Lyubov' Cherevan – postgraduate, senior lecturer, Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Saint Petersburg, Russia, 4erevanb@mail.ru

Vyacheslav Tozik – PhD, Associate Professor, Head of Intersectoral Institute for Advanced Training, Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Saint Petersburg, Russia, tozikvt@mail.ru

УДК 004.056.53

ЭКСПЛУАТАЦИОННЫЕ ХАРАКТЕРИСТИКИ РИСКА НАРУШЕНИЙ БЕЗОПАСНОСТИ ИНФОРМАЦИОННОЙ СИСТЕМЫ

К.А. Щеглов[а, б], А.Ю. Щеглов[а, б]

[а] Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, Санкт-Петербург, Россия, info@npp-itb.spb.ru

[б] ЗАО «НПП «Информационные технологии в бизнесе», Санкт-Петербург, Россия, info@npp-itb.spb.ru

Рассмотрены известные подходы к оценке эффективности и к проектированию средств защиты информационных систем, в результате чего выявлен их общий недостаток, заключающийся в том, что данные подходы не позволяют рассматривать угрозы информационной безопасности в качестве эксплуатационной характеристики информационной системы. Это не позволяет построить адекватную модель информационной системы, в том числе защищенной, как системы с отказами и восстановлениями, характеризуемыми возникновением и устранением угроз в процессе эксплуатации системы, не позволяет выявить и соответствующие взаимосвязи, в том числе временные, между угрозами и их совокупностями, эксплуатируемыми атаками, ввести количественные меры актуальности угроз и эффективности атак как эксплуатационных характеристик системы и, как следствие, количественно оценить уровень эксплуатационной безопасности информационной системы и эффективности средств защиты. Предложены основы теории эксплуатационной информационной безопасности, введены основные эксплуатационные параметры и характеристики, предложен подход к оцениванию эксплуатационных характеристик информационной системы, введены эксплуатационные характеристики риска нарушений безопасности информационной системы, в том числе риска потерь.

Ключевые слова: информационная безопасность, защита информации, информационная система, угроза, атака, эксплуатационная безопасность, уровень безопасности, несанкционированный доступ, оценка эффективности, теория массового обслуживания, теория надежности, средство защиты, риск.

OPERATIONAL CHARACTERISTICS OF INFORMATION SYSTEM SECURITY THREATS RISK

K. Shcheglov[c, d], A. Shcheglov[c, d]

[c] Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Saint Petersburg, Russia, info@npp-itb.spb.ru

[d] JSC "Information Technologies in Business", Saint Petersburg, Russia, info@npp-itb.spb.ru

The paper deals with widely used methods for effectiveness evolution and information systems security tools development. Their general disadvantage that consists in not giving the possibility to consider security threats as information system operational characteristic is revealed. Therefore, an adequate information system model can't be created, including secure information system like system with failure and recovery parameters, characterized by threats appearing and elimination during system operation. Also it doesn't make it possible to identify appropriate relationship (including time ones), between threats and their exploited aggregates and doesn't give the possibility to introduce quantitative measures of the threats relevance and attacks effectiveness (like system operational characteristic). Consequently, it doesn't make it possible to estimate the level of information system operational security and security tools effectiveness. The principles of operational security theory are suggested, base operational parameters and characteristics are introduced, a method for operational characteristics of information system estimation is proposed, operational characteristics of risks for information system security threats and loss risks are stated.

Keywords: information security, information system, threat, attack, operational security, security level, unauthorized access, effectiveness evaluation, queueing theory, reliability theory, security tools, risk.

Введение

Под математической теорией риска формально понимается совокупность моделей и методов теории вероятностей, применяемых к анализу случайных величин и их распределений [1]. Это обусловлено тем, что величина возможного ущерба в стохастической ситуации до осуществления этой ситуации неизвестна и потому случайна. Риском же называют совокупность значения возможного ущерба в некоторой стохастической ситуации и его вероятности [1, 2].

Применению теории риска в области информационной безопасности, в частности, для оценки эффективности средств защиты информации и при проектировании защищенной информационной системы (далее ИС), ввиду актуальности данной задачи, в настоящее время посвящено достаточно много работ,

например [3–8]. Общий недостаток известных подходов состоит в том, что они не позволяют рассматривать угрозы информационной безопасности в качестве эксплуатационной характеристики ИС. Это не позволяет построить адекватную модель ИС, в том числе защищенной, как системы с отказами и восстановлениями, характеризуемыми возникновением и устранением угроз в процессе эксплуатации системы, не позволяет выявить и соответствующие взаимосвязи, в первую очередь, временные, между угрозами и их совокупностями, эксплуатируемыми атаками, ввести количественные меры актуальности угроз и эффективности атак как эксплуатационных характеристик системы и, как следствие, адекватно количественно оценить эксплуатационные характеристики риска нарушений безопасности ИС.

