На правах рукописи

Ботвинко Анатолий Юрьевич

РАЗРАБОТКА МОДЕЛИ ДЛЯ АНАЛИЗА ПОКАЗАТЕЛЕЙ ЭФФЕКТИВНОСТИ МЕЖСЕТЕВОГО ЭКРАНА С РАНЖИРОВАНИЕМ ПРАВИЛ ФИЛЬТРАЦИИ

Специальность 05.13.17 - теоретические основы информатики

АВТОРЕФЕРАТ

на соискание ученой степени кандидата
физико-математических наук

Москва – 2021

Работа выполнена на кафедре прикладной информатики и теории вероятностей Российского университета дружбы народов.

Научный руководитель: доктор технических наук, профессор, заведующий кафедрой прикладной информатики и теории вероятностей Российского университета дружбы народов (РУДН)

Самуйлов Константин Евгеньевич

Официальные оппоненты: доктор технических наук, член-корреспондент Российской академии наук, главный научный сотрудник – заведующий лабораторией мультимедийных систем и технологий Московского физико-технического института
Дворкович Александр Викторович

доктор технических наук, профессор, заведующий кафедрой сети связи и системы коммутации Ордена Трудового Красного Знамени федерального государственного бюджетного образовательного учреждения высшего образования «Московский технический университет связи и информатики»
Степанов Сергей Николаевич

доктор технических наук, профессор, заведующий кафедрой сетей связи и передачи данных Санкт-Петербургского государственного университета телекоммуникаций им. проф. М.А. Бонч-Бруевича (СПБГУТ)
Кучерявый Андрей Евгеньевич

Защита диссертации состоится «03» декабря 2021 г. в 16 час. 30 мин. на заседании диссертационного совета ПДС 0200.001 на базе Российского университета дружбы народов (117198, г. Москва, ул. Миклухо-Маклая, д.6).

С диссертацией можно ознакомиться в научной библиотеке Российского университета дружбы народов по адресу: 117198, Москва, ул. Миклухо-Маклая, дом 6 (отзывы на автореферат просьба направлять по указанному адресу) или на официальном сайте диссертационных советов РУДН по адресу: http://dissovet.rudn.ru

Автореферат разослан «____» октября 2021 г.

Ученый секретарь диссертационного совета
ПДС 0200.001, доцент  А.В. Демидова

Общая характеристика работы

Актуальность темы исследования

Устойчивое функционирование информационной инфраструктуры, в том числе для автоматизированных систем (АС) специального назначения, бесперебойное функционирование которых критически важно для обеспечения безопасности и обороноспособности государства в условиях лавинообразного роста объема информационных потоков сетей общего пользования, высокой неоднородности и изменчивости характеристик сетевого трафика, широкого использования мультимедийных протоколов, чувствительных к длительности задержки передачи данных, а также значительного увеличения количества компьютерных атак, требует обеспечения высокой производительности межсетевых экранов (МЭ).

Одним из основных факторов, влияющих на время поиска правил фильтрации, а значит и на производительность МЭ, служит порядок расположения правил в наборах, представляющих собой линейные списки большой размерности. Это связано с тем, что время поиска правила, соответствующего фильтруемым данным, пропорционально количеству проверенных правил, а время фильтрации информационного потока, удовлетворяющего условиям, содержащимися в конце набора большой размерности, будет значительно больше времени необходимого для фильтрации данных, удовлетворяющих условиям, содержащимся в начале набора правил.

В качестве методов оптимизации набора правил фильтрации большинство авторов, рассматривают методы, предполагающие статическую оптимизацию правил или использующие специализированные структуры данных. Недостатком данных методов является слабая универсальность. Кроме того, они плохо приспособлены для работы с распределенными и меняющимися во времени информационными потоками.

В диссертационной работе разработан метод оптимизации набора правил фильтрации (метод ранжирования правил), учитывающий изменение характеристик информационных потоков. Повышение эффективности фильтрации трафика достигается периодическим ранжированием правил фильтрации в порядке убывания их весов, полученных в соответствии с характеристиками фильтруемых информационных потоков.

Особенностью разработанного подхода является использование непараметрического метода локальной аппроксимации (МЛА) при оценивании характеристик фильтруемых информационных потоков. В процессе ранжирования набора правил учитываются текущие характеристики и динамика изменений характеристик информационных потоков. При этом отсутствует необходимость подбора параметрической модели, приемлемой для всех оцениваемых значений характеристик информационных потоков.

Применение МЛА обеспечило адаптивность метода, а также высокую скорость реакции на изменение характеристик фильтруемых информационных потоков за счет специфики построения оценок МЛА:

– использование локально-параметрической модели со скользящий областью постоянства параметров и управляемым параметром локальности, определяющим размер области локальности;

– использование специальной функции локальности для задания ценности предыдущих значений при вычислении оценок характеристик фильтруемых информационных потоков.

Таким образом, актуальной задачей является разработка метода ранжирования набора правил фильтрации, обеспечивающего повышение эффективности фильтрации информации, ее поддержание на уровне достаточном для обеспечения устойчивой работоспособности АС, а также построение математических моделей для вычисления показателей эффективности МЭ.

Степень разработанности темы

Теоретические и прикладные основы исследования опираются на фундаментальные труды в области теории массового обслуживания, теории телетрафика, теории идентификации и оценивания, теории нелинейного программирования таких ученых, как Г. П. Башарин, П. П. Бочаров, В. М. Вишневский, Ю. В. Гайдамака, А. В. Дворкович, В. Я. Катковник, Е. А. Крук, А. Е. Кучерявый, Е. А. Кучерявый, С. П. Моисеева, Д. А. Молчанов, А. А. Назаров, Ю. Н. Орлов, А. П. Пшеничников, К. Е. Самуйлов, С. Н. Степанов, И. И. Цитович, F. Baskett, E. Gelenbe, W. Hardle, F. P. Kelly, L. Kleinrock и ряда других исследователей.

Исследованиям в области повышения эффективности фильтрации информационных потоков посвящены работы ученых и специалистов различных стран: Н. Ф. Бахарева, К. В. Иванов, В. Н. Тарасов, П. И. Тутубалин, H. B. Acharya, E. Al-Shaer, J. M. A. Calero, Q. Duan, B. Feng, K. Ghoudi, M. Huang, G. Marchetto, M. M. Masud, P. Nanda, B. J. Oommen, K. Ramli, A. Shameli-Sendi, Z. Trabelsi, A. Yazidi, S. Zeidan, L. Zhang.

Цели и задачи исследования

Целью исследования является повышение производительности МЭ за счет использования разработанного метода ранжирования набора правил фильтрации.

Для достижения цели в диссертационной работе решаются **следующие задачи**.

