УДК 004.056.53 **DOI:** 10.15622/sp.47.1

А.В. Федорченко, Д.С. Левшун, А.А. Чечулин, И.В. Котенко

АНАЛИЗ МЕТОДОВ КОРРЕЛЯЦИИ СОБЫТИЙ БЕЗОПАСНОСТИ В SIEM-СИСТЕМАХ. ЧАСТЬ 1

Федорченко А.В., Левшун Д.С., Чечулин А.А., Котенко И.В. **Анализ методов корреляции событий безопасности в SIEM-системах. Часть 1.**

Аннотация. Статья посвящена анализу методов корреляции событий безопасности в системах управления информацией и событиями безопасности (SIEM-системах). Процесс корреляции событий безопасности рассматривается в виде многоуровневой иерархии этапов, цель каждого из которых заключается в выполнении определенных операций над обрабатываемыми данными безопасности. На основе результатов проведенного анализа в работе приводится описание каждого этапа процесса корреляции и схемы их взаимодействия.

Ключевые слова: процесс корреляция данных, события безопасности, анализ событий безопасности, системы оценки защищенности, SIEM-системы.

Fedorchenko A.V., Levshun D.S, Chechulin A.A., Kotenko I.V. **An Analysis of Security Event Correlation Techniques in Siem-Systems. Part 1.**

Abstract. The paper is devoted to the analysis of security event correlation methods in Security Information and Event Management (SIEM) systems. The correlation process is considered to be a multilevel hierarchy of stages. The goal of each stage consists in executing appropriate operations on security data being processed. Based on this analysis we outline each correlation stage and their interaction scheme.

Keywords: data correlation process; security event; security event analysis; computer network security evaluation systems; SIEM systems.

1. Введение. В настоящее время все больше внимания уделяется обеспечению безопасности информации как в крупных учреждениях и компаниях, так и в средних и малых организациях. Защищаемые объекты имеют различные уровни доступа, всевозможные варианты развертывания вычислительных сред и разнообразные топологии сетевого взаимодействия. Задача обеспечения безопасности с помощью универсальных средств обнаружения и предотвращения атак усложняется, в том числе за счет стремительного роста числа пользователей и разнообразия типов устройств, использования облачных технологий и многократного увеличения объема и скорости передачи и обработки информации. Одним из классов средств, позволяющих обеспечивать безопасность систем любого уровня и набора устройств, являются системы управления информацией и событиями безопасности (Security Information and Event Management, SIEM) [1, 2]. Преимущество данных решений заключается в гибкости применения и независимости от набора спецификаций и платформ для конечной защищаемой инфраструктуры.

Схема взаимодействия компонент SIEM-систем в общем виде всегда содержит модуль корреляции информации и событий безопасности, который является основополагающим элементом применяемых механизмов безопасности. Работа компонентов корреляции в широком смысле направлена на обнаружение атак, вредоносной активности и нарушений политики безопасности. В узком смысле компонент корреляции предназначен для поиска связей, зависимостей и причинно-следственных отношений между событиями безопасности и другой информацией безопасности. Данный компонент выполняет как функции корреляции событий, так и их пред- пост обработку в зависимости от конкретной реализации системы.

Учитывая необходимость оперирования событиями и информацией безопасности, компонент корреляции можно рассматривать с разных точек зрения. Корреляция представляется либо с точки зрения рассмотрения ее как процесса, либо с точки зрения множеств входных и выходных данных, преобразуемых внутри модуля корреляции. В первом случае корреляция представляет собой последовательность операций над событиями, специально определенную для получения конкретного решения и являющуюся непрерывной относительно работы всей системы. Во втором случае корреляция задается через множество типов событий [3], преобразуемых таким образом, что несколько событий могут образовывать одно более сложное событие и восприниматься системой как неделимое [4-6].

Стоит отметить, что в процесс корреляции можно представить с использованием двух основных типов операций (функций): (1) *комбинирования событий безопасности в одно мета-событие*; (2) *идентификации и удаления (или обозначения) ложных или бесполезных событий безопасности* [7]. Важными операциями корреляции также являются представление событий безопасности и обучение (переобучение) системы корреляции во время работы. Вместе с тем данные типы операций на разных этапах процесса корреляции выполняют отличающиеся по своему характеру действия.

Главным образом данная *работа посвящена* анализу методов корреляции за счет их детального изучения, выявления принципов работы основных компонентов модуля корреляции, классификации и последующего сравнения компонентов друг с другом по основным характеристикам. *Целью работы* является задание основополагающих этапов корреляции и определение достоинств и недостатков различных методов корреляции, используемых в мировой практике для реализаций современных SIEM-решений.

Данный анализ производился за счет изучения научно-технической литературы, включающей описания как отдельных методов корреляции и их общих обзоров, так и особенностей реализации компонентов корреляции в конкретных решениях открытых продуктов данного класса. *Новизна* статьи заключается в предложении собственной схемы этапов процесса корреляции и оценке применения в них различных методов. Качественное сравнение методов корреляции достигается за счет определения их общих и частных характеристик, на основе которых предлагается новая система классификации.

Результаты работы представлены в двух статьях. Данная статья является перовой частью описания проводимых исследований, где корреляция событий безопасности рассматривается как процесс, разделенный на отдельные этапы. В разделе 2 приводятся релевантные работы по тематике описания процесса корреляции. В них рассматриваются прототипы систем корреляции предупреждений для систем обнаружения вторжений (СОВ). В данных работах также анализируются различные варианты выполнения базовых этапов в процессе корреляции. В разделе указываются преимущества и недостатки используемых подходов и схем. В разделе 3 описывается процесс корреляции событий безопасности с точки зрения оптимального позиционирования этапов работы модуля корреляции и распределения обработки данных по уровням. В данном разделе раскрываются задачи, возлагаемые на каждый этап, необходимость их использования и возможные способы реализации.

2. Релевантные работы. Тема представления процесса корреляции событий безопасности в виде последовательно выполняемых этапов активно исследуется на протяжении последних 20 лет. За это время были предложены различные методики обработки разнородных данных и преобразования в процессе корреляции низкоуровневых событий к высокоуровневым, а также рассмотрены возможные схемы, описывающие сам процесс корреляции. Вместе с развитием данной тематики публиковались работы, посвященные классификации этапов процесса корреляции событий безопасности [8-13].

В [8] предлагается разбиение процесса корреляции на выполняемые задачи. Выделяются следующие составляющие процесса: *сжатие* (*compression*), *счет* (*count*), *подавление* (*suppression*), *логическая замена* (*boolean*) и *обобщение* (*generalization*). Под *сжатием* подразумевается преобразование в одно событие безопасности нескольких одинаковых событий. *Счет* представляет собой замену похожих событий безопасности одним новым событием, а *подавление* — осуществление задержки обработки событий безопасности с низким

приоритетом до окончания обработки события безопасности с более высоким приоритетом. Процесс *логической замены* заключается в преобразовании некоторого множества событий в новое событие, удовлетворяющее определенному логическому шаблону. В результате выполнения *обобщения* производится перевод события безопасности к высокоуровневому представлению (суперклассу) для удовлетворения необходимой важности уведомления. Несомненным достоинством работы является возможность добавления статистических методов в описанную модель корреляции событий безопасности, несмотря на то, что она основана на строго детерминированных подходах. Недостатки представленной работы заключаются в отсутствии среди указанных задач элемента *предупреждения ошибок (fault prediction)* и элемента *предупреждения нарушений (preventive maintenance)*.

