БЕЗОПАСНОСТЬ СЛОЖНЫХ ИНФОРМАЦИОННО-КОММУНИКАЦИОННЫХ СИСТЕМ

УДК 004.056
DOI: 10.17586/0021-3454-2018-61-11-997-1004

АНАЛИЗ СВОЙСТВ СОБЫТИЙ БЕЗОПАСНОСТИ ДЛЯ ОБНАРУЖЕНИЯ ИНФОРМАЦИОННЫХ ОБЪЕКТОВ И ИХ ТИПОВ В НЕОПРЕДЕЛЕННЫХ ИНФРАСТРУКТУРАХ

А. В. ФЕДОРЧЕНКО

Санкт-Петербургский институт информатики и автоматизации РАН,
199178, Санкт-Петербург, Россия
E-mail: fedorchenko@comsec.spb.ru

Исследуется область корреляции событий для систем управления информацией и событиями безопасности. Цель исследований заключается в определении типов информационных объектов с помощью анализа журнала событий безопасности целевой инфраструктуры. Предлагаемый механизм корреляции событий основан на определении связей между эквивалентными свойствами событий по признаку взаимной используемости. Результатом исследования полученных связей является определение анализируемой инфраструктуры в виде типов высокоуровневых объектов. Описываются результаты эксперимента по структурному анализу событий журнала безопасности ОС Windows, а также случаи нестабильности предлагаемого механизма и их возможные причины. Приводится оценка и интерпретация полученных результатов, свидетельствующих о возможности применения представленного подхода на практике.

Ключевые слова: *события безопасности, корреляция событий, структурный анализ данных, мониторинг безопасности, SIEM-системы*

Введение. Современные информационно-коммуникационные инфраструктуры, как правило, являются достаточно масштабными и имеют сложные уникальные архитектуры. Задача обеспечения безопасности подобных инфраструктур является актуальной, поскольку они реализованы в финансовых и правительственных организациях, крупных компаниях и критически важных объектах военного, промышленного и энергетического назначения. Однако по мере информатизации и глобализации этих сфер деятельности выполнение задач безопасности за счет существующих подходов становится все более затруднительным.

Одним из классов средств защиты различных критических инфраструктур являются системы управления информацией и событиями безопасности (Security Information and Event Management — SIEM) [1—4]. Последующее развитие подобных систем подразумевает преодоление таких сложностей, как гетерогенность исходных данных и их большой объем, высокая уникальность и изменчивость анализируемых объектов. В свою очередь, основополагающим компонентом архитектуры SIEM, предназначенным для реализации данного функционала, является механизм корреляции.

В настоящей статье представлен механизм корреляции событий, основанный на структурном анализе журнала безопасности. В частности, описывается вид отношений между свойствами событий по признаку их взаимной используемости. Предполагается, что различные события задаются разными действиями, которые выполняются между определенными типами информационных объектов, а свойства событий являются характеристиками данных типов объектов. В результате вычисления „силы“ связей, реализующих данный вид отношений между свойствами событий, теоретически предполагается и практически подтверждается получение групп свойств, определяющих тип информационных объектов анализируемой инфраструктуры. Новизна предлагаемого подхода заключается в адаптации механизма корреляции событий к неопределенной целевой инфраструктуре за счет автоматизированного принятия решения о составе характеристик типов ее объектов. Применение такого подхода позволяет минимизировать затраты человеческих и временных ресурсов на конфигурирование механизма корреляции для работы в целевой инфраструктуре.

Релевантные работы. Механизм корреляции — определение отношений между событиями, а также причинно-следственных зависимостей между ними — позволяет обнаруживать вредоносные и аномальные действия, определять источник и цель компьютерной атаки, выявлять многошаговые атаки и зависит от конкретной реализации [5, 6].

Несмотря на разнообразие механизмов и техник корреляции событий [7—9] в настоящее время наиболее распространенным остается правило-ориентированный метод [7—10]. Основным недостатками этого метода являются сложность и длительность составления правил администратором безопасности. Даже при использовании предопределенных наборов правил из соответствующих шаблонов конфигурация механизма корреляции применительно к каждой отдельно взятой инфраструктуре является достаточно трудоемким процессом и требует высокого уровня квалификации специалиста. При этом эффективность реализации механизма корреляции с помощью правило-ориентированного метода напрямую зависит от качества составления словаря правил.

