УДК 004.056, 519.816

А.В. КРАВЧУК

МОДЕЛЬ ПРОЦЕССА УДАЛЕННОГО АНАЛИЗА ЗАЩИЩЕННОСТИ ИНФОРМАЦИОННЫХ СИСТЕМ И МЕТОДЫ ПОВЫШЕНИЯ ЕГО РЕЗУЛЬТАТИВНОСТИ

Кравчук А.В. **Модель процесса удаленного анализа защищенности информационных систем и методы повышения его результативности.**
Аннотация. Рассмотрены подходы к проведению анализа защищенности информационных систем. Предложена модель процесса анализа защищенности информационных систем на основе теории принятия решений. Рассмотрены существующие методы решения проблемы марковских процессов принятия решений в условиях частично наблюдаемой среды.
Ключевые слова: анализ защищенности, тестирование на проникновение, компьютерная атака, марковские процессы, принятие решений, частичная наблюдаемость.

Kravchuk A.V. **The Model of Process of Remote Security Analysis of Information Systems and Methods of Improving it's Performance.**
Abstract. This article considers approaches to remote security analysis of information systems. The model of process of remote security analysis of information systems using decision making theory is proposed. Existing methods to solve partially observable Markov decision processes problem are reviewed.
Keywords: information security analysis, penetration testing, computer attack, Markov processes, decision making, partially observability.

1. Введение. Существующие средства удаленного анализа защищенности информационных систем (ИС) можно условно разделить на 2 класса:

– средства сбора сведений о сетях и обнаружения уязвимостей (Nmap, NetCat, Nessus, MaxPatrol и другие);

– средства тестирования на проникновение (Core Impact, Immunity Canvas, Metasploit Framework и другие).

Соответственно, анализ защищенности ИС предполагает выполнение двух проверок:

1. Проверка на наличие уязвимостей в ИС. Результаты данной проверки характеризуются высоким уровнем ошибок I рода (ложные срабатывания средства сбора сведений о сети и/или анализатора уязвимостей) [1]. Также имеют место ошибки II рода (пропуск уязвимых состояний ИС).

2. Для повышения достоверности анализа защищенности ИС проводится тестирование на проникновение (ТнП), позволяющее оценить информационную безопасность компьютерных сетей посредством формирования и исполнения различных компьютерных атак. Результаты ТнП также характеризуются ошибками II рода, которые обусловлены следующими факторами:

– неполнотой баз данных с эксплойтами (всегда существует вероятность существования эксплойта нулевого дня);

– ситуациями, в которых соотношения между уязвимостями, конфигурациями сети или составляющими её узлами позволяют нанести ИС более серьезный ущерб, нежели использование одиночных уязвимостей. Другими словами, ошибки II рода в данном случае обусловлены не двойками <уязвимость, эксплойт>, а составными компьютерными атаками (СКА). Под СКА будем понимать последовательные многоэтапные действия злоумышленника, включающие как действия по применению эксплойтов, так и удаленный сбор информации о компьютерной сети для формирования последующих этапов СКА, а также различные синтаксические преобразования вредоносного исполняемого кода [2, 3].

Одной из основных проблем при проведении анализа защищенности ИС является повышение результативности проводимого анализа защищенности, которое может быть достигнуто за счет автоматизации рассматриваемого процесса, а также за счет «интеллектуального» моделирования всех возможных сценариев организации СКА злоумышленником и выявления недостатков конфигурирования средств защиты информации (СЗИ). Целью ТнП является выявление максимального количества уязвимостей ИС и недостатков конфигурации СЗИ, последнее из которых возможно за счет использования рационального соотношения действий по сканированию сети и выявлению её уязвимых состояний, с одной стороны, и непосредственному применению эксплойтов, с другой стороны.

Таким образом, целью данной работы является построение модели процесса анализа защищенности ИС, которая соответствует процессу проведения СКА злоумышленником (далее – «модель СКА»). Модель СКА позволит адекватно отразить элементы компьютерной атаки и взаимосвязи между ними, а предлагаемые методы позволят повысить результативность анализа защищенности ИС.

2. Описание предлагаемой модели. Проведенный сравнительный анализ подходов к моделированию (СЗИ) и компьютерных атак позволил выделить следующие основные классы моделей, построенных с использованием теорий вероятностей; нечетких множеств; игр; графов; автоматов; сетей Петри; случайных процессов. Сравнительный анализ существующих моделей проводился в соответствии с основными фундаментальными проблемами системно-кибернетических исследований:

– построения исследовательской модели «система-среда»;

– анализа свойств системы;

– наблюдения системы;

– выбора альтернативных вариантов [4].

Пусть некоторый программный агент (ПА) управляет проведением СКА на целевую систему ЦС, под которой будем понимать компьютерную сеть, состоящую, по крайней мере, из одного узла. Рассмотрим ЦС, которая в произвольный момент времени может находиться в одном из N различных состояний, $S = \{s_0, s_1, \dots, s_N, \}$, где S – пространство состояний. Для дискретных систем запись уравнений в пространстве состояний основывается не на дифференциальных, а на разностных или рекуррентных уравнениях.

Состояние ЦС для ПА будет являться дискретной случайной величиной, то есть в результате проведения атакующих воздействий он может принять то или иное предположение о состоянии ЦС, причем неизвестно заранее, какое именно. С практической точки зрения пространство состояний представляет собой совокупность конфигураций узлов и топологии сети, а состояние ЦС может описываться следующими переменными: $Node = \{M_0, M_1, \dots, M_N\}$, $Port\_num = \{1 \dots 65535\}$, $Port\_state = \{open, closed, filtered\}$, $OS = \{WinXP, Linux - 2.2.6 \dots\}$, $Sys\_state = \{stable, vuln, compromised, updated\}$ и т.д.

ПА не имеет достаточной информации для того, чтобы сделать вывод о реальном состоянии ЦС. Он имеет возможность выполнять доступные ему действия. В каждый дискретный момент времени t_i в распоряжении ПА имеются множество допустимых действий $A = \{a_0, a_1, \dots a_M\}$, которое состоит из трех подмножеств:

– множества действий по применению эксплойтов, A_{exp};

– множества действий по сканированию сети, идентификации ОС, сервисов и их уязвимых состояний, A_{scan};

– множества действий по модификации синтаксических характеристик вредоносного кода, $A_{obfuscate}$.