В [9] производится обзор работ в области корреляции предупреждений для систем обнаружения вторжений (Intrusion Detection System, IDS). В частности, рассматриваются этапы и операции процесса корреляции, описывается модель данных формата обмена сообщениями обнаружения вторжений (Intrusion Detection Message Exchange Format, IDMEF), а также приводится пример процесса корреляции для обнаружения типовой атаки. Процесс корреляции в данной работе условно делится на три этапа: (1) предобработка (*preprocessing*); (2) анализ предупреждений (*alarm analysis*); (3) корреляция предупреждений (*alarm correlation*). На втором этапе выделяются такие методы и этапы процесса корреляции, как измерение схожих признаков (*similarity measures*), кластеризация, интеллектуальный анализ данных (*data mining*), удаление, редукция и слияние. Стоит отметить, что в результате выполнения каждого из трех этапов формируются простые события, мета-события и сценарии атак соответственно, а по окончании выполнения процесса корреляции формируется отчет.

В [10] выделяется шесть этапов процесса корреляции: нормализация (*normalization*), агрегация (*aggregation*), корреляция (*correlation*), отсеивание ложных срабатываний (*false alert reduction*), анализ стратегии атаки (*attack strategy analysis*) и приоритизация (*prioritization*). Описываются четыре основных метода корреляции: (1) на основе сценариев атак; (2) ориентированного на правила; (3) статистического и (4) временного. Главное отличие [10] от аналогичных работ заключается в точном связывании этапов процесса корреляции с используемыми в них конкретными методами. В данной работе также выделены отдельные группы методов корреляции.

В [11-13] авторы выделяют пять подходов к процессу корреляции событий безопасности, основанных на: (1) *подобии* (сходстве) (*similarity based*); (2) *предопределении сценариев атак* (*predefined attack scenarios based*); (3) *многоуровневых вычислениях* (*multi-stage*) *на базе предпосылок и последствий*; (4) *использовании множества источников информации* (*multiple information sources*); (5) *фильтрации* (*filter based*).

Первый подход заключается в вычислении величины подобия двух событий безопасности на основе атрибутов, ассоциируемых с этими событиями. События, величина подобия которых достаточно велика, группируются.

Второй подход заключается в объединении в последовательность связанных этапов проведения атак на основе заранее определенных шаблонов сценариев атак. Данный подход применяется для получения агрегированного и более высокоуровневого взгляда на угрозы безопасности.

Третий подход основывается на формировании сценариев атак путем связывания отдельных этапов их проведения при условии, что один из этапов является необходимым условием для проведения другого.

Четвертый подход направлен на приоритизирование и классификацию потоков событий безопасности в зависимости от источника информации о событии безопасности.

Пятый подход основан на удалении из процесса корреляции событий по заранее определенным правилам (фильтрам). Решение об удалении события из процесса корреляции принимается на основе значений одного или нескольких его атрибутов.

Похожая классификация приводится в [14]. Приведенная авторами модель корреляции событий безопасности состоит из двух частей: (1) подхода, *основанного на графах атак* (*an attack graph-based*) и (2) подхода, *основанного на подобии* (*similarity-based*). При этом первый подход используется для корреляции событий безопасности, вызванных известными атаками, а также для построения гипотез о вероятно необнаруженных или упущенных событиях безопасности. В свою очередь, второй подход применяется для корреляции событий безопасности, вызванных неизвестными атаками, а также для уточнения известных графов атак.

В [7, 15, 16] выделяются и описываются отдельные этапы и уровни процесса корреляции.

В [7] Крюгел, Валеур и Вигна раскрывают методы и подходы к корреляции предупреждений в зависимости от фазы процесса, указывают на их достоинства и недостатки, а также спорные моменты. В работе

процесс корреляции предупреждений представлен в виде этапов, которые преобразуют оповещения сенсоров в отчеты о вторжениях и направлены на разные аспекты процесса корреляции (рисунок 1).

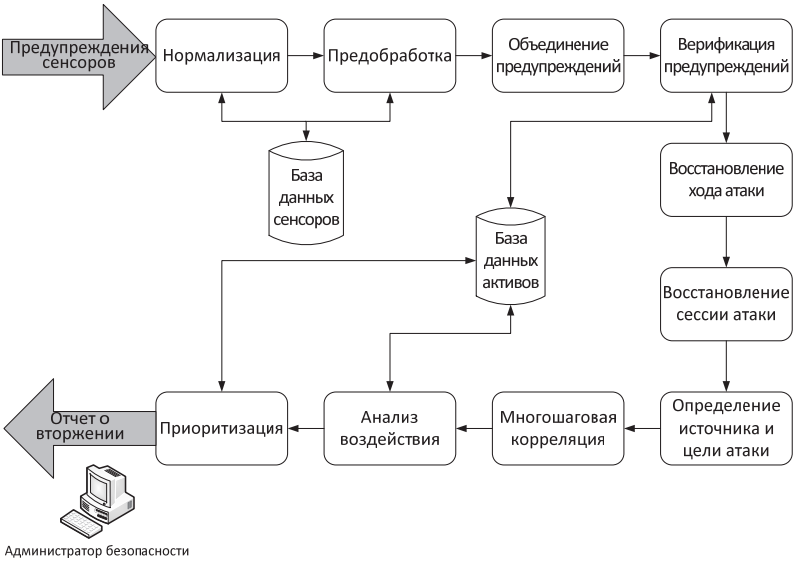


Рис. 1. Представление процесса корреляции предупреждений в [7]

Именно данная схема взята за основу во многих других работах (в том числе и в настоящей статье) для представления этапов процесса корреляции и их последовательности. В [7] рассматриваются методы корреляции предупреждений в системах обнаружения вторжений (Intrusion Detection System, IDS), которые также применимы в SIEM-системах, несмотря на то, что предупреждения являются только отдельным типом событий. Иными словами, за счет расширения типов обрабатываемых событий SIEM-системы расширяют системы обнаружения вторжений в рамках корреляции. В данной работе также производится структурное деление этапов процесса корреляции предупреждений в зависимости от их целей, таких как: сбор (*collection*), агрегация и верификация (*aggregation and verification*), анализ высокоуровневых структур (*high-level structures*), крупномасштабная корреляция (*large-scale correlation*), оценка (*evaluation*). В рамках детектирования атак приводятся следующие виды систем: (1) на основе злоупотреблений (*misuse-based*; в основу положена база знаний; атака идентифицируется при соответствии записи базы с параметрами входных

данных); (2) на основе аномалий (*anomaly-based*; текущее состояние сравнивается с эталонным с помощью оценки вероятности отклонения).

В [15] процесс корреляции предупреждений представлен в виде логических блоков. Авторы выделяют следующие блоки: *нормализация данных* (*Data Normalization Unit*), *корреляция на основе фильтрации* (*Filter-based Correlation Unit*), *редукция данных* (*Data Reduction Unit*), *анализ намерений* (*Intention Recognition*) и *анализ воздействий* (*Impact Analysis*).

