МОДЕЛЬ И ПАРАМЕТРИЧЕСКАЯ ОПТИМИЗАЦИЯ ПРОАКТИВНОЙ ЗАЩИТЫ СЕРВИСА ЭЛЕКТРОННОЙ ПОЧТЫ ОТ СЕТЕВОЙ РАЗВЕДКИ

Горбачев А.А.[1]

Цель исследования: *повышение защищенности сервиса электронной почты информационных систем в условиях сетевой разведки.*

Используемые методы: *для достижения цели исследования использованы методы математической статистики, исследования случайных процессов, математического программирования, алгоритмы эвристической оптимизации.*

Результат исследования: *разработана полумарковская модель проактивной защиты сервиса электронной почты от сетевой разведки, позволяющая определять вероятностно-временные характеристики процесса передачи сообщений электронной почты. На основе анализа трафика произведена проверка статистических гипотез о видах распределений времени наступления событий, под воздействием которых исследуемая система эволюционирует в дискретном множестве состояний, выполнена точечная и интервальная оценка значений параметров данных распределений. Решение системы линейных интегральных уравнений Вольтерра с ядрами разностного типа осуществлялось с использованием численных методов преобразования Лапласа. Решена задача векторной оптимизации по определению оптимальных параметров конфигурирования сообщений электронной почты, позволяющих максимизировать результативность защиты сервиса электронной почты, робастность моделируемой системы и минимизировать накладные затраты при соответствующих ограничениях.*

Научная новизна: *заключается в разработке модели и решении задачи оптимизации параметров сервиса электронной почты в условиях сетевой разведки с применением математического аппарата полумарковских процессов, численных методов преобразования Лапласа, параметрической оценкой статистических характеристик модели, скаляризацией многокритериальной оптимизационной задачи методом идеальной точки и поиском экстремума целевых функций с использованием алгоритма роя частиц.*

Ключевые слова: *полумарковский процесс, векторная оптимизация, сетевая ловушка, робастность, преобразование Лапласа, биоинспирированный алгоритм.*

DOI:10.21681/4311-3456-2022-3-69-81

Введение

Сервис электронной почты (далее – СЭП) информационно-аналитических систем различного назначения, как процесс получения, обработки, хранения, пересылки электронных сообщений, получил широкое распространение при реализации информационного обмена между физическими лицами, организациями и различными ведомствами.

В связи с этим, нарушение конфиденциальности, целостности и доступности информации, передаваемой с использованием СЭП, является целью злоумышленников на различном уровне, от физических лиц (хакеров) до разведок иностранных государств, имеющих высокий потенциал и возможности по реализации соответствующих угроз безопасности информации [1]. Основными угрозами безопасности информации применительно к СЭП являются: распространение нежелательных сообщений (спам, фишинг), распространение вредоносного программного обеспечения (ВПО), проведение массовых и целевых компьютерных атак типа «отказ в обслуживании» (*DOS*, *DDOS*-атаки), нарушение конфиденциальности и целостности информации (перехват и анализ содержимого трафика) [2, 3].

Процесс сетевой разведки связан с реализацией всех вышеуказанных угроз, в частности, с попытками непосредственной передачи сообщения электронной почты через целевые почтовые серверы (рис. 1).

1 Горбачев Александр Александрович, адъюнкт Краснодарского высшего военного орденов Жукова и Октябрьской Революции Краснознаменное училища имени генерала армии С.М. Штеменко, г. Краснодар, Россия. E-mail: infosec23.00@ mail.ru.

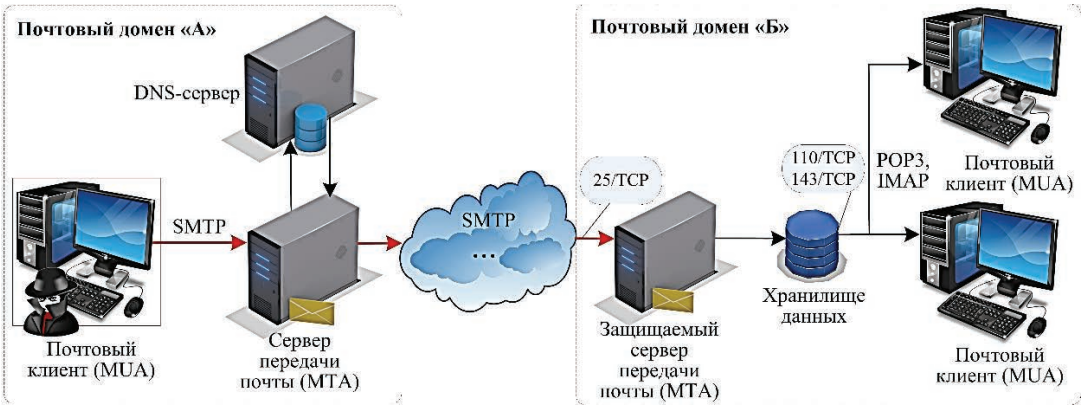


Рис. 1. Клиент-серверная архитектура сервиса электронной почты в условиях сетевой разведки

В целях предотвращения реализации вышеуказанных угроз в информационных системах, содержащих СЭП, применяются меры защиты, основанные на разграничении *IP*-адресов и доменных имен легитимных и нелегитимных отправителей электронной почты («серые», «белые», «черные» списки), средства фильтрации содержимого электронных сообщений, средства антивирусной защиты, средства обнаружения вторжений, основанные на технологиях машинного обучения, а также средства криптографической защиты информации [4-6]. Описанные способы защиты имеют ограниченную эффективность по отношению к вредоносным сообщениям с обфускацией кода и угрозам нулевого дня и являются «реактивными» по отношению к воздействию злоумышленника, то есть реагирующими на попытки и признаки вредоносных воздействий.

Одним из подходов по защите сервиса электронной почты, не исключающим, а дополняющим вышеуказанные, является применение так называемых средств «проактивной» (превентивной) защиты. Под проактивными понимают широкий класс способов защиты информационных систем от сетевой разведки и компьютерных атак, которые возможно применять заблаговременно без значительных затрат на анализ и актуализацию баз данных о признаках деятельности злоумышленников. Средства проактивной защиты могут быть реализованы на различных уровнях эталонной модели взаимодействия открытых систем (ЭМВОС). В частности, могут быть использованы способы управления потоком данных протоколами транспортного и прикладного уровней, фрагментация пакетов данных на уровнях *IP/TCP*, изменение (мутация) *IP/MAC*-адресов, номеров портов и других структурно-функциональных характеристик узлов информационных систем [7-10].

В приведенном исследовании под проактивной защитой понимается способ защиты сервиса электронной почты, основанный на использовании различных механизмов имитации канала связи низкого качества на прикладном и транспортном уровне ЭМВОС, между почтовым клиентом и сервером. Данные механизмы предусматривают отправку почтовым сервером множества служебных многострочных откликов, а также фрагментов ответных откликов через случайные промежутки времени [7]. С целью снижения информативности демаскирующих признаков генерация ответных откликов почтового сервера и их фрагментов на различных этапах почтовой транзакции осуществляется через экспоненциально распределенные промежутки времени с параметрами распределений, подлежащими оптимизации.

