МЕТОДИКА ФОРМИРОВАНИЯ ЗНАЧИМОГО МНОЖЕСТВА ПРАВИЛ КОРРЕЛЯЦИИ ДЛЯ ВЫЯВЛЕНИЯ РАСПРЕДЕЛЕННЫХ СОБЫТИЙ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

METHODOLOGY FOR FORMING A SIGNIFICANT SET OF THE RULES OF CORRELATION TO IDENTIFY DISTRIBUTED EVENTS OF INFORMATION SECURITY

A. Gaynov
I. Zavodtsev

Summary. In this script a technique for forming a significant set of the rules for SIEM-systems is proposed, which allows to identify and eliminate possible conflicts in the process of forming correlation rules in case of simultaneously setting complementary, parallel or interrelated relations between different security events. In general it allows reducing the number of information security incidents, which are not detected by other methods.

Keywords: an information security incident, a SIEM-system, a log-file.

Гайнов Артур Евгеньевич
Соискатель, Кубанский институт информационной защиты, г. Краснодар, ArturGaynov@mail.ru,
Заводцев Илья Валентинович
К.т.н., доцент, Кубанский институт информационной защиты, г. Краснодар, nilrs@mail.ru

Аннотация. В работе предложена методика формирования значимого множества правил для SIEM-систем, которая позволяет выявлять и устранять возможные конфликты в процессе формирования правил корреляции при одновременном задании в них дополняющих, параллельных или взаимосвязанных отношений между различными событиями безопасности. Это позволяет в целом снизить количество необнаруженных с помощью других способов инцидентов информационной безопасности.

Ключевые слова: инцидент информационной безопасности, SIEM-система, лог-файл.

В условиях интенсивного развития и внедрения информационных и телекоммуникационных технологий особое внимание уделяется вопросам обеспечения безопасности критически важных объектов. Выведение таких объектов из строя может привести к тяжелым и даже катастрофическим последствиям [1].

В связи с тем, что в настоящий момент ни одна из существующих SIEM-систем не обеспечивает полноценного управления инцидентами информационной безопасности (ИИБ) при постоянно увеличивающихся требованиях к защите информационных ресурсов, возникает необходимость разработки новых технических решений по построению компонентов SIEM-систем на основе других принципов.

При этом, в такой системе должно быть реализовано упреждающее управление безопасностью, а так же надежный и устойчивый сбор данных о событиях информационной безопасности (СИБ). Дополнительными требованиями к таким системам выступают следующие подходы [2, 3]:

- межуровневая корреляция СИБ и их многоуровневое моделирование;
- интеллектуальный (упреждающий) мониторинг безопасности.

Решение этой проблемы предполагается достигать за счет эффективных процедур кластеризации и корреляции СИБ, поступающих из различных слоев, а также в обеспечении требуемой полноты представления данных обо всех потенциальных критических СИБ.

В общем случае реализация алгоритма продукционного вывода в SIEM-системах включает следующую последовательность действий [3, 5]:

- формирование исходного множества правил для выполнения корреляционного анализа;
- задание фоновых условий и уровня глубины анализа для каждого правила;
- отбор значимых правил в действующее множество, на основании которых и будет выполняться поиск взаимосвязей между распределенными СИБ;
- выявление и устранение конфликтов во вновь сформированном множестве;
- проверка для каждого правила из действующего множества соответствие фактической глубины анализа заданной.

Однако у большинства существующих SIEM-систем процедура формирования новых правил корреляции выполняется вручную, с помощью экспертных методов. Что ставит качество обнаружения и расследования рас-

пределенных по времени и месту ИИБ в зависимость от уровня знаний конкретного эксперта и его навыков в выявлении взаимосвязей между разнородными СИБ.

Поэтому целесообразна разработка специальной процедуры отбора для корреляции только значимых правил, а также устранения возможных конфликтов между ними, что позволит расширить спектр обнаруживаемых ИИБ путем расследования не только простых событий деструктивного характера, но и составных, включающих разнесенные по времени и по месту СИБ. Эффективность выявления корреляционных связей между распределенными событиями будет обеспечиваться методикой формирования специального множества значимых правил корреляции и реализованными в ней возможностями разрешения конфликтов при отборе таких правил.