Отличительной особенностью данной работы является представление модели процесса корреляции (рисунок 2), которая снижает количество обрабатываемых событий безопасности так рано, как это только возможно. Это осуществляется путем вывода из процесса корреляции незначимых или ложных событий безопасности еще на начальных этапах процесса. Авторы ввели дополнительный компонент для работы с некоррелируемыми данными. По итогам эксперимента на наборах данных DARPA 2000 авторам удалось добиться процента редукции 99,38 % (в среднем).

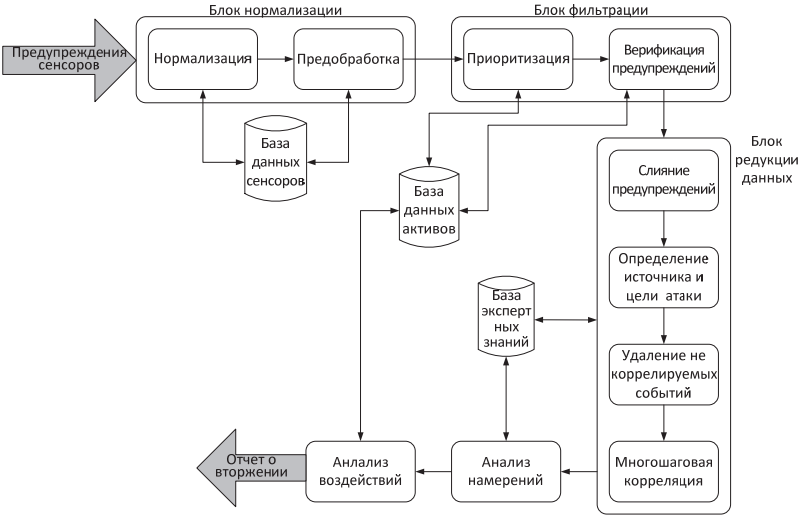


Рис. 2. Модель процесса корреляции событий безопасности [15]

Важно отметить, что данная модель не лишена недостатков. Во-первых, в блоке фильтрации этапы верификации и приоритезации безусловно будут обрабатывать в том числе и дубликаты событий безопасности, так как этап слияния предупреждений находится на более высоком уровне (блок редукции данных). Учитывая ресурсоемкость процесса проверки событий безопасности на подлинность, данное

решение подлежит дополнительному рассмотрению. Во-вторых, этап удаления из процесса корреляции событий безопасности данных, которые не могут быть коррелированы (*Удаление не коррелируемых событий*), не оставляет процессу права на ошибку. При увеличении среднего показателя редукции событий безопасности и облегчении дальнейшего анализа открытым остаётся вопрос гарантии того, что из процесса корреляции не удаляются важные события. И, в-третьих, задача модуля анализа воздействия заключается в исключении из процесса корреляции сценариев атак, влияние которых на инфраструктуру сети незначительно или невозможно, то есть в улучшении коэффициента редукции, при этом модуль анализа воздействия в блок редукции не входит.

В [16] предлагается разделение методов процесса корреляции событий безопасности по следующим уровням обработки данных (рисунок 3): (1) *первичные (сырые) данные* (*raw data*); (2) *события* (*events*); (3) *отчеты* (*reports*). В зависимости от конкретного уровня, данные подвергаются соответствующей обработке для дальнейшего использования на более высоких уровнях, а в итоге — для принятия решения о возможных контрмерах и визуализации результатов.

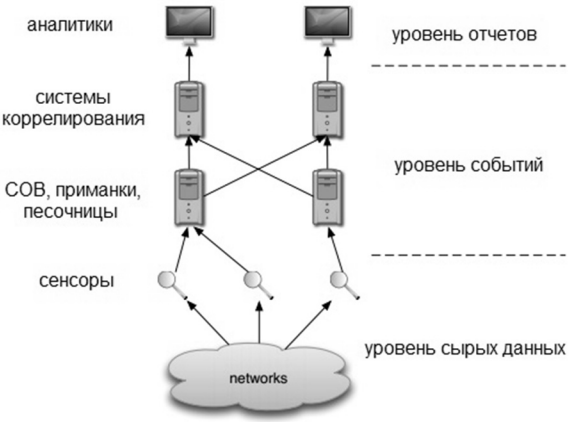


Рис. 3. Уровни процесса корреляции событий безопасности [16]

Уровень первичных данных состоит в основном из сетевых сенсоров, которые принимают необработанные данные и осуществляют их первичный анализ. Источниками данных для системы корреляции событий безопасности на данном уровне также могут быть системы журналирования различных сетевых приложений. На уровне первичных данных реализуются следующие процессы: *отбор пакетов* (*packet sampling*), *вероятностный анализ* (*probabilistic analysis*),

обнаружение аномальной активности (*attack detection*), *обнаружение сканирования портов* (*detection of port scans*), *идентификация приложений* (*application identification*) и *анализ полезной нагрузки пакетов* (*payload analysis*). На уровне событий выполняется локальный процесс корреляции событий безопасности и распределенный процесс корреляции событий безопасности, включая соблюдение конфиденциальности данных. На уровне отчетов специализированные приложения визуализируют результаты процесса корреляции наиболее подходящим образом. Уровень отчетов выполняет генерацию возможных активных контрмер и верификацию событий безопасности.

В результате проведенного анализа была сформирована собственная модель процесса корреляции, в основу которой вошли необходимые с точки зрения авторов этапы обработки данных. Данные этапы были логически сгруппированы и связанны в зависимости от функциональной задачи каждого этапа. Описанные результаты раскрываются в разделе 3.

3. Процесс корреляции событий безопасности. Процесс корреляции событий безопасности является сложной задачей. Поэтому предлагается разбить ее на подзадачи с помощью декомпозиции. Применение декомпозиции к системе, реализующей процесс корреляции событий безопасности, позволит представить систему в виде простых функциональных модулей. Такой подход обеспечит рассмотрение каждого модуля независимо друг от друга. При этом важно четко определить функциональную нагрузку каждого модуля и порядок их взаимодействия.

Кроме того, предлагается применить многоуровневый подход к декомпозиции системы корреляции событий безопасности. Многоуровневый подход подразумевает разбиение на группы множества модулей системы корреляции, а также упорядочивание групп модулей по уровням, образующим иерархию. В соответствии с принципом иерархии, для каждого промежуточного уровня следует указать непосредственно примыкающие к нему соседний вышележащий и нижележащий уровни. При этом, с одной стороны, каждая группа модулей одного уровня должна быть сформирована таким образом, чтобы все модули этой группы для выполнения своих функций использовали результаты обработки событий безопасности, полученные на соседнем нижележащем или текущем уровне. С другой стороны, результаты обработки данных каждого модуля, отнесенного к некоторому уровню, могут быть переданы только модулям соседнего вышележащего уровня или среди модулей текущего уровня. Такой подход позволит определить задачи, выполняемые на каждом из

уровней, принципы их взаимодействия, а также типы данных, передаваемые между ними. Кроме того, декомпозиция системы, реализующей процесс корреляции, упрощает разработку, отладку и тестирование отдельных уровней или модулей.