1. Анализ влияния МЭ на время установления сессии по протоколу установления сессий SIP (англ. Session Initiation Protocol).

2. Разработка модели системы массового обслуживания (СМО) для анализа производительности МЭ с учетом с порядка правил фильтрации.

3. Разработка метода ранжирования набора с применением метода локальной аппроксимации при фильтрации информационных потоков. Разработка имитационной модели и проведение численного эксперимента для анализа показателей эффективности МЭ.

Научная новизна работы

1. Построена математическая модель установления сессии по протоколу SIP между двумя пользователями. В отличие от ранее известных моделей в процедуру установления сессии по протоколу SIP внесена задержка фильтрации информационных потоков, проходящих через МЭ.

2. Процесс обслуживания заявок в разработанных моделях МЭ описан с помощью функции распределения времени обслуживания заявки фазового типа, зависящей от порядка правил фильтрации, что ранее не применялось в известных моделях.

3. Для ранжирования набора правил фильтрации МЭ применен МЛА. В отличие от известных ранее, метод ранжирования использует непараметрическую модель для вычисления весов правил фильтрации.

Теоретическая и практическая значимость работы

Разработанный метод ранжирования правил фильтрации может быть использован для решения задач в области обеспечения сетевой безопасности в целях повышения производительности МЭ. Аналитическая и имитационная модель может быть применена при расчете показателей эффективности МЭ для принятия решения о необходимости ранжирования текущего набора правил фильтрации.

Результаты работы получили практическую реализацию в области организационно-технических мероприятий по обеспечению надежности ракетно-космических систем и изделий ракетно-космической техники при выпуске национального стандарта Российской Федерации ограниченного распространения ГОСТ РО 1410-001-2020.

Методология и методы исследования

В диссертации применены методы теории вероятностей, теории марковских случайных процессов, теории массового обслуживания, теории телетрафика, теории идентификации и оценивания, теории нелинейного программирования.

Положения, выносимые на защиту

1. Модель установления сессии по протоколу SIP в виде неоднородной сети массового обслуживания (СеМО) с учетом передаваемых данных через МЭ. Метод расчета и численный анализ влияния МЭ на время установления сессии.

2. Модель СМО с обслуживанием фазового типа, учитывающая порядок правил фильтрации в наборе для анализа показателей эффективности МЭ. Распределение фазового типа учитывает порядок правил фильтрации в наборе.

3. Имитационная модель и анализ показателей эффективности МЭ с учетом ранжирования правил и фильтрации МЭ информационных потоков.

Степень достоверности и апробация результатов

Степень достоверности основных результатов, полученных в диссертации, подтверждается их апробацией на научных конференциях и семинарах:

– VIII и IX международной отраслевой научно-технической конференции «Технологии информационного общества», Москва, МТУСИ, 2014, 2015.

– VI и VII всероссийской конференции «Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем», Москва, РУДН, 2016, 2017.

– XI международной отраслевой научно-технической конференции «Технологии информационного общества», Москва, Федеральное агентство связи, 2017.

– III международной научно-техническая конференции студентов, аспирантов и молодых ученых «Интернет вещей и 5G, INTHITEN 2017», Санкт-Петербург, 2017.

– XXI и XXIII международной конференции «Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь» (DCCN-2018, DCCN-2020), Москва, РУДН, 2018, 2020.

– Доклад на семинаре кафедры прикладной информатики и теории вероятностей Российского университета дружбы народов.

Степень достоверности также подтверждается научными публикациями. Основные результаты по теме диссертационного исследования изложены в 15 печатных работах, из которых [1-5] – в рецензируемых изданиях, рекомендованных ВАК РФ, издания [6-7] в зарубежных научных изданиях, входящих в Scopus и/или Springer, [8-15] – в трудах международных и всероссийских научных конференций.

Диссертация состоит из введения, трех глав, заключения, библиографии из 129 наименований на русском и английском языках. Научная работа изложена на 132 страницах текста, содержит 65 рисунков и 13 таблиц.

Основное содержание работы

Во введении обоснована актуальность темы диссертации, определены цели и задачи исследований, сформулирована теоретическая и практическая значимость работы, представлены основные положения работы, выносимые на защиту.

В первой главе исследованы теоретические основы и определения, относящиеся к межсетевому экранированию, описаны наиболее типичные проблемы построения набора правил фильтрации МЭ. Получена оценка влияния МЭ на длительность задержки передачи данных в АС с мультимедийными протоколами передачи данных. Оценка вычислена для чувствительной к временным задержкам услуге установления сессии между двумя абонентами по протоколу SIP с транспортным протоколом UDP (англ. User Datagram Protocol) без потерь и искажения пакетов.

Для оценки влияния МЭ на время установления сессии между двумя абонентами предложена математическая модель в виде открытой экспоненциальной СеМО. Показано, что время пребывания в СеМО равно времени установления сессии. Последовательная оценка времени ожидания и времени обслуживания для каждого функционального блока позволила, выполнив суммирование, получить время установления сессии.

СеМО состоит из шести узлов, каждый из которых является СМО, моделирующей работу соответствующего функционального блока в процедуре установления сессии. Узлы $i = 1,4,6$, соответствующие оборудованию пользователей и сети IP/MPLS, представлены с помощью многолинейных СМО $M|M|\infty$ без накопителя с числом приборов большим максимального количества заявок, поступающих в узлы, и дисциплиной обслуживания без ожиданий. Оставшиеся узлы $i = 2,3,5$ заданы в виде однолинейных СМО $M|M|1|\infty$ с накопителем неограниченной емкости, дисциплиной обслуживания в порядке поступления и экспоненциальной функцией распределения длительности обслуживания заявок. Функция распределения длительности обслуживания не зависит от класса заявок и имеет экспоненциальное распределение

$B_i(t) = 1 - e^{\mu_i t}, t \geq 0, i = 1,2,...,6$ с интенсивностью μ_i. Поток сообщений SIP в СеМО является неоднородным пуассоновским потоком с суммарной интенсивностью Λ, независящим от состояния сети и представляющим собой сумму четырех пуассоновских потоков сообщений с интенсивностью λ_0.

Использовав формулы для вычисления среднего времени пребывания в системах $M|M|1|\infty$, $M|M|\infty$, найдена средняя задержка запроса на установление сессии T_{SRD} и среднее время установления сессии T_S.

