V. S. Matveeva

Information Entropy and Its Application for Information Security Tasks

Keywords: information entropy, detecting file types, detecting encrypted data

The article deals with information entropy application in information security and its efficiency. The article provides an analysis of entropy properties and dependences and their impact on results of solving modern information security tasks.

В. С. Матвеева

ЭНТРОПИЯ И ЕЕ ИСПОЛЬЗОВАНИЕ ДЛЯ РЕШЕНИЯ ЗАДАЧ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ

В 1948 г. в своей работе [1] Клод Шеннон ввел понятие «энтропия информации». Согласно К. Шеннону, энтропия удовлетворяет следующим свойствам:

1) Энтропия непрерывна по вероятности появления каждого из символов, то есть изменение значения вероятности на малую величину должно вызывать малое результирующее изменение энтропии;

2) Если вероятность появления каждого из символов равна 1/n, где n — количество возможных значений символов, то энтропия — монотонная возрастающая функция от n;

3) Энтропия характеризуется свойством аддитивности, то есть если есть возможность сделать выбор из двух промежуточных вариантов, то результирующая энтропия должна быть взвешенной суммой энтропий каждого из промежуточных вариантов.

Для удовлетворения всем этим свойствам К. Шенноном введена формула:

$$H = -K \sum_{i=1}^{n} p(i) * \log p(i), \quad (1)$$

где К — положительная константа.

Для энтропии информации введена формула:

$$H = -\sum_{i=0}^{255} p(i) \log_2 p(i), \quad (2)$$

где основание логарифма соответствует количеству возможных значений бита, и соответственно энтропия является мерой количества бит, необходимых для описания байт в измеряемом фрагменте информации и темпа роста количества этих бит с ростом размера данных [2].

Легко показать, что она удовлетворяет всем трем свойствам:

График функции $f(\rho(i)) = \rho(i) * log_2\, p(i)$, где $\rho(i) \in [0.1]$, показан на рис. 1 и является непрерывным.

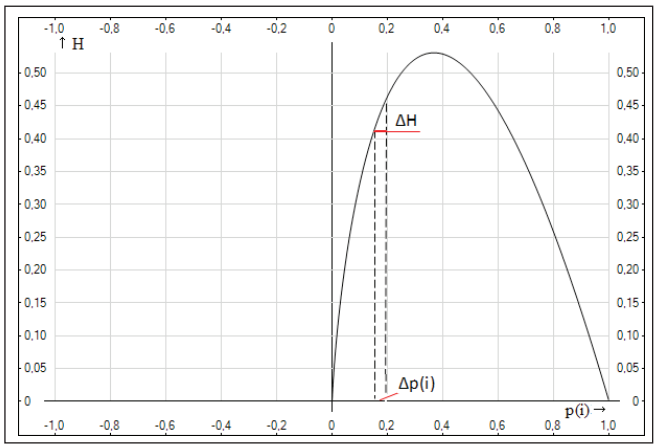

Рис. 1. График функции f(ρ(i))

Любое сколь угодно малое изменение вероятности ρ(i) вызывает изменение значения энтропии.

При $\rho(i)=1/n$ для каждого i, где n — количество значений символов, которые могут принимать символы в последовательности данных, энтропия H принимает максимальное значение, равное $\log n$. Доказательство этого приведено в [3]. При измерении информации в битах максимальное значение энтропии равно $\log_2 2^8 = 8$.

Логарифм введен с целью получения эффекта аддитивности количества информации ($I(s_i) = \log_2 1/p(i)$), получаемого при появлении двух независимых символов. Для получения вероятности таких событий надо перемножить их вероятности [3]:

$$I(s_1) + I(s_2) = \log_2 \frac{1}{p(1)} + \log_2 \frac{1}{p(2)} = \log_2 \frac{1}{p(1)*p(2)} = I(s_1, s_2).$$

Также К. Шеннон доказал [2], что при эффективном переводе последовательности символов в последовательность бит для отправки по каналу каждой последовательности s^n можно поставить в соответствие последовательность бит $l(s^n)$, причем:

$l(s^n) = \log_2 1/p(s^n)$.