Таким образом, $A = A_{\text{exp}} \cup A_{scan} \cup A_{obfuscate}$.

Фактическое состояние ЦС является скрытым от непосредственного наблюдения ПА. Отсутствие возможности прямого наблюдения состояния ЦС обусловлено следующими факторами:

1. Несмотря на различные реализации сетевых протоколов (в том числе стека протоколов в различных ОС) и клиент-серверной архитектуры, которые предоставляют широкие возможности по идентификации ОС, сервисов и их уязвимых состояний, в ходе протокольного обмена данными с ЦС даже при отсутствии влияния среды, под которой будем понимать СЗИ, имеет место априорная неопределенность идентификации состояния ЦС.

2. Существенное влияние на проведение СКА оказывает среда, представленная СЗИ (межсетевые экраны, системы обнаружения атак, антивирусное ПО, установленное на узлах сети, штатные средства ЗИ уровня ОС). По своим характеристикам среда является частично наблюдаемой.

При этом СКА должна закончиться в случаях:

– наступления некоторого события, которое его остановит (срабатывание СЗИ, неудачное применение эксплойта, обусловленное неудовлетворительными результатами идентификации уязвимых состояний ЦС);

– достижения цели СКА. Под целью СКА понимается перевод ЦС в некоторое терминальное (скомпрометированное) состояние и последующее выполнение деструктивных действий.

Каждое действие, выполняемое ПА, достигает намеченной промежуточной цели с вероятностью p_n. Значения p_n достигают максимальных значений при условии, что СЗИ не произвело ни одной записи об атаке, и минимальных, если СЗИ заблокировало сеанс связи (сессию). Для обозначения вероятности достижения ЦС состояния s_{t+1}, если в состоянии s_t ПА было выполнено действие $a \in A$, будем использовать выражение $T(s_t, a, s_{t+1}) = P(s_{t+1} | s_t, a)$ и называть его моделью перехода или функцией перехода. Модель перехода представляет собой трехмерную таблицу переходных вероятностей и может быть представлена динамической байесовской сетью.

В качестве допущения примем, что процесс проведения СКА является марковским процессом первого порядка, то есть текущее состояние ЦС зависит только от предыдущего состояния и не зависит от каких-либо более ранних состояний. Таким образом, вероятность перевода ЦС в состояние $s' = s_t$ из состояния $s = s_{t-1}$ зависит только от s и выполненного ПА действия a, а не от всей истории состояний и действий. Для оценивания фактического состояния ЦС ПА имеет возможность анализировать поступающие в ответ на выполняемые им действия символы наблюдения (реакции ЦС). На основании оценок состояния ЦС ПА принимает решение о дальнейших действиях. Для обозначения вероятности получения символа наблюдения $o \in O$ при нахождении ЦС в состоянии s после выполнения ПА действия a на предыдущем шаге будем использовать выражение $Z(s', a, o) = P(o_t = o \mid s_t = s', a_{t-1} = a)$ и называть его моделью наблюдения или функцией наблюдения.

Для учета результативности и скрытности СКА целесообразно использовать функцию вознаграждения (reward function) и функцию полезности (utility function). Функция полезности отображает последо-

вательность полученных символов наблюдений, выполненных ПА действий и состояний ЦС на вещественное число, которому в данной работе соответствует степень достижения цели программным агентом или результативность СКА. Она будет являться критерием или целевой функцией.

ПА должен принимать рациональные решения на основании анализа поступаемых символов наблюдения таким образом, чтобы максимизировать функцию полезности. В общем случае доход, полученный за несколько шагов, является случайной величиной, зависящей от начального состояния ЦС и принимаемых в каждый момент времени решений [5].

Условия ТнП накладывают ограничения на время проведения СКА. Будем считать, что временной интервал принятия решений конечен, то есть существует такое фиксированное критическое время $t_{\text{кр}}$, после которого проведение СКА не имеет смысла.

Конечность временного интервала принятия решений, с одной стороны, делает модель СКА более адекватной, а, с другой, значительно усложняет её, т.к. в данном случае оптимальное действие для конкретного состояния ЦС со временем может измениться. При выборе конечного временного интервала принятия решений оптимальная стратегия будет нестационарной. В случае бесконечного временного интервала оптимальная стратегия будет стационарной, т.к. нет смысла проводить различные действия для одного и того же состоянии ЦС в разное время. Следует отметить, что понятие «бесконечного интервала принятия решений» говорит лишь о том, что для выполнения действий не устанавливаются фиксированные сроки.

Несмотря на то, что временной интервал принятия решений конечен, может сложиться ситуация, когда $t_{\text{кр}} \longrightarrow \infty$ и в результате действий ПА ЦС не достигнет терминального состояния, то общая полезность, связанная с аддитивными вознаграждениями будет бесконечной. Для решения данной проблемы целесообразно использовать поправочный коэффициент $\gamma \in [0,1)$, называемый также коэффициентом обесценивания или коэффициентом переоценки. Он описывает предпочтение ПА текущих вознаграждений перед будущими вознаграждениями. Если коэффициент переоценки γ близок к 0, то вознаграждения, которые должны быть получены в отделенном будущем, рассматриваются как малозначащие. Он представляет модель изменения предпочтений злоумышленника во времени. Использование коэффициента обесценивания будет гарантировать, что значение функции полезности будет конечным. В реальных условиях проведения СКА количество шагов – конечно, однако введение коэффициента обесценивания по-

зволяет повысить производительность алгоритма поиска «оптимальной» стратегии.