Цель статьи — разработка специализированной методики формирования правил корреляции, позволяющая учитывать в продукционных правилах причинные, дополняющие, параллельные или взаимосвязанные отношения между различными СИБ и, одновременно, выявлять и устранять возможные конфликты в процессе отборе таких правил в значимое множество правил корреляции.

Задание правил обычно реализуется на интуитивно понятном уровне, но выработка корректного набора правил по отношению к конкретной задаче в общем случае достаточно затруднительна. Это связано с субъективностью задания правил администратором безопасности, необходимостью учета в разрабатываемых им правилах различных фоновых условий, а также невозможностью применять с прежней эффективностью готовые правила при возникновении новой (нестандартной) ситуации в информационной системе (ИС). При этом администратор безопасности должен описать столько правил (сигнатур), сколько необходимо для эффективной работы средства анализа, однако количество случайных событий в ИС огромно, а количество возможных ИИБ постоянно растет. Все это приводит к конфликтам внутри самого множества правил, когда при последовательной обработке правил могут выдаваться незапланированные директивы, появляться пропуски или, наоборот, управляющий алгоритм будет попадать в петлю.

Поэтому в настоящем методике предлагаются дополнительные действия [4]:

- задание фоновых условий и исходного уровня глубины выполняемого анализа правилами;
- формирование исходного множества правил для выполнения корреляционного анализа;
- отбор значимых правил в действующее множество;
- выявление и устранение конфликтов среди отобранных правил;
- проверка для каждого правила из действующего множества соответствие фактической глубины анализа заданной.

В качестве фоновых условий определяют обстоятельства, влияющие на учет (рассмотрение) тех или иных признаков СИБ при проверке правил корреляции. Каждому из фоновых условий администратор безопасности присваивает коэффициент уверенности CFi.

Примером набора фоновых условий может служить применение на сетевом устройстве (сервере) операционной системы Linux (коэффициент уверенности $CF_1=0,1$) с командным интерпретатором bash (коэффициент уверенности $CF_2=0,2$). При чем, если используется интерпретатор bash с версией 4.2 и 4.3 (коэффициент уверенности $CF_3=0,3$), а в операционной системе отсутствует соответствующий ему патч (коэффициент уверенности $CF_4=0,4$), то общий коэффициент уверенности CF_S для этого набора фоновых условий составит:

$$CF_s = CF_1+CF_2+CF_3+CF_4 = 1, \quad (1)$$

в противном случае, при изменении любого из фоновых условий, соответствующий ему коэффициент уверенности CF_i устанавливается равным 0. В данном примере максимальный вес всех коэффициентов уверенности имеет CF_4, так как именно он оказывает самое существенное влияние возможность эксплуатации злоумышленником уязвимости CVE: 2014–6278. Фоновые условия, также как и признаки включаются в структуру правила (сигнатуру).

Кроме фоновых условий для каждого правила задают параметр глубины анализа $H_i = (V_i, T_i)$, определяющий временной интервал T, в течение которого с указанного источника данных собирается информация о СИБ, и объем данных V, содержащий информацию об этих событиях. Использование параметра глубины анализа основано на той теоретической предпосылке, что одно событие, происходящее в течение определенного временного промежутка, может являться причиной другого события.

Как правило, единица измерения временного интервала T составляет 1 минуту, а типовой временной интервал сбора данных составляет 1 сутки. Ограничение временного интервала связано со следующим: около 70% правил корреляции работают с событиями, которые произошли в течение суток, 20% — до одной недели, 5% — не более месяца. Оставшиеся — в интервале квартал или полгода.

Таблица 1. Анализ данных из журнала безопасности Windows 7, где использовано два значимых поля ($V_i = 2$)

№ п/п	Используемые поля из log-журнала (объем данных, V)	
	Код события	Ключевые слова
1	4624, Logon	Аудит успеха
2	4768, Account Logon	Kerberos Ticket Events
3	4688, Detailed Tracking	Process Creation

Таблица 2. Расширенный состав полей журналов регистрации

№ п/п	Используемые поля из log-журнала (объем данных, V)				
	Код события	Ключевые слова	Системное время	Пользователь	ID процесса
1	4672, Special Logon	Аудит успеха	2015–10–12T07:06:49.816368900Z	S-1–5–18	940

Объем данных V для правила определяется количеством значимых полей из журналов регистрации этого источника, используемых для корреляционного анализа признаков распределенных событий безопасности. На начальном этапе V задают минимального размера. Например, анализ данных из журнала безопасности операционной системы Windows 7 обычно имеет $V_i = 2$ (табл. 1), что зачастую достаточно для анализа локальных ИИБ.