Элементы теории случайных процессов широко используются при описании взаимодействия информационных систем в условиях деструктивных воздействий и конфликтов. Исследование предметной области показало, что использование математических моделей марковских случайных процессов в рамках исследования сложных систем [11-13] принимается в соответствии с существующим предположением о том, что большинство реальных потоков событий, являются простейшими. При использовании полумарковских случайных процессов [14-18] либо ограничиваются рассмотрением стационарных вероятностных характеристик для произвольных распределений времени ожидания событий, либо оценка вероятностно-временных характеристик производится только для показательного закона распределения, который обеспечивает соблюдение марковского свойства в любой момент времени. Тем не менее, гипотеза о марковости случайного процесса не всегда является состоятельной, что обосновывает необходимость использования полумарковских процессов и проверку статистических характеристик модели.

Таблица 1

Характеристика дискретного пространства состояний полумарковского процесса передачи сообщения электронной почты

Порядковый номер	Качественная характеристика состояний
1	Ожидание сервером поступления от клиента (злоумышленника) на порт 25 *TCP* пакетов с флагом *SYN* на установление сетевого соединения (клиент и сервер находятся в состоянии простоя)
2	Ожидание сервером поступления от клиента *SMTP* команд *EHLO* (первый этап почтовой транзакции)
3	Ожидание сервером поступления от клиента *SMTP* команд *MAIL* (второй этап почтовой транзакции)
4	Ожидание клиентом окончания поступления от сервера множества промежуточных откликов через изменяемые интервалы времени (с последующим переходом ко второму этапу почтовой транзакции)
5	Ожидание сервером поступления от клиента *SMTP* команд *RCPT* (второй этап почтовой транзакции)
6	Ожидание клиентом окончания поступления от сервера фрагментированных откликов, содержащих код ошибки о временной недоступности сервера (с последующим переходом на третий этап почтовой транзакции)
7	Ожидание сервером поступления от клиента *SMTP* команд *DATA* (третий этап почтовой транзакции)
8	Ожидание клиентом поступления от сервера откликов, содержащих код о невозможности передачи сообщения, разделенных на фрагменты, направляемые через изменяемые интервалы времени (без перехода на четвертый этап почтовой транзакции)
9	Ожидание клиентом поступления отклика «250 *OK*» о подтверждении получения сервером сообщения электронной почты (четвертый этап почтовой транзакции)

Модель проактивной защиты СЭП от сетевой разведки

Функционирование сервиса электронной почты рассматривается на уровне детализации, представленном в стандарте (*RFC* 5321) *SMTP* протокола, как последовательность этапов почтовой транзакции, определяющих дискретное пространство состояний процесса информационного обмена.

Моделируемая система эволюционирует (переходит из одного состояний в другое) под воздействием потока случайных событий, вероятностные характеристики которого в общем случае неизвестны. То есть система с некоторого начального момента времени пребывает в состоянии i в течение случайного промежутка времени, распределенного по произвольному закону, и после появления события, инициирующего эволюцию системы, осуществляет переход в состояние j с переходной вероятностью p_{ij}.

Случайный процесс представляет собой полумарковский процесс, который определяется следующими характеристиками:

- дискретное конечное множество $\{n\}$ состояний системы размерностью n (табл. 1);
- функции распределения $\{F_{ij}(t)\}$ непрерывных случайных величин $\{T_{ij}\}$ длительности ожидания перехода системы из соответствующих состояний (табл. 2);
- переходные вероятности полумарковской цепи $\{p_{ij}\}$;
- произведение переходных вероятностей p_{ij} на соответствующие функции распределения $F_{ij}(t)$ представляют собой элементы $Q_{ij}(t)$ полумарковской матрицы Q.

Исходными данными для модели являются функции $F_{ij}(t)$ с соответствующими значениями статистических параметров.

Процесс передачи сообщений электронной почты с учетом использования средств проактивной защиты от сетевой разведки, моделируемый полумарковским процессом, однозначно идентифицируется ориентированным графом (рис. 2).

Искомыми вероятностно-временными характеристиками полумарковского процесса являются вероятности $P_{ij}(t)$ пребывания системы в состоянии j в момент времени t, при условии, что в момент времени t=0, система находилась в состоянии i (интервально-переходные вероятности).

Также определяются функции распределения $G_{ij}(t)$ времени первого посещения системой состояния j, при условии, что в момент времени t=0, система на-

Таблица 2

Характеристика функций распределения $F_{ij}(t)$ длительностей ожидания случайных событий, инициирующих эволюцию случайного процесса

Переменная	Характеристика случайного события
$F_{12}(t)$	Поступление от клиента (злоумышленника) на порт 25 почтового сервера *TCP* пакетов с флагом *SYN* на установление сетевого соединения
$F_{23}(t)$	Поступление от клиента (злоумышленника) *SMTP* команд *EHLO*
$F_{34}(t)$	Направление почтовым сервером многострочного ответного отклика, разделенного на множество промежуточных ответных откликов, направляемых через регулируемые промежутки времени
$F_{35}(t)$	Поступление от клиента (злоумышленника) *SMTP* команд *MAIL*
$F_{41}(t)$	Поступление от клиента (злоумышленника) *TCP* пакетов с флагом *RSET* на разрыв сетевого соединения и прекращение приема многострочного отклика почтового сервера
$F_{45}(t)$	Поступление от клиента (злоумышленника) *SMTP* команд *MAIL* после получения клиентом многострочного отклика сервера
$F_{56}(t)$	Направление почтовым сервером ответного отклика с кодом ошибки о невозможности передачи сообщений электронной почты
$F_{57}(t)$	Поступление от клиента (злоумышленника) *SMTP* команд *RCPT*
$F_{61}(t)$	Поступление от клиента (злоумышленника) *TCP* пакетов с флагом *RSET* на разрыв сетевого соединения после получения ответного отклика, содержащего код временной ошибки
$F_{67}(t)$	Поступление от клиента (злоумышленника) *SMTP* команд *RCPT* после получения клиентом отклика с ошибкой о невозможности передачи сообщений электронной почты
$F_{78}(t)$	Направление почтовым сервером фрагментов ответного отклика с кодом постоянной ошибки о невозможности передачи сообщений электронной почты
$F_{79}(t)$	Поступление от клиента (злоумышленника) *SMTP* команд *DATA*
$F_{81}(t)$	Поступление от клиента (злоумышленника) *TCP* пакетов с флагом *RSET* на разрыв сетевого соединения после приема фрагментов ответного отклика
$F_{88}(t)$	Направление почтовым сервером повторных фрагментов ответного отклика с кодом ошибки о невозможности передачи сообщений электронной почты
$F_{91}(t)$	Поступление от почтового сервера *SMTP* откликов «250 *OK*» о подтверждении получения сервером сообщения электронной почты