Помимо подхода, базирующегося на использовании логических конструкций — правил, множество сигнатурных механизмов (методов) корреляции событий, предупреждений и инцидентов безопасности основано на использовании таких моделей, как шаблоны (сценарии) [10], графы [7, 11], конечные автоматы [7,12], состояния [13, 14] и др.

К другому виду механизмов корреляции — эвристическим, относятся методы, в которых анализ поведения объектов осуществляется на основе заранее составленной интеллектуальной модели. Объектами подобных моделей могут служить узлы сети, сетевой трафик, активы, ресурсы и другие элементы инфраструктуры. Эвристические методы основаны преимущественно на интеллектуальном анализе данных и машинном обучении, а также являются самообучающимися. К данным методам относятся: байесовские сети [7, 10, 15], иммунные сети [7, 15], искусственные нейронные сети [7, 15, 16] и др.

Постановка задачи. Процесс выявления корреляции является непрерывным и должен выполняться в реальном масштабе времени. Место и роль этого процесса в SIEM-системах обусловливаются его задачами:

1) определение взаимосвязей между разнородной информацией безопасности;

2) определение информационных объектов целевой инфраструктуры и их характеристик;

3) группировка низкоуровневых событий в высокоуровневые метасобытия;

4) выявление инцидентов безопасности на основе анализа поведения разноуровневых информационных объектов инфраструктуры.

Таким образом, механизм корреляции (процесс установления взаимосвязи) осуществляет обработку данных от момента их поступления из гетерогенных источников и до формирования отчета о текущем состоянии защищенности анализируемой инфраструктуры.

Глобальной задачей исследований является разработка механизма корреляции разнородных событий безопасности с его автоматизированной адаптацией к работе в неопределенной и динамически изменяемой инфраструктуре. В настоящей статье рассматривается выполнение второй из указанных задач на основе структурного анализа журнала событий безопасности.

Определение типов информационных объектов. Исходными данными для выполнения структурного анализа является множество событий безопасности E журнала L:

$$E^L = \{e_1, e_2, \ldots, e_n\}.$$

Под событием понимается факт либо результат какого-либо действия на любом из его этапов: попытка (событие отказа, событие старта), промежуточный результат действия (события, описывающие продолжительные по времени действия), конечный результат выполнения действия (завершено корректно, завершено с ошибкой). Каждое событие e состоит из множества свойств p и их значений v, которые характеризуют определенное действие:

$$e = \{p : v\}, P = \{p_1, p_2, \ldots, p_d\}, V = \left\{V^{p_1}, V^{p_2}, \ldots, V^{p_d}\right\},$$

где P и V^p — множества свойств и множества их значений соответственно, d — общее количество свойств.

В каждый момент времени неопределенная инфраструктура I состоит из множества информационных объектов O:

$$O^I = \{o_1, o_2, \ldots, o_s\},$$

существующих во времени (т.е. имеющих некоторую продолжительность жизненного цикла) и состояние которых описывается с помощью одной или нескольких характеристик x из множества характеристик X:

$$o \in O^I, \ x \in X, \ X = \{x_1, x_2, \ldots, x_r\}, \ |X| \geq 1,$$

где r — общее число характеристик объекта.

Каждый информационный объект обязательно является частью более высокоуровневого объекта и (или) содержит более низкоуровневые объекты. Связь между объектами определяется их непосредственным взаимодействием.

Предполагается, что набор характеристик X^o, описывающих информационный объект o, однозначно определяет тип t информационного объекта:

$$X^o = t, \ o \in O, \ X^o \subset X, \ t \in T, \ T = \{t_1, t_2, \ldots, t_m\},$$

где m — количество типов информационных объектов множества T.