С учетом сказанного выше функция полезности СКА примет вид:

$$U_h\left(s_0, s_1, \dots, s_{t_{\text{кр}}};\ a_0, a_1, \dots, a_{t_{\text{кр}}}\right) = R(s_0, a_0) + \gamma R(s_1, a_1) + \gamma^2 R(s_2, a_2) \dots + \gamma^{t_{\text{кр}}} R(s_{i-1}, a_{i-1}),\ i = \overline{0, t_{\text{кр}}}. \quad (1)$$

Тогда любая оптимальная стратегия в обобщенном представлении удовлетворяет следующему соотношению:

$$\pi^* = \underset{\pi}{argmax}\, M\left[\sum_{t=0}^{t_{кр}-1} (\gamma^t R(s_t, a_t) \mid \pi, s_0 = s)\right]. \quad (2)$$

Последним, не описанным свойством СКА недостающим для формирования её модели, является неопределенность среды. Если абстрагироваться от неопределенности среды, то представляется возможным рассматривать получаемые символы наблюдений, $o_k \in O$, в качестве состояний ЦС, $s_i \in S$.

Однако результаты проведенных экспериментов показывают, что данный подход неприемлем, поскольку ЦС может находиться в таких скрытых состояниях, что при одинаковых символах наблюдения, получаемых в ответ на атакующие воздействия ПА, перевод ЦС в другое состояние потребует различных действий. Данное противоречие между теорией и практикой возникает вследствие допущения о марковских свойствах процесса проведения СКА и неопределенностями среды межсетевого взаимодействия, обусловленных функционированием СЗИ и структурно-функциональными характеристиками используемых протоколов передачи данных.

Таким образом, возникает проблема потери информации об истории предыдущих действий и наблюдений. В пошаговом представлении история может быть представлена в виде $h_t = \{\langle a_0\rangle, \langle o_1, a_1\rangle, \dots, \langle o_{t-1}, a_{t-1}\rangle, \langle o_t\rangle\}$. Хранение истории предыдущих действий и наблюдений требует больших затрат памяти и приводит к усложнению и без того информационно насыщенной модели. Вместо этого представляется возможным свести всю необходимую информацию о предыдущих атакующих воздействиях и полученных символах наблюдений в доверительные состояния, которые позволят решить проблему потери информации об истории предыдущих действий и наблюдений.

Доверительное состояние, b – дискретное распределение апостериорных вероятностей на множестве фактических состояний ЦС S, сопоставляющее каждому состоянию ЦС вероятность нахождения в нём. Доверительное состояние b есть стохастический вектор в момент времени t:

$$b_t = \langle b(s_0), b(s_1), \dots, b(s_{N-1}) \rangle, \quad (3)$$

где $N = |S|$ - количество возможных состояний ЦС, а $b_t(s_i)$ представляет собой вероятность нахождения ЦС в состоянии $s_i, i = \overline{1,N}$ в момент времени t. В соответствии с основным постулатом теории вероятностей $0 \leq b_t(s_i) \leq 1 \; \forall \; s_i \in S$. Позиция элемента (значения вероятности) в стохастическом векторе соответствует номеру состояния ЦС. Во избежание понятийной путаницы, отметим, что $b_t \neq b_t(s_i)$ – в первом случае, b_t – это случайный вектор, сформированный из значений вероятностей нахождения ЦС в состояниях $s_i \in S$, а во втором, $b_t(s_i)$ – конкретное значение апостериорной вероятности нахождения ЦС в состоянии $s_i \in S$ в момент времени t.

Каждый элемент доверительного состояния можно представить в следующей форме:

$$b_t(s) = P(s_t = s | h_t, b_0), \quad (4)$$

В работе [6] было доказано, что в любой момент времени t доверительное состояние b_t является достаточной статистикой последовательности полученных ПА символов наблюдений и выполненных на их основании действий до момента времени t. Доверительное состояние можно рассматривать как текущее состояние информации о внутреннем состоянии ЦС или как информационный вектор ПА, на основании которого он принимает решения (проводит СКА).

Таким образом, история выполненных действий и полученных символов наблюдения может быть представлена в виде информационного стохастического вектора ПА или апостериорного распределения вероятностей, каждый элемент которого есть:

$$b_t(s) = P(s_t = s | o_t, a_{t-1}, o_{t-1}, \dots, a_0). \quad (5)$$

Динамика изменения информационного вектора b является критическим фактором для проведения ТнП.

Введение доверительных состояний позволяет произвести декомпозицию проблемы рационального выбора атакующих воздействий с учетом «оптимальной» стратегии и блока оценивания фактического состояния ЦС на основании получаемых символов наблюдений.

С учетом введенных параметров и характеристик предметной области и математического аппарата МППРЧНС представляется целесообразным представить реализацию решения научной задачи поиска рациональной стратегии проведения СКА в виде интеллектуального модуля, представленного на рисунке 1.

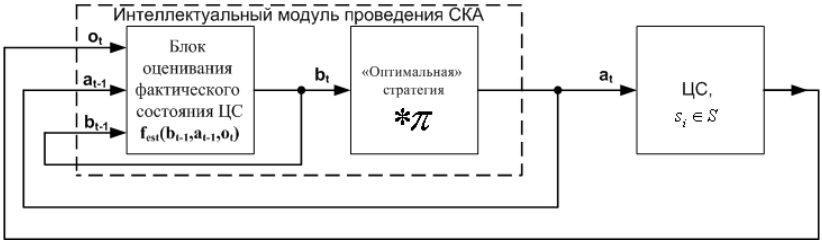

Рис. 1. Интеллектуальный модуль проведения СКА

Первый элемент интеллектуального модуля проведения СКА представлен блоком оценивания фактического состояния ЦС (БОФСЦС). Он отвечает за актуализацию распределения вероятностей нахождения ЦС в скрытых (ненаблюдаемых) состояниях, составляющих доверительного состояние ЦС. Поддержание распределения вероятностей в актуальном состоянии позволяет корректно отражать фактическое состояние ЦС и выполнять действия рациональным способом.

В качестве входных данных БОФСЦС принимает предыдущее доверительное состояние b_{t-1}, предыдущее атакующее воздействие a_{t-1} и текущий символ наблюдения o_t. На выходе БОФСЦС рекурсивным способом генерируется обновленное доверительное состояние или информационный вектор программного агента.

