DOI: 10.18372/2410-7840.19.11438
УДК 681.61

КОРРЕЛЯЦИЯ СОБЫТИЙ В SIEM-СИСТЕМАХ НА ОСНОВЕ НЕМОНОТОННОГО ВЫВОДА

Юрий Самохвалов, Сергей Толюпа

Рассмотрен подход к созданию корреляционных механизмов в SIEM-системах. В качестве логической основы таких механизмов предложено использование немонотонных правил умолчаний в сочетании с выводом резолюционного типа. Данный подход расширяет возможности классических корреляционных механизмов SIEM-систем за счет возможности одновременного использования как общезначимых продукционных правил, так и правил умолчаний, которые позволяют описывать типичные ситуации. Это дает возможность обрабатывать исключения без их предварительной идентификации и создавать более гибкие механизмы корреляции.
Ключевые слова: *SIEM-система, корреляция, сигнатура, продукция, немонотонный вывод, правила умолчаний.*

Введение. Разнообразная природа бизнес-процессов современных предприятий и необходимость ответа на возникающие вызовы их безопасности требуют организации системы защиты реального времени. Для решения данной задачи в настоящее время широко внедряются системы управления событиями и инцидентами информационной безопасности (ИБ), строящиеся на основе систем класса Security Information and Event Management (SIEM). Эти системы позволяют администраторам безопасности сфокусироваться на реальных угрозах, обеспечивая их средствами, позволяющими оперативно реагировать на угрозы безопасности информационной системы (ИС).

Основным инструментом SIEM является корреляционный анализ, средствами которого могут быть выявлены угрозы, описанные правилами корреляции, либо типовые угрозы на базе шаблона, а также отклонения от стандартных настроек конфигурации и политики безопасности, кроме того, могут быть определены причинно-следственные связи произошедших событий, имеющих отношение к информационной безопасности.

Существуют различные методы корреляции, которые подразделяют на две большие группы – сигнатурные и бессигнатурные [1-3]. Первые подразумевают наличие неких правил, определяемых пользователем, по которым выявляется инцидент. А вторые – это методы с обучением, то есть такие методы настроены производителями SIEM на определенные действия согласно обучающему набору данных. Поскольку эти методы разрабатываются и внедряются производителями SIEM-систем поэтому такими методами невозможно управлять, что является их основным недостатком. В свою очередь, сигнатурные методы отличаются гибкостью и эффективностью обнаружения угроз безопасности информации, что в современных условиях может значительно повышать эффективность систем защиты информации.

Анализ сигнатурных методов [4-6] показывает, что эти методы основаны на классических общезначимых продукционных правилах, которые позволяют получать только монотонные, общезначимые выводы. В реальных условиях процесс корреляции осуществляется, как правило, в условиях неполноты информации об угрозах. Поэтому в таких условиях строгие рассуждения становятся непригодными и на их место приходят, так называемые правдоподобные, предположительные заключения. Это положение ставит в разряд актуальных задачу разработки корреляционного механизма в SIEM-системах с возможностями немонотонного вывода, которое определяет *цель* и основное содержание статьи.

Изложение основного материала. Корреляционные механизмы являются основой подсистемы распознавания инцедентов SIEM-систем. Под *инцидентом* информационной безопасности понимается нежелательное событие (или совокупность событий), которое может скомпрометировать бизнес-процессы компании или непосредственно угрожает ее информационной безопасности [7]. Любое событие ИБ характеризуется некоторым множеством информационных признаков, которые делятся на две категории:

– Security Information – информация, связанная с безопасностью, поступающая от серверных и пользовательских приложений, от операционных систем, подсистем информационной безопасности;

– Security Events – информация, поступающая непосредственно от сетевого и телекоммуни-

кационного оборудования – коммутаторов, аппаратных брандмауэров, фильтров защиты от атак извне.

Далее, любой признак, в общем, с одной стороны, может иметь насколько непосредственных причин его возникновения, а с другой – каждая причина может индуцировать появление нескольких ассоциированных признаков. Кроме этого каждая причина может быть ассоциированным признаком более общей причины и т.д. Это положение позволяет на множестве информационных признаков построить причинно-следственную иерархическую структуру распознавания событий ИБ.