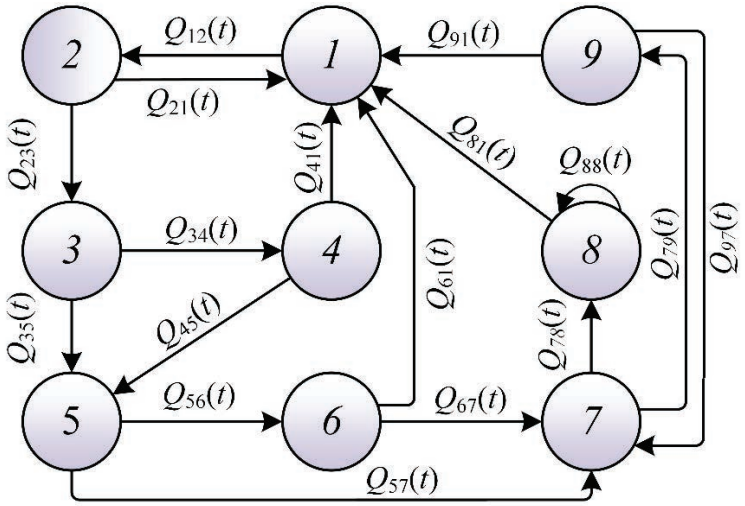


Рис. 2. Ориентированный граф состояний полумарковского процесса передачи сообщений электронной почты в условиях сетевой разведки

ходилась в состоянии *i*. Оценка вероятностей $P_{ij}(t)$ осуществляется на основе решения системы интегральных уравнений Вольтерра 2 рода с интегральным ядром разностного типа (типа «свертки») по известному выражению:

$$P_{ij}(t) = \delta_{ij}\Psi_i(t) + \sum_{k=1}^{n} p_{ik} \int_0^t f_{ik}(t) P_{jk}(t-\tau)\, dt, \quad (1)$$

где δ_{ij} — символ Кронекера: δ_{ij}=1 при $i=j$ и δ_{ij}=0 при $i\neq j$, p_{ij} – переходные вероятности полумарковского процесса, $\Psi_i(t)$ — вероятность того, что система не покинет соответствующие состояние в момент времени *t*. Как правило, решение подобных уравнений осуществляется с использованием преобразования Лапласа. Вычислительная сложность нахождения изображений и оригиналов от результатов вышеуказанных преобразований, отсутствие аналитической формы оригиналов для сложных функций, длительное время ограничивало широкое применение математического аппарата полумарковских процессов. Развитие вычислительной техники и численных методов преобразования Лапласа позволяет осуществлять вышеуказанные расчеты на относительно доступной аппаратной базе за разумное время. Наиболее распространенными современными методами численной аппроксимации преобразования Лапласа являются алгоритмы Фурье-Эйлера, Гавера-Вина, Гавера-Штефеста, Исегера, Талбота[2].

Оценка функций распределения $G_{ij}(t)$ производится из следующего выражения в матричной форме[3]:

$$\mathbf{G}(t) = \int_0^{\infty} L^{-1}\left\{\mathbf{p}\cdot\mathbf{f}(s)\cdot\left(\mathbf{I}-\mathbf{p}\cdot\mathbf{f}(s)\right)^{-1}\cdot\left[\mathbf{I}\times\left(\mathbf{I}-\mathbf{p}\cdot\mathbf{f}(s)\right)^{-1}\right]^{-1}\right\} dt, \quad (2)$$

где: L^{-1} – обратное преобразование Лапласа, «×» – произведение Адамара.

Финальные вероятности полумарковского процесса определяются из выражений:

$$P_{ij} = P_j = \frac{P_j^* \cdot \langle T_j \rangle}{\sum_{j=1}^{n} P_j^* \cdot \langle T_j \rangle}, \quad (3)$$

$$\mathbf{P}^* = \mathbf{B}^{-1}\mathbf{b}, \quad (4)$$

$$\mathbf{B} = \begin{vmatrix} -1 & p_{21} & \dots & p_{91} \\ p_{12} & -1 & \dots & p_{92} \\ \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & 1 \end{vmatrix} \quad (5)$$

где: **B** – матрица переходных вероятностей с учетом условия нормировки; **P*** - вектор финальных вероятностей вложенной цепи Маркова полумарковского процесса; P_j – *j*-й элемент вектора финальных вероятностей полумарковского процесса; T_j – среднее значение безусловного времени пребывания системы в *j*-м состоянии.

Указанные характеристики могут выступать в качестве метрик результативности защиты сервиса электронной почты в условиях сетевой разведки при использовании средств проактивной защиты.

Параметрическая идентификация математической модели

С целью получения экспериментальных данных для статистической оценки параметров модели в статье рассматривается сценарий работы СЭП со средней нагрузкой почтового трафика 2 Гбит/с посредством снятия выборки («дампа») трафика с порта 25 за интервал времени, обеспечивающий репрезентативность данных (1,5 часа). Процесс предварительной обработки дампа трафика (рис. 3) включает в себя фильтрацию трафика по идентификаторам (признакам) в соответствии с характеристиками модели, содержащимися в таблице 2. Временной ряд (рис. 3) не содержит явного тренда и гетероскедастичности (непостоянной дисперсии). Количественная оценка стационарности временного ряда может быть проведена с использованием специфических параметрических тестов или статистических «тестов единичного корня» (*Unit root test*), которые позволяют оценить стационарность временных рядов, описываемых авторегрессионными моделями различного порядка (табл. 3) [19].

Результаты определения значений критериев согласия χ^2 и Холландера-Прошана, точечных и доверительных оценок параметров распределения, характер кривых распределений (рис. 5) указывают на то, что гипотеза об экспоненциальности $F_{12}(t)$, является менее предпочтительной, чем гипотеза о распределении Вейбулла, которая не противоречит данным выборки при уровне значимости α = 0,01.

Теоретическая функция распределения времени ожидания выхода системы из состояния 1 в состояние 2 имеет вид:

$$F_{12}(t) = 1 - e^{-\left(\frac{t}{\beta_{12}}\right)^{\nu_{12}}} = 1 - e^{-1{,}10085 t^{0{,}5411}}, \quad (6)$$

2 Escobar F.H., Leguizamo F.A., Cantillo J.H. Comparsion of Stehgest's and Iseger's algorithms for Laplacian inversion in pressure well tests. ARPN Journal of Engineering and Applied Sciences. 2014. vol. 9(6). pp. 919-922.

3 Warr R.L., Collins D.H. An Introduction to Solving for Quantities of Interest in Finite-State Semi-Markov Processes. 2012. pp. 1-18. // arXiv, 2012. URL: https://arxiv.org/abs/1212.1440/ (дата обращения 20.05.2022).