Это связано с тем, что в среднестатистической системе аудита нормальным считается поток событий равный 8000–10000 событий в секунду (EPS), при этом общее количество данных от 50–80 источников с учетом разных типов событий и набора учитываемых полей может достигать десятков тысяч EPS, что оказывает существенную нагрузку на систему.

В тоже время для выявления распределенных ИИБ зачастую необходимо рассматривать расширенный состав полей журналов регистрации, объем данных может быть увеличен, например, до V=5 (табл. 2).

Параметр H для всех правил задается администратором безопасности. Основная цель — определить *средние* значения количества данных о распределенных событиях безопасности, которые необходимо получать для анализа от разных источников в различные временные интервалы (рабочий день, ночь, выходные и т.д.). При необходимости на дальнейших этапах обработки данных глубина анализа для каждого правила может быть увеличена вплоть до максимальных значений, определяемых наибольшим количеством полей в журналах регистрации и наибольшим периодом времени за который отслеживаются СИБ в ИС.

В алгоритме формирования значимого множества правил корреляции (рисунок 1) на этапе развертывания системы защиты администратор безопасности, исходя из своих знаний, задает исходное множество правил для выявления признаков деструктивных СИБ. Для этого формируется первоначальный список правил корреляции, содержащих совокупность возможных признаков (p) обнаруживаемого события или совокупности нескольких СИБ.

В общем случае, правила корреляции строятся на основе закономерностей и представляют собой выражения в виде:

$$B = A_1 \, AND \ldots AND \, A_n, \qquad (2)$$

где $A_1, \ldots, A_n, B$ — предикаты, при этом предикат B является целевой (THEN ACTION) частью, $A_1 \, AND \ldots AND \, A_n$ — условной (IF) частью, объединяющей признаки различных событий (совокупности событий) и фоновые условия для данной ИС.

В качестве признаков событий безопасности, подлежащих обнаружению, определяют признаки тех событий, которые влияют на общую защищенность всей информационной системы или отдельного ее элемента. Например, к учитываемым в правилах корреляции событиям безопасности относят данные:

- о попытках изменения полномочий учетных записей;
- о входе одного пользователя под разными учетными записями;
- о превышении среднего времени соединения между узлами;
- о большом количестве узлов в сети организации, пытающихся соединится с одним тем же внешним ресурсом.

В исходное (начальное) множество включают правила трех видов.