БОФСЦС позволяет рекурсивно рассчитать доверительное состояние за счет применения теоремы Байеса к двум базовым моделям: модели переходов $T(s,a,s')$ и модели наблюдения $Z(s,a,o)$. Если доверительное состояние не может быть рассчитано БОФС, то вероятность перехода $b_{t-1} \longrightarrow b_t$ будет равна нулю. При этом, учитывая введенные доверительные состояния, доверительная модель перехода $T(s,a,s')$ примет вид:

$$\begin{gathered}\tau(b,a,b') = P(b_t|a_{t-1},b_{t-1}) = P(b_t = b'|a_{t-1} = a,\ b_{t-1} = b) = \\ = \sum_{o\in O} P(b'|a,b,o) * P(o|a,b\ ) == \sum_{o\in O} P(b_t = b'|a_{t-1} = \\ a, b_{t-1} = b, o_t = o) * P(o_t = o|a_{t-1} = a, b_{t-1} = b\ ),\end{gathered} \quad (6)$$

где:

$$P(b'|a,b,o) = P(b_t|a_{t-1},b_{t-1},o_t) = \begin{Bmatrix} 1, f_{est}(b_{t-1},a_{t-1},o_t) = b_t \\ \\ 0, \text{в противном случае} \end{Bmatrix}. \quad (7)$$

Стоит отметить, что при выполнении атакующих воздействий $A_{scan} \in A$ переход ЦС из одного состояния в другое отсутствует. Данный класс действий влияет только на динамику изменения информационного стохастического вектора ПА.

Тогда каждый элемент обновленного доверительного состояния b_t в момент времени t может быть вычислен рекурсивно на основании апостериорной вероятности нахождения ЦС в состоянии $s \in S,$ действия a_{t-1} и символа наблюдения o_t, полученного в результате выполнения действия a_{t-1}.

В соответствии с работами [7,8] обновленный элемент доверительного состояния или обновленную апостериорную вероятность нахождения ЦС в состоянии $s' = s_t$ можно вычислить по формуле:

$$\begin{gathered} b_t(s') = P(s' \mid b_{t-1}, a_{t-1}, o_t) = \frac{P(s', b_{t-1}, a_{t-1}, o_t)}{P(b_{t-1}, a_{t-1}, o_t)} = \\ = \frac{P(o_t \mid s', b_{t-1}, a_{t-1}) P(s' \mid b_{t-1}, a_{t-1}) P(b_{t-1}, a_{t-1})}{P(o_t \mid a_{t-1}, b_{t-1}) P(b_{t-1}, a_{t-1})} \\ = \frac{P(o_t \mid s', b_{t-1}, a_{t-1}) P(s' \mid b_{t-1}, a_{t-1})}{P(o_t \mid a_{t-1}, b_{t-1})} = \\ = \frac{1}{P(o_t \mid a_{t-1}, b_{t-1})} Z(s', a_{t-1}, o_t) \sum_{s \in S} T(s, a_{t-1}, s') * b_{t-1}(s), \end{gathered} \tag{8}$$

где $P(o_t \mid a_{t-1}, b_{t-1})$ –полная вероятность получения символа наблюдения o_t в момент времени t после выполнения действия a_{t-1} из доверительного состояния b_{t-1}, представляемая как нормирующий коэффициент [9]. Суммирование производится по всем состояниям, из которых выполнение программным агентом действия a_{t-1} приводит к переводу ЦС в состояние s'.Нормирующий коэффициент определяется как:

$$P(o_t \mid a_{t-1}, b_{t-1}) = \sum_{s' \in S} Z(s', a_{t-1}, o_t) \sum_{s \in S} T(s, a_{t-1}, s') * b(s_t). \tag{9}$$

Нормирующий множитель вводится для того, чтобы сумма вероятностей обновленного доверительного состояния b_{t+1} была равна 1.

Выражение (8) задает функцию перехода из информационного состояния b_{t-1} в информационное состояние b_t и представляет собой Байесовский фильтр. Результатом рекурсивного применения формулы (8) ко всем состояниям $s \in S$ станет обновленное доверительное состояние, представленное стохастическим вектором b.

Таким образом, введение БОФСЦС позволяет ПА производить принимать решения на основании доверительного состояния или информационного вектора вместо истории атакующих воздействия и полученных символов наблюдений.

Вторым элементом является «оптимальная» (рациональная) стратегия, отображающая множество доверительных состояний в множество атакующих воздействий ПА:

$$^{*}\pi: \mathcal{B} \longrightarrow A. \tag{10}$$

Данный элемент отвечает за формирование атакующих воздействий и представляет собой функцию от доверительного состояния ЦС, отображающую текущее доверительное состояние в атакующее воздействие. Другими словами, после того, как ПА вычисляет доверительное состояние ЦС, он должен выбрать атакующее воздействие на основании доверительного состояния ЦС.

Рассмотренные выше параметры и характеристики задачи последовательного принятия решений в случае частично наблюдаемой среды с моделью перехода, моделью наблюдения, доверительными состояниями и мгновенными вознаграждениями позволяют сформировать модель СКА на основе математического аппарата марковских процессов принятия решений в условиях частично наблюдаемой среды (МППРЧНС). СКА задается следующим кортежем из 7 элементов:

$$\text{СКА} = \langle S, A, O, T(s,a,s'),\ Z(s,a,o)R(s,a), b_0 \rangle, \tag{11}$$

где:

1. $S = \{s_0, s_1, \ldots, s_N\}$ – конечное множество фактических (скрытых от непосредственного наблюдения) состояний целевой системы. Под множеством состояний ЦС понимается множество конфигураций ИТКС и составляющих её узлов.

2. $A = \{a_0, a_1, \ldots, a_M\}$– конечное множество действий, доступных ПА.

3. $O = \{o_0, o_1, \ldots, o_K\}$ – конечное множество символов наблюдений, получаемых ПА при нахождении ЦС в состояниях $S = \{s_0, s_1, \ldots, s_N\}$.

4. $T(s,a,s'): S \times A \times S \longrightarrow [0,1]$ – модель перехода, задающая условные переходные вероятности между состояниями, то есть $T(s,a,s') = P(s_t = s' |\ s_{t-1} = s, a_{t-1} = a)$ представляет вероятность того, что ЦС будет переведена из состояния $s_{t-1} = s$ в состояние $s_t = s'$ в результате выполнения ПА действия $a_{t-1} = a \in A$. Поскольку T является распределением условных переходных вероятностей, то:

$$\sum_{s' \in S} T(s,a,s') = 1, \forall (s,a). \tag{12}$$

В данной работе в качестве допущения принимается, что модель перехода $T(s, a, s')$ инвариантна по времени, то есть стохастическая матрица T не изменяется со временем. Для учёта переходных вероятностей, зависящих от времени, переменная состояния s должна включать в себя привязанную ко времени переменную.