Утверждение 1 Средняя задержка запроса на установление сессии и среднее время установления сессии с учетом фильтрации сигнальных сообщений в МЭ вычисляются по формулам (1) и (2).

$$T_{SRD} = 2\mu_1^{-1} + \frac{2}{\mu_2 - 3\lambda_0} + \frac{2}{\mu_3 - 3\lambda_0} + 2\mu_4^{-1} + \frac{2}{\mu_5 - 2\lambda_0} + \mu_6^{-1}. \quad (1)$$

$$T_S = 2\mu_1^{-1} + \frac{3}{\mu_2 - 5\lambda_0} + \frac{3}{\mu_3 - 5\lambda_0} + 3\mu_4^{-1} + \frac{3}{\mu_5 - 4\lambda_0} + 2\mu_6^{-1}. \quad (2)$$

Утверждение 2 Доля времени фильтрации пакетов МЭ в средней задержке запроса установления сессии и среднем времени установления сессии вычисляется по формулам (3) и (4).

$$\Delta_{T_F - T_{SRD}} = \frac{\frac{2}{\mu_2 - 5\lambda_0}}{2\mu_1^{-1} + \frac{2}{\mu_2 - 3\lambda_0} + \frac{2}{\mu_3 - 3\lambda_0} + 2\mu_4^{-1} + \frac{2}{\mu_5 - 2\lambda_0} + \mu_6^{-1}}. \quad (3)$$

$$\Delta_{T_F - T_S} = \frac{\frac{3}{\mu_2 - 5\lambda_0}}{2\mu_1^{-1} + \frac{3}{\mu_2 - 5\lambda_0} + \frac{3}{\mu_3 - 5\lambda_0} + 3\mu_4^{-1} + \frac{3}{\mu_5 - 4\lambda_0} + 2\mu_6^{-1}}. \quad (4)$$

На рисунках 1-2 приведены графики, показывающие долю времени фильтрации сигнальных сообщений в МЭ при установлении сессии по протоколу SIP.

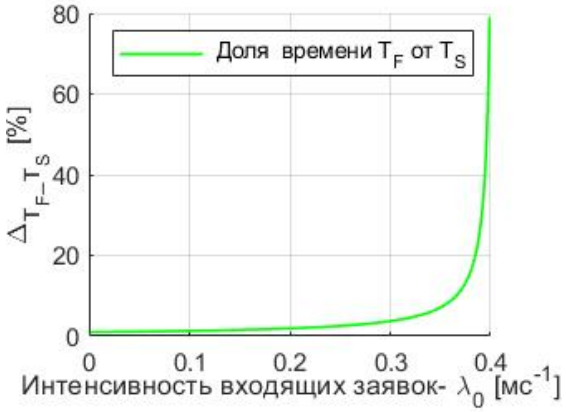

Рис. 1 Доля времени фильтрации во времени установлении сессии

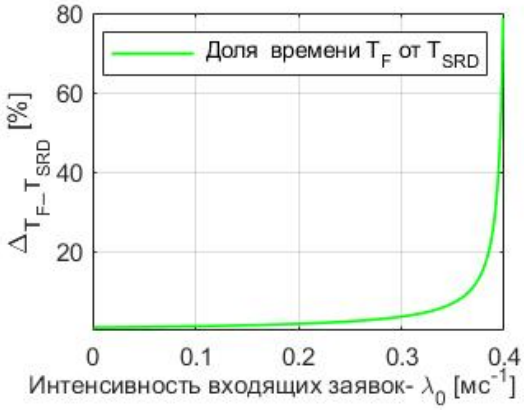

Рис. 2 Доля времени фильтрации в средней задержке запроса на установление сессии

Анализ задержек T_S и T_{SRD} показал результаты, удовлетворяющие рекомендованным требованиям к показателям качества восприятия услуги. При этом значения T_S и T_{SRD} при $\lambda_0 < 0.35$ [мс$^{-1}$] близки к значению минимально воспринимаемой задержки. При значениях интенсивности λ_0, близких к пределу существования стационарного режима в СеМО, T_S достигает значения 760 [мс], T_{SRD} – 560 [мс] и не превышает рекомендованных 2 [с], согласно требованиям качества обслуживания для услуги установления сессии.

При высокой нагрузке и приближении к пределу существования стационарного режима в СеМО, происходит резкое увеличение времени пребывания сигнальных сообщений в МЭ до 80%. Полученные результаты показали целесообразность уменьшения времени нахождения заявок в МЭ для снижения значений показателей качества восприятия и обслуживания для услуги установления сессии по протоколу SIP в условиях нормальной и высокой нагрузки.

Во <u>второй главе</u> разработана математическая модель МЭ с ранжированием правил фильтрации, исследованы наиболее важные для построения модели особенности процесса фильтрации пакетов МЭ. Описаны показатели эффективности МЭ, сформулирована задача нахождения оптимального порядка правил, приведено ее решение для случая отсутствия ошибок несогласованности и избыточности в наборе правил фильтрации, приведено описание разработанного метода ранжирования набора правил фильтрации, основанного на ранжировании в соответствии с оценками МЛА.

Модель является сложной стохастической системой, поэтому для ее построения использовался агрегативный подход, который представляет систему в виде агрегата $M(k) = \{Z, L, T, \Phi, G, X, Y\}$ (рисунок 3), имеющего входные $\mathbf{x}_k = (x_1^k, ..., x_N^k) \in X$ и выходные сигналы $\mathbf{y}_k = (q_k, v_k, w_k, u_k) \in Y$, где k номер интервала функционирования системы $[t_{k-1}, t_k)$, в течение которого осуществляется фильтрация одной пачки пакетов, ранжирование набора правил не производится, а вектор состояния $\mathbf{z}_k = (\mathbf{r}_k, \mathbf{d}_k) \in Z$ остается неизменным, $\mathbf{r}_k = (r_1^k, ..., r_N^k)$ – набор правил в котором компонента r_i^k является правилом стоящим на i-ом месте в наборе, $\mathbf{d}_k = (d_1^k, ..., d_N^k)$ – вектор времен обслуживания пакетов, в котором d_i^k соответствует времени обработки пакетов i-го типа на интервале $[t_{k-1}, t_k)$. Компонента $x_i^k, i = 1, ..., N$ вектора $\mathbf{x}_k$ – случайная величина, характеризующая число пакетов i-го типа соответствующих r_i^k-му правилу, а компоненты q_k, v_k, w_k, u_k соответствуют значениям показателям эффективности на интервале времени $[t_{k-1}, t_k)$: среднему числу пакетов в накопителе, времени обслуживания, времени ожидания начала обслуживания и времени пребывания в системе соответственно.