Пусть $P = \{ p_i \,|\, i = \overline{1, n} \}$ – множество информационных признаков. Первый уровень непосредственных причин представим множеством

$$E_1 = \{ E_j^1 \mid E_j^1 = (p_{j1},\ p_{j2}, ..., p_{jm}),\ j = \overline{1, l_1} \},$$

где $p_{im} \in P$ – ассоциированные признаки причины E_j^1. Следующие уровни формируются рекурсивно:

$$E_i = \{ E_j^i \mid E_j^i = (E_{j1}^{i-1},\ E_{j2}^{i-1}, ..., E_{jm}^{i-1}),\ j = \overline{2, l_i} \}.$$

Тогда всю совокупность признаков инцидентов и возможных причин их возникновения можно представить, например, следующей 3-х уровневой структурой, которую можно интерпретировать как классификатор инцидентов (рис. 1).

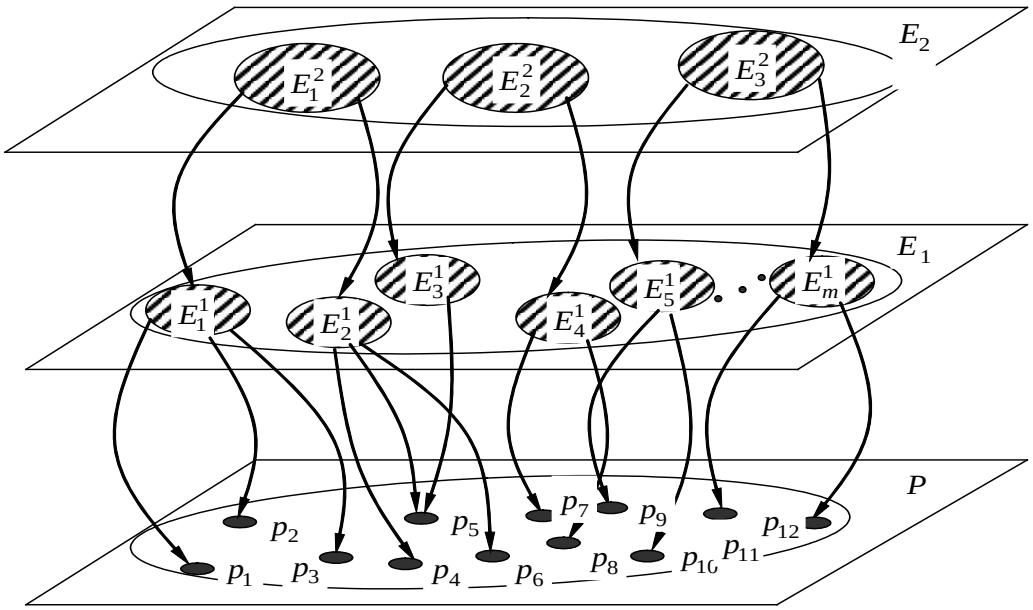

Рис. 1. Пример классификатора инцидентов

Учитывая выше сказанное модель распознавания инцидентов формально представим следующим кортежем:

$$< K, P_H, R, E >,$$

где K – классификатор признаков; $P_H \subset P$ – множество наблюдаемых признаков; $R = \{ R_i \}$ – множество правил распознавания, E – инцидент.

Процесс распознавания начинается с первого уровня и заканчивается последним по продукционным правилам:

$$R_1 : (K, P_H) \to H_1, R_2 : (K, P_H, H_1) \to H_2, ..., \quad R_n : (K, P_H, H_1, ..., H_{n-1}) \to E, \qquad (1)$$

где H_i – промежуточные гипотезы.

В традиционных продукционных системах правила (1) являются классическими правилами продукций и применяются для формализации строго корректные рассуждения. В условиях неполноты, неточности или изменчивой информацией, к которых, как правило, проявляются угрозы, наши рассуждения часто предположительны, всего лишь правдоподобны и должны подвергаться пересмотру. Такие рассуждения называются модифицируемыми, логика рассуждений которых следующая: вывести *p* из множества посылок *A* и отказаться от *p*, как только информация *q*, не подтверждающая *p*, будет добавлена к *A*. Например, зная, что большинство птиц может летать и что Тити — птица, можно заключить, что Тити может летать. Между тем этот вывод не является абсолютно корректным и общезначимым, ибо не учитывает возможных исключений. Следовательно, он неточен и подлежит пересмотру. Если уточнено, что Тити — страус, то утверждение «Тити может летать» отвергается [8].