5. $Z(s', a, o): A \times S \times O \longrightarrow [0,1]$ – модель наблюдения. Для каждого результирующего состояния $s' \in S$, каждого действия $a \in A$, каждого символа наблюдения $o \in O$, модель наблюдения $Z(s', a, o)$ определяет вероятность получения символа наблюдения $o_t = o \in O$ в результате выполнения ПА действия $a_{t-1} = a \in A$, приводящего к переходу ЦС в результирующее состояние s'. Другими словами $Z(s', a, o) = P(o_t = o \mid s_t = s', a_{t-1} = a)$. Данная условная вероятность определена для всех троек вида (s', a, o), для которых имеет место:

$$\sum_{o \in O} Z(s', a, o) = 1, \forall (s', a). \tag{13}$$

6. $R(s, a): S \times A \longrightarrow \mathbb{R}$ – функция мгновенного вознаграждения, присваивающая некоторое числовое значение, являющееся оценкой результативности и скрытности и называемое доходом или вознаграждением, за выполнение программным агентом атакующего воздействия a при нахождении ЦС в состоянии s. Вознаграждение $r_t = R(s, a)$ в момент времени t может принимать как положительные, так и отрицательные значения. В качестве допущения в данной работе предполагается, что вознаграждение ограничено сверху и снизу, $R_{min} < R < R_{max}$. Целью ПА является максимизация суммы вознаграждений, получаемых в процессе проведения СКА в течение некоторого времени $t < t_{\text{кр}}$. В общем виде цель ПА может быть представлена как максимизация математического ожидания суммарного вознаграждения:

$$M\left[\sum_{t=0}^{t_{\text{кр}}-1} \gamma^t r_t\right], \tag{14}$$

где $M[\ ]$ – математическое ожидание, r_t – мгновенное вознаграждение в момент времени t, и γ: $0 \leq \gamma < 1$ – коэффициент переоценки, введение которого гарантирует ограниченность вознаграждений, определяет предпочтения ПА по непосредственному применению эксплойтов, либо по увеличению точности идентификации ЦС посредством действий по её сканированию, а также позволяет повысить производительность алгоритма поиска «оптимальной» стратегии проведения СКА. Учитывая, что фактические состояния ЦС являются непосредственно ненаблюдаемыми, ПА принимает решения на основе информационно-

го вектора b_t. Тогда функция мгновенного вознаграждения, в которой вознаграждение зависит от текущего доверительного состояния $b_t = b \in \mathcal{B}$ и действия $a_t = a \in A$, выполненного ПА, из данного доверительного состояния, примет вид:

$$R_B(b,a) = \sum_{s \in S} b(s) * R(s,a),\ \forall b(s): \sum_{s \in S} b(s) = 1. \quad (15)$$

Приведенная формула означает, что ПА получает вознаграждение$r = R_B(b,a)$ за предположение о том, что ЦС пребывает в некотором состоянии. Данное допущение оправдано, поскольку значение функции оценивания фактического состояния ЦС $f_{est}(b_t, a_t, o_{t+1})$ вычисляется на основании получения символа наблюдения и модели перехода состояний ЦС. С формальной точки зрения мгновенное вознаграждение для случая доверительных состояний, $r = R_B(b,a)$, есть не что иное, как математическое ожидание. Тогда задача выбора последовательности атакующих воздействий, составляющих СКА, сводится к максимизации математического ожидания дохода, получаемого в $t_{\text{кр}}$-шаговом процессе проведения СКА при заданном начальном доверительном состоянии b_0.

7. b_0 – начальное дискретное распределение вероятностей на множестве фактических состояний ЦС S, сопоставляющее каждому состоянию из множества S вероятность нахождения в нём ЦС. Для решения проблемы частичной наблюдаемости среды программный агент должен осуществлять выбор атакующих воздействий на основании множества доверительных состояний ЦС $\mathcal{B} = \{b_0, b_1, \dots, b_n\}$, а не её фактических состояний S.

Таким образом, все рассмотренные элементы модели: состояния ЦС S, атакующие воздействия A, символы наблюдений O, вознаграждения R и три распределения вероятностей T, O, b_0 формируют вероятностную модель процесса проведения СКА. При этом четыре последние элемента задаются с помощью соответствующих матриц.

С учетом приведенных рассуждений, графическое представление модели СКА на основе марковских процессов принятия решений в условиях частично наблюдаемой среды примет вид, приведенный на рисунке 2.

Модель, представленная на рисунке 2 требует дополнительных пояснений. В момент времени $t-1$ ЦС пребывает в некотором состоянии $s_i \in S$, где $i = \overline{1,N}$. ПА, зная распределение вероятностей нахождения ЦС в состояниях $s_i \in S$, b_{t-1} выполняет действие $a_{t-1} \in A$, что приводит к мгновенному получению вознаграждения $r_{t-1} = R_B(b_{t-1}, a_{t-1})$ и вызывает переход в некоторое новое (или в то же) состояние $s \in S$ с вероятностью $P(s_t = s' | s_{t-1} = s, a_{t-1} = a)$.