При этом каждая характеристика x из множества X принадлежит только одному типу информационного объекта:

$$t = X^t, \ X^t \subset X, \ X^{t_i} \cap X^{t_j} = \varnothing \ \ \forall i, j \in \{1, 2, \ldots, m\}, \ i \neq j.$$

Допускается, что изменение состояния целевой инфраструктуры I описывается событиями E в журнале L. Связь между событиями безопасности E и объектами инфраструктуры O строится на следующем утверждении: каждое свойство p события e является характеристикой x объекта o:

$$p^e = x^o.$$

Другими словами, каждое событие описывает минимум один информационный объект инфраструктуры.

Отношения между свойствами событий безопасности могут определяться различными признаками, такими как взаимная используемость свойств событий и используемость по типам событий, взаимная уникальность значений свойств, коррелируемость значений свойств и др.

Предполагается, что различные события безопасности E характеризуют различные действия, выполняемые объектами-источниками над объектами-целями из множества O и описываемые с помощью характеристик X объектов. Функция определения отношений эквивалентности свойств по их совместной используемости опирается на следующую гипотезу: исключительно совместное использование свойств p_1 и p_2, задающих характеристики x_1 и x_2 одного или нескольких объектов из множества O, свидетельствует о принадлежности описываемых объектов к одному или нескольким типам T объектов.

Функция определения связи по взаимному использованию f свойств p_1 и p_2 событий описывается следующим выражением:

$$f\left(p_1, p_2\right) = \frac{\left|E^{p_1} \cap E^{p_2}\right|}{\left|E^{p_1} \cup E^{p_2}\right|},\ \{p_1, p_2\} \subset P,\ \left\{E^{p_1}, E^{p_2}\right\} \subset E,$$

где E^p — множество событий, каждый элемент которого содержит определенное свойство p.

Результат вычисления отношения между событиями имеет диапазон значений [0,1]. Следовательно, для установления факта наличия связи между свойствами (перевода к бинарному результату) необходимо задавать некий порог lim, преодоление которого в большую сторону позволяет считать связь существенной.

Эксперименты и результаты. Исходными данными для проведения экспериментов были события системного журнала безопасности хоста под управлением ОС Windows 8. Подсистема ведения журнала (logging) в операционной системе была настроена на сбор максимального количества типов событий безопасности. После нормализации характеристики анализируемого журнала следующие:

— количество событий ~ 6 700 000;

— количество заявленных типов событий — 44 (из более чем 250 заявленных разработчиками [17]);

— количество свойств событий — 110;

— размер данных журнала — 7Гб в формате XML, 1,25 Гб в формате CSV;

— время записи журнала — 36 дней.

Выше было сформулирована гипотеза, определяющая типы объектов по свойствам, эквивалентным по взаимной используемости. На рис. 1 представлена гистограмма использования свойств относительно общего количества событий. Отчетливо видно, что отдельные группы свойств имеют равный либо очень близкий по значению показатель общей используемости (a). Таким образом, выдвигаемая гипотеза предварительно подтверждается.

В результате эксперимента было выделено 18 групп свойств, а общее количество свойств равно 60, при этом $\lim \to 1$. Другими словами, свойства одной группы с высокой вероятностью используются в событиях исключительно совместно. Наиболее значимые типы объектов представлены на рис. 2, также для каждого типа отмечено *среднее* значение показателя общей используемости составляющих группу свойств, а сами группы упорядочены по данному показателю.

Отдельный интерес представляют свойства нулевой группы, которые наблюдаются во всех типах событий. Следует отметить, что факт наличия нулевой группы в анализируемой инфраструктуре можно установить путем проверки показателя общей используемости, поскольку для подобных свойств он должен быть равен 1.