Исследование модифицируемых рассуждений и их применение в вычислительных системах является предметом так называемых немонотонных логик, из которых наиболее известной является логика умолчаний Рейтера. Логика умолчаний позволяет формализовать такие рассуждения в виде правил умолчаний:

$$\frac{\alpha(x) : M\beta_1(x), ..., M\beta_m(x)}{\gamma(x)}, \qquad (2)$$

где $\alpha(x), \beta_1(x), ..., \beta_m(x)$ и $\gamma(x)$ – формулы языка L предикатов первого порядка, $\alpha(x)$ называется требованием умолчания, $\beta_i(x)$ – обоснованием умолчания, $\gamma(x)$ – следствием умолчания, а M – некий символ метаязыка.

Правило (2) гарантировано позволяет получить заключение только тогда, когда его следствие и обоснование совпадают, т.е. когда это правило является нормальным правилом умолчаний:

$$\frac{\alpha(x) : M\beta(x)}{\beta(x)}. \qquad (3)$$

Семантика иго такова. Если $\alpha(x)$ истинно и если $\beta(x)$ выполнимо, то $\beta(x)$ выводимо.. Это правило типичных ситуаций, позволяющее обрабатывать исключения без их предварительной идентификации.

Уточним понятие выполнимости $\beta(x)$ в данном правиле. Известно, что формула выполнима тогда и только тогда, когда она не является противоречивой. Таким образом, задача проверки выполнимости следствия $\beta(x)$ эквивалентная задаче доказательства его непротиворечивости в аксиоматической системе данной предметной области.

Далее, умолчание (3), по сути, представляет продукционное правило $\alpha(x) \to \beta(x)$ с дополнительной проверкой заключения $\beta(x)$ на выполнимость. С другой стороны, общезначимое продукционное правило $\alpha(x) \to \beta(x)$ можно рассматривать как правило умолчаний, в котором $\beta(x)$ всегда выполнимо. Поэтому в дальнейшем под правилами умолчаний будем понимать также и классические продукционные правила.

В логике умолчаний доказательство определяется следующим образом [8]. Пусть $D = \{d_i\}$ – множество правил умолчаний вида (3), а F – множество формул (аксиом) языка L. Далее, пусть дана формула $f \in F$. Тогда последовательность конечных подмножеств $D_0, D_1, D_2, ..., D_k$ из D есть доказательство для f в системе (D, F) тогда и только тогда, когда

$F \bigcup \{ KC(D_0) \} \vdash f$,
$F \bigcup \{ KC(D_i) \} \vdash KT(D_{i-1})$ для $i = 1, 2, ..., k$,
$D_k = \varnothing$,
$F \bigcup \{ KC(D_i) \mid 0 \le i \le k \}$ выполнимо,

где $KC(D_i)$ и $KT(D_i)$ соответственно конъюнкции следствий и требований умолчаний из D_i.

Это доказательство можно интерпретировать следующим образом. Первое подмножество D_k выбирается пустым. Последовательно строятся $D_{k-1}, ..., D_1, D_0$. Множество основных аксиом с добавленной к нему конъюнкцией следствий из всех умолчаний D_0 должно обеспечивать доказательство f классическим образом. Из построения подмножества D_{i-1} вытекает, что множество F (с добавленной к нему конъюнкцией следствий из D_i) должно позволять доказывать требования из D_{i-1} и, следовательно, гарантировать применимость умолчаний из D_{i-1}. Глобальная применимость всех умолчаний устанавливается проверкой выполнимости объединения F и конъюнкций следствий всех использованных умолчаний.

Данное доказательство, по сути, является процессом установления выводимости f из (D, F) или, что то же самое, доказательством теоремы $(D, F) \vdash f$, в которой система (D, F) выступает в качестве посылки теоремы, а формула f – ее заключением. Такое доказательство можно построить следующим алгоритмом.

Если $F \vdash f$, то процесс завершаем – формула f выводима. В противном случае в D ищутся умолчания $d_i, (i = \overline{1,k})$, следствия которых позволят доказать f, т.е. $(F \bigcup_i \beta_i(x)) \vdash f$. Если таких умолчаний нет, то процесс завершаем – формула f не имеет доказательства в системе (D, F). В противном случае если все $\alpha_j(x), (j = \overline{1,k})$ выводимы, а все $\beta_j(x)$ выполнимы в $(F \bigcup_i \beta_i(x))$, то процесс завершаем, формула f выводима в системе (D, F).

Таким образом, центральным моментом рассмотренного алгоритма является установление выводимости $\alpha_j(x)$ и выполнимости $\beta_j(x)$. С учетом сказанного для создания эффективных механизмов реализации этих операций может быть использован метод проблемно-ориентированного доказательства [9], который обеспечивает корректные выводы в этих условиях.