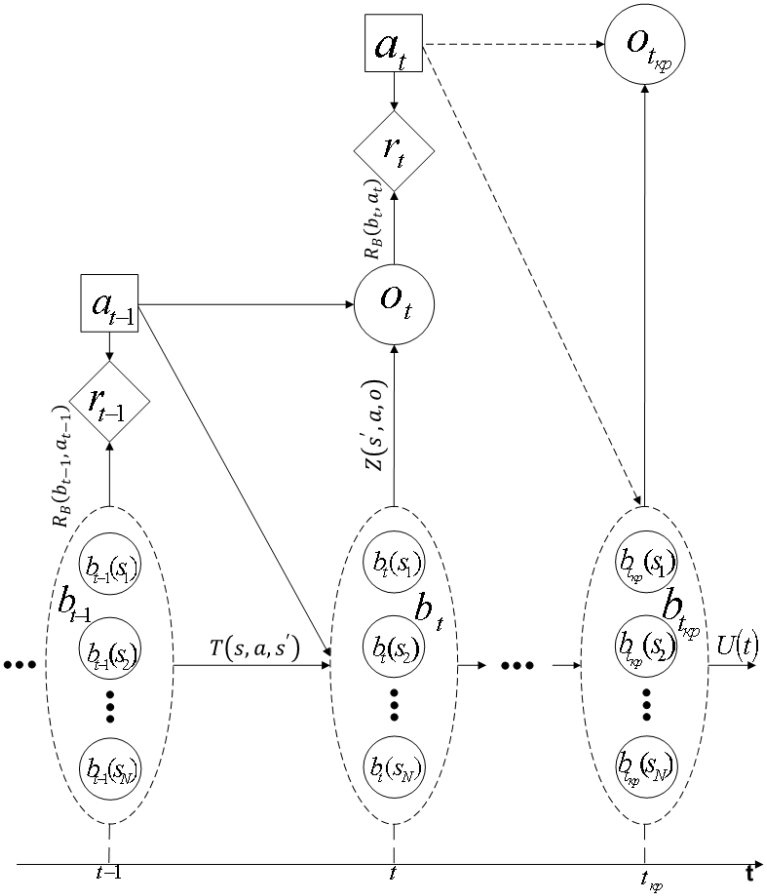

Рис. 2. Графическое представление модели СКА на основе МППРЧНС

Три компоненты модели S, A и T формируют ядро марковского процесса принятия решений и определяют динамику марковского процесса принятия решений в условиях частично наблюдаемой среды. В отличие от обычного марковского процесса принятия решений ПА не имеет возможности непосредственного наблюдения состояния ядра МППР в процессе принятия решений. Вместо этого ему предоставляется возможность получить символ наблюдения $o_t \in O$ с вероятностью $P(o_t = o \mid s_{t-1} = s, a_{t-1} = a)$.

В отличие от скрытой марковской модели, функционирующей автономно, МППРЧНС управляется действиями, выбираемыми про-

граммным агентом. Стоит отметить, что допущение о марковских свойствах процесса проведения СКА не всегда соответствует действительности. Для устранения этого недостатка были введены доверительные состояния. Соответственно, решение о выполнении действия ПА принимает на основании информации о доверительном состоянии ЦС.

Решением задачи МППРЧНС является нахождение оптимальной стратегии, то есть последовательности атакующих воздействий $\{A_t\}_{t=0,1,\dots,t_{\text{кр}}}$, приводящей к достижению цели СКА. Качество стратегии оценивается функцией полезности, являющейся математической функцией от мгновенных вознаграждений. Целью ПА является оптимизация функции полезности, которая в свою очередь является функцией параметров скрытности и результативности СКА.

Тем не менее, из-за частичной наблюдаемости, поиск строго оптимальной стратегии влечет чрезмерные затраты вычислительных мощностей, а в большинстве случаев нахождение строго оптимальной стратегии невозможно. Кроме того, возникает ряд проблем при применении аппарата МППРЧНС в сфере информационной безопасности в целом и в ТнП в частности. Во-первых, это необходимость корректно задать пространство состояний S, которое будет представлено конфигурационными параметрами сети и составляющих ее узлов. Во-вторых, необходимость задать на основании экспериментальных данных 2 таблицы с вероятностями $T(s, a, s')$, $Z(s', a, o)$, и одну $R(s, a)$ экспертным способом, а также начальное доверительное состояние b_0. И, наконец, в третьих, вычислять после каждого выполненного действия новое доверительное состояние. Другими словами, решение задачи МППРЧНС упирается в «проклятие размерности».

3. Анализ методов решения задачи МППРЧНС. Решением задачи МППРЧНС является «оптимальная» стратегия, обеспечивающая максимум ожидаемого суммарного дохода. Стратегия $\pi: \mathcal{B} \longrightarrow A$ задает действие a для каждого доверительного состояния $b \in \mathcal{B}$ и порождает функцию ценности (value-function) $V(b, \pi)$, которая определяет ожидаемую сумму вознаграждений за выполнение стратегии π, начиная с доверительного состояния b.

Функция ценности может быть вычислена по следующей формуле:

$$V(b, \pi) = \mathrm{M}\left[\sum_{t=0}^{t_{\text{кр}}-1} \gamma^t R(s_t, a_t) | b, \pi\right]. \quad (16)$$

Графически стратегия может быть представлена как граф, узлами которого являются доверительные состояния, а дуги – действиями ПА.

Для решения задачи МППРЧНС, то есть для вычисления оптимальной стратегии выбора действий, за последние десятилетия было предложено множество алгоритмов, которые можно разделить на два класса: алгоритмы времени, близкого к реальному, и автономные алгоритмы.

Автономные алгоритмы [10-13] определяют наилучшее действие, подлежащее выполнению, для всех возможных ситуаций, до начала проведения процесса принятия решения, то есть до начала ТнП. Использование перечисленных алгоритмов для проведения ТнП сопровождается определенными трудностями, основной из которых является необходимость повторного перерасчета всей стратегии при малейшем изменении в конфигурации компьютерной сети. Наиболее перспективными среди автономных алгоритмов на данный момент являются точечные алгоритмы (point-based), которые обеспечивают перерасчет функции ценности только для некоторых выбранных доверительных состояний $b \in \mathcal{B}$.

Алгоритмы времени, близкого к реальному [14-18], разработаны для преодоления «проклятия размерности» и осуществляют планирование только для текущего информационного вектора ПА или доверительного состояния. Другими словами, позволяют вычислять около оптимальные локальные стратегии на каждом шаге принятия решений во время исполнения алгоритма. В ряде работ, например в [19], они называются агентно-ориентированными алгоритмами поиска. Их основными недостатками является ограничения, связанные с работой в реальном времени. Наиболее перспективным среди алгоритмов времени, близкого к реальному, является детерминированный разреженный алгоритм частично наблюдаемого дерева (Determinized Sparse Partially Observable Tree, DESPOT).