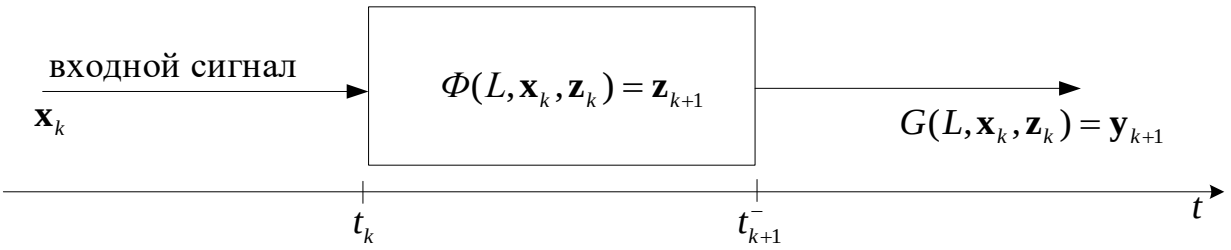

Рис. 3 Схема модели МЭ с ранжированием правил фильтрации

Параметры системы $L=(\mu_0,\mu,N,C)$: μ_0 – интенсивность обслуживания при первичной обработке пакета сетевой картой МЭ, μ – интенсивность обслуживания при проверке соответствия пакета одному правилу набора фильтрации, C – емкость накопителя системы, N – число правил в наборе фильтрации. Переход состояний системы происходит в момент времени t_{k+1}^-, оператор переходов состояний системы Φ вычисляет веса правил $\mathbf{p}_k=(p_1^k,...,p_N^k)$ в соответствии с характеристиками информационных потоков $\widehat{\mathbf{x}}_k$. Оператор определяет состояние системы $\mathbf{z}_{k+1}=(\mathbf{r}_{k+1},\mathbf{d}_{k+1})$, где вектор $\mathbf{r}_{k+1}$ вычисляется ранжированием набора правил $\mathbf{r}_k$ в соответствии с весами $\mathbf{p}_k$, а вектор времен обслуживания пакетов $\mathbf{d}_{k+1}$ вычисляется в соответствии с полученным набором $\mathbf{r}_{k+1}$ и интенсивностями обслуживания μ_0,μ. G – оператор выходов системы в момент времени t_{k+1}^- определяет показатели эффективности системы на интервале времени $[t_{k-1},t_k)$.

Для расчета показателей эффективности, рассматривается функционирование агрегата $M(k)$ на интервале $[t_{k-1},t_k),k\geq 1$. Времена обслуживания заявки на каждой фазе независимы между собой и распределены экспоненциально. Схематично процесс обслуживания заявки изображен на рисунке 4.

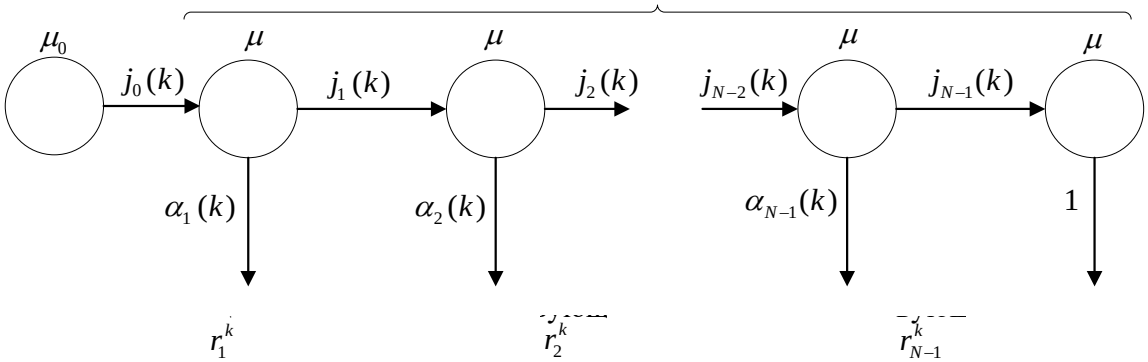

Рис. 4 Фазовое представление процесса обслуживания заявки в модели межсетевого экрана

Вероятность перехода заявки с i-ой фазы на следующую $i+1$-ю фазу для k-го интервала времени определяется в виде

$$j_i(k)=\begin{cases}1, i=0,\\ 1-\lambda_i(k)\big/\lambda_{\square}(k), i\in 1,...,N-1.\end{cases} \quad (5)$$

Обозначим $\alpha_i(k)=1-j_i(k)$ – вероятность завершения обслуживания заявки на i-ой фазе для k-го интервала времени. Тогда агрегат $M(k)$ можно представить в виде однолинейной СМО с накопителем ограниченной емкости, неоднородным пуассоновским входящим потоком и обслуживанием заявок с функцией распределения (ФР) $B_k(t)$ длительности обслуживания заявок фазового типа, зависящей от порядка правил фильтрации.

$$B_k(t)=\begin{cases}j_0(k)\mathrm{E}(1,\mu_0)+\sum_{i=1}^{N-1} j_i(k)\mathrm{E}(i,\mu), k\in 1,...,N-1,\\ \mathrm{E}(N,\mu), k=N,\end{cases} \quad (6)$$

где $E(i,\mu_0)$ – распределение Эрланга i-го порядка.

В СМО поступает поток заявок, являющийся суперпозиций N независимых пуассоновских потоков заявок. По классификации Башарина-Кендалла данная СМО обозначается $\mathrm{M_N}$/PH/1/C.

В диссертационной работе разработан алгоритм вычисления показателей эффективности МЭ для экспоненциального распределения времени обслуживания. Алгоритм основан на решении системы уравнений глобального баланса марковского процесса $\{\eta(t), t\ge 0\}$, представленной в виде матрично-рекуррентных формул. Применение матрично-рекуррентных формул позволило эффективно произвести расчеты характеристик СМО.

Определим случайный процесс (СП) $\{\eta(t), t\ge 0\}$ на множестве состояний $X=\{(0);(h,i), h=1,...,C+1, i=0,...,N\}$. Состояния множества X имеют следующий смысл. Если в некоторый момент времени $\eta(t)=0$, то в системе нет заявок, если $\eta(t)=(h,i)$, то в системе находится h заявок, а обслуживаемая заявка находится на i-ой фазе. Построенный таким образом СП $\{\eta(t), t\ge 0\}$ является однородным марковским процессом (МП).

Используя матричное представление, запишем ниже ФР времени обслуживания заявки $B_k(t)$. Далее везде для краткости индекс k интервала времени агрегата $M(k)$ опускается.

$$B_k(t)=1-\boldsymbol{\beta}^{\mathrm{T}} e^{\mathbf{M}t}\mathbf{1}, t>0, \boldsymbol{\beta}^{\mathrm{T}}\mathbf{1}=\mathbf{1}, \quad (7)$$

где пара $(\boldsymbol{\beta},\mathbf{M})$ – PH представление порядка $N+1$, $\boldsymbol{\beta}^{\mathrm{T}}=(\beta_0,...,\beta_N)$ – вектор вероятностей направления заявки на обслуживание в фазы $0,1,...,N$ в момент времени t_k, где компонента $\beta_i, \mathrm{i}\in 0,...,N$ соответствует вероятности начала обслуживания заявки на i-ой фазе в момент t_k, $\mathbf{M}$ – инфинитезимальная матрица.