Предлагаемое обобщенное представление системы корреляции событий безопасности приведено на рисунке 4.

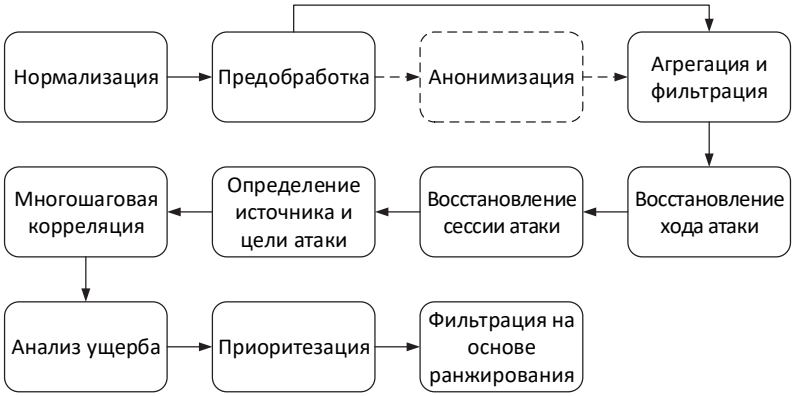


Рис. 4. Представление системы корреляции в виде модулей

В основу компонентов данной схемы легли этапы процесса корреляции, описанные в [7]. Отметим, что в рамках указанного представления данные, обработанные одним модулем, используются в качестве входных данных для следующего модуля. Тем не менее прохождение каждого события безопасности последовательно через одни и те же модули не является необходимым. Реальный процесс корреляции событий безопасности в зависимости от конкретной задачи может быть представлен гораздо более сложными схемами. Например, обработка потока событий безопасности может включать циклические операции между некоторыми модулями системы, а сами модули в некоторых ситуациях могут использоваться параллельно.

Важно отметить, что в рамках модульного представления системы корреляции событий безопасности не отражена реализация процесса верификации или проверки источников событий безопасности на подлинность. Если система корреляции событий безопасности будет получать в качестве входных данных события безопасности от любых источников, ничто не помешает злоумышленнику сгенерировать поток ошибок второго рода, притворившись одним из сенсоров. Наличие подобного потока событий значительно ухудшает качество процесса корреляции событий безопасности и может привести к обнаружению

сценариев атак, которых не существуют. Задача верификации источников событий безопасности ложится на этап сбора данных, поэтому в рамках системы корреляции событий безопасности каждое событие безопасности уже считается верифицированным, то есть полученным от разрешенного источника событий безопасности.

Далее модули системы корреляции событий рассматриваются более подробно.

Нормализация. Из-за того что источники данных могут поставлять информацию в разном формате, возникает необходимость преобразования формата каждого события безопасности в некоторый нормализованный формат, который был бы понятен всем модулям обработки. Данное преобразование, или *нормализация*, означает, что синтаксис и семантика события безопасности прозрачны и беспрепятственно определяемы.

Предобработка. После нормализации обработанные события безопасности нуждаются в дополнительной *предобработке*, так как часть источников может пропускать некоторые поля данных, важные для процесса корреляции (например, время начала, время окончания и источник события).

Анонимизация. Данный модуль системы корреляции необходим, если производится работа с событиями минимум от двух источников, между которыми не установлено доверительное отношение. Анонимизация применяется для удаления или сокрытия конфиденциальной (или важной с юридической точки зрения) информации из событий безопасности. Существует две операции данного модуля: *анонимизация* и *псевдоанонимизация* [17, 18]. Анонимизация препятствует восстановлению конфиденциальных данных, в то время как псевдоанонимизация — обратима, а значит, оригинальные данные могут быть восстановлены доверительной стороной. В общем случае желательно проводить псевдоанонимизацию, так как это позволяет получить доступ к оригинальной информации в ситуациях, когда необходим дальнейший анализ. Однако данное решение накладывает значительные вычислительные ограничения. Ключевой задачей модуля анонимизации является как сохранение необходимых свойств для анализа безопасности, так и способность их сокрытия от нежелательных сторон. Извлечение подобных свойств предполагает *деанонимизацию* (обратный процесс) данных, что возможно только при использовании псевдоанонимизирующих методов.

Агрегация и фильтрация. Задача модуля фильтрации и агрегации заключается в удалении из системы корреляции событий по заранее определенным правилам (фильтрам), и в объединении данных, которые возникли в результате независимого обнаружения одного и того же события различными источниками. Решение об удалении события из

системы корреляции принимается на основе значений одного или нескольких его атрибутов. Не прошедшие фильтрацию события безопасности больше в процессе корреляции не участвуют, а оставшиеся — переходят на этап агрегации. Решение об агрегировании двух событий безопасности принимается на основе содержащихся в них данных. При идентичности значений заранее определенных атрибутов событий, а также удовлетворении временных характеристик событий заданному интервалу, такие события безопасности агрегируют в одно мета-событие.

Восстановление хода атаки. Задача модуля восстановления хода атаки ограничена объединением событий безопасности, вызванных активностью одного злоумышленника по отношению к одной цели (рисунок 5). Восстановление хода атаки построено на объединении событий безопасности с совпадающими атрибутами цели и источника атаки, временные параметры которых попадают в заданный временной интервал. Требование к временным параметрам заключается в том, чтобы время окончания более раннего события, характерного для конкретной атаки, было достаточно и определенно близко ко времени старта другого события, продолжающего соответствующую атаку.

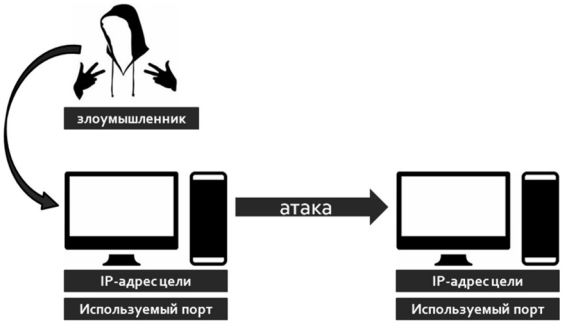


Рис. 5. Восстановление хода атаки

Восстановление сессии атаки. Цель данного модуля — поиск связи между *сетевыми* (*network-based*) и *системными* (*host-based*) событиями безопасности (рисунок 6). Это необходимо для объединения ряда событий безопасности, вызванных злоумышленником при тестировании различных эксплоитов против определенной программы или запуском одного и того же эксплоита несколько раз для подбора правильных значений определенных параметров (например, смещений и адресов памяти для переполнения буфера). Процесс поиска связи между событиями усложняется за счет отличающегося предоставления информации в сетевых и системных событиях. Сетевые сенсоры могут

предоставить информацию, характеризирующую обнаруженные атаки, например, IP-адреса источника и цели, используемые порты. Данные, поступающие от системных сенсоров, с другой стороны, содержат информацию об объекте, который был атакован, и субъекте, которым была осуществлена данная атака. Именно обнаружение связи между сетевыми и системными событиями позволяет определить отдельную сессию производимой атаки.