Несмотря на большое количество существующих алгоритмов поиска около оптимальной стратегии выполнения действий, ни один из них, взятый по отдельности не может быть использован для моделирования ТнП для средних и больших сетей (50 и больше узлов). Для решения данной проблемы предполагается разработать комбинированный алгоритм.

4. Заключение. В работе рассмотрен новый подход к моделированию компьютерных атак с позиции злоумышленника или, другими словами, для моделирования процесса удаленного анализа защищенности информационных систем. Подход был проверен с использованием двух свободно распространяемых пакетов программ – SARSOP и DESPOT [20].

Стоит отметить, что проблема нахождения около оптимальной стратегии является NP-трудной [21]. На данный момент применение марковских процессов принятия решений в условиях частично наблюдаемой среды для моделирования ТнП ограничивается малыми сетями (до 20 узлов). Для его применения к сетям большей размерности требуются дополнительные способы аппроксимации как по количеству вариантов, из которых надо выбирать решения (множества доверительных состояний $\mathcal{B}$), так и по количеству итераций рекуррентного алгоритма вычисления максимального значения функции ценности.

Таким образом, для использования методов МППРЧНС для проведения удаленного анализа защищенности информационных систем в средних и больших компьютерных сетях требуется проведение дальнейших исследований по повышению производительности поиска рациональной (около оптимальной) стратегии.

Литература

1. *Кравчук А.В., Еремеев М.А.* Анализ методов распознавания вредоносных программ // Вопросы защиты информации. 2014. №3. С. 44–51.
2. *Кравчук А.В., Еремеев М.А., Потерпеев Г.Ю.* Модель и методы дистанционного контроля мобильных персональных устройств // Материалы 22-й НТК «Методы и технические средства обеспечения безопасности информации». СПб. 2013. С. 24–26.
3. *Кравчук А.В., Еремеев М.А.* Подход к моделированию компьютерных атак // Материалы 23-й НТК «Методы и технические средства обеспечения безопасности информации». СПб. 2014. С. 69–71.
4. *Калинин В.Н., Резников Б.А., Варакин Е.И.* Теория систем и оптимального управления. В 2 ч. Ч.1. Основные понятия, математические модели и методы анализа систем // Л.:ВИКИ имени А.Ф. Можайского. 1979. 319 с.
5. *Макаров И.М., Виноградская Т.М., Рубчинский А.А.* Теория выбора и принятия решений: Учебное пособие // М.: Наука. 1982. 328 с.
6. *Astrom K.J.* Optimal control of Markov processes with incomplete state information // Journal of mathematical analysis and applications. 1965. no. 10. pp. 174-205.
7. *Стратонович Р. Л.* Условные марковские процессы и их применение к теории оптимального управления // М.: Изд-во МГУ. 1966. 319 с.
8. *Jazwinski A.H.* Stochastic processes and filtering theory // New York: Academic Press. 1970. 391 p.
9. *Ross S., Pineau J.* Online Planning Algorithms for POMDPs // Journal of Artificial Intelligence Research. 2008. №32. pp. 663–704.
10. *Hauskrecht M.* Value-function approximations for partially observable Markov decision processes // Journal of Artificial Intelligence Research. 2000. №13. pp. 33–94.
11. *Pineau J., Gordon G., Thrun S.* Point-based value iteration: an anytime algorithm for POMDPs // Proceedings of the International joint conference on artificial intelligence (IJCAI-03). 2003. pp. 1025–1032.
12. *Braziunas D., Boutilier C.* Stochastic local search for POMDP controllers // Proceedings of the 19-th National conference on artificial intelligence (AAAI-04). 2004. pp. 690–696.
13. *Smith T., Simmons R.* Point-based POMDP algorithms: improved analysis and implementation // Proceedings of the 21th conference on uncertainty in artificial intelligence (UAI-05). 2005. pp. 542–547.

14. *Satia J.K., Lave R.E.* Markovian decision processes with probabilistic observation of states // Management Science. 1973. vol. 20(1). pp. 1–13.
15. *Barto A.G., Bradtke S.J., Singhe S.P.* Learning to act using real-time dynamic programming // Artificial Intelligence. 1995. vol. 72 (1). pp. 81–138.
16. *Washington R.* BI-POMDP: bounded, incremental partially observable Markov model planning. // Proceedings of the 4th European conference on planning. 1997. pp. 440–451.
17. *McAllester D., Singh S.* Approximate planning for factored POMDPs using belief state simplification // Proceedings of the 15th annual conference on uncertainty in artificial intelligence (UAI-99). 1999. pp. 409–416.
18. *Shani G., Brafman R., Shimony S.* Adaptation for changing stochastic environments through online POMDP policy learning // Proceedings of the workshop on reinforcement learning in non-stationary environments, ECML. 2005. pp. 61–70.
19. *Koenig S.* Agent-centered search // AI Magazine. 2001. vol. 22(4). pp. 109–131.
20. Approximate POMDP planning software. URL: http://bigbird.comp.nus.edu.sg/pmwiki/farm/appl/ (дата обращения 18.01.2015).
21. *Lusena C., Goldsmith J., Mundhenk M.* Nonapproximability results for partially observable Markov decision processes // Journal of artificial intelligence research. 2001. vol. 14. pp. 83–103.