В основе этого метода лежит вывод резолюционного типа, использующий семантику решаемой задачи в качестве стратегии управления.

Процессу доказательства предшествуют некоторые подготовительные операции, заключающиеся в аксиоматизации посылок (фактов и правил) и заключения теоремы, заданных на естественном языке, то есть в построении правильно построенных формул (ППФ) логики предикатов и приведении этих формул в скулемовскую стандартную форму. В результате теорема будет представлена множеством дизъюнктов *S*.

Пусть задана теорема $T =< M, F, G >$, где $M = \{m_i \mid i = \overline{1,m}\}$ – множество дизъюнктов, которые обозначают факты теоремы (описание угрозы и аксиомы предметной области), $F = \{f_j \mid j = n\}$ – множество хорновских дизъюнктов, представляющих правила вывода, а *G* – множество дизъюнктов (целевая формула), которые описывают заключение теоремы, в качестве которой выступают формулы $\alpha_j(x)$ и $\beta_j(x)$. Если необходимо установить выводимость $\alpha_j(x)$ тогда требуется доказать, что

$$(M \cup F \cup \neg G) \vdash \square,$$

где $\square$ – пустой дизъюнкт.

Если необходимо установить выполнимость $\beta_j(x)$, в этом случае требуется доказать, что $(M \cup F \cup \neg G) \nvdash \square$, В общем случае для множества $S = \{M \cup F \cup G\}$ можно следующим образом описать процесс доказательства.

Проверяется выводимость формулы *G* в множестве *M*. Если $M \vdash G$, то процесс завершается – теорема доказана. В противном случае формируется так называемый незавершенный вывод $\Im$ дизъюнкта $\square$, который представляет собой множество $\{M \cup F \cup (\text{все промежуточные выводы})\}$. Далее ищется правила вывода, применение которого позволило бы продолжить этот вывод. Если такого правила нет, то процесс завершается – множество *S* не имеет решения.

Пусть таким правилом будет правило f_1. В этом случае формируется новая теорема

$$T_1 =< M, F, f_1^A >,$$

где f_1^A – антецедент правила f_1. Если $M \vdash f_1^A$, то проверяется выполнимость консеквента f_1^K этого правила. Если формула f_1^K не выполнима, процесс завершается – множество *S* не имеет решения. В противном случае консеквента f_1^K правила f_1 переводится в разряд фактов $M = M \cup f_1^K$ и ищется следующее правила, применение которого позволило бы продолжить вывод $\Im$. Если $M \nvdash f_1^A$, то опять формируется незавершенный вывод $\Im_1$ и процесс реверсируется. И так повторяется до тех пор, пока вывод $\Im$ не закончится пустым дизъюнктом.

Выводы. Рассмотренный подход расширяет возможности классических корреляционных механизмов SIEM-систем за счет возможности одновременного использования как общезначимых продукционных правил, так и правил умолчаний, которые позволяют описывать типичные ситуации. Это дает возможность обрабатывать исключения без их предварительной идентификации и создавать более гибкие механизмы корреляции. Такие механизмы могут быть созданы на основе анализа журнала событий ОС Windows, что позволит оперативно получать информацию об источнике угрозы.

ЛІТЕРАТУРА

[1]. Hanemann, A., Marcu, P. Algorithm design and application of service-oriented event correlation. [Электронный ресурс] URL: http://www.researchgate.net/publication/221033552_Algrithm_design_and_application_of_service-oriented_event_correlation (дата обращения 25.05.2014).

[2]. Muller, A. Event Correlation Engine. [Электронный ресурс] URL: ftp://ftp.tik.ee.ethz.ch/pub/students/2009-FS/MA-2009-01.pdf (дата обращения 25.05.2014).

[3]. Шелестова, О. Корреляция SIEM – это просто. Сигнатурные методы. [Электронный ресурс] URL: http://www.securitylab.ru/analytics/431459.php (дата обращения 30.03.2014).

[4]. Олеся Шелестова. Корреляция SIEM. Сигнатурные методы //исследовательский центр Positive Research [Электронный ресурс] 2012. URL: http://www.securitylab.ru/analytics/431459.php.