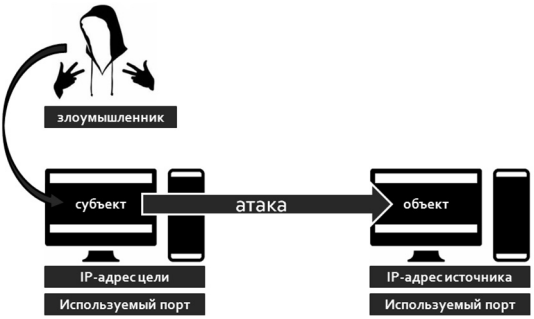
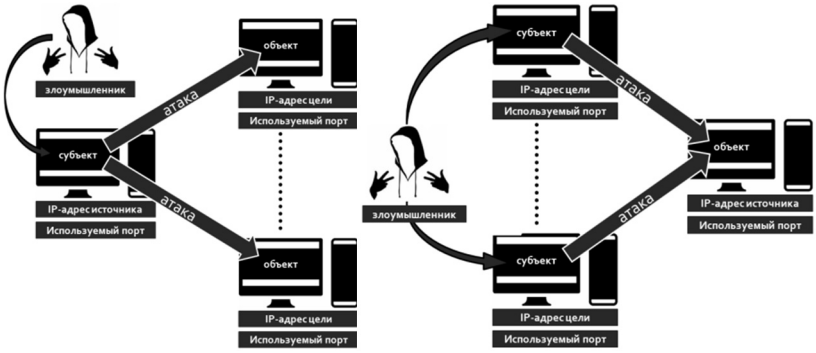


Рис. 6. Восстановление сессии атаки

Определение цели и источника атаки. Задача этого модуля — идентификация хостов, которые являются либо источником, либо целью обнаруживаемых атак. Данный модуль объединяет события безопасности, ассоциируемые с отдельным хостом, который атакует несколько жертв (сценарий *один-ко-многим* (*one2many*) (рисунок 7а)), и с несколькими хостами, которые атакуют одну жертву (сценарий *многие-к-одному* (*many2one*) (рисунок 7б)).



а) б)

Рис. 7. Определение цели и источника атаки: а) сценарий один-ко-многим; б) сценарий многие-к-одному

Многошаговая корреляция. Модуль используется для распознания сложных сценариев, которые состоят из нескольких отдельных атак (рисунок 8).

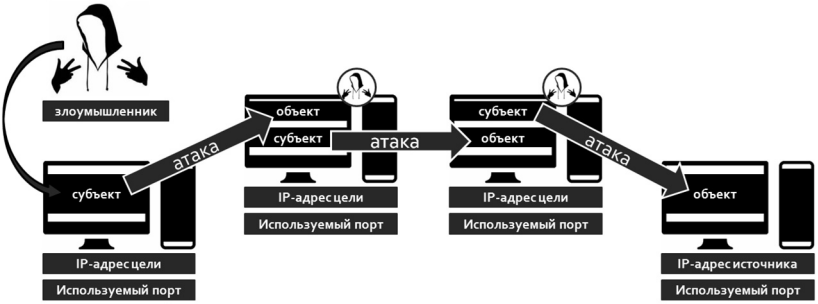


Рис. 8. Схема многошаговой корреляции

Обычно эти сценарии определяются с использованием той или иной формы экспертных знаний [19-21]. Модуль многошаговой корреляции также можно использовать для верификации высокоуровневых событий безопасности. При этом определяются сценарии атак, которые заведомо не имеют значения. Это позволяет удалить из процесса корреляции последовательности событий, которые коррелированны неверно. Например, при опросе сетевого окружения на предмет наличия какого-либо оборудования (принтера или сканера) последовательность событий по количеству запросов будет выглядеть как сканирование сети, однако реально данная последовательность незначительна и фактически не является атакующими действиями. К подобному роду ошибок также можно отнести действия приложений, использующих в работе пиринговые (peer-to-peer, P2P) сети.

Анализ ущерба. Помимо аналитической информации, полученной на предыдущих этапах, данный модуль использует стороннюю (не содержащуюся в событиях) информацию для анализа сценариев атак с точки зрения ущерба от их влияния на инфраструктуру сети или используемые ресурсы. На основе данных об ущербе модуль назначает более высокую степень важности сценариям атак, которые угрожают более значимым активам сети. Информация о сети и соответствующих ресурсах хранится в базе данных активов, которая содержит подробности об используемых сетевых сервисах, зависимостях между ними, а также их важности для функционирования сети.

Приоритизация. Модуль приоритизации должен учитывать политику безопасности и требования безопасности инфраструктуры, в которой развернута система корреляции. Фактически, модуль

ориентирован на пожелания пользователя, использующего систему корреляции. Поэтому его основной задачей является выделение сценариев атак в соответствии с их приоритетом для пользователя.

Фильтрация на основе ранжирования. Этот модуль используется для снижения количества рассматриваемых сценариев атак, критичность реализации которых мала для корректной работы защищаемой инфраструктуры. Данный модуль должен учитывать политику безопасности и требования безопасности инфраструктуры, в которой развернута система корреляции. Удаление из процесса корреляции сценариев атак с низким рангом (ущерб от влияния которых на анализируемую сеть отсутствует или незначителен) снижает количество ошибок второго рода, увеличивая точность работы процесса корреляции.

На следующем этапе исследования системы, реализующей процесс корреляции событий безопасности, была произведена группировка описанных модулей (рисунок 9).

Модуль нормализации преобразует события безопасности от разнородных источников в формат, понятный системе корреляции событий безопасности. Преобразование происходит без потерь и дополнений. Модуль предобработки дополняет нормализованные события безопасности значениями атрибутов, которые могли быть пропущены источником события безопасности. Модуль анонимизации работает со значениями атрибутов предобработанных событий безопасности, преобразуя или удаляя их.

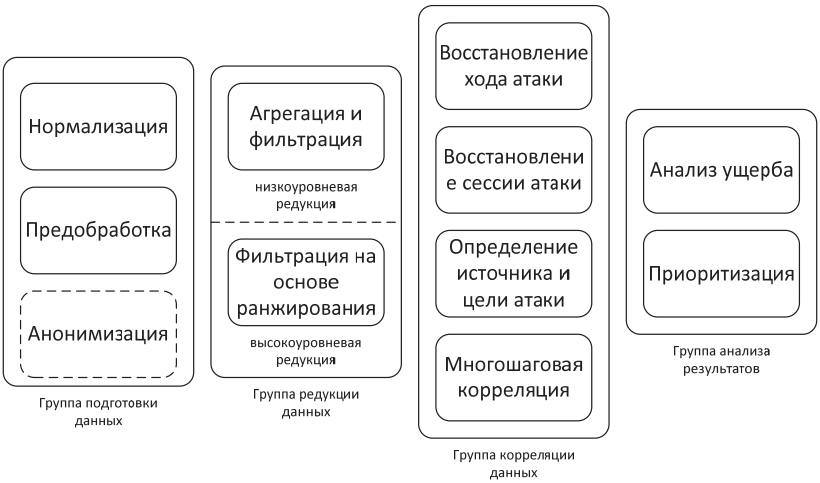


Рис. 9. Группировка модулей

Перечисленные выше модули были объединены в *группу подготовки данных*, которая непосредственно обрабатывает данные, полученные от разнородных источников, с целью их подготовки для модулей более высокого уровня.