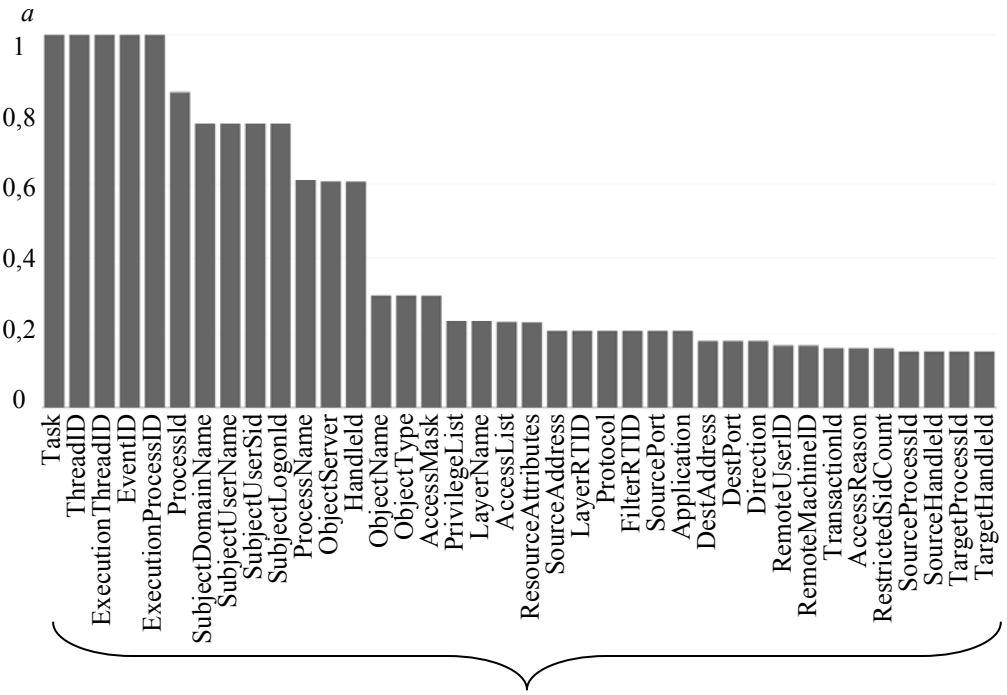
a
1
0,8
0,6
0,4
0,2
0
Task
ThreadID
ExecutionThreadID
EventID
ExecutionProcessID
ProcessId
SubjectDomainName
SubjectUserName
SubjectUserSid
SubjectLogonId
ProcessName
ObjectServer
HandleId
ObjectName
ObjectType
AccessMask
PrivilegeList
LayerName
AccessList
ResourceAttributes
SourceAddress
LayerRTID
Protocol
FilterRTID
SourcePort
Application
DestAddress
DestPort
Direction
RemoteUserID
RemoteMachineID
TransactionId
AccessReason
RestrictedSidCount
SourceProcessId
SourceHandleId
TargetProcessId
TargetHandleId
Свойства