Тогда система уравнений глобального баланса в матричной форме для СМО, описывающая процесс функционирования МЭ с учетом фильтрации имеет вид.

$$
\begin{gathered}
-\lambda_{\square} p_0 + \mathbf{p}_1^{\mathrm{T}} \boldsymbol{\mu} = 0, \\
\mathbf{p}_1^{\mathrm{T}} (-\lambda_{\square} \mathbf{I} + \mathbf{M}) + \lambda_{\square} \boldsymbol{\beta}^{\mathrm{T}} p_0 + \mathbf{p}_2^{\mathrm{T}} \boldsymbol{\mu} \boldsymbol{\beta}^{\mathrm{T}} = \mathbf{0}^{\mathrm{T}}, \\
\mathbf{p}_h^{\mathrm{T}} (-\lambda_{\square} \mathbf{I} + \mathbf{M}) + \lambda_{\square} \mathbf{p}_{h-1}^{\mathrm{T}} + \mathbf{p}_{h+1}^{\mathrm{T}} \boldsymbol{\mu} \boldsymbol{\beta}^{\mathrm{T}} = \mathbf{0}^{\mathrm{T}}, h = 2,3...,C, \\
\mathbf{p}_{C+1}^{\mathrm{T}} \mathbf{M} + \lambda_{\square} \mathbf{p}_C^{\mathrm{T}} = \mathbf{0}^{\mathrm{T}},
\end{gathered}
\tag{8}
$$

где $\boldsymbol{\mu}^T = (0, \alpha_1 \mu, ..., \alpha_{N-1}\mu, \mu)$ – вектор интенсивностей завершения обслуживания заявки в СМО, p_0 – стационарная вероятность отсутствия заявок в системе, $p_{i,j}$ – стационарная вероятность обслуживания заявки на j-ой фазе и нахождения i заявок в системе, $\mathbf{p}_h^{\mathrm{T}} = (p_{h,0}, p_{h,1}, ..., p_{h,N}), h = 1,..,C+1$ – вектор стационарных вероятностей.

Решение СУР (8) позволяет вычислить показатели эффективности для стационарного режима работы СМО. Использование данного подхода позволило эффективно произвести расчеты показателей эффективности МЭ при фильтрации информационных потоков близких к реальным.

Далее ниже сформулирована задача нахождения оптимального порядка правил фильтрации. Под оптимальным порядком правил $(r_1^k, ..., r_N^k)$ набора $\mathbf{r}_k$ будем понимать такой порядок правил в переупорядоченном векторе $\tilde{\mathbf{r}}_k = (r_{\gamma_1(k)}^k, ..., r_{\gamma_N(k)}^k)$, при котором целевая функция примет минимальное значение:

$$
\begin{gathered}
f(\gamma) = \sum_{i=1}^{N} p_i^k \cdot \gamma_i(k) \xrightarrow[\gamma_i(k)]{} \min, k \geq 1, \\
\gamma_i(k) \in \{1, ..., N\}, i = 1, 2, ..., N.
\end{gathered}
\tag{9}
$$

Предполагается, что веса p_i^k отражают частоту соответствия пакетов r_i^k-му правилу на интервале времени $[t_{k-1}, t_k)$. Решение задачи нахождения оптимального порядка правил фильтрации при отсутствии ошибок несогласованности и избыточности в наборе правил является тривиальным и находится простой перестановкой правил в порядке убывания их весов, т.е.

$$
\gamma_j(k) = \begin{cases}
\underset{i \in \{1,...,N\}}{\arg\max}\, p_i^k, \; j = 1, \\
\underset{i \in \{1,...,N\} / \{\gamma_1(k), ..., \gamma_j(k)\}}{\arg\max}\, p_i^k, 1 < j < N, \\
\gamma_N(k) = \gamma_N(k-1) = ... = \gamma_N(0) = N, \; j = N.
\end{cases}
\tag{10}
$$

Общая схема получения оценок МЛА, основана на решении специальных экстремальных задач, определяется следующими соотношениями:

$$
\begin{gathered}
\hat{x}_m(t, \delta) = \hat{c}_N \varphi(0), \hat{c}_m = \underset{c}{\arg\min}\, J_m(t, c, \delta), \\
J_m(t, c, \delta) = \frac{1}{m} \sum_{S=1}^{m} p\left(\frac{t - t^{(S)}}{\delta} \right) F(z^{(S)} - c^{\mathrm{T}} \varphi(t - t^{(S)})),
\end{gathered}
\tag{11}
$$

где m – число наблюдений, $\varphi(t) \in R_m$ – вектор координатных функций, $c \in R_m$ – вектор неопределенных коэффициентов, $\hat{c}_m$ – оценка этого вектора, определенная минимизацией функционала J_m по c, $F(\cdot)$ – неотрицательная функция потерь, $p(u)$ – функция локальности, имеющая максимум в нуле и стремящаяся к нулю при $\|u\| \to \infty$, $\hat{x}_m(t,\delta)$ – оценка значения $x(t)$ в точке t.

МЛА по существу является скользящим методом наименьших квадратов (МНК) с областью локальности, зависящей от параметра δ.

Описано применение МЛА для решения задачи оценки средних значений количества пакетов, прошедших фильтрацию в соответствии с условиями правил фильтрации, которая сведена к одной из прикладных задач идентификации – получения оценок значений функции по конечной совокупности экспериментальных данных. Оценки строятся в соответствии с локально-линейной моделью МЛА (моделью первого порядка) с квадратичной функцией потерь.

Утверждение 3 Оценка числа пакетов i-го типа, соответствующих r_i^1 правилу фильтрации МЭ в момент t при фиксированном i – является решением задачи (11) и вычисляется по формуле (12).

$$\hat{x}_i(t,\delta) = \frac{\sum_{S=1}^{m} p((t-t_i^{(S)})\delta^{-1})[B_2(t) - B_1(t)(t-t_i^{(S)})] \cdot x_i^{(S)}}{B_2(t)B_0(t) - B_1^2(t)}, \qquad (12)$$

где $B_l(t) = \sum_{S=1}^{m} p(\,(t-t_i^{(S)})\delta^{-1})(t-t_i^{(S)})^l$, $l = 0,1,2$, m – число точек экспериментальных данных, $t_i^{(S)}$ – момент времени поступления опытных значений величины $x_i^{(S)}$, характеризующей число пакетов i-го типа, соответствующих r_i^k-му правилу.