Модули восстановления хода атаки, восстановления сессии атаки, определения цели и источника атаки и многошаговой корреляции в совокупности направлены на решение задачи поиска причинно-следственных связей между событиями. Данные модули выделены в *группу корреляции данных*.

Важной задачей процесса корреляции событий безопасности является редукция — операция уменьшения общего количества событий до необходимого и достаточного для проведения анализа, не превосходя вычислительные способности инфраструктуры [22, 23].

По признаку осуществления редукции общего потока информации, модули агрегации и фильтрации и фильтрации на основе ранжирования объединены в *группу редукции данных*. Стоит отметить, что модуль агрегации и фильтрации оперирует с событиями безопасности, в отличие от модуля фильтрации на основе ранжирования, который работает со сценариями атак. Поэтому условно разделим группу редукции данных на две части: *низкоуровневой редукции* и *высокоуровневой редукции*.

Группа высокоуровневой редукции нуждается в специальной информации, зависящей от ущерба при реализации атаки на активы инфраструктуры с учетом критичности потери данных активов. Данные результаты используются в модулях анализа ущерба и приоритизации, которые предлагается выделить в *группу анализа результатов*. При формировании из описанных групп обобщенной системы корреляции, с учетом разбиения на уровни, была получена следующая иерархическая схема (рисунок 10).

На первом уровне, оперирующем непосредственно с событиями безопасности от различных источников, находится группа подготовки данных. Нормализованные, предобработанные и при необходимости анонимизированные данные поступают на второй уровень, где целесообразно расположить группу низкоуровневой редукции для первичной агрегации и фильтрации данных. Дальнейшие операции для увеличения общего показателя редукции производятся над данными более высокого уровня (сложными событиями, сценариями атак). Необходимый переход от простых событий безопасности к сложным событиям, атакам и сценариям атак осуществляется на третьем уровне, который соответствует группе корреляции данных. После прохождения третьего уровня данные дополнительно обрабатываются на четвертом

уровне, который соответствует группе анализа результатов корреляции. После этого данные передаются на пятый уровень, где расположена группа высокоуровневой редукции, осуществляющая фильтрацию сценариев атак.

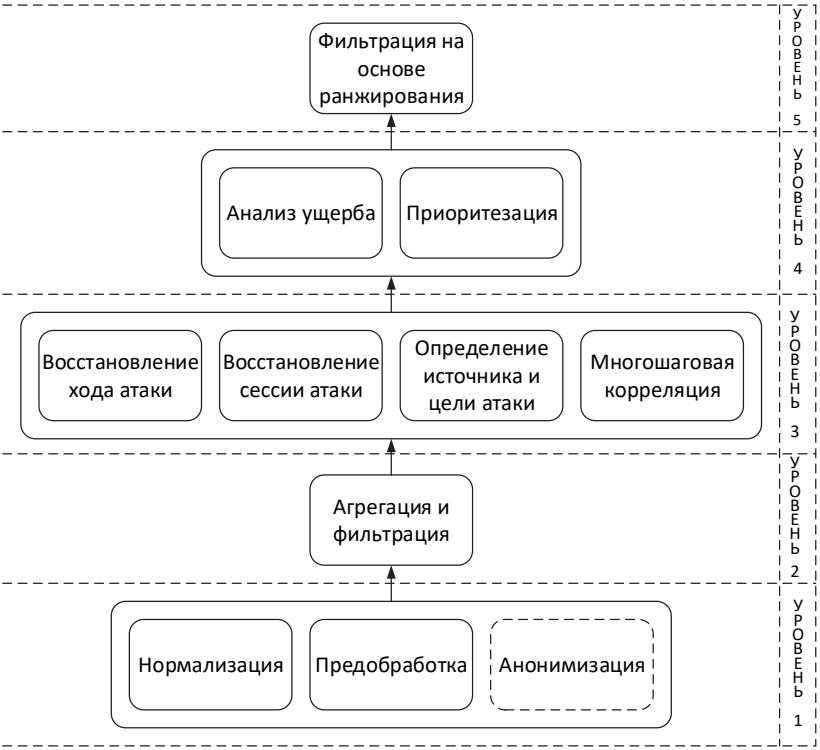


Рис. 10. Многоуровневая иерархия групп модулей

5. Заключение. В данной статье указано место и роль процесса корреляции в SIEM-системе. Описаны основные этапы корреляции, их роль и необходимость использования, а также произведен их анализ с точки зрения реализации системы оценки защищенности компьютерных инфраструктур. По итогам исследования получена четырехуровневая модель процесса корреляции с детализацией каждого уровня.

В результате проделанной работы были однозначно разделены этапы процесса корреляции. Это позволит в дальнейшем более детально исследовать этапы отдельно, а также выделить конкретные методы для эффективного применения на каждом из этапов. В следующей части описания проведенных исследований будут непосредственно

рассмотрены и классифицированы методы корреляции событий безопасности. Также будет произведена оценка применимости представленных методов на различных этапах процесса корреляции.

Литература

1. *Kotenko I.V., Chechulin A.A.* A Cyber Attack Modeling and Impact Assessment Framework // Proceedings of 5th International Conference on Cyber Conflict 2013 (CyCon 2013). 2013. pp. 119–142.
2. *Kotenko I.V., Polubelova O.V., Saenko I.V.* The Ontological Approach for SIEM Data Repository Implementation // 2012 IEEE International Conference on Green Computing and Communications (GreenCom). IEEE Computer Society. 2012. pp. 761–766.
3. *Liu G., Mok A.K., Yang E.J.* Composite Events for Network Event Correlation // IEEE/IFIP International Symposium on Integrated Network Management. 1999. pp. 247–260.
4. *Котенко И.В., Саенко И.Б., Полубелова О.В., Чечулин А.А.* Технологии управления информацией и событиями безопасности для защиты компьютерных сетей // Проблемы информационной безопасности. Компьютерные системы. 2012. № 2. С. 57–68.
5. *Котенко И.В., Саенко И.Б., Полубелова О.В., Чечулин А.А.* Применение технологии управления информацией и событиями безопасности для защиты информации в критически важных инфраструктурах // Труды СПИИРАН. 2012. Вып. 1 (20). С. 27–56.
6. *Котенко И.В., Саенко И.Б., Чечулин А.А.* Проактивное управление информацией и событиями безопасности в информационно-телекоммуникационных системах // Вопросы радиоэлектроники. 2014. Вып 3. № 1. С. 170–180.
7. *Kruegel C., Valeur F., Vigna G.* Intrusion Detection and Correlation: Challenges and Solutions // University of California, Santa Barbara, USA: Springer. 2005. pp. 29-33.
8. *Jakobson G., Weissman M.D.* Alarm correlation // IEEE Network. 1993. vol. 7(6). pp. 52–59.
9. *Zurutuza U., Uribeetxeberria R.* Intrusion Detection Alarm Correlation: A Survey // Proceedings of IADAT International Conference on Telecommunications and computer Networks. 2004. pp. 1–3.
10. *Sadoddin R., Ghorbani A.* Alert Correlation Survey: Framework and Techniques // Proceedings of 2006 International Conference on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services (PST`06). 2006. Article no. 37.
11. *Dadkhah S., Shoja M.R.K., Taheri H.* Alert Correlation through a Multi Components Architecture // International Journal of Electrical and Computer Engineering (IJECE). 2013. vol. 3. no. 4. pp. 461–466.
12. *Elshoush H.T., Osman I.M.* Alert correlation in collaborative intelligent intrusion detection systems — A survey // Applied Soft Computing. 2011. pp. 4349–4365.
13. *Ning P., Xu D.* Correlation analysis of intrusion alerts // Intrusion Detection Systems: series Advances in Information Security. Springer. 2008. vol. 38. pp. 65–92.
14. *Ahmadinejad S.H., Jalili S., Abadi M.* A hybrid model for correlating alerts of known and unknown attack scenarios and updating attack graphs // Computer Networks. 2011. no. 55. pp. 2221–2240.
15. *Elshoushand H.T., Osman I.M.* An improved framework for intrusion alert correlation // Proceedings of World Congress on Engineering 2012 (WCE 2012). 2012. vol. 1. pp. 518–524.