Путем преобразования формулы (2) с учетом формулы $p(i) = кол(i)/n$, где $кол(i)$ — это количество байт со значением i, n — общий размер данных, получаем следующую формулу для энтропии информации:

$$H = \log_2 n - \frac{1}{n} \sum_{i=0}^{255} кол(i) * \log_2 кол(i) = \log_2 n - \sum_{i=0}^{255} p(i) * \log_2 кол(i). \quad (3)$$

Эти формулы легче оценивать, так как диапазон значений в формулах не на интервале [0,1], а на интервале [0,+∞].

Из формул (3) видна зависимость энтропии информации от размера данных (n) и распределения значений байт в этих данных (то есть самих данных, частота встречаемости каждого символа ($кол(i)$). Поясним эти зависимости:

1) Установить общий закон изменения энтропии в зависимости от размера данных невозможно, так как $кол(i)$ также зависит от размера данных и от формата содержимого. Однако логично предположить, что энтропия данных небольшого размера будет отличаться от энтропии данных такого же формата большого размера, так как количество информации в небольшом объеме данных (а значит, и количество бит, необходимых для описания данных) меньше, чем в большом объеме данных, где зависимость установить сложнее. Зависимость энтропии от размера показана на основании экспериментов в ряде статей [4, 5]. На практике, используя небольшую выборку файлов одного формата содержимого, можно также показать эту зависимость.

С помощью программного модуля «SharpAESCrypt» версии 1.0, который производит зашифрование/расшифрование файлов по алгоритму AES, произведено шифрование 1320 файлов различных размеров и сделан подсчет значений энтропии для каждого из них. Так как различные данные зашифрованы с использованием одного алгоритма шифрования, то статистические свойства их содержимого одинаковые. В результате сортировки полученных результатов выявлены следующие закономерности в зависимости от размеров файлов:

Размер файлов в диапазоне 101709–114702 байт, значения энтропии > 7,997

Размер файлов в диапазоне 114703–238445 байт, значения энтропии > 7,998

Размер файлов в диапазоне 238446–2037517 байт, значения энтропии > 7,999

Размер файлов в диапазоне 2037518–21006925 байт, значения энтропии > 7,9999

Таким образом, показано, что значения энтропии зависят от размера данных при условии одинакового характера содержимого данных. На более маленьких размерах данных значение энтропии не может достигать максимальных значений.

2) Распределение значений байт в данных обуславливается применяемым алфавитом и/или способом преобразования данных (формат содержимого), если данные представлены как результат работы алгоритмов сжатия, шифрования и т. д. Если производить расчет энтропии двух текстовых файлов размером 1 Мб, то при содержимом в виде 1048576 символов со значением 1 энтропия файла равна 0, а при содержимом в виде 1048576 байт данных, зашифрованных по алгоритму AES, энтропия файла равна 7,999850.

Таким образом, показано, что энтропия отображает характер содержимого данных.

Как было сказано выше, энтропия является характеристикой данных, так как рассчитывается на основании распределения значений байт в них. В связи с этим после введения этого понятия К. Шенноном энтропия начала активно использоваться для решения задач информационной безопасности (обеспечение конфиденциальности, целостности, доступности).

Сегодня энтропию информации наиболее часто используют для решения следующих вопросов в рамках задач информационной безопасности:

1) Определение границ файла, отклонений в самом файле [6, 7, 8]. Принцип основан на эмпирическом расчете диапазонов значений энтропий для интересующих фрагментов файлов и сравнении энтропии неизвестного фрагмента файла с полученными значениями в ходе эксперимента. Неизвестный фрагмент относят к тому или иному известному в случае максимальной близости его значения энтропии к экспериментальному для конкретного типа фрагментов.

Актуальность этого направления использования энтропии подтверждена на конференции ZeroNights 2013 Антоном Дорфманом, который выступил с докладом «Реверсинг форматов данных: о чем могут рассказать данные». В докладе он изложил принципы классификации форматов файлов в зависимости от их содержимого. Для определения границ заголовков для разных форматов файлов докладчиком предложено использовать энтропию.

Недостатки использования энтропии при определении границ файла, отклонений в самом файле совпадают с теми, что изложены для п. 2.