[5]. Борисов В. И., Шабуров А. С. О Применении сигнатурных методов анализа информации в SIEM-системах/ Безопасность в информационной сфере № 3(17) / 2015 С. 23-27.
[6]. Федорченко А.В., Левшун Д.С., Чечулин А.А., Котенко И.В. Анализ методов корреляции событий безопасности в SIEM-системах. Часть 1. // Труды СПИИРАН. 2016. Вып. 47. С. 5-27.
[7]. Стандарт ISO/IEC TR 18044:2004 "Information technology - Security techniques - Information security incident management".
[8]. Тей А., Грибомон П., Луи Ж., Лог Ж. Логический подход к искусственному интеллекту. - М.: Мир, 1990. - 429 с.
[9]. Самохвалов Ю.Я. Метод проблемно-ориентированного доказательства в нечеткой логике // Кибернетика и системный анализ. - 1995. - № 5. - С. 58-68.

REFERENCES

[1]. Hanemann, A., Marcu, P. Algorithm design and application of service-oriented event correlation. [Electronic resource] URL: http://www.researchgate.net/publication/221033552_Algrithm_design_and_application_of_service-oriented_event_correlation (date of the application 25.05.2014).
[2]. Muller, A. Event Correlation Engine. [Electronic resource] URL: ftp:// ftp. tik. ee. ethz. ch / pub / students / 2009-FS / MA-2009-01. pdf (date of the application 25.05.2014).
[3]. Shelestova O. SIEM correlation – it's simple. Signature methods. [Electronic resource] URL: http:// www. securitylab. ru / analytics/ 431459.php (date of the application 30.03.2014).
[4]. Shelestova O. SIEM correlation. Signature methods // research center Positive Research [Electronic resource] 2012. URL: http: // www. securitylab. ru / analytics / 431459.php.
[5]. Borisov V.I., Shaburov A.S. About application of Signature methods of information analysis in SIEM-systems / Security in the information sphere № 3(17) / 2015 P. 23-27.
[6]. Fedorchenko A.V., Levshun D.S., Chechulin A.A., Kotenko I.V. Analysis of security events correlation methods in SIEM-systems. Part 1. // Proceedings of SPIIRAS. 2016. № 47. P. 5-27.
[7]. Standard ISO/IEC TR 18044:2004 "Information technology - Security techniques - Information security incident management".
[8]. Tay A., Grybomon P., Lui G., Log G. A logical approach to artificial intelligence. - M.: Mir, 1990. - 429 pp.
[9]. Samokhvalov Y. Method of problem-oriented proof in fuzzy logic // Cybernetics and system analysis.- 1995. - № 5. P. 58-68.

КОРЕЛЯЦІЯ ПОДІЙ В SIEM-СИСТЕМАХ НА ОСНОВІ НЕМОНОТОННОГО ВИВЕДЕННЯ

Розглянуто підхід до створення кореляційних механізмів в SIEM-системах. У якості логічної основа таких механізмів запропоновано використання немонотонних правил умовчань у поєднанні з виводом резолюційного типу. Даний підхід розширює можливості класичних кореляційних механізмів SIEM-систем за рахунок можливості одночасного використання як загальнозначущих продукційних правил, так і правил умовчань, які дозволяють описувати типові ситуації. Це дає можливість обробляти виключення без їх попередньої ідентифікації і створювати більш гнучкі механізми кореляції.

Ключові слова: SIEM-система, кореляція, сигнатура, продукція, немонотонний вивід, правила умовчань.

EVENTS CORRELATION IN THE SIEM-SYSTEMS BASED ON UNMONOTONOUS OUTPUT

Going near creation of cross-correlation mechanisms is considered in the SIEM-systems. As logical basis of such mechanisms the use non-monotonic rules of silences is offered in combination with the conclusion of resolution type. This approach extends possibilities of classic cross-correlation mechanisms of the SIEM-systems due to possibility of the simultaneous use of both rules of products and rules of silences which allow to describe typical situations. It enables to process exceptions without their preliminary authentication and create more flexible mechanisms of correlation.

Keywords: The SIEM-system, correlation, signatura, products, unmonotonous conclusion, rules of silences.

Самохвалов Юрий Яковлевич, д.т.н., профессор, Киевский национальный университет имени Тараса Шевченко.
E-mail: yu1953@ukr.net
Самохвалов Юрій Якович, д.т.н, професор, Київський національній університет імені Тараса Шевченка.
Samokhvalov Yuri, Doctor of technical science, professor, Taras Shevchenko Kyiv National University.

Толюпа Сергей Васильевич, д.т.н., профессор, Киевский национальный университет имени Тараса Шевченко.
E-mail: tolupa@i.ua
Толюпа Сергій Васильович, д.т.н, професор, Київський національній університет імені Тараса Шевченка.
Toliupa Sergey, Doctor of technical science, professor, Taras Shevchenko Kyiv National University.