16. *Limmer T., Dressler F.* Survey of event correlation techniques for attack detection in early warning systems. Tech report // University of Erlangen, Dept. of Computer Science. 2008. 37 p.
17. *Flegel U.* Pseudonymizing Unix Log Files // Infrastructure Security. 2002. LNCS 2437. pp. 162–179.
18. *Pang R., Paxson V.* A high-level programming environment for packet trace anonymization and transformation // Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications. 2003. pp. 339–351.
19. *Kotenko I.V., Chechulin A.A.* Computer Attack Modeling and Security Evaluation based on Attack Graphs // Proceedings of 7th International Conference on "Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications" (IDAACS'2013). 2013. pp. 614–619.
20. *Kotenko I.V., Chechulin A.A.* Fast Network Attack Modeling and Security Evaluation based on Attack Graphs // Journal of Cyber Security and Mobility. 2014. vol. 3. no. 1. pp. 27–46.
21. *Котенко И.В, Степашкин М.В., Дойникова Е.В.* Анализ защищенности автоматизированных систем с учетом социо-инженерных атак // Проблемы информационной безопасности. Компьютерные системы. 2011. № 3. С. 40–57.
22. *Файзуллин Р. Р., Васильев В. И.* Метод оценки защищенности сети передачи данных в системе мониторинга и управления событиями информационной безопасности на основе нечеткой логики // Вестник УГАТУ. 2013. Том 17. № 2(55). С. 150–156.
23. *Karlzen H.* An Analysis of Security Information and Event Management: The Use of SIEMs for Log Collection, Management and Analysis // Department of Computer Science and Engineering, University of Gothenburg. 2009. 45 p.

References

1. Kotenko I.V., Chechulin A.A. A Cyber Attack Modeling and Impact Assessment Framework. Proceedings of 5th International Conference on Cyber Conflict 2013 (CyCon 2013). 2013. pp. 119–142.
2. Kotenko I.V., Polubelova O.V., Saenko I.V. The Ontological Approach for SIEM Data Repository Implementation. 2012 IEEE International Conference on Green Computing and Communications (GreenCom). IEEE Computer Society. 2012. pp. 761–766.
3. Liu G., Mok A.K., Yang E.J. Composite Events for Network Event Correlation. IEEE/IFIP International Symposium on Integrated Network Management. 1999. pp. 247–260.
4. Kotenko I.V., Saenko I.V., Polubelova O.V., Chechulin A.A. [Methods for security event and information management for computer networks protection]. *Problemy informacionnoj bezopasnosti. Komp'juternye sistemy – Problems of information security. Computer systems*. 2012. vol. 3. pp. 57–68 (In Russ.).
5. Kotenko I.V., Saenko I.V., Polubelova O.V., Chechulin A.A. [Application of the methods of security event and information management for information protection in the critical infrastructures]. *Trudy SPIIRAN–SPIIRAS Proceedings*. 2012. vol. 1 (20). pp. 27–56. (In Russ.).
6. Kotenko I.V., Saenko I.V., Chechulin A.A. [Proactive management of security information and events in the information and telecommunication systems]. *Voprosy radiojelektroniki –Questions of Radioelectronics*. 2014. vol. 3. no. 1. pp. 170–180 (In Russ.).
7. Kruegel C., Valeur F., Vigna G. Intrusion Detection and Correlation: Challenges and Solutions. University of California, Santa Barbara, USA: Springer. 2005. pp. 29-33.

8. Jakobson G., Weissman M.D. Alarm correlation. *IEEE Network*. 1993. vol. 7(6). pp. 52–59.
9. Zurutuza U., Uribeetxeberria R. Intrusion Detection Alarm Correlation: A Survey. Proceedings of IADAT International Conference on Telecommunications and computer Networks. 2004. pp. 1–3.
10. Sadoddin R., Ghorbani A. Alert Correlation Survey: Framework and Techniques. Proceedings of 2006 International Conference on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services (PST`06). 2006. Article no. 37.
11. Dadkhah S., Shoja M.R.K., Taheri H. Alert Correlation through a Multi Components Architecture*. International Journal of Electrical and Computer Engineering (IJECE).* 2013. vol. 3. no. 4. pp. 461–466.
12. Elshoush H.T., Osman I.M. Alert correlation in collaborative intelligent intrusion detection systems — A survey. *Applied Soft Computing.* 2011. pp. 4349–4365.
13. Ning P., Xu D. Correlation analysis of intrusion alerts. *Intrusion Detection Systems: series Advances in Information Security.* Springer. 2008. vol. 38. pp. 65–92.
14. Ahmadinejad S.H., Jalili S., Abadi M. A hybrid model for correlating alerts of known and unknown attack scenarios and updating attack graphs. *Computer Networks*. 2011. vol. 55. pp. 2221–2240.
15. Elshoushand H.T., Osman I.M. An improved framework for intrusion alert correlation. Proceedings of World Congress on Engineering 2012 (WCE 2012). 2012. vol. 1. pp. 518–524.
16. Limmer T., Dressler F. Survey of event correlation techniques for attack detection in early warning systems. Tech report. University of Erlangen. Dept. of Computer Science. 2008. 37 p.
17. Flegel U. Pseudonymizing Unix Log Files. *Infrastructure Security*. 2002. LNCS 2437. pp. 162–179.
18. Pang R., Paxson V. A high-level programming environment for packet trace anonymization and transformationю Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications. 2003. pp. 339–351.
19. Kotenko I.V., Chechulin A.A. Computer Attack Modeling and Security Evaluation based on Attack Graphs. Proceedings of 7th International Conference on “Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications” (IDAACS'2013). 2013. pp. 614–619.
20. Kotenko I.V., Chechulin A.A. Fast Network Attack Modeling and Security Evaluation based on Attack Graphs. *Journal of Cyber Security and Mobility*. 2014. vol. 3. no. 1. pp. 27–46.
21. Kotenko I.V., Stepashkin M.V., Dojnikova E.V. [Security analysis of information systems taking into account social engineering attacks]. *Problemy informacionnoj bezopasnosti Kompyuternye sistemy – Problems of information security. Computer systems*. 2011. vol. 3. pp. 40–57 (In Russ.).
22. Fajzullin R.R., Vasil'ev V.I. [Protectability assessment method of a data-transmission network in security information and event management system on a basis of fuzzy logic]. *Vestnik UGATU – Proceedings USATU*. 2013. vol. 17. no. 2 (55). pp. 150–156 (In Russ.).
23. Karlzen H. An Analysis of Security Information and Event Management: The Use of SIEMs for Log Collection, Management and Analysis. Department of Computer Science and Engineering, University of Gothenburg. 2009. 45 p.