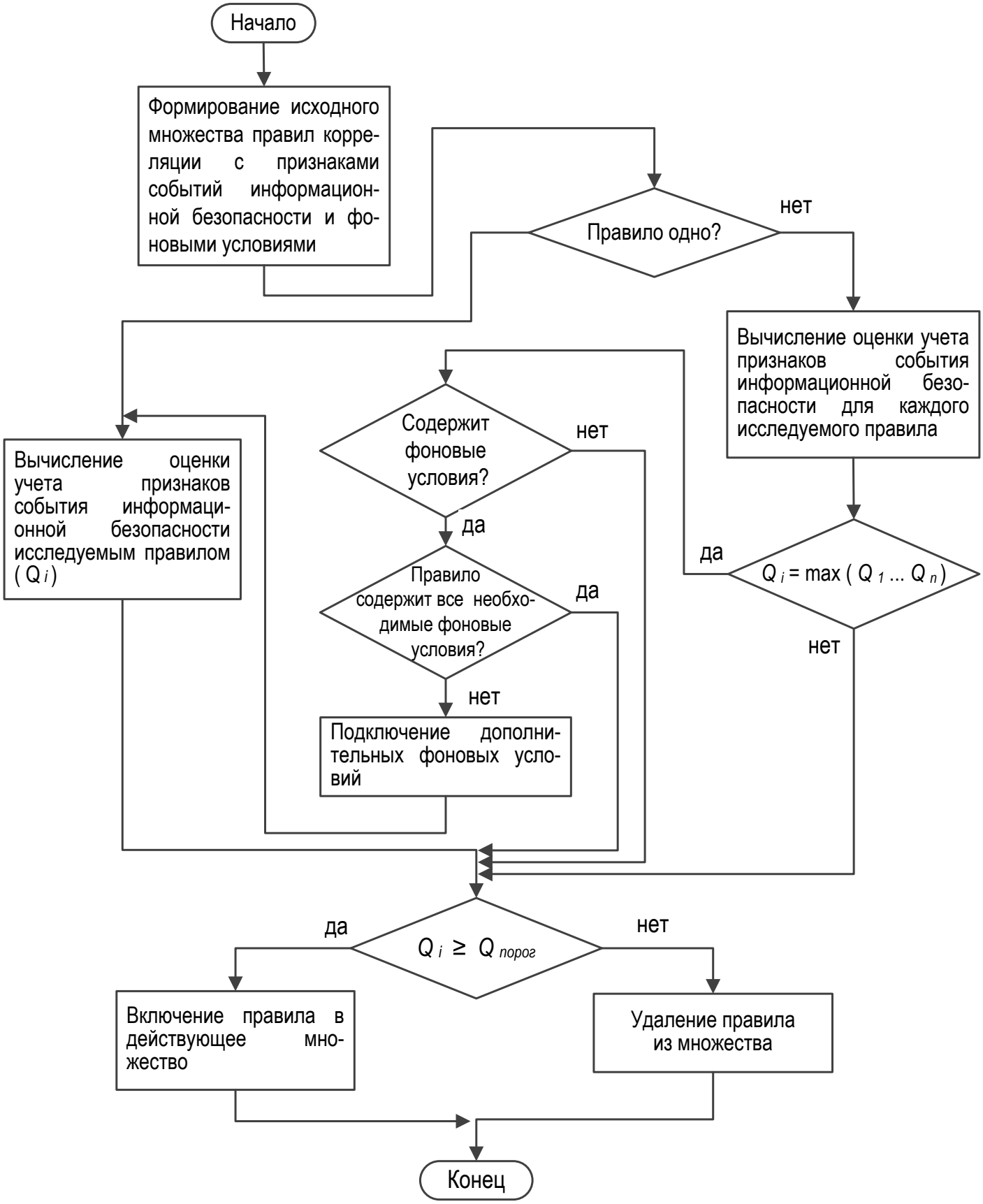

Рис. 1. Алгоритм формирования значимого множества правил корреляции

1. Правила, которые описывают признаки ИИБ, состоящего из одиночного СИБ. Например, правило осуществляет выдачу сигнала об опасности, если выполнена остановка (пауза) критичной службы на сервере:

Alert Event1 *= device action {stoppet, paused}*

AND matches filter {Windows/System Services and Auditing/Critical Services}

OR device vendor {Microsoft}, EventID {7036},

где *stoppet, paused* — признак полной или частичной остановки контролируемого сервиса,

Windows/System Services and Auditing/Critical Services — признак выявления совпадения с сервисом, который поставлен на контроль,

device vendor {Microsoft}, EventID {7036} — признак, указывающий на успешное завершение действия (переход в другое состояние) с учетом фонового условия, т.е. действие подлежит обязательному контролю, если работа ведется в среде ОС *Windows.*

2. Правила, которые описывают признаки ИИБ, состоящего из нескольких последовательных СИБ, произошедших за определенный период времени. Например, если получают сигнал от средства антивирусной защиты и выявляют последующее сканирование сети с того устройства (хоста), на котором сработал антивирус. Чтобы идентифицировать такой ИИБ, включающий распределенные СИБ, в формируемом правиле необходимо связать признаки сканирования сети и обнаружения вируса:

Alert Event2 *= select current_timestamp {'Critical' severity}, host_virus.host_ip*

AND host_scan {src_ip = host_virus.host_ip}

AND timer: within {1 minute},

где *'Critical' severity* — признак срабатывания средства антивирусной защиты,

src_ip = host_virus.host_ip — признак выявления совпадения между инфицированным вирусом хостом, и АРМ, осуществляющим сканирование сети,

timer: within {1 minute} — признак учета временного интервала, в течение которого, после инфицирования хоста, может начаться процесс сканирования локальной вычислительной сети.

3. Правила, которые идентифицируют признаки ИИБ на основе выявления отклонений (аномалий) от средних значений активности того или иного устройства (программы) за определенный период времени. Например, правило отслеживает превышение среднего показателя срабатываний антивируса за квартал:

Alert Event3 *= Current_Infected_Hosts {Host_Count_Threshold}*

OR Current_Virus_Count {Virus_Count_Threshold},

где *Host_Count_Threshold* — признак, показывающий текущее значение «среднего показателя» срабатываний антивируса за квартал,

Virus_Count_ Threshold — признак, определяющий заданное администратором безопасности на основе статистики за предыдущие периоды среднее значение этого показателя.