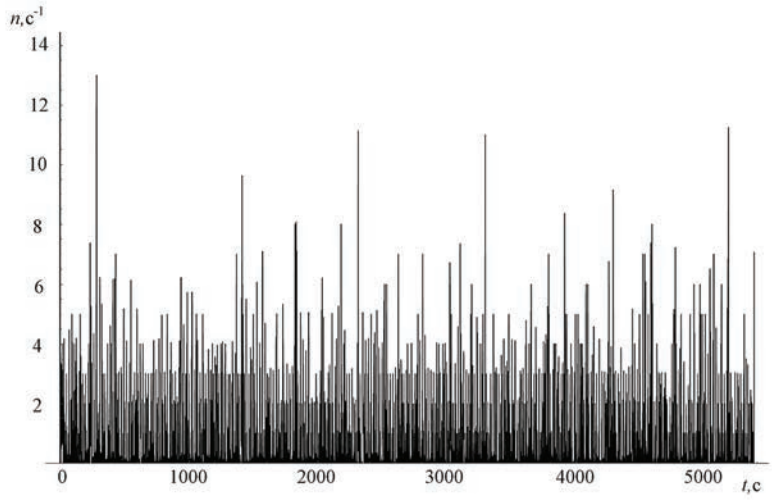


Рис.3. Временной ряд секундных частот регистрации TCP пакетов с флагом SYN на установление сетевого соединения через порт 25

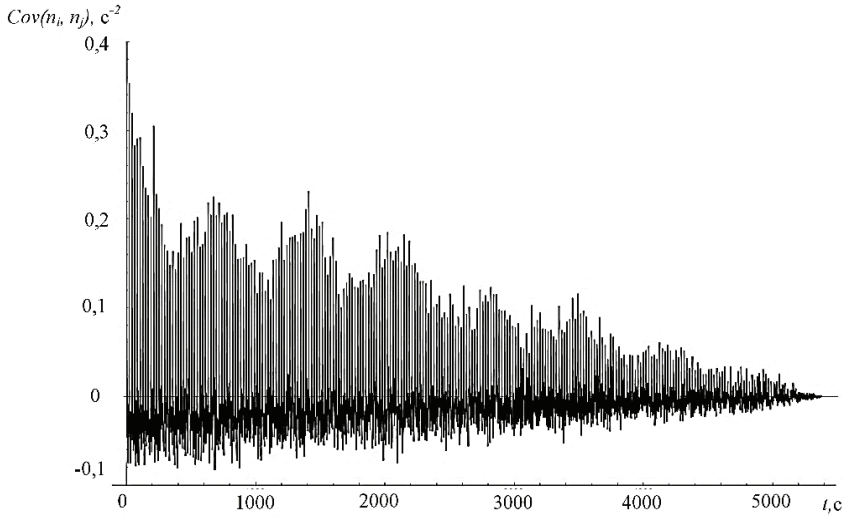
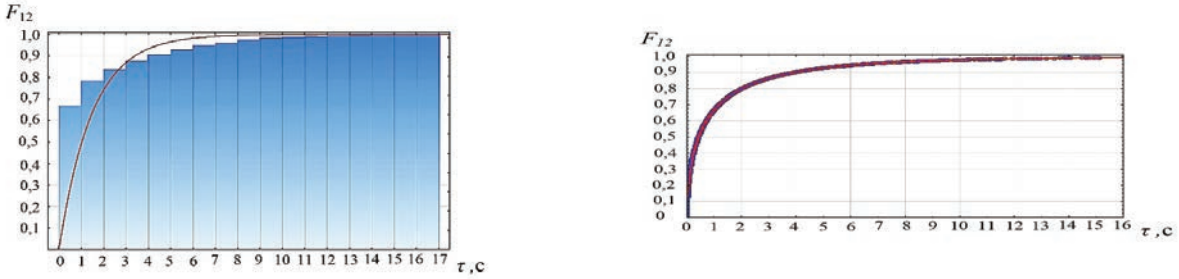


Рис.4. Автоковариационная функция процесса регистрации TCP пакетов с флагом SYN на установление сетевого соединения через порт 25

Таблица 3

Результаты оценки стационарности исследуемого временного ряда с использованием тестов единичного корня

Наименование теста (критерия)	**Значение статистики**	**Критические значения t-статистики при α=0,01**	**Уровень значимости**	**Интерпретация результата теста**
Тест Дики-Фуллера	-2139,35	-3,4632	$7{,}29 \cdot 10^{-20}$	временной ряд стационарен
Тест Филипса-Перрона	-60,570	-3,4314	0,0001	временной ряд стационарен



а) *б)*

Рис. 5. Эмпирическая кумулятивная и теоретическая функция распределения исследуемого временного ряда для: а) экспоненциального распределения; б) распределения Вейбулла

Контролируемые характеристики процесса проактивной защиты тестового сервера с целью снижения вероятности идентификации демаскирующих признаков также распределены экспоненциально, со средними значениями λ_{45}, λ_{67}, λ_{88}, учитывающими средние количественные параметры фрагментации ответных откликов защищаемого почтового сервера, в соответствии с представленным способом защиты СЭП:

$$\lambda_{45} = (d_{45}T_{45})^{-1}, \lambda_{67} = (d_{67}T_{67})^{-1}, \lambda_{88} = (d_{88}T_{88})^{-1}, \quad (7)$$

где: T_{ij} – среднее значение времени задержки между промежуточными откликами почтового сервера на соответствующем этапе почтовой транзакции;

с; d_{ij} – среднее количество ответных промежуточных откликов/фрагментов, направляемых средству сетевой разведки, шт.

В работе целевая функция, характеризующая накладные расходы на формирование дополнительных ответных откликов почтового сервера, определяется как доля ρ от выделенной скорости передачи данных V:

$$\rho = \frac{\gamma}{V}\left(\frac{d_{45}}{T_{45}} + \frac{d_{67}}{T_{67}} + \frac{d_{88}}{T_{88}}\right), \quad (8)$$

где: γ = 40 байт – величина дополнительных служебных данных на 1 дополнительный промежуточный отклик/фрагмент, [бит]; V – выделенная скорость передачи данных, [бит/с].

В условиях отсутствия применения средств проактивной защиты вероятностно-временные характеристики $P_{ij}(t)$ процесса функционирования сервиса электронной почты имеют вид, представленный на рисунке 6 (а).

Интервально-переходные вероятности $P_{ij}(t)$ определены исходя из условия начала полумарковского процесса в состоянии 1, то есть, при $P_{11}(0)$ = 1. Характер переходного процесса показывает, что время установления стационарных вероятностей случайного процесса составляет 2 секунды, а значения финальных вероятностей пребывания системы P_{ij} распределены так, что почтовый сервер практически все время находится в незагруженном состоянии ожидания поступления команд от легитимных клиентов. Полученные результаты указывают на тот факт, что для такого низкого значения потока поступающих заявок от легитимных клиентов отдельные этапы процесса почтовой транзакции занимают относительно малые промежутки времени, обусловленные производительностью почтового сервера. При случайной задержке отправки ответных откликов в среднем на 1 секунду характер вероятностно-временных характеристик процесса реализации почтовой транзакции кардинально меняется в связи с ростом интервально-переходных и финальных вероятностей пребывания системы в состояниях удержания средства сетевой разведки (рис. 6, б).

Функции $G_{ij}(t)$ позволяет оценить вероятности достижения соответствующих состояний впервые к конкретному моменту времени. Так для случая с обслуживанием легитимных клиентов (рис. 7, а) переход из состояния 1 в любое другое состояние происходит практически одновременно (за малый промежуток времени) для всех состояний, в связи с чем все кривые функций распределений на графике фактически совпадают. Если рассматривать условие случайной задержки при отправке ответных откликов в среднем на 1 секунду, то в данном случае длительность первого посещения системой некоторых состояний с определенной вероятностью увеличится на несколько порядков (ось абсцисс на рис. 7, б построена в логарифмическом масштабе). Так с вероятностью 0,8 система впервые посетит состояние 9 из состояния 1 к моменту времени t = 100 с.