2) Классификация по форматам файлов [4, 6, 9, 10, 11, 12, 13, 14]. Принцип основан на эмпирическом расчете диапазонов значений энтропий для интересующих форматов файлов и сравнении энтропии неизвестного формата файла с полученными значениями в ходе эксперимента. Неизвестный формат относят к тому или иному известному в случае максимальной близости его значения энтропии к экспериментальному для конкретного фрагмента.

На основе проведенных экспериментов в работах [6, 10] показано, что результат правильного определения формата файла при описанном принципе от 0 до 100 % и не может служить самостоятельным решением этой задачи.

В работах [4, 6, 9] показано, что диапазоны значений энтропий файлов существующих форматов файлов значительно пересекаются. Пример гистограммы из работы [9] приведен на рис. 2.

Таким образом, описанный способ не может использоваться для корректного определения формата файла, но может использоваться в совокупности с другими для достижения высоких показателей определения формата файла [11, 12, 13, 14].

Помимо этого, при классификации файлов с высокой энтропией (файлы-архивы, pdf-файлы, MP3-файлы и т. д.) получается высокий процент ложных срабатываний [15].

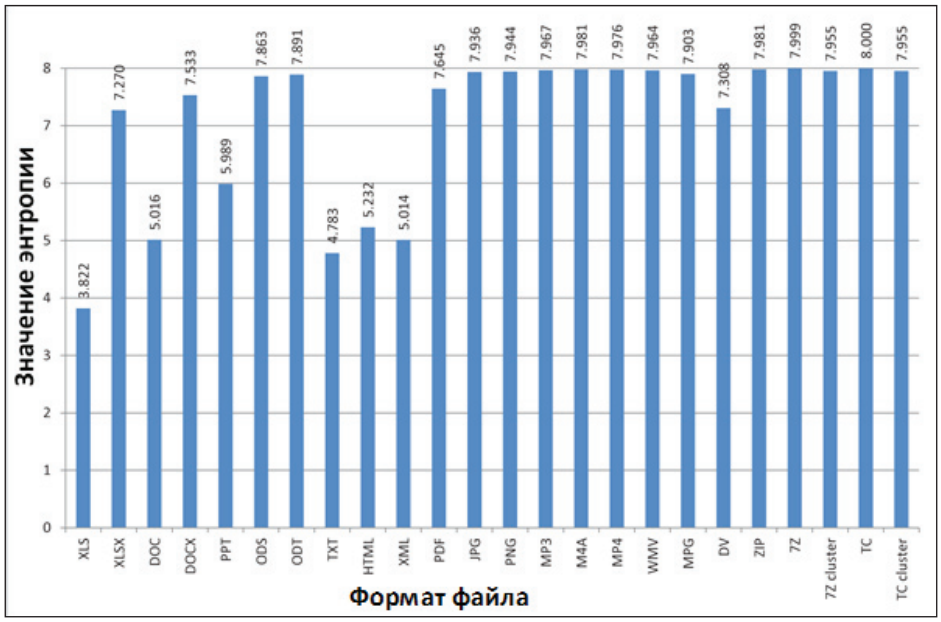

Рис. 2. Гистограмма, отображающая значение энтропии для различных форматов файлов

3) Определение упакованных, зашифрованных и случайных данных [4, 16, 17, 18, 19, 20]. Применение энтропии для решения этой задачи дает неоднозначный результат при определении зашифрованных данных, так как диапазон значений энтропии для существующих форматов данных пересекается с диапазоном значений для шифрованной и случайной информации.

В [15] приведены результаты применения различных подходов для классификации файлов. Все подходы дают низкий процент корректного определения сжатых/зашифрованных файлов.

На практике был проведен эксперимент. На небольшой выборке файлов различных форматов в количестве 100 экземпляров файлов на каждый формат показано, что диапазоны возможных значений энтропии для этих форматов пересекаются. Выборки в количестве 100 экземпляров для каждого формата хватило, чтобы продемонстрировать пересечение диапазонов возможных значений энтропий. Полученные результаты приведены в таблице 1.