В качестве функции локальности выбрана экспоненциальная функция вида

$$p_i(t,\delta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-t_i^{(S)})^2}{2\delta^2}}. \qquad (13)$$

Для использования формулы (13) необходимо задать значение параметра локальности δ. Определим его в зависимости ширины скользящего окна. Обозначим ширину скользящего окна $L_i = t_i - t_i^{(s)}$ и найдем δ из (13):

$$\delta = \frac{L_i}{\sqrt{-2\ln(p_i\sqrt{2\pi})}}. \qquad (14)$$

Изменение ширины скользящего окна позволяет избежать равнозначной оценки исходной информации и учесть как последние, так и прошлые тенденции развития процесса. Проведенные расчеты по опытной информации показали, что для увеличения точности наиболее позднюю информацию необходимо оценивать выше, чем информацию, характеризующие прошлые тенденции. Такая оценка проводится путем взвешивания или дисконтирования, т.е. придания большего удельного веса по степени информативности более поздней информации. Эта задача решается путем выбора значений параметра δ.

Утверждение 4 Ширина скользящего окна для экспоненциальной функции локальности вида (13) рассчитывается по формуле:

$$L_i = \frac{\vartheta(t_i^{(m)} - t_i^{(1)} + \Delta)}{n}. \quad (15)$$

где Δ – шаг сетки, m – число узлов, ϑ – число уровней информации, имеющих больший вес. Рекомендуемое число уровней информации при $n < 20$ составляет 30-40% от общего объема выборки, при $n < 100$ – 10-25%.

Приведено описание разработанного метода ранжирования в соответствии с оценками МЛА. Под ранжированием набора правил понимается сортировка правил в порядке убывания их весов. Для вычисления весов для каждого правила составляется временной ряд, отражающий число пакетов, прошедших фильтрацию в соответствии с условиями правила в течении нескольких временных интервалов. С помощью МЛА вычисляются оценки средних значений количества пакетов, прошедших фильтрацию в соответствии с условиями правил для последующих интервалов времени.

В третьей главе построена имитационная модель МЭ с ранжированием набора правил фильтрации.

Имитационная модель предназначена для вычисления показателей эффективности МЭ с ранжированием правил. Использование методов имитационного моделирования позволило снять ограничения на вид ФР входящего потока заявок и обслуживания заявок модели МЭ при вычислении показателей эффективности.

Разработан алгоритм функционирования имитационной модели МЭ. Алгоритм представлен в виде псевдокода близкого по синтаксису к операторам системы матричных вычислений MATLAB.

Проведен качественный и количественный анализ численных результатов, полученных с помощью аналитической и имитационной модели. Сравнительный анализ показал достаточную степень адекватности аналитической модели имитационному моделированию.

Выполнена оценка эффективности разработанного метода ранжирования набора правил фильтрации при различных параметрах модели, МЛА и характеристиках поступающего трафика.

При моделировании рассмотрено два типа условий функционирования системы, соответствующих нормальному режиму функционирования системы и функционированию МЭ в условиях перегрузки. Под нормальными условиями функционирования понимается такое соотношение интенсивностей поступления сетевых пакетов, интенсивностей обслуживания заявок и емкости накопителя, которое не приводит к функционированию системы с нагрузкой $\rho > 0.9$. Функционирование системы в условиях перегрузки рассмотрено при поступающем трафике с гармонически изменяющейся интенсивностью входящего потока, при котором система функционирует с перегрузкой или близка к пределу потери пакетов. Данные условия можно приближенно считать моделированием функционирования МЭ с перегрузкой, например при осуществлении DDoS атак.

В зависимости от значений параметров МЛА и выбранных весов рассмотрены следующие методы ранжирования правил фильтрации:

1. Ранжирование с адаптивным δ – проводилось согласно оценкам МЛА 1-го порядка при параметре локальности δ, рассчитываемым по формулам 14-15 с числом уровней информации $\vartheta = 5$.

2. Ранжирование МНК – проводилось согласно оценкам МЛА 1-го порядка при $\delta = \infty$. Полученные оценки являются оценками МНК.

3. Ранжирование без сглаживания – проводилось согласно оценкам МЛА 1-го порядка при $\delta \to 0$.

4. Ранжирование не проводится. Веса правил $p_i^k = 0, i = 1, ..., N$.

Моделирование процесса фильтрации трафика в нормальных условиях функционирования системы (рисунок 5) показало, что применение ранжирования позволило уменьшить среднее время обслуживания (рисунок 6) на 0.007 мс (около 26.6%), уменьшить нагрузку в системе и значимо снизить значения остальных показателей эффективности, за исключением средней длины очереди, что объясняется малым числом заявок в очереди.

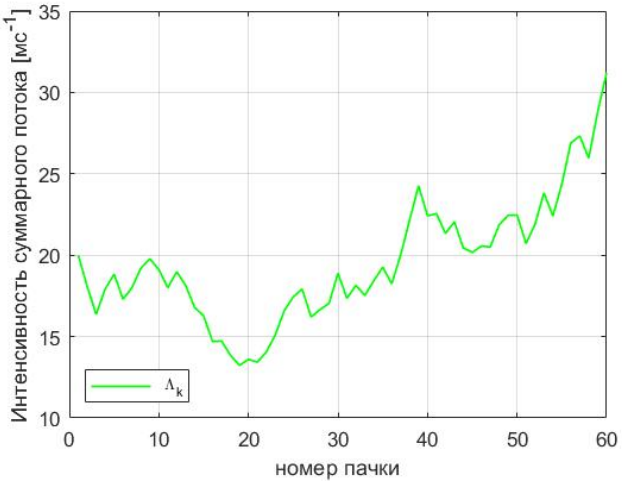

Рис. 5 Суммарная интенсивность входящего потока

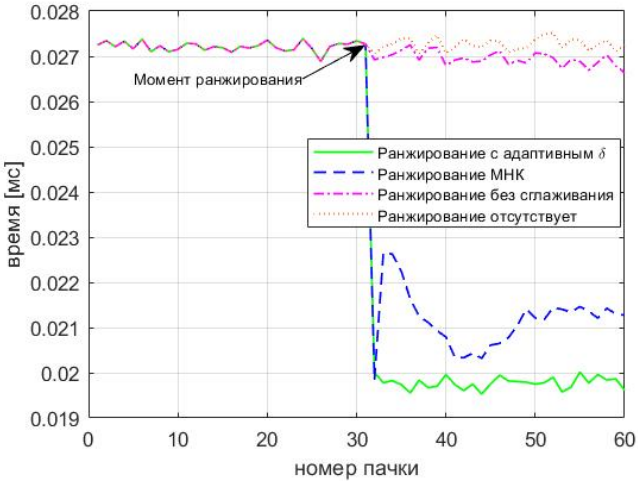

Рис. 6 Среднее время пребывания в системе

Моделирование процесса фильтрации трафика с гармонически изменяющейся интенсивностью входящего потока (рисунок 7) показало, что использование ранжирования позволило уменьшить среднее время обслуживания (рисунок 8) на 0.007 мс (около 26.5%), что привело к снижению нагрузки в системе и значимому снижению значений всех показателей эффективности.