Рис. 1

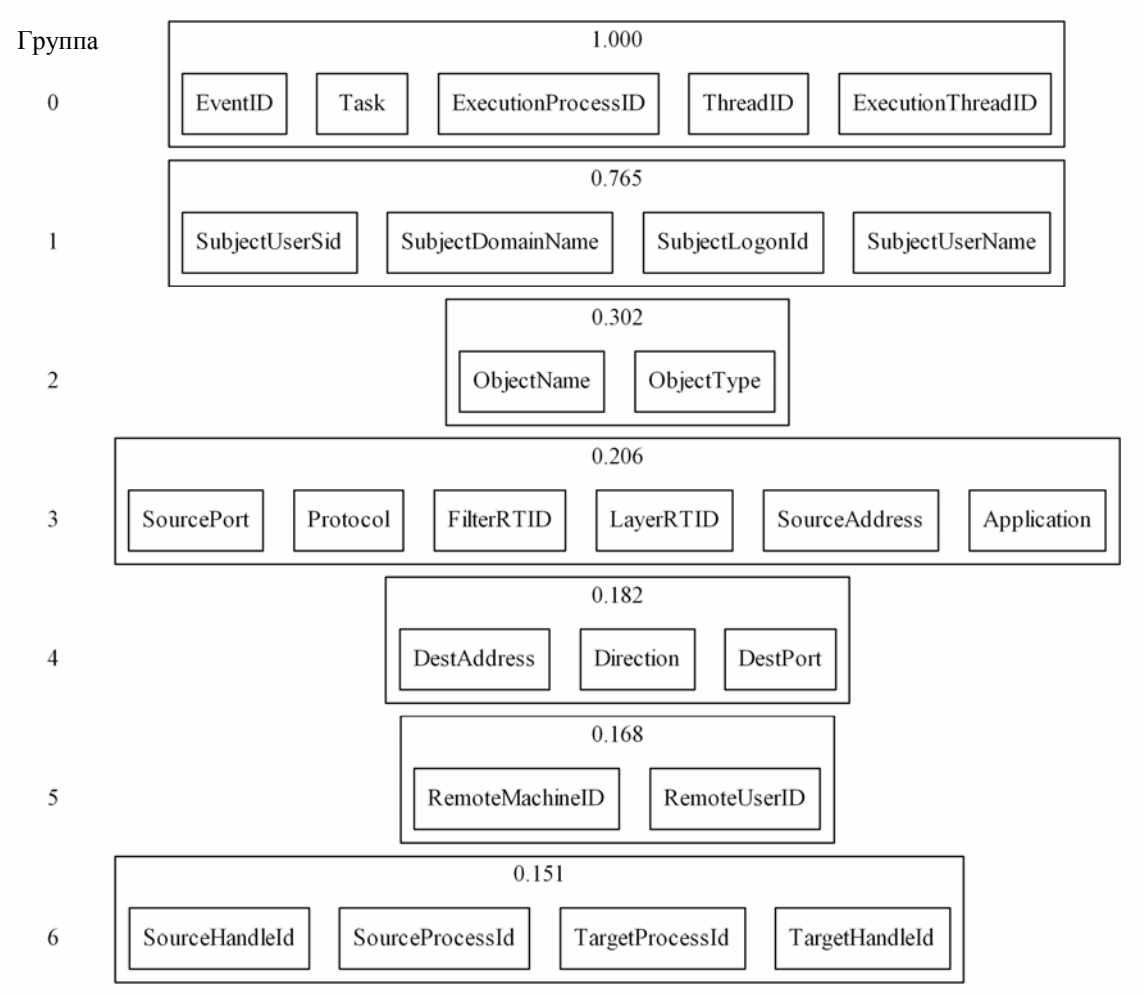
Группа
0
1.000
EventID
Task
ExecutionProcessID
ThreadID
ExecutionThreadID
1
0.765
SubjectUserSid
SubjectDomainName
SubjectLogonId
SubjectUserName
2
0.302
ObjectName
ObjectType
3
0.206
SourcePort
Protocol
FilterRTID
LayerRTID
SourceAddress
Application
4
0.182
DestAddress
Direction
DestPort
5
0.168
RemoteMachineID
RemoteUserID
6
0.151
SourceHandleId
SourceProcessId
TargetProcessId
TargetHandleId

Рис. 2

Общее рассмотрение сформированных типов объектов позволяет заключить, что большинство их них являются достаточно семантически выраженными. Так, свойства групп 1 и 2 описывают типы объектов „Subject“ и „Object“ соответственно; наряду с этим группа 6 явно содержит два типа объектов — „Source“ и „Target“. Из этого следует, что подход к определению типов информационных объектов инфраструктуры необходимо расширить для случая, когда показатель используемости двух типов объектов одинаков.

Заключение. Рассмотрен механизм корреляции событий, предназначенный для выявления типов информационных объектов целевой инфраструктуры за счет структурного анализа журнала безопасности. Проведенные эксперименты не только подтвердили предположение о типизации информационных объектов, но и практически доказали возможность предложенного подхода в SIEM-системах для задачи обнаружения их типов.

Исследования выполнены при поддержке Президента Российской Федерации, грант № МК-314.2017.9.

СПИСОК ЛИТЕРАТУРЫ

1. *Kotenko I. V., Chechulin A. A.* A Cyber attack modeling and impact assessment framework // Proc. of 5th Intern. Conf. on Cyber Conflict (CyCon 2013). 2013. P. 119—142.