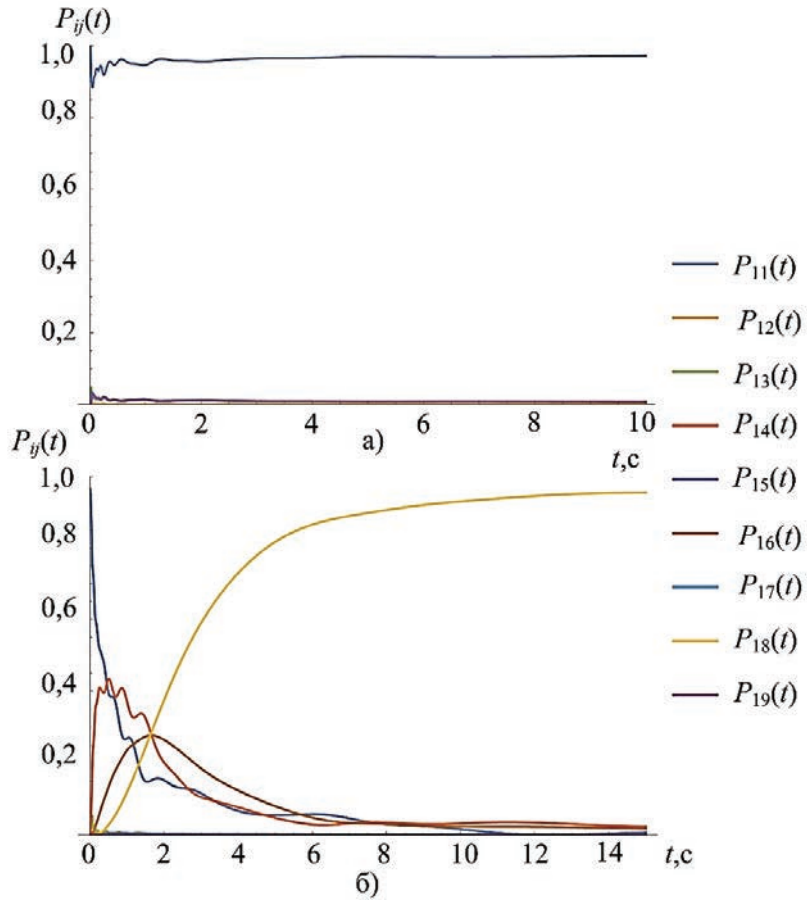


Рис.6. Зависимость интервально-переходных вероятностей $P_{ij}(t)$ от времени для процесса функционирования СЭП без использования (а) и с использованием (б) конфигурирования параметров передачи сообщений

Корректность результатов моделирования

С одной стороны, статистическая оценка параметров математической модели всегда связана с неточностью, обусловленной принятым уровнем значимости и мощностью критериев согласия статистических гипотез. С другой стороны, полученные приближенные оценки используются в последовательности операций по вычислению необходимых вероятностно-временных характеристик исследуемой системы. В случае низкой обусловленности линейных преобразований над матрицами, в частности, над матрицей переходных вероятностей, полученные выходные характеристики модели могут оказаться некорректными, вследствие высокой чувствительности (низкой робастности) к вариации элементов соответствующих матриц. В работе для количественной оценки обусловленности (робастности) операций над матрицей переходных вероятностей **B** используется число обусловленности. Число обусловленности определяется как произведение l_1 нормы матрицы **B** и l_1 нормы обратной матрицы $\mathbf{B}^{-1}$, то есть:

$$\mu(\mathbf{B}) = \|\mathbf{B}\|_1 \cdot \|\mathbf{B}^{-1}\|_1 \quad (9)$$

$$\|\mathbf{B}\|_1 = \max_j \sum_i^n |\mathbf{B}_{ij}|, \quad (10)$$

где: i – номер строки матрицы p, j – номер столбца матрицы p.

Выбор l_1 нормы матрицы по сравнению со спектральной нормой обусловлен снижением вычислительной сложности. Высокие значения (более 100) числа обусловленности свидетельствуют о низкой («плохой») обусловленности матрицы, вызванной близостью матрицы к сингулярной (вырожденной). Так как величина $\mu(\mathbf{B})$ принимает значения из интервала $[1; +\infty)$, то с целью нормализации значений и последующего использования методов скаляризации, в качестве показателя корректности оценок оптимальных значений факторов-аргументов выступает величина $\mu^{-1}(\mathbf{B})$.

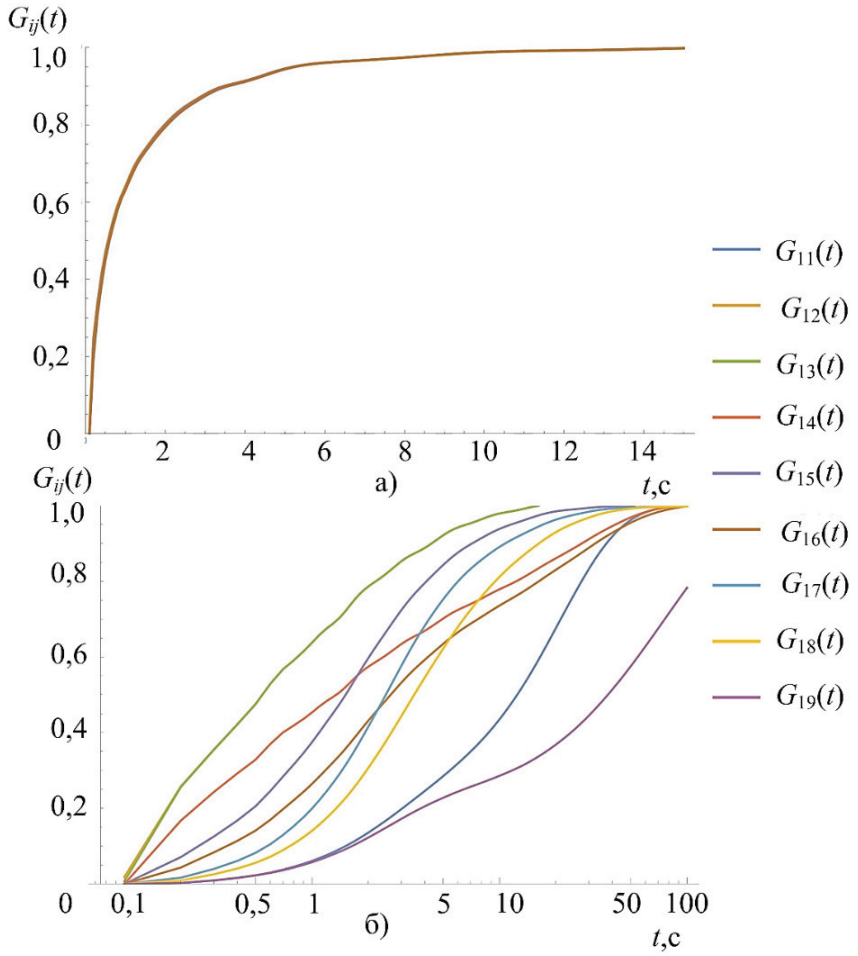


Рис.7. Функции распределения $G_{ij}(t)$ времени первого посещения состояний процесса для процесса функционирования СЭП без использования (а) и с использованием (б) конфигурирования параметров передачи сообщений