По сравнению с результатами, полученными при моделировании процесса фильтрации трафика с нормальными условиями функционирования, абсолютные погрешности показателей эффективности имеют более высокие значения, за исключением среднего времени обслуживания, так как оно не зависит от нагрузки системы.

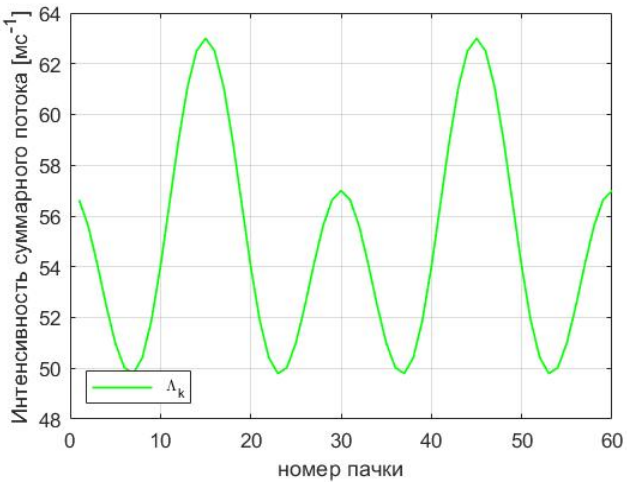

Рис. 7 Суммарная интенсивность входящего потока

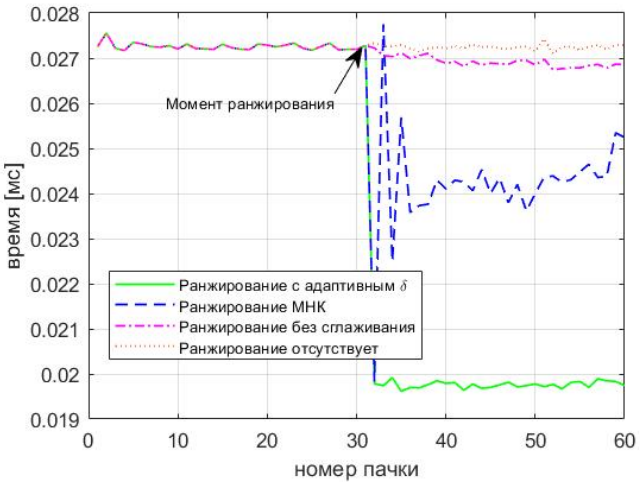

Рис. 8 Среднее время пребывания в системе

Сделанное предположение о повышении производительности МЭ за счет использования разработанного метода ранжирования набора правил, адаптивного к изменяющимся характеристикам информационных потоков, подтвердилось. В среднем ранжирование с адаптивным δ для нормальных условий и условий перегрузки системы показало большую эффективность по сравнению с остальными методами. Это объясняется меньшей ошибкой аппроксимации у МЛА при адаптивном δ по сравнению с МНК.

Заключение

Ниже приведены основные выводы и результаты диссертационной работы.

1. Построена математическая модель для анализа влияния МЭ на время установления сессии по протоколу SIP в виде СеМО. Полученные результаты говорят о целесообразности уменьшения времени пребывания заявок в МЭ для уменьшения значений показателей качества восприятия и обслуживания услуги установления сессии по протоколу SIP в условиях нормальной и высокой нагрузки.

2. Разработана и исследована математическая модель МЭ, отражающая особенности фильтрации информационных потоков и ранжирования набора правил в виде однолинейной СМО с накопителем ограниченной емкости, неоднородным пуассоновским входящим потоком и обслуживанием заявок с функцией распределения длительности обслуживания заявок фазового типа, зависящей от порядка правил.

3. Разработан алгоритм вычисления показателей эффективности для экспоненциального распределения времени обслуживания.

4. Разработан метод ранжирования набора правил фильтрации в соответствии с оценками МЛА средних значений количества пакетов, прошедших фильтрацию в соответствии с условиями правил. Дана рекомендация и метод выбора управляющих параметров МЛА для дисконтирования информации.

5. Разработана имитационная модель для анализа показателей эффективности МЭ при фильтрации информационных потоков. Численный эксперимент

на близких к реальным данным показал эффективность разработанного метода ранжирования набора правил фильтрации при различных параметрах модели, МЛА и характеристиках поступающего трафика.

Список работ, опубликованных по теме диссертации

Статьи в рецензируемых изданиях, рекомендованных ВАК РФ:

1. Ботвинко, А. Ю. Математическая модель работы межсетевого экрана для мультимедийного трафика / А. Ю. Ботвинко, К. Е. Самуйлов // T-Comm – Телекоммуникации и Транспорт. – 2015. – Т.9. – №.12. – С. 56–60. – Текст : непосредственный.
2. Ботвинко, А. Ю. Оценка времени установления сессии между пользователями при наличии межсетевого экрана / К.Е. Самуйлов, А. Ю. Ботвинко, Э. Р. Зарипова // Вестник Российского университета дружбы народов, серия: Математика, информатика, физика. – Москва : РУДН – 2016. – Т.1. – С. 59–66. – Текст : непосредственный.
3. Ботвинко, А. Ю. Анализ схемы управления доступом в сети CDMA с резервированием для мягких хэндовер-вызовов / А. Ю. Ботвинко, К. Е. Самуйлов // T-Comm – Телекоммуникации и транспорт. – 2014. – т. 8 . – №. 9 . – С. 22–25. – Текст : непосредственный.
4. Botvinko, A. Evaluation of firewall performance when ranging a filtration rule set / A. Botvinko, K. Samouylov // Discrete and Continuous Models and Applied Computational Science. – 2021. – V. 29. – № 3. – P. 230–241. – DOI 10.22363/2658-4670-2021-29-3-230-241.
5. Botvinko, A. Evaluation of the firewall influence on the session initiation by the sip multimedia protocol / A. Botvinko, K. Samouylov // Discrete and Continuous Models and Applied Computational Science. – 2021. – № 3. – P. 221–229. – DOI 10.22363/2658-4670-2021-29-3-221-229.