Таблица 1. Результаты расчета значений энтропии для различных форматов файлов

Формат файла	**Минимальное значение энтропии**	**Максимальное значение энтропии**
RAR	6,709057	7,999813
Docx	1,992581	7,998396
PDF	5,083566	7,998874
EXE	2,572076	7,999985
AVI	4,966052	7,999895
MP3	2,926377	7,998412
AES-криптоконтейнер	7,997363	7,999992

Таким образом, энтропия не может служить эффективной мерой для идентификации зашифрованных/случайных данных.

Согласно [21], 80 % всех вредоносных программ являются упакованными. В [22] используют энтропию для анализа упакованности/зашифрованности секций исполняемых файлов. Для анализа

рассчитываются пороговые значения с учетом применяемых злоумышленниками техник упаковки/шифрования/запутывания кода в секциях. В статье приводятся ограничения использования этого подхода для больших файлов и при сокрытии фактов запутывания/шифрования.

В [23] с учетом этой идеи предложен подход по локализации оригинальной точки входа исполняемого файла. При этом отмечен достаточно высокий процент ложных срабатываний (28 %).

4) Оценка стойкости паролей и алгоритмов [24, 25]. На основании подсчета энтропии данных, полученных с помощью различных алгоритмов шифрования (алгоритмов генерации пароля), производится их сравнение с точки зрения объема информации, который дает каждый из них. Недостатком подхода для определения криптостойкости алгоритмов шифрования является то, что энтропия не учитывает сложность формирования структуры данных, и если данные зашумлены, то все равно формально они обладают большим количеством информации [24].

Для определения стойкости паролей скорее используется количество информации, а не само значение энтропии.

5) Детектирование аномалий в трафике [26, 27, 28, 29, 30, 31] и классификация трафика [28, 30]. Изменение энтропии полей пакетов в трафике (IP-адреса отправителя/получателя, порты отправителя/получателя) применяются для отслеживания следующих атак:

— DoS/DDoS, что характеризуется увеличением количества входящих пакетов одного типа (TCP SYN, UDP, ICMP и т. д.). Вызывает значительное уменьшение энтропии пакетов на порт и IP-адрес получателя;

— распространение вредоносной программы по сети (сетевой червь), что характеризуется отправкой небольших пакетов большому числу компьютеров. Вызывает значительное уменьшение энтропии порта отправителя и IP-адреса получателя;

— сканирование портов, что характеризуется отправкой от одного компьютера пробных пакетов на широкий диапазон портов другого компьютера. Вызывает значительное уменьшение энтропии порта отправителя и IP-адресов отправителя и получателя;

— большое количество легитимных одновременных обращений к заданному сервису (загрузка файла), что характеризуется увеличением и входящего и исходящего трафика. Вызывает незначительное уменьшение энтропии портов и IP-адресов отправителя и получателя; и др.

Способ является достаточно эффективным для небольшого объема трафика, а именно при применении его для каждого хоста в отдельности. Сложность в применении этого способа может составлять большой объем трафика (трафик большой организации) или адаптивная атака (изменение трафика происходит постепенно), а также большие затраты ресурсов и времени для проведения таких подсчетов для трафика на лету.

Основной недостаток всех предложенных подходов состоит в том, что для их применения необходимо вырабатывать пороговые значения для выполнения условий проверки или составления эталонных образов, что является достаточно уязвимым местом в связи с возможной ограниченностью и неполнотой результатов эксперимента.

Тем не менее значение энтропии зависит от содержимого данных, поэтому может использоваться как дополнительная характеристика при оценке данных, но не основная, так как имеет следующие недостатки [19]:

— неточность, так как не учитывает структуру содержимого файла;

— непостоянство, зависимость от размера файлов;

— низкая эффективность, при подсчете учитывается каждый байт, что весьма ресурсозатратно при реализации.

В связи с вышеизложенным и с усложнением структур данных и распространением шифрования энтропия уже не может использоваться как самостоятельное средство для оценки данных.