Определение оптимальных параметров проактивной защиты СЭП в условиях сетевой разведки

С учетом приведенной модели функционирования СЭП задача векторной оптимизации (1) имеет вид: факторы аргументы: множество факторов аргументов *X* представляет собой множества контролируемых параметров проактивной защиты СЭП, посредством которых изменяются значения целевых функций F_i:

$$X = \{d_{45}, d_{67}, d_{88}, T_{45}, T_{67}, T_{88}\} \quad (11)$$

параметры: множество параметров A включает в себя неконтролируемые параметры функционирования СЭП в условиях сетевой разведки, влияющих на значения целевых функций F_i:

$$A = \{F_{12}(t), F_{23}(t), F_{34}(t), F_{35}(t), F_{41}(t), F_{56}(t), F_{57}(t), F_{61}(t), F_{78}(t), F_{79}(t), F_{81}(t), F_{91}(t)\} \quad (12)$$

целевые функции (критерии): с учетом разработанной модели общая задача векторной оптимизации имеет вид:

$$\begin{cases} F_1(X, A) = P_{14} + P_{16} + P_{18}, \\ F_2(X, A) = \rho, \\ F_3(X, A) = \mu^{-1}(\mathbf{B}), \end{cases} \quad (13)$$

$$\begin{cases} F_1(X, A) \to \max\limits_{X, A \in Q}, \\ F_2(X, A) \to \min\limits_{X, A \in Q}, \\ F_3(X, A) \to \max\limits_{X, A \in Q}, \end{cases} \quad (14)$$

допустимое множество: величины задержек T_{ij} ограничены исходя из средних значений тайм-аутов, используемых для разрыва соединения при идентификации средств защиты.

$$Q:\begin{cases} l_{12}\ge 0, l_{23}\ge 0, l_{34}\ge 0, l_{35}\ge 0, l_{41}\ge 0, l_{56}\ge 0, \\ l_{57}\ge 0, l_{61}\ge 0, l_{78}\ge 0, l_{79}\ge 0, l_{81}\ge 0, l_{91}\ge 0, \\ 1\le d_{45}\le 1000, 1\le d_{67}\le 1000, 1\le d_{88}\le 1000, \\ 0\le T_{45}\le 30, 0\le T_{67}\le 10, 0\le T_{88}\le 10, \\ 0\le r_{:@}\le 1, \\ 0\le F_1(X,A)\le 1, \\ 0\le F_2(X,A)\le r_{:@} \\ 0< F_3(X,A)\le 1. \end{cases} \quad (15)$$

Прямое решение задачи векторной оптимизации представляет собой множество параметров, оптимальных по Парето. Исходя из равнозначности критериев, характера достижимого критериального пространства и поставленной цели исследования выбор одного оптимального набора параметров функционирования средств проактивной защиты принято осуществлять с использованием метода идеальной точки. За «идеальную точку» принимается состояние СЭП, которое характеризуется экстремальными (идеальными) значениями целевых функций F_i.

Тогда скалярная целевая функция $R(X, A)$ имеет физический смысл евклидовой метрики (расстояния) между «идеальной точкой» и точкой фронта Парето, а выбор оптимального набора значений исходных целевых функций и факторов аргументов соответствует минимальному значению указанного расстояния, а скалярная целевая функция имеет вид:

$$\begin{cases} R(X,A)=\sqrt{(F_1-1)^2+(F_2)^2+(F_3-0,09)^2} \\ R(X,A)\to \min\limits_{X,A\in Q}, \end{cases} \quad (16)$$

С целью оценки характера зависимостей критериев друг от друга и нахождения Парето-фронта целесообразно реализовать визуализацию критериального пространства (рис. 8).

Так как, целевые функции F_1, F_2, F_3 нелинейные, то решение указанной оптимизационной задачи (нахож-

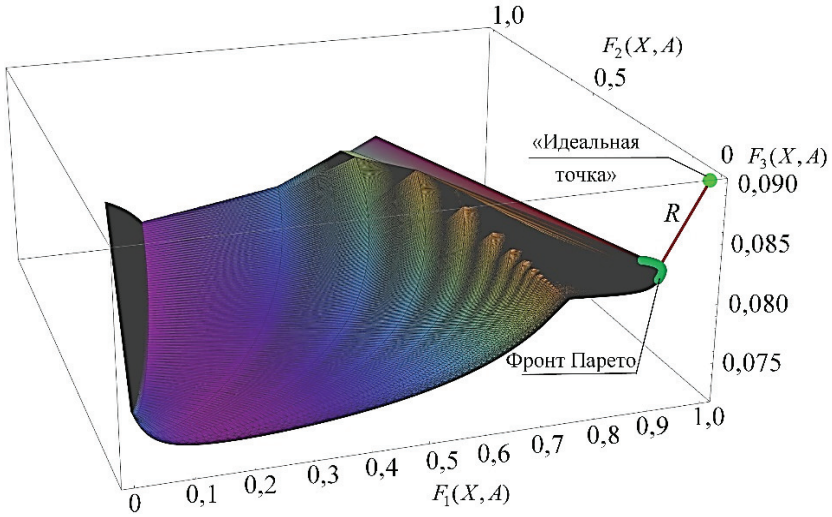


Рис. 8. Графическая интерпретация достижимого критериального пространства, определения фронта Парето и метода идеальной точки

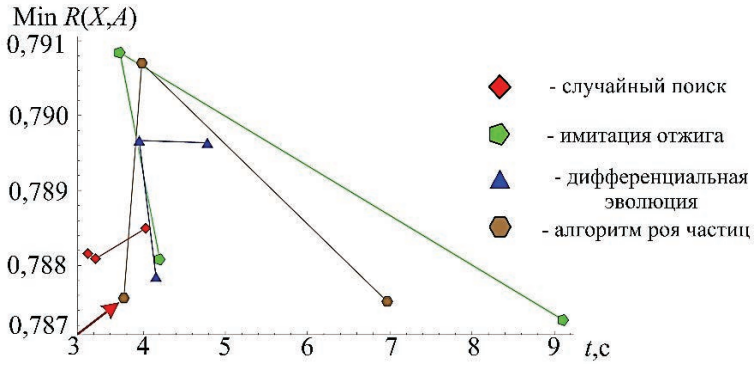


Рис. 9. Сравнение стохастических алгоритмов нелинейной оптимизации на примере скалярной целевой функции R(X, A) по значению найденного минимума и временной сложности вычислений

дение экстремумов целевых функций) осуществляется с использованием одного из известных алгоритмов нелинейной оптимизации. Для сравнения было выбрано четыре алгоритма нелинейной численной оптимизации: эвристические алгоритмы дифференциальной эволюции, имитации отжига и роя частиц, а также чисто случайный поиск. Сравнительная оценка алгоритмов производилась на примере целевой функции $R(X, A)$:

Для каждого алгоритма оптимизации выполнено несколько расчетов с разным количеством случайных точек поиска (от 5 до 100). При решении задачи минимизации скалярной целевой функции перечисленные алгоритмы определяют экстремальное значение в диапазоне от 0,787 до 0,790, имеют сравнимую временную сложность от 3 до 9 с при фиксированных исходных данных ($\rho_{кр}$ = 0,001, V=1 Мбит/с), причем сходимость стохастических алгоритмов обеспечивается уже при значении n = 5-10. Таким образом, рассматриваемые алгоритмы пригодны для решения поставленной задачи, так как обеспечивают достаточную точность и удовлетворительную временную сложность, при подгонке гиперпараметров. В данной работе используется алгоритм роя частиц, который имитирует поведение некоторых биологических видов (рой пчел) при решении задач оптимизации в природных условиях.