Публикации в сборниках материалов конференций, представленных в зарубежных научных изданиях, входящих в Scopus и/или Springer:

6. Botvinko, A. Firewall Simulation Model with Filtering Rules Ranking / A. Botvinko, K. Samouylov. In: V. Vishnevskiy, K Samouylov., D. Kozyrev (eds) // Communications in Computer and Information Science : Distributed Computer and Communication Networks, DCCN 2020. – Cham : Springer, 2020. – Vol 1337. – P. 533-545. – DOI: 10.1007/978-3-030-66242-4_42.
7. Transmission Latency Analysis in Cloud-RAN / E. Sopin, A. Botvinko, A. Darmolad [et al.] In: V. Vishnevskiy, K. Samouylov, D. Kozyrev (eds) // Distributed Computer and Communication Networks, DCCN 2020, Lecture Notes in Computer Science. – Cham : Springer, 2020. – Vol 12563. – P. 77-86. DOI: 10.1007/978-3-030-66471-8_7.

Материалы международных, всероссийских, молодежных научных конференций

8. Ботвинко, А. Ю. Оптимизация набора правил фильтрации в межсетевых экранах / А. Ю. Ботвинко // Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем : материалы

Всероссийской конференции с международным участием, Москва, 24–28.04.2017. – Москва : РУДН. – 2017. – С. 13–14. – Текст : непосредственный.
9. Ботвинко, А. Ю. Разработка и исследование методов динамической оптимизации правил фильтрации межсетевых экранов / А. Ю. Ботвинко, К. Е. Самуйлов // «Интернет вещей и 5G, INTHITEN 2017» : материалы 3-ей международной научно-технической конференции студентов, аспирантов и молодых ученых, Санкт-Петербург, 20.12.2017. – Санкт-Петербург : СПбГУТ им. проф. М. А. Бонч-Бруевича. – 2017. – С. 28–33. – Текст : непосредственный.
10. Ботвинко, А. Ю. Анализ схемы управления доступом в сети CDMA с резервированием для мягких хэндовер вызовов / А. Ю. Ботвинко., К. Е. Самуйлов // Технологии информационного общества : материалы VIII Международной отраслевой научно-технической конференции, Москва, 20-21 февраля 2014 г. – Москва : МТУСИ. – 2014 – С. 22. – Текст : непосредственный.
11. Самуйлов, К. Е. Математическая модель работы межсетевого экрана для мультимедийного трафика / К. Е. Самуйлов, И. С. Зарядов, А. А. Щербанская, А. Ю. Ботвинко // «Технологии информационного общества» : материалы IX Международной отраслевой научно-технической конференции, Москва, 24 марта 2015 г. – Москва : МТУСИ. – 2015 – С. 30. – Текст : непосредственный.
12. Влияние межсетевого экрана на среднее время установления сессии / Е. И. Балыка, А. Ю. Ботвинко, Э. Р. Зарипова, Д. А. Саитов // Информационно-телекоммуникационные технологии и математическое моделирование высокотехнологичных систем : материалы Всероссийской конференции с международным участием, Москва, 18–22.04.2016. – Москва : РУДН. – 2016. – С. 77-78. – Текст : непосредственный.
13. Зарипова, Э. Р. Анализ среднего времени фильтрации пакетов межсетевым экраном / Э. Р. Зарипова, О. В. Чехонина, Д. А. Лазарев, А. Ю. Ботвинко // Сборник трудов XI Международной отраслевой научно-технической конференции «Технологии информационного общества», Москва, 15- 16 марта 2017 года. – Москва : ООО «ИД Медиа Паблишер» . – 2017. – С. 43–44. – Текст : непосредственный.
14. Ботвинко, А. Ю. Адаптивное ранжирование набора правил межсетевого экрана методом локальной аппроксимации / А. Ю. Ботвинко., К. Е. Самуйлов // «Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь (DCCN-2018) : материалы XXI Международной научной конференции, Москва, 17–21.09.2018. – Москва : РУДН. – 2018. – С. 334-341. – Текст : непосредственный.
15. Об оценке среднего времени фильтрации мультимедийного трафика в межсетевом экране / Э. Р. Зарипова, А. Ю. Ботвинко, Е. И. Балыка, Д. А. Саитов // // Телекоммуникационные и вычислительные системы. Международный форум информатизации (МФМ-2016); Международный конгресс “Коммуникационные технологии сети”. – Москва : Брис-М. – 2016. – С. 27-28. – Текст : непосредственный.

Ботвинко Анатолий Юрьевич (Россия)

Разработка модели для анализа показателей эффективности межсетевого экрана с ранжированием правил фильтрации

В диссертации построена математическая модель для анализа влияния МЭ на время установления сессии по протоколу SIP в виде СеМО. Полученные результаты говорят о целесообразности уменьшения времени пребывания заявок в МЭ для уменьшения значений показателей качества восприятия и обслуживания для услуги установления сессии по протоколу SIP в условиях нормальной и высокой нагрузки. Разработана и исследована математическая модель МЭ, отражающая особенности фильтрации информационных потоков и ранжирования набора правил в виде однолинейной СМО с накопителем ограниченной емкости, неоднородным пуассоновским входящим потоком и обслуживанием заявок с функцией распределения длительности обслуживания заявок фазового типа, зависящей от порядка правил. Разработан алгоритм вычисления показателей эффективности для экспоненциального распределения времени обслуживания. Разработан метод ранжирования набора правил фильтрации в соответствии с оценками МЛА средних значений количества пакетов, прошедших фильтрацию в соответствии с условиями правил. Дана рекомендация и метод выбора управляющих параметров МЛА для дисконтирования информации. Разработана имитационная модель для анализа показателей эффективности МЭ при фильтрации информационных потоков. Численный эксперимент на близких к реальным данным показал эффективность разработанного метода ранжирования набора правил фильтрации при различных параметрах модели, МЛА и характеристиках поступающего трафика.

Botvinko Anatoliy (Russia)

Development of a model for analyzing firewall efficiency indicators with ranking of filtering rules

In this thesis, a mathematical model in the form of a SeMO has been constructed to analyze the influence of the FW on the session setup time using the SIP protocol. The results of studies indicate the expediency of reducing the stay time networks packets in the FW to reduce the values QoE and QoS for the SIP under normal and high load conditions. The mathematical model of the FW reflecting the features of filtering data flows and ranking a set of rules has been developed and investigated. The model is a single-line queuing model with a limited capacity storage device, an inhomogeneous Poisson incoming flow and phase-type request service duration distribution function that depends on the order of the rules. An algorithm for calculating performance estimates for the exponential distribution of service time has been developed. The method has been developed for ranking a set of filtering rules in accordance with the MLA estimates of the average values of the number of network packets. The recommendation for selecting the control parameters of the MLA and method for discounting information have been given. A simulation model has been developed to analyze the efficiency indicators of the FW. A numerical experiment based on close to real data has been showed the efficiency method for ranking a set of filtering rules for various parameters of model, parameters of MLA and characteristics of incoming traffic.