СПИСОК ЛИТЕРАТУРЫ:

1. *Shennon C.* A Mathematical Theory of communication // The Bell System Technical Journal. 1948. V. 27. P. 379–423, 623–656.
2. *Falcioni M., Loreto V., Vulpiani A.* Kolmogorov's Legacy about Entropy, Chaos, and Complexity // The Kolmogorov Legacy in Physics. 2003. P. 85–108.
3. *Хэмминг Р. В.* Теория кодирования и теория информации. М.: Радио и связь. 1983. С. 176.
4. *Jozwiak I., Kedziora M., Melinska A.* Theoretical and Practical Aspects of Encrypted Containers Detection — Digital Forensics Approach // Dependable Computer Systems. 2011. V. 97. P. 75–85.
5. *Wu Y., Zhou Y., Saveriades G., [etc.]* Local Shannon entropy measure with statistical tests for image encryption // Information Sciences: an International Journal. 2013. V. 222. P. 323–342.
6. Digital Forensics File Carving Advances [Электронный ресурс]: KoreLogic DFRWS-2006 Project. 2006. URL: http://www.korelogic.com/Resources/Projects/dfrws_challenge_2006/DFRWS_2006_File_Carving_Challenge.pdf (дата обращения: 12.08.2014).
7. *Cardoso A., Crespo R., Kokol P.* Assessing Software Structure by Entropy and Information Density // ACM SIGSOFT Software Engineering Notes. 2004. V. 29. I. 2. P. 2.
8. *Sorokin I.* Comparing files using structural entropy // Journal in Computer Virology. 2011. V. 7. I. 4. P. 259–265.
9. *Weston P., Wolthusen S.* Forensic Entropy Analysis of Microsoft Windows Storage Volume // Information Security for South Africa. 2013. P. 1–7.
10. *Hall G., Davis W.* Sliding Window Measurement for File Type Identification // Technical report, Computer Forensics and Analysis Group, ManTech Security and Mission Assurance. 2006.
11. *Roussev V.* Data fingerprinting with similarity digests // Advances in Digital Forensics VI. 2010. P. 207–226.
12. *Sportiello L., Zanero S.* Context-based file block classification // Advances in Digital Forensics VIII. 2012. P. 67–82.
13. *Sportiello L., Zanero S.* Context-Based File Block Classification // Advances in Digital Forensics VIII. 2012. Part 2. P. 67–82.
14. *Veenman Cor J.* Statistical Disk Cluster Classification for File Carving // IAS'07. Proceedings of the Third International Symposium on Information Assurance and Security. 2007. P. 393–398.
15. *Qiming L.* A Novel Support Vector Machine Approach to High Entropy Data Fragment // Proceedings of South African Information Security Multi-Conference. Digital Forensics & Incident Analysis. 2010. P. 236–247.
16. *Davis T.* Utilizing Entropy to Identify Undetected Malware [Электронный ресурс]: Guidance Software.2009. URL: http://image.lifeservant.com/siteuploadfiles/VSYM/99B5C5E7-8B46-4D14-A53EB8FD1CEEB2BC/43C34073-C29A-8FCE-4-B653DBE35B934F7.pdf (дата обращения 12.08.2014).
17. *Shannon M.* Forensic Relative Strength Scoring: ASCII and Entropy Scoring // International Journal of Digital Evidence. 2004. V. 2. I. 4. P. 1–19.
18. *Jyźwiak I., Kędziora M., Melińska A.* Methods for Detecting and Analyzing Hidden FAT32 Volumes Created with the Use of Cryptographic Tools // Proceedings of the 8th International Conference on Dependability and Complex Systems DepCoS-RELCOMEX. 2013. P. 237–244.
19. *Wu Y., Zhou Y., Saveriades G. [etc.]* Local Shannon entropy measure with statistical tests for image randomness // Information Sciences: an International Journal. 2013. V. 222. P. 323–342.
20. *Salomon D.* Data Privacy and Security. Springer New York, 2003. — 100 p.
21. *Guo F., Ferrie P., Chiueh T.-C.* A study of the packer problem and its solutions // InRAID'08: Proceedings of the 11th international symposium on Recent Advances in Intrusion Detection. 2008. P. 98–115.
22. *Lyda R., Hamrock J.* Using Entropy Analysis to Find Encrypted and Packed Malware // IEEE Security and Privacy. 2007. V. 5. I. 2. P. 40–45.
23. *Jeong G., etc.* Generic Unpacking using Entropy Analysis // Malicious and Unwanted Software (MALWARE). 5th International Conference. 2010. P. 98–105.
24. *Сен Н. Д., Котляров В. П., Григорьев Я. Ю.* Применение оценок на основе энтропии для сравнения криптостойкости алгоритмов шифрования // Современные наукоемкие технологии. 2013. № 2. С. 105–106.
25. NIST SP800-63, Information Security. NIST. 2006. — 65 p.
26. *Androulidakis G., Chatzigiannakis V., Papavassiliou S.* Network Anomaly Detection and Classification via Opportunistic Sampling // IEEE Network: The Magazine of Global Internetworking — Special issue title on recent developments in network intrusion detection. 2009. V. 23. I. 1. P. 6–12.
27. *Lall A., Sekar V., Ogihara M., [etc.]* Data Streaming Algorithms for Estimating Entropy of Network Traffic // SIGMETRICS'06/Performance'06. Proceedings of the Joint International Conference on Measurement and Modeling of Computer Systems. 2006. P. 145–156.
28. *Lakhina A., Crovella M., Diot C.* Mining Anomalies Using Traffic Feature Distributions // SIGCOMM'05. Proceedings of the 2005 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications. 2005. P. 217–228.
29. *Wagner A., Plattner B.* Entropy Based Worm and Anomaly Detection in Fast IP Networks // WETICE'05. Proceedings of the 14th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprise. 2005. P. 171–177.
30. *Xu K., Zhang Z.-L., Bhattacharyya S.* Profiling Internet Backbone Traffic: Behavior Models and Applications // SIGCOMM'05. Proceedings of the 2005 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications. 2005. P. 169–180.
31. *Feinstein L., etc.* Statistical Approaches to DDoS Attack Detection and Response // DARPA Information Survivability Conference and Exposition. 2003. Proceedings. 2003. V. 1. P. 303–314.