2. *Kotenko I. V., Polubelova O. V., Saenko I. B.* The ontological approach for siem data repository implementation // Proc. IEEE Intern. Conf. on Green Computing and Communications, Conference on Internet of Things, and Conference on Cyber, Physical and Social Computing. 2012. P. 761—766.

3. *Дойникова Е. В., Котенко И. В.* Методики и программный компонент оценки рисков на основе графов атак для систем управления информацией и событиями безопасности // Информационно-управляющие системы. 2016. № 5. С. 54—65.

4. *Котенко И. В., Дойникова Е. В.* Методика выбора контрмер на основе комплексной системы показателей защищенности в системах управления информацией и событиями безопасности // Информационно-управляющие системы. 2015. № 3. С. 60—69.

5. *Kruegel C., Valeur F., Vigna G.* Intrusion Detection and Correlation: Challenges and Solutions. Springer, 2004.

6. *Федорченко А. В., Левшун Д. С., Чечулин А. А., Котенко И. В.* Анализ методов корреляции событий безопасности в SIEM-системах. Ч. 1 // Тр. СПИИРАН. 2016. Вып. 47. С. 5—27.

7. *Muller A.* Event Correlation Engine: Master`s Thesis / Swiss Federal Institute of Technology. Zurich. 2009. 165 p.

8. *Limmer T., Dressler F.* Survey of Event Correlation Techniques for Attack Detection in Early Warning Systems: Tech. report / University of Erlangen, Germany. 2008. 37 p.

9. *Федорченко А. В., Левшун Д. С., Чечулин А. А., Котенко И. В.* Анализ методов корреляции событий безопасности в SIEM-системах. Ч. 2 // Тр. СПИИРАН. 2016. Вып. 49. С. 208—225.

10. *Sadoddin R., Ghorbani A.* Alert correlation survey: framework and techniques // Proc. of the Intern. Conf. on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services (PST`06). 2006. Art. no. 37.

11. *Ning P., Xu D.* Correlation analysis of intrusion alerts // Intrusion Detection Systems: Ser. Advances in Information Security. 2008. Vol. 38. P. 65—92.

12. *Ghorbani A. A., Lu W., Tavallaee M.* Network Intrusion Detection and Prevention. Springer, 2010. 224 p.

13. *Hasan M. A.* Conceptual framework for network management event correlation and filtering systems // Proc. of the 6th IFIP/IEEE Intern. Symp. on Integrated Network Management. 1999. P. 233—246.

14. *Zurutuza U., Uribeetxeberria R.* Intrusion detection alarm correlation: A survey // Proc. of IADAT Intern. Conf. on Telecommunications and Computer Networks. 2004. P. 1—3.

15. *Guerer D. W., Khan I., Ogler R., Keffer R.* An Artificial Intelligence Approach to Network Fault Management / SRI International, CA, USA. 1996. 10 p.

16. *Elshoush H. T., Osman I. M.* Alert correlation in collaborative intelligent intrusion detection systems — a survey // Applied Soft Computing. 2011. P. 4349—4365.

17. Windows Security Log Events [Электронный ресурс]: <https://www.ultimatewindowssecurity.com/securitylog/encyclopedia/Default.aspx>, 30.05.2018.

Сведения об авторе

Андрей Владимирович Федорченко — аспирант; СПИИРАН; лаборатория проблем компьютерной безопасности; мл. научный сотрудник; E-mail: fedorchenko@comsec.spb.ru

Поступила в редакцию
27.08.18 г.

Ссылка для цитирования: *Федорченко А. В.* Анализ свойств событий безопасности для обнаружения информационных объектов и их типов в неопределенных инфраструктурах // Изв. вузов. Приборостроение. 2018. Т. 61, № 11. С. 997—1004.