Федорченко Андрей Владимирович — младший научный сотрудник лаборатории проблем компьютерной безопасности, Федеральное государственное бюджетное

учреждение науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: безопасность компьютерных сетей, обнаружение вторжений, вредоносные программы. Число научных публикаций — 14. fedorchenko@comsec.spb.ru, http://comsec.spb.ru/ru/staff/fedorchenko; 14-я линия В.О., 39, ком. 205, Санкт-Петербург, 199178; р.т.: +7-(812)-328-71-81.

Fedorchenko Andrey Vladimirovich — junior researcher of computer security problems laboratory, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Science (SPIIRAS). Research interests: computer network security, intrusion detection, malware. The number of publications — 14. fedorchenkoandrei28@rambler.ru, http://comsec.spb.ru/ru/staff/fedorchenko; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7-(812)-328-71-81.

Левшун Дмитрий Сергеевич — программист лаборатории проблем компьютерной безопасности, Федеральное государственное бюджетное учреждение науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: компьютерная безопасность, защита встроенных устройств, системы киберфизической безопасности, безопасность распределённых систем, корреляция событий безопасности. Число научных публикаций — 5. levshun@comsec.spb.ru, http://comsec.spb.ru/levshun; 14-я линия В.О., 39, ком. 205, Санкт-Петербург, 199178; р.т.: +7-(812)-328-71-81.

Levshun Dmitry Sergeevich — software developer of computer security problems laboratory, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Science (SPIIRAS). Research interests: distributed system security, embedded devices, event correlation, cyber-physical security systems. The number of publications — 5. levshun@comsec.spb.ru, http://comsec.spb.ru/levshun; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7-(812)-328-71-81.

Чечулин Андрей Алексеевич — к-т техн. наук, старший научный сотрудник лаборатории проблем компьютерной безопасности, Федеральное государственное бюджетное учреждение науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: безопасность компьютерных сетей, обнаружение вторжений, анализ сетевого трафика, анализ уязвимостей. Число научных публикаций — 150. andreych@bk.ru, http://comsec.spb.ru/ru/staff/chechulin; 14-я линия В.О., 39, ком. 205, Санкт-Петербург, 199178; р.т.: +7-(812)-328-71-81.

Chechulin Andrey Alexeevich — Ph.D., senior researcher of computer security problems laboratory, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Science (SPIIRAS). Research interests: computer network security, intrusion detection, analysis of the network traffic, vulnerability analysis. The number of publications — 150. andreych@bk.ru, http://comsec.spb.ru/ru/staff/chechulin; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7-(812)-328-71-81.

Котенко Игорь Витальевич — д-р техн. наук, профессор, заведующий лабораторией проблем компьютерной безопасности, Федеральное государственное бюджетное учреждение науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН). Область научных интересов: безопасность компьютерных сетей, в том числе управление политиками безопасности, разграничение доступа, аутентификация, анализ защищенности, обнаружение компьютерных атак, межсетевые экраны, защита от вирусов и сетевых червей, анализ и верификация

протоколов безопасности и систем защиты информации, защита программного обеспечения от взлома и управление цифровыми правами, технологии моделирования и визуализации для противодействия кибер–терроризму. Число научных публикаций — 450. ivkote@comsec.spb.ru, http://www.comsec.spb.ru; 14-я линия В.О., 39, Санкт-Петербург, 199178; р.т.: +7-(812)-328-71-81, Факс: +7(812)328–4450.

Kotenko Igor Vitalievich — Ph.D., Dr. Sci., professor, head of computer security problems Laboratory, St. Petersburg Institute for Informatics and Automation of Russian Academy of Sciences (SPIIRAS). Research interests: computer network security, including security policy management, access control, authentication, network security analysis, intrusion detection, firewalls, deception systems, malware protection, verification of security systems, digital right management, modeling, simulation and visualization technologies for counteraction to cyber terrorism. The number of publications — 450. ivkote@comsec.spb.ru, http://www.comsec.spb.ru; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7-(812)-328-71-81, Fax: +7(812)328–4450.

Поддержка исследований. Работа выполнена при финансовой поддержке РФФИ (проекты №14-07-00697, 14-07-00417, 15-07-07451, 16-37-00338, 16-29-09482 офи_м), при частичной поддержке бюджетных тем № 0073-2015-0004 и 0073-2015-0007, а также гранта РНФ 15-11-30029 в СПИИРАН

Acknowledgements. This research is supported by RFBR (projects No. 14-07-00697, 14-07-00417, 15-07-07451, 16-37-00338, 16-29-09482 офи_м), in part by the budget (projects No. 0073-2015-0004 and 0073-2015-0007) and by the grant of RSF 15-11-30029 in SPIIRAS

РЕФЕРАТ

Федорченко А.В., Левшун Д.С., Чечулин А.А., Котенко И.В. **Анализ методов корреляции событий безопасности в SIEM-системах. Часть 1.**

Данная статья посвящена анализу различных схем, реализующих процесс корреляции событий безопасности в SIEM-системах. Исследуемая предметная область изучается в мировом сообществе на протяжении более чем двух десятилетий и в данном исследовании были учтены многие устоявшиеся аспекты корреляции. Рассмотрен ряд работ, направленных преимущественно на модульное представление процесса корреляции и применение различных подходов на его отдельных этапах.

В ходе проделанного анализа была получена и обоснована собственная схема процесса корреляции, состоящая из нескольких уровней и отдельных этапов на каждом из них, а также из установленных связей как между уровнями, так и между этапами. В работе подробно описывается каждый из этапов процесса корреляции, а также определенные для него функции.

SUMMARY

Fedorchenko A.V., Levshun D.S., Chechulin A.A., Kotenko I.V. **An Analysis of Security Event Correlation Techniques in Siem-Systems. Part 1.**

This paper provides an analysis of various schemes that can be applied for the security events correlation in SIEM systems. This field has been studied in the international community for over two decades, so in this paper we take into account established aspects of correlation. The paper contains a description of several techniques of information correlation that use a modular representation of correlation process and apply different approaches at separate stages of correlation.

During the investigation of correlation techniques the new scheme of correlation process was obtained and proved. This scheme contains several levels which, in their turn, consist of independent elements. This scheme also contains links between levels and between elements. In addition to the general scheme, the paper provides a description of each correlation stage and its specific functions.