REFERENCES:

1. *Shennon C. A* Mathematical Theory of communication / C. Shannon // The Bell System Technical Journal. 1948. V. 27. P. 379–423, 623–656.
2. *Falcioni M.* Kolmogorov's Legacy about Entropy, Chaos, and Complexity / M. Falcioni, V. Loreto, A. Vulpiani // TheKolmogorov Legacy in Physics. 2003. P. 85–108.
3. *Hjemming R. V.* Teorija kodirovanija i teorija informacii. M.: Radio i svjaz'. 1983. P. 78–86.
4. *Jozwiak I.* Theoretical and Practical Aspects of Encrypted Containers Detection – Digital Forensics Approach / I. Jozwiak, M. Kedziora, A. Melinska // Dependable Computer Systems. 2011. P. 75–85.
5. *Wu Y.* Local Shannon entropy measure with statistical tests for image encryption / Y. Wu, Y. Zhou, G. Saveriades [etc.] // Information Sciences: an International Journal. 2013. V. 222. P. 323–342.
6. Digital Forensics File Carving Advances: KoreLogic DFRWS-2006 Project. 2006. URL: http://www.korelogic.com/Resources/Projects/dfrws_challenge_2006/DFRWS_2006_File_Carving_Challenge.pdf.
7. *Cardoso A.* Assessing Software Structure by Entropy and Information Density / A. Cardoso, R. Crespo, P. Kokol // ACM SIGSOFT Software Engineering Notes. 2004. V. 29. I. 2. P. 22.
8. *Sorokin I.* Comparing files using structural entropy / I. Sorokin // Journal in Computer Virology. 2011. V. 7. I. 4. P. 259–265.
9. *Weston P.* Forensic Entropy Analysis of Microsoft Windows Storage Volume / P. Weston, Wolthusen S. // Information Security for South Africa. 2013. P. 1–7.
10. *Hall G.* Sliding Window Measurement for File Type Identification / G. Hall, W. Davis // Proceedings of IEEE Workshop on Information Assurance Workshop. 2006.
11. *Roussev V.* Data fingerprinting with similarity digests / V. Roussev // Advances in Digital Forensics VI. 2010. P. 207–226.
12. *Sportiello L.* Context-based file block classification / L. Sportiello, S. Zanero // Advances in Digital Forensics VIII. 2012. P. 67–82.
13. *Sportiello L.* Context-Based File Block Classification / L. Sportiello, Zanero S. // Advances in Digital Forensics VIII. 2012. Part 2. P. 67–82.
14. *Veenman Cor J.* Statistical Disk Cluster Classification for File Carving / Cor J. Veenman // IAS '07 Proceedings of the Third International Symposium on Information Assurance and Security. 2007. P. 393–398.
15. *Qiming L.* A Novel Support Vector Machine Approach to High Entropy Data Fragment / L. Qiming, etc. // *5th International Annual Workshop on Digital Forensics & Incident Analysis. 2010.* P. 236–247.
16. *Davis T.* Utilizing Entropy to Identify Undetected Malware / T. Davis // Guidance Software. 2009.
17. *Shannon M.* Forensic Relative Strength Scoring: ASCII and Entropy Scoring / M. Shannon // International Journal of Digital Evidence. 2004. V. 2. – I. 4.