При условии $\rho_{кр}$ = 0,01, V = 1 Гбит/с значение скалярной целевой функции составляет величину R = 0,9191, а отдельные критерии имеют значения F_1 = 0,8755, F_2 = 0,0894, F_3 = 0,00086, множество оптимальных параметров проактивной защиты СЭП представляет собой X = {29; 179; 1; 9,199; 9,993; 2,1 804}.

Выводы

Параметрическая оптимизация проактивной защиты СЭП позволяет добиться результативности защиты в метрике F_1 в диапазоне 0,85-0,92, при снижении на величину до 0,1 % полезной скорости передачи данных на реализацию механизмов защиты (при V = 1 Мбит/с) и обеспечении достаточной робастности процесса функционирования проактивных средств защиты относительно погрешностей исходных данных и оценок статистических характеристик полумарковского процесса.

Описанный подход к противодействию реализации непосредственной сетевой разведке через СЭП обеспечивают максимизацию результативности защиты узлов информационной системы, за счет имитации канала связи низкого качества и удержания средств злоумышленника в состоянии ожидания информационного обмена. При этом эксплуатация средств проактивной защиты задействует незначительную часть ресурсов сети передачи данных.

Научная новизна представленной модели и решения задачи параметрической оптимизации СЭП заключается в применении математического аппарата полумарковских случайных процессов с дискретным пространством состояний, параметрической оценкой статистических характеристик модели, определением выходных характеристик модели с использованием численных методов преобразования Лапласа, скаляризацией многокритериальной оптимизационной задачи выбора оптимального режима функционирования средства защиты методом идеальной точки и поиском экстремума целевых функций с использованием алгоритма роя частиц.

Полученные результаты могут быть использованы при реализации технологий проактивной защиты СЭП информационных систем различного назначения критически важной информационной инфраструктуры.

Научный руководитель: *Максимов Роман Викторович, доктор технических наук, профессор, профессор Краснодарского высшего военного орденов Жукова и Октябрьской Революции Краснознаменное училища имени генерала армии С.М. Штеменко, г. Краснодар, Россия. E-mail: rvmaxim@yandex.ru*

Литература

1. Maximov R.V., Sokolovsky S.P., Telenga A.P. Model of client-server information system functioning in the conditions of network reconnaissance. CEUR Workshop Proceeding. 2019. pp. 44-51.
2. Kumari A., Agrawal N., Umesh L. Attack over Email System: Review. International Journal of Scientific Research and Engineering Trends. 2017. vol. 3. pp. 200-206.
3. Schneider M., Shulman H., Sidis A., Sidis R., Waidner M. Diving into Email Bomb Attack. 50th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN). 2020. pp. 286-293. DOI: 10.1109/DSN48063.2020.00045.
4. Castillo D.P., Regidor F.M., Higuera J.B., Higuera J.R., Montalvo J.A. A new mail system for secure data transmission in cyber physical systems. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. 2020. vol. 28(2). pp. 23-48. DOI: 10.1142/S0218488520400127.

5. Zobal L., Kolar D., Kroustek J. Exploring current e-mail cyber threats authenticated SMTP honeypot. 17-th International Conference on Security and Cryptography. Paris. 2020. pp.253-262.
6. Douzi S., AlShahwan F.A., Lemoudden M., Ouahidi B. Hybrid email spam detection model using artificial intelligence. International Journal of Machine Learning and Computing. 2020. vol. 10(2). pp. 316-322. DOI: 10.18178/ijmlc.2020.10.2.937.
7. Соколовский С. П., Горбачев А. А. Способ проактивной защиты почтового сервера от нежелательных сообщений электронной почты // Вопросы оборонной техники. Серия 16: Технические средства противодействия терроризму. 2021. № 3-4 (153-154). С. 31-40.
8. Walla S., Rossow C. MALPITY: Automatic identification and Exploitation of Tarpit Vulnerabilities in Malware. 2019 IEEE European Symposium on Security and Privacy (EuroS&P). 2019. pp. 590-605. DOI: 10.1109/EuroSP.2019.00049.
9. Maximov R.V., Sokolovsky S.P., Telenga A.P. Methodology for substantiating the characteristics of false network traffic to simulate information system. CEUR Workshop Proceeding. 2021. pp. 115-124.
10. Maximov R.V., Sokolovsky S.P., Telenga A.P. Honeypots network traffic parameters modelling. CEUR Workshop Proceeding. 2021. pp. 229-239.
11. Петров М.Ю., Фаткиева Р.Р. Модель синтеза распределенных атакующих элементов в компьютерной сети // Труды учебных заведений связи. 2020. № 2. С. 113-120. DOI:10.31854/ 1813-324X-2020-6-2-113-120.
12. Евневич Е.Л., Фаткиева Р.Р. Моделирование информационных процессов в условиях конфликтов // Вопросы кибербезопасности. 2020. № 2(36). С. 42-49. DOI:10.21681/2311-3456-2020-2-42-49.
13. Будников С.А., Бутрик Е.Е., Соловьев С.В. Моделирование APT-атак, эксплуатирующих уязвимость ZEROLOGON // Вопросы кибербезопасности. 2021. № 6(46). С. 47-61. DOI:10.21681/2311-3456-2021-6-47-61.
14. Фаткиева Р.Р. Комплекс моделей для оценивания сетевой безопасности автоматизированных систем управления предприятием // Труды СПИИРАН. 2020. № 3. С. 621-643. DOI: 10.15622/sp.2020.19.3.6.
15. Андрещев И.А., Будников С.А., Гладков А.В. Полумарковская модель оценки конфликтной устойчивости информационной инфраструктуры // Вестник ВГУ. Серия: системный анализ и информационные технологии. 2017. № 1. С. 10-17.
16. Горин А.Н., Будников С.А. Применение программной среды Matlab для имитационного моделирования сложных систем военного назначения // Системы управления, связи и безопасности. 2019. № 1. С. 221-251. DOI: 1024411/2410-9916-2019-10114.
17. Дровникова И.Г., Паринов М.Л., Овчинникова Е.С. Аналитическая модель оценки вероятности реализации сетевых атак в динамике конфликтного взаимодействия с системой защиты информации в автоматизированных системах органов внутренних дел // Вестник Воронежского института МВД России. 2021. № 4. С. 43-56.
18. Колосок И.Н., Гурина Л.А. Оценка показателей киберустойчивости систем сбора и обработки информации в ЭЭС на основе полумарковских моделей // Вопросы кибербезопасности. 2021. № 6 (46). С. 2-11. DOI:10.21681/2311-3456-2021-6-2-11.
19. Arltova M., Fedorova D. Selection of Unit Root Test on the Basis of Length of the Time Series and Value of AR(1) Parameter. Statistika. 2016. vol. 96(3). pp. 47-64.

