характеристик сетей массового обслуживания под воздействием атак;
   - обнаружения низкоактивных распределенных атак типа «отказ в обслуживании» в компьютерных сетях на основе оценки вероятности потерь заявок в сети.
6. Произведена экспериментальная оценка эффективности и сравнительный анализ разработанного программно-аналитического комплекса для обнаружения атак типа «отказ в обслуживании» в компьютерных сетях.

ПУБЛИКАЦИИ ПО ТЕМЕ ДИССЕРТАЦИИ

1. Щерба М.В. Методика разработки системы защиты информации комплекса муниципальных информационных систем / М.В. Щерба //

Информационные технологии моделирования и управления. – 2009. – Выпуск 6(58). – С. 850-854.
2. Щерба М.В. Некоторые аспекты построения системы защиты информации комплекса муниципальных информационных систем / М.В. Щерба // Омское время - взгляд в будущее: матер. регион. молодеж. науч.-техн. конф. – Омск: Изд-во ОмГТУ, 2010. – Кн. 1. – С. 294-296.
3. Щерба М.В. Электронная информационная система для определения реабилитационного потенциала и прогноза у детей-инвалидов / Ю.В. Наумова, Е.В. Щерба, М.В. Щерба // Россия молодая: передовые технологии – в промышленность: матер. III Всерос. молодежн. науч.-техн. конф. – Омск: Изд-во ОмГТУ, 2010. – Кн. I. – С. 283-286.
4. Щерба М.В. Методика разработки системы защиты информации комплекса муниципальных информационных систем / М.В. Щерба, О.Т. Данилова // Международный и региональный опыт построения информационного общества: сборник материалов Международного информационного конгресса «МИК-2010». – Омск: Правительство Омской области, 2011. – Ч. 1. – С. 403-406.
5. **Щерба М.В. Анализ комплексного подхода к защите информации при её передаче в распределенных беспроводных сетях / В.И. Никонов, Е.В. Щерба, М.В. Щерба // Омский научный вестник. Серия «Приборы, машины и технологии». - 2011. - №2(100). - С. 193-197.**
6. Щерба М.В. Методика оценки надёжности и защищенности распределенных компьютерных сетей / М.В. Щерба // Перспективы развития информационных технологий: сборник материалов III Международной научно-практической конференции. – Новосибирск: Издательство НГТУ, 2011. – Ч. 1. – С. 229-233.
7. Щерба М.В. Методика оценки надёжности и защищенности распределенных компьютерных сетей / М.В. Щерба // Информационные системы и технологии: сборник материалов Международной научно-технической Интернет-конференции, г. Орёл, апрель-май 2011. – Орёл: ФГОУ ВПО «Госуниверситет-УНПК», 2011. – Т. 1. – С. 193-196.
8. Щерба М.В. Разработка системы анализа надёжности и защищенности распределенных компьютерных сетей / М.В. Щерба // Новые информационные технологии: тезисы докладов XIX Международной студенческой конференции-школы-семинара. – Москва: МИЭМ, 2011. – С. 274-275.
9. **Щерба М.В. Система анализа устойчивости распределенных компьютерных сетей к атакам на «отказ в обслуживании» / М.В. Щерба // Омский научный вестник. Серия «Приборы, машины и технологии». - 2012. - №1(106). - С. 282-286.**