18. *Jyźwiak I.* Methods for Detecting and Analyzing Hidden FAT32 Volumes Created with the Use of Cryptographic Tools/ I. Jyźwiak, M. Kędziora, A Melińska // Proceedings of the 8th International Conference on Dependability and Complex Systems DepCoS-RELCOMEX. 2013. P. 237–244.
19. *Wu Y.* Local Shannon entropy measure with statistical tests for image randomness / Y. Wu, Y. Zhou, G. Saveriades [etc.] // Information Sciences: an International Journal. 2013. V. 222. P. 323–342.
20. *Salomon D.* Data Privacy and Security / D. Salomon - Springer New York. 2003. P. 100.
21. *Guo F.* A study of the packer problem and its solutions // F. Guo, P. Ferrie, T.-C. Chiueh // InRAID '08: Proceedings of the 11th international symposium on Recent Advances in Intrusion Detection. 2008. P. 98–115.
22. *Lyda R.* Using Entropy Analysis to Find Encrypted and Packed Malware / R. Lyda, J. Hamrock // IEEE Security and Privacy. 2007. V. 5. – I. 2. P. 40–45.
23. *Jeong G.* Generic Unpacking using Entropy Analysis / G. Jeong, etc. // Malicious and Unwanted Software (MALWARE), 2010 5th International Conference on. 2010. P. 98–105.
24. *Sen N. D.*, PRIMENENIE OCENOK NA OSNOVE JENTROPII DLJA SRAVNENIJA KRIPTOSTOJKOSTI ALGORITMOV SHIFROVANIJA / N. D. Sen, V. P. Kotljarov, JA.JU. Grigor'ev // Sovremennye naukoemkie tehnologii. 2013. № 2. P. 105–106.
25. NIST SP800-63, Information Security. NIST, 2006. – 65 p.
26. *Androulidakis G.* Network Anomaly Detection and Classification via Opportunistic Sampling / G. Androulidakis, V. Chatzigiannakis, S. Papavassiliou // IEEE Network: The Magazine of Global Internetworking - Special issue title on recent developments in network intrusion detection. 2009. V. 23. – I. 1. P. 6–12.
27. *Lall A.* Data Streaming Algorithms for Estimating Entropy of Network Traffic / A. Lall, V. Sekar M. Ogihara, [etc.] // SIGMETRICS '06/Performance '06 Proceedings of the joint international conference on Measurement and modeling of computer systems. P. 145–156.
28. *Lakhina A.* Mining Anomalies Using Traffic Feature Distributions / A. Lakhina, M. Crovella, C. Diot // SIGCOMM '05 Proceedings of the 2005 conference on Applications, technologies, architectures, and protocols for computer communications. 2005. P. 217–228.
29. *Wagner A.* Entropy Based Worm and Anomaly Detection in Fast IP Networks / A. Wagner, B. Plattner // WETICE '05 Proceedings of the 14th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprise. 2005. P. 171–177.
30. *Xu K.* Profiling Internet Backbone Traffic: Behavior Models and Applications / K. Xu, Z.-L. Zhang, S. Bhattacharyya // SIGCOMM '05 Proceedings of the 2005 conference on Applications, technologies, architectures, and protocols for computer communications. 2005. P. 169–180.
31. *Feinstein L.* Statistical Approaches to DDoS Attack Detection and Response / L. Feinstein, etc. // DARPA Information Survivability Conference and Exposition, 2003. Proceedings. 2003. V. 1. P. 303–314.