На следующем шаге из исходного множества производят отбор значимых правил в действующее множество, на основании которых и будет выполняться поиск взаимосвязей между признаками СИБ.

Для этого выполняют оценку количества правил в исходном множестве. Если в списке только одно правило, то для него сразу вычисляют оценку Q, отражающую степень учета признаков СИБ именно этим правилом (2):

$$\begin{cases} Q_i = \frac{p_i}{\sum_{i=1}^{n} p_i} \cdot \sum_{k=1}^{p_i} w_k^i \cdot CF_s^i, \\ \sum_{k=1}^{m} w_k = 1, \\ i = \overrightarrow{1,n},\ k = \overrightarrow{1,m} \end{cases}$$

где n — общее количество правил корреляций, включенных в действующее множество,

m — общее количество признаков СИБ, учитываемых правилами из действующего множества,

p_i — количество признаков СИБ, учитываемых i-м правилом,

w_k^i — весовой коэффициент k-*го* признака, учитываемого i-м правилом,

CF_s^i — суммарный коэффициент уверенности для всех учитываемых правилом фоновых условий.

Весовые коэффициенты признаков определяются заранее экспертным методом при составлении правил администратором безопасности в зависимости от ис-

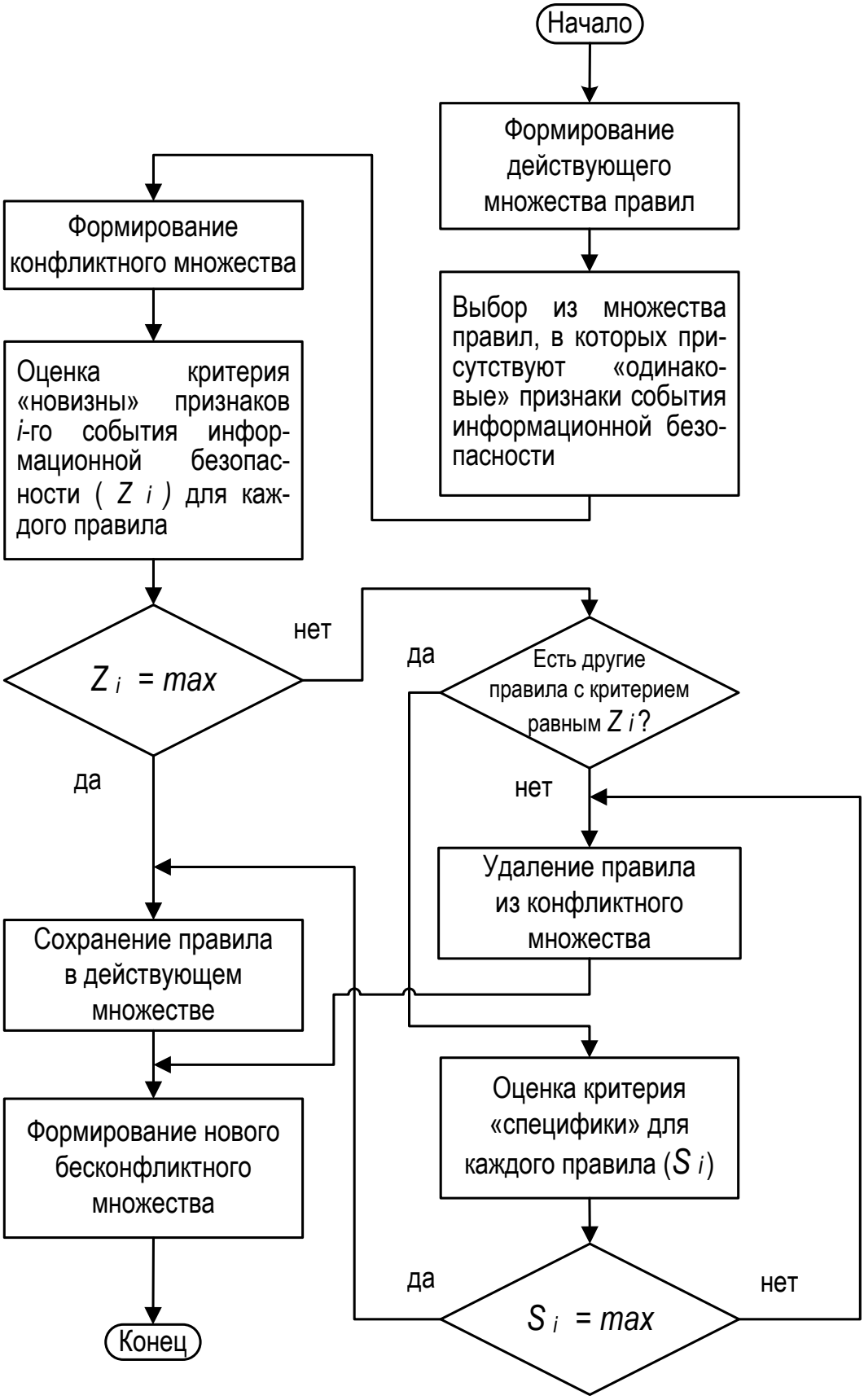

Рис. 2. Алгоритм выявления и устранения возможных конфликтов в множестве правил

пользуемых моделей защищенности ИС и описаний уязвимостей программного обеспечения. Таким образом, чем большее количество признаков учитывает конкретное правило, и чем они значимее, тем более весомую оценку Q будет иметь данное правило.

Если в список включено несколько правил, то аналогичную оценку Q вычисляют для каждого из них, но для последующего отбора выбирают правило, обладающее наивысшей оценкой среди проверяемых правил

$$Q_i = max\ (Q_1,\ \ldots,\ Q_n) \qquad (4)$$

Остальные правила дополнительно проверяют на наличие фактического учета в сигнатуре правила фоновых условий. Если в правиле использованы фоновые условия, то их проверяют на корректность и полноту описания для данной ИС. В случае неполноты учета требуемых фоновых условий сигнатурой правила, производят подключение дополнительных фоновых условий и повторное вычисление оценки Q. Подключение дополнительных фоновых условий осуществляет администратор безопасности экспертным методом.