MODEL AND PARAMETRIC OPTIMIZATION OF PROACTIVE PROTECTION OF THE EMAIL SERVICE FROM NETWORK INTELLIGENCE

Gorbachev A.A.[4]

The purpose of the study: *to increase the security of the e-mail service of information systems in the conditions of network intelligence.*

Methods used: *methods of mathematical statistics, random processes research, mathematical programming, heuristic optimization algorithms were used to achieve the research goal.*

The result of the study: *a semi-Markov model of proactive protection of the e-mail service from network intelligence has been developed, which allows determining the probabilistic and temporal characteristics of the process of transmitting e-mail messages. Based on traffic analysis, statistical hypotheses about the types of distributions of the time of occurrence of events, under the influence of which the system under study evolves in a discrete set of states, are verified, point and interval estimates of the values of the parameters of these distributions are performed. The solution of the system of linear integral Volterra equations with integral kernels of the difference type was carried out*

4 Alexander A. Gorbachev, Applicant for academic degree of candidate of Ph.D., Krasnodar, Krasnodar Higher Military Orders of Zhukov and the October Revolution Red Banner School named after General of the Army S.M. Shtemenko, Krasnodar, Russia. E-mail: infosec23.00@mail.ru

using numerical methods of the Laplace transform. The problem of vector optimization is solved to determine the optimal parameters for configuring e-mail messages, allowing to maximize the effectiveness of the protection of the e-mail service, the robustness of the simulated system and minimize overhead costs under appropriate restrictions. The extremum of the objective functions was found using a bioinspired particle swarm algorithm. Scalarization of Pareto-optimal estimates was carried out using the ideal point method.

Scientific novelty: *it consists in developing a model and solving the problem of optimizing the parameters of the e-mail service in the conditions of network intelligence using the mathematical apparatus of semi-Markov processes, numerical methods of Laplace transformation, parametric evaluation of statistical characteristics of the model, scalarization of a multi-criteria optimization problem by the ideal point method and search for the extremum of objective functions using the particle swarm algorithm.*

Keywords: *semi-Markov process, vector optimization, network trap, robustness, Laplace transform, bioinspired algorithm.*

References

1. Maximov R.V., Sokolovsky S.P., Telenga A.P. Model of client-server information system functioning in the conditions of network reconnaissance. CEUR Workshop Proceeding. 2019. pp. 44-51.
2. Kumari A., Agrawal N., Umesh L. Attack over Email System: Review. International Journal of Scientific Research and Engineering Trends. 2017. vol. 3. pp. 200-206.
3. Schneider M., Shulman H., Sidis A., Sidis R., Waidner M. Diving into Email Bomb Attack. 50th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN). 2020. pp. 286-293. DOI: 10.1109/DSN48063.2020.00045.
4. Castillo D.P., Regidor F.M., Higuera J.B., Higuera J.R., Montalvo J.A. A new mail system for secure data transmission in cyber physical systems. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. 2020. vol. 28(2). pp. 23-48. DOI: 10.1142/S0218488520400127.
5. Zobal L., Kolar D., Kroustek J. Exploring current e-mail cyber threats authenticated SMTP honeypot. 17-th International Conference on Security and Cryptography. Paris. 2020. pp.253-262.
6. Douzi S., AlShahwan F.A., Lemoudden M., Ouahidi B. Hybrid email spam detection model using artificial intelligence. International Journal of Machine Learning and Computing. 2020. vol. 10(2). pp. 316-322. DOI: 10.18178/ijmlc.2020.10.2.937.
7. Sokolovskij S. P., Gorbachev A. A. Sposob proaktivnoj zashhity pochtovogo servera ot nezhelatel'nyh soobshhenij jelektronnoj pochty // Voprosy oboronnoj tehniki. Serija 16: Tehnicheskie sredstva protivodejstvija terrorizmu. 2021. № 3-4 (153-154). S. 31-40.
8. Walla S., Rossow C. MALPITY: Automatic identification and Exploitation of Tarpit Vulnerabilities in Malware. 2019 IEEE European Symposium on Security and Privacy (EuroS&P). 2019. pp. 590-605. DOI: 10.1109/EuroSP.2019.00049.
9. Maximov R.V., Sokolovsky S.P., Telenga A.P. Methodology for substantiating the characteristics of false network traffic to simulate information system. CEUR Workshop Proceeding. 2021. pp. 115-124.
10. Maximov R.V., Sokolovsky S.P., Telenga A.P. Honeypots network traffic parameters modelling. CEUR Workshop Proceeding. 2021. pp. 229-239.
11. Petrov M.Ju., Fatkieva R.R. Model' sinteza raspredelennyh atakujushhih jelementov v komp'juternoj seti // Trudy uchebnyh zavedenij svjazi. 2020. № 2. S. 113-120. DOI:10.31854/ 1813-324X-2020-6-2-113-120.
12. Evnevich E.L., Fatkieva R.R. Modelirovanie informacionnyh processov v uslovijah konfliktov // Voprosy kiberbezopasnosti. 2020. № 2(36). S. 42-49. DOI:10.21681/2311-3456-2020-2-42-49.
13. Budnikov S.A., Butrik E.E., Solov'ev S.V. Modelirovanie APT-atak, jekspluatirujushhih ujazvimost' ZEROLOGON // Voprosy kiberbezopasnosti. 2021. № 6(46). S. 47-61. DOI:10.21681/2311-3456-2021-6-47-61.
14. Fatkieva R.R. Kompleks modelej dlja ocenivanija setevoj bezopasnosti avtomatizirovannyh sistem upravlenija predprijatiem // Trudy SPIIRAN. 2020. № 3. S. 621-643. DOI: 10.15622/sp.2020.19.3.6.
15. Andreshhev I.A., Budnikov S.A., Gladkov A.V. Polumarkovskaja model' ocenki konfliktnoj ustojchivosti informacionnoj infrastruktury // Vestnik VGU. Serija: sistemnyj analiz i informacionnye tehnologii. 2017. № 1. S. 10-17.
16. Gorin A.N., Budnikov S.A. Primenenie programmnoj sredy Matlab dlja imitacionnogo modelirovanija slozhnyh sistem voennogo naznachenija // Sistemy upravlenija, svjazi i bezopasnosti. 2019. № 1. S. 221-251. DOI: 1024411/2410-9916-2019-10114.
17. Drovnikova I.G., Parinov M.L., Ovchinnikova E.S. Analiticheskaja model' ocenki verojatnosti realizacii setevyh atak v dinamike konfliktnogo vzaimodejstvija s sistemoj zashhity informacii v avtomatizirovannyh sistemah organov vnutrennih del // Vestnik Voronezhskogo instituta MVD Rossii. 2021. № 4. S. 43-56.
18. Kolosok I.N., Gurina L.A. Ocenka pokazatelej kiberustojchivosti sistem sbora i obrabotki informacii v JeJeS na osnove polumarkovskih modelej // Voprosy kiberbezopasnosti. 2021. № 6 (46). S. 2-11. DOI:10.21681/2311-3456-2021-6-2-11.
19. Arltova M., Fedorova D. Selection of Unit Root Test on the Basis of Length of the Time Series and Value of AR(1) Parameter. Statistika. 2016. vol. 96(3). pp. 47-64.