References

1. Kravchuk A.V., Eremeev M.A. [Analysis of malware recognition methods]. *Voprosy zaschity informatcii – The question of information protection*. 2014. vol. 3. pp. 44–51. (In Russ.).
2. Kravchuk A.V., Eremeev M.A., Poterpeev G.J. [Model and methods of remote control of mobile personal devices]. *Trydy 22-oi naychno-tehnicheskoi konferencii Metody i tehnicheskie sredstva obespecheniya informacionnoi bezopasnosti* [Proceedings of the 22-th Scientific and Technical Conference on Methods and Technical Security Facilities]. SPb. 2013. pp. 24–26. (In Russ.).
3. Kravchuk A.V., Eremeev M.A. [Approach to modeling computer network attacks]. *Trydy 23-ei naychno-tehnicheskoi konferencii Metody i tehnicheskie sredstva obespecheniya informacionnoi bezopasnosti* [Proceedings of 23-th Scientific and Technical Conference on Methods and Technical Security Facilities]. SPb. 2014. pp. 69–71. (In Russ.).
4. Kalinin V.N., Reznikov B.A., Varakin E.I. *Teorija sistem i optimal'nogo ypravlenia. V 2 ch. Ch.1. Osnovnye ponyatia, matematicheskie modeli i metody analiza sistem* [Theory of systems and optimal control. Main definitions, mathematical models and system analysis methods]. L.:VIKI imeni A.F. Mozhayskogo. 1979. 319 p. (In Russ.).
5. Makarov I.M., Vinogradskaya T.M., Rubchinsky A.A. *Teoria vybora i prinyatia reshinii: Uchebnoe posobie* [The theory of choice and decision-making: tutorial]. M.: Nayka. 1982. 328 p. (In Russ.).
6. Astrom K.J. Optimal control of Markov processes with incomplete state information. Journal of mathematical analysis and applications. 1965. no. 10. pp. 174–205.
7. Stratonovich R. L. Conditional Markov Processes and Their Application to the Theory of Optimal Control. M.: MGU. 1966. 319 p.
8. Jazwinski A.H. Stochastic processes and filtering theory. New York: Academic Press. 1970. 391 p.
9. Ross S., Pineau J. Online Planning Algorithms for POMDPs. Journal of Artificial Intelligence Research. 2008. vol. 32. pp. 663–704.
10. Hauskrecht M. Value-function approximations for partially observable Markov decision processes. Journal of Artificial Intelligence Research. 2000. vol. 13. pp. 33–94.

11. Pineau J., Gordon G., Thrun S. Point-based value iteration: an anytime algorithm for POMDPs. Proceedings of the International joint conference on artificial intelligence (IJCAI-03). 2003. pp. 1025–1032.
12. Braziunas D., Boutilier C. Stochastic local search for POMDP controllers. Proceedings of the 19-th National conference on artificial intelligence (AAAI-04). 2004 pp. 690–696.
13. Smith T., Simmons R. Point-based POMDP algorithms: improved analysis and implementation. Proceedings of the 21th conference on uncertainty in artificial intelligence (UAI-05). 2005. pp. 542–547.
14. Satia J.K., Lave R.E. Markovian decision processes with probabilistic observation of states. Management Science. 1973. vol. 20(1). pp. 1–13.
15. Barto A.G., Bradtke S.J., Singhe S.P. Learning to act using real-time dynamic programming. Artificial Intelligence. 1995. vol. 72(1). pp. 81–138.
16. Washington R. BI-POMDP: bounded, incremental partially observable Markov model planning. Proceedings of the 4th European conference on planning. 1997. pp. 440–451.
17. McAllester D., Singh S. Approximate planning for factored POMDPs using belief state simplification. Proceedings of the 15th annual conference on uncertainty in artificial intelligence (UAI-99). 1999. pp. 409–416.
18. Shani G., Brafman R., Shimony S. Adaptation for changing stochastic environments through online POMDP policy learning. Proceedings of the workshop on reinforcement learning in non-stationary environments, ECML. 2005. pp. 61–70.
19. Koenig S. Agent-centered search. AI Magazine. 2001. vol. 22 (4). pp. 109–131.
20. Approximate POMDP planning software. Available at: http://bigbird.comp.nus.edu.sg/pmwiki/farm/appl/ (accessed 18.01.2015).
21. Lusena C., Goldsmith J., Mundhenk M. Nonapproximability results for partially observable Markov decision processes. Journal of artificial intelligence research. 2001. vol. 14. pp. 83–103.

Кравчук Алексей Владимирович — адъюнкт кафедры систем сбора и обработки информации, Военно-космическая академия имени А.Ф. Можайского. Область научных интересов: защита информации, теория принятия решений. Число научных публикаций — 10. kvazikrav@yandex.ru; ул. Ждановская, д.13, Санкт-Петербург, 197198; р.т. +7(812)237-19-60.

Kravchuk Aleksey Vladimirovich — Ph.D. student of the information acquisition and data processing department, Mozhaisky Military Space Academy. Research interests: information security, decision making theory. The number of publications — 10. kvazikrav@yandex.ru; Zdanovskaya str.13, 197198, Saint-Petersburg, Russia; office phone +7(812)237-19-60.

РЕФЕРАТ

Кравчук А.В. **Модель процесса удаленного анализа защищенности информационных систем и методы повышения его результативности.**

Статья посвящена рассмотрению нового подхода к моделированию процесса удаленного анализа защищенности информационных систем и методам повышения его результативности. В основе предлагаемого подхода лежит теория принятия решений. Непосредственное моделирование реализовано с помощью марковских процессов принятия решений в условиях частично наблюдаемой среды (МППРЧНС). Для повышения результативности удаленного анализа защищенности информационных систем предлагается обзор существующих методов решения задачи МППРЧНС, их основных характеристик и возможностей по применению к тестированию информационных систем на проникновение (ТнП).

Моделирования анализа защищенности на основе МППРЧНС является новым подходом. Его сильной стороной является наиболее адекватное использование в модели характеристик предметной области. К основному недостатку, требующему разрешения, следует отнести необходимость подбора/разработки аппроксимирующих методов, позволяющих использовать модель МППРЧНС в компьютерных сетях средних и больших размеров.

SUMMARY

Kravchuk A.V. **The Model of Process of Remote Security Analysis of Information Systems and Methods of Improving it's Performance.**

The article is devoted to approval of new approach to modeling of process of remote security analysis of information systems and methods of improving it's performance. The heart of proposed approach is the decision theory. Direct modeling is performed using partially observable Markov decision processes (POMP). Review of existing methods to solve POMDP problem is carried out in order to improve the performance of remote security analysis of information systems. Also main characteristics and capabilities of these methods for using in penetration testing domain are provided.

Modeling of security analysis on POMDP basis represents a new approach. Main advantage is what it allows for using main characteristics of penetration testing process. It's main drawback, which have to be overcome, resides in necessity to select/develop approximation technique allowing for using POMDP model in computer networks of medium or big size.