Однако некоторые правила могут включать в себя небольшое количество признаков, что может быть обусловлено как особенностью самой сигнатуры расследуемого СИБ, так и малозначимостью самого правила. В последнем случае это приводит к необоснованному увеличению общего объема правил и, как следствие, снижению скорости анализа данных.

Поэтому для уменьшения в сформированном множестве количества малозначимых правил оценку всех отобранных правил Q дополнительно сравнивают с пороговым уровнем $Qпорог$. Задание пороговой оценки производится администратором безопасности в пределах 40–70% от уровня максимально возможного значения $Qmax$, что позволяет ограничить рост потребляемой памяти — ведь каждое анализируемая сигнатура СИБ требует порождения отдельного процесса для своего обслуживания.

Также с пороговым уровнем сравнивают оценку правил, если их сигнатура вообще не требует использования фоновых условий или уже корректно учитывает все требуемые для данной ИС фоновые условия.

При удовлетворении оценкой степени учета признаков СИБ для i-го правила корреляции заданному требованию, его включают в действующее множество правил, в противном случае удаляют из исходного множества. На следующем этапе между отобранными из исходного множества правил корреляции выявляют и устраняют возможные конфликты (рисунок 2). Это связано с необходимостью нивелирования субъективности администратора безопасности при задании правил, а также для отбора в множество наиболее эффективных правил из числа сформированных ранее с учетом изменения конфигурации и структуры ИС.

Для этого из полученного после оценки учета признаков множества выбирают те правила, которые одновременно содержат «одинаковые» признаки деструктивного события или совокупности СИБ, и формируют конфликтное множество правил.

Для каждого i-го правила конфликтного множества оценивают коэффициент «новизны» Z_i, отражающий близость признаков СИБ, рассматриваемых этим правилом, к признакам, учитываемых всеми другими правилами из этого множества. Для этого вычисляют произведение попарных коэффициентов совпадений признаков i-го правила с признаками каждого из l правил, включенных в конфликтное множество:

$$Z_i = \prod_{j=1}^{l} \frac{2p_{(j,i)}^{сов}}{p_i + p_j}, j = \overrightarrow{1, l},$$

где p_i и p_j — количество признаков СИБ, учитываемых i-м и j-м правилом, соответственно,

$p_{(j,i)}^{сов}$ — количество совпадающих признаков, включенных одновременно в j-е и i-е правила корреляции,

l — количество правил корреляции, отобранных в конфликтное множество.