ANALYSIS OF SECURITY EVENTS PROPERTIES FOR DETECTION OF INFORMATION OBJECTS AND THEIR TYPES IN UNCERTAIN INFRASTRUCTURES

A. V. Fedorhcenko

St. Petersburg Institute for Informatics and Automation of the RAS,
199178, St. Petersburg, Russia
E-mail: fedorchenko@comsec.spb.ru

The field of event correlation for systems of security information and event management systems is investigated. The purpose of the research is to determine the types of information objects by analyzing the security event log of the infrastructure under study. A correlation approach based on definition of relationships between equivalent events properties by their mutual utilizing is proposed. The study of revealed relationships results in definition of the analyzed infrastructure in the form of types of high-level objects. Results of an experiment on the structural analysis of the Windows security events log are presented. The cases of unstable work of the proposed approach and their possible causes are described. The evaluation and interpretation of the obtained results testifying to the possibility of application of the presented approach in practice are given.

Keywords: security events, events correlation, structural data analysis, security monitoring, SIEM-systems

REFERENCES

1. Kotenko I.V., Chechulin A.A. *Proc. of 5th Intern. Conf. on Cyber Conflict (CyCon 2013)*, 2013, pp. 119–142.
2. Kotenko I.V., Polubelova O.V., Saenko I.B. *Proc. 2012 IEEE Intern. Conf. on Green Computing and Communications, Conference on Internet of Things, and Conference on Cyber, Physical and Social Computing*, 2012, pp. 761–766.
3. Doynikova E.V., Kotenko I.V. *Informatsionno-upravliaiushchie sistemy* (Information and Control Systems), 2016, no. 5, pp. 54–65. (in Russ.)
4. Kotenko I.V., Doynikova E.V. *Informatsionno-upravliaiushchie sistemy* (Information and Control Systems), 2015, no. 3, pp. 60–69. (in Russ.)
5. Kruegel C., Valeur F., Vigna G. *Intrusion Detection and Correlation: Challenges and Solutions*, Springer, 2004.
6. Fedorchenko A.V., Levshun D.S., Chechulin A.A., Kotenko I.V. *Trudy SPIIRAN* (SPIIRAS Proceedings), 2016, no. 47, pp. 5–27. (in Russ.)
7. Muller A. *Event Correlation Engine*, Master`s Thesis, Swiss Federal Institute of Technology, Zurich, 2009, 165 p.
8. Limmer T., Dressler F. *Survey of Event Correlation Techniques for Attack Detection in Early Warning Systems*, Tech. report, University of Erlangen, Germany, 2008, 37 p.
9. Fedorchenko A.V., Levshun D.S., Chechulin A.A., Kotenko I.V. *Trudy SPIIRAN* (SPIIRAS Proceedings), 2016, no. 49, pp. 208–225. (in Russ.)
10. Sadoddin R., Ghorbani A. *Proc. of the Intern. Conf. on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services (PST`06)*, 2006, Art. no. 37.
11. Ning P., Xu D. *Intrusion Detection Systems: series Advances in Information Security*, 2008, vol. 38, pp. 65–92.
12. Ghorbani A.A., Lu W., Tavallaee M. *Network Intrusion Detection and Prevention*, Springer, 2010, 224 p.
13. Hasan M.A. *Proc. of the 6th IFIP/IEEE Intern. Symp. on Integrated Network Management*, 1999, pp. 233–246.

14. Zurutuza U., Uribeetxeberria R. *Proc. of IADAT Intern. Conf. on Telecommunications and Computer Networks*, 2004, pp. 1–3.
15. Guerer D.W., Khan I., Ogler R., Keffer R. *An Artificial Intelligence Approach to Network Fault Management*, SRI International, CA, USA, 1996, 10 p.
16. Elshoush H.T., Osman I.M. *Applied Soft Computing*, 2011, pp. 4349–4365.
17. *Windows Security Log Events,* https://www.ultimatewindowssecurity.com/securitylog/encyclopedia/Default.aspx.

Data on author

Andrey V. Fedorchenko — Post-Graduate Student; St. Petersburg Institute for Informatics and Automation of the RAS, Laboratory of Cyber-Security Problems; Junior Scientist; E-mail: fedorchenko@comsec.spb.ru

For citation: Fedorhcenko A. V. Analysis of security events properties for detection of information objects and their types in uncertain infrastructures. *Journal of Instrument Engineering*. 2018. Vol. 61, N 11. P. 997—1004 (in Russian).

DOI: 10.17586/0021-3454-2018-61-11-997-1004