Чем выше у правила значение коэффициента Z, тем больше в нем учитываемых признаков совпадает признаками, учитываемыми другими правилами. Приоритет отдают правилам с наивысшим значением коэффициента Z, эти правила доминируют над остальными и сохраняются в формируемом множестве.

В случае несоответствия правила данному требованию или когда несколько правил имеют равный приоритет, осуществляют сравнение правил по критерию «специфики» S_i, отражающему количество признаков СИБ, загружаемых в рабочую память для проверки:

$$S_i = \frac{p_i}{max(p_1,\ldots,p_n)}, \qquad (6)$$

где p_i — количество признаков СИБ, учитываемых i-м правилом,

$max(p_1,\ldots,p_n)$ — наибольшее количество признаков, учитываемых одним из правил, входящим в действующее множество.

Рис. 3. Алгоритм корректировки фактического значения глубины анализа для правил

Здесь предпочтение отдают тому правилу, применение которого требует проверки наибольшего количества признаков СИБ, остальные правила удаляются из множества. Из отобранных правил формируют действующее (бесконфликтное) множество правил корреляции.

После устранения конфликтов в сформированном множестве осуществляют корректировку фактического значения глубины анализа H для всех его правил (рисунок 3). Так как параметр H задается администратором безопасности и определяет средние значения количества данных о распределенных СИБ, получаемых для анализа от разных источников в различные временные интервалы, то такая процедура позволяет контролировать и, при необходимости, уточнять приоритеты в расследовании распределенных СИБ и, следовательно, повысить уверенность в обнаружении ИИБ.

На этом этапе выполняют проверку фактического значения глубины анализа Hi для i-го правила. Если правило имеет глубину анализа соответствующую заданному уровню, то его включают в действующее множество правил корреляции. Если проверяемое правило имеет глубину анализа менее заданного уровня, администратором безопасности увеличивается глубина анализа на один шаг и повторяют проверку фактического значения глубины анализа Hi. При необходимости на дальнейших этапах обработки данных глубина анализа для каждого правила может увеличиваться вплоть до максимальных значений.

Затем для проведения корреляционного анализа данных случайным равновероятным способом из всего действующего множества правил выбирают одно значимое правило и проверяют взаимосвязанность признаков СИБ как в пределах одного набора данных, так и из различных (в том числе и временных) наборов. Если обнару-

жены скрытые отношения между распределенными СИБ, выдают сигнал оповещения об обнаружении ИИБ.

Таким образом, в статье предложена методика формирования значимого множества правил для SIEM-систем, которая позволяет выявлять и устранять возможные конфликты в процессе формирования правил корреляции при одновременном задании в них дополняющих, параллельных или взаимосвязанных отношений между различными СИБ. Использование методики позволит в целом снизить количество необнаруженных с помощью других способов ИИБ.

ЛИТЕРАТУРА

1. Аналитический отчет «Обзор инцидентов информационной безопасности АСУ ТП зарубежных государств» (по материалам Интернет-изданий за 2008–2010 гг.) М.: НТЦ «Станкоинформзащита». [Электронный ресурс]. — Режим доступа: http://itdefence.ru
2. Котенко И. В., Саенко И. Б. Построение системы интеллектуальных сервисов для защиты информации в условиях кибернетического противоборства // Труды СПИИРАН. СПб.: Наука, 2012. Вып. 3(22). С. 84–100.
3. Заводцев И. В., Гайнов А. Е. Анализ требований, предъявляемых к современным средствам управления инцидентами информационной безопасности: Информационная безопасность — актуальная проблема современности. Совершенствование образовательных технологий подготовки специалистов в области ИБ / И. В. Заводцев, А. Е. Гайнов // сборник трудов VI–VII Всероссийской научно-технической конференции. — Краснодар: ФВАС, 2013. — Т1. — 365 с.
4. Патент РФ № 2610395 C1, 2017 г. МПК G06F21/55, Заводцев И. В., Гайнов А. Е.
5. Kruegel Ch., Valeur F. Intrusion Detection and Correlation. Challenges and Solutions. Springer Science + Business Media, Inc., 2005. ISBN: 0–387–23398–9. [Режим доступа]: http://link.springer.com/book/10.1007%2Fb101493.

Журнал «Современная наука: актуальные проблемы теории и практики»