Эксплуатационные характеристики риска

Прежде всего, в рамках основ излагаемой теории эксплуатационной информационной безопасности введем основные понятия, определим ключевые эксплуатационные параметры и характеристики ИС.

Под *эксплуатационной информационной безопасностью* понимается свойство ИС функционировать в безопасном состоянии – без обнаруженных в ней либо внесенных в нее в процессе эксплуатации и не устраненных «каналов» несанкционированного доступа к информации.

Под *«каналом» несанкционированного доступа к информации* понимается появившийся в ИС по каким-либо причинам способ осуществления злоумышленником несанкционированного доступа с какой-либо целью к обрабатываемой в системе информации во время эксплуатации ИС.

Под *нарушением эксплуатационной информационной безопасности* понимается обнаружение в ИС либо внесение в нее в процессе эксплуатации «каналов» несанкционированного доступа к информации.

Под *угрозой безопасности ИС* понимается нарушение ее эксплуатационной информационной безопасности в результате обнаружения в ИС либо внесения в нее в процессе эксплуатации «каналов» несанкционированного доступа к информации.

Замечание. Нарушение эксплуатационной информационной безопасности (угроза безопасности ИС) – это не есть осуществление несанкционированного доступа к информации, это возникновение условий для успешной атаки на ИС с использованием существующего (выявленного) в системе неустраненного «канала» («каналов») несанкционированного доступа к информации.

Введем основные эксплуатационные параметры и характеристики ИС, которые будут далее нами использованы для определения характеристики риска как эксплуатационной характеристики ИС, после чего определим эксплуатационные характеристики риска, в том числе риска потерь.

Под *интенсивностью нарушения эксплуатационной информационной безопасности (возникновения угроз)* λ будем понимать количество обнаружений в ИС либо внесений в нее в процессе эксплуатации «каналов» несанкционированного доступа к информации в единицу времени.

Под *интенсивностью восстановления эксплуатационной информационной безопасности (нейтрализации угроз)* μ будем понимать количество устранений выявленных в ИС либо внесенных в нее «каналов» несанкционированного доступа к информации в единицу времени.

Под *уровнем эксплуатационной информационной безопасности ИС* будем понимать количественную оценку готовности ИС к безопасной эксплуатации – значение вероятности P_0 того, что информационная система готова к безопасной эксплуатации.

Данный подход к оцениванию уровня эксплуатационной безопасности ИС, $P_0 = f(\lambda, \mu)$, рассмотренный нами впервые в [9], позволяет применить математический аппарат теории массового обслуживания для получения требуемой количественной оценки [9, 10]. При этом важным является тот факт, что при расчете характеристики эксплуатационной информационной безопасности используются не некие гипотетические, заданные экспертным путем, значения параметров ИС, а реальные значения параметров, полученные на основании анализа соответствующей статистики угроз. Это позволяет говорить об адекватности получаемых подобным образом моделей безопасности информационных систем.

Под *атакой на информационную систему* понимается воздействие злоумышленником на ИС посредством использования (эксплуатации) обнаруженных в ИС либо внесенных в нее в процессе эксплуатации «каналов» несанкционированного доступа к информации (угроз безопасности ИС).

В общем случае атака многоступенчата и предполагает последовательное использование атакой некой совокупности угроз, присутствующих в системе. Формальное описание атаки (атаки, реализуемой с определенной целью) предлагается представлять взвешенным ориентированным графом атаки на ИС.

Под *графом атаки на информационную систему* понимается взвешенный ориентированный граф (или орграф), вершинами которого являются угрозы, характеризуемые («взвешенные») вероятностью наличия обнаруженных или внесенных в систему и не устраненных «каналов» несанкционированного доступа к информации, создающих угрозу, которые должны эксплуатироваться (использоваться) злоумышленником при осуществлении атаки; дуги графа определяют последовательность использования злоумышленником угроз (направленного перехода между вершинами) при осуществлении атаки.

Пример графа атаки на ИС приведен на рис. 1. На рис. 1 P_{0r}, где $r = 1,\ldots,R$, обозначает вероятность отсутствия в системе r-й угрозы (информационная система готова к безопасной эксплуатации в отноше-

нии *r*-й угрозы), одной из *R* угроз, последовательно эксплуатируемых (используемых) злоумышленником при осуществлении описываемой графом атаки. Эту характеристику, как отмечали ранее, мы можем рассчитать с использованием аппарата теории массового обслуживания.

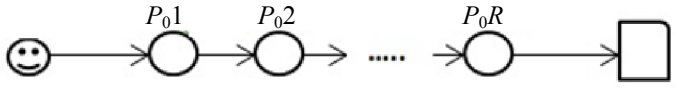

Рис. 1. Пример графа атаки на ИС

С учетом сказанного может быть сделан крайне важный вывод: каждая последующая угроза, используемая при осуществлении атаки, может рассматриваться, в терминологии теории надежности [11], в качестве параллельно включенного в систему резерва (для осуществления атаки необходимо использовать всю совокупность угроз, приведенных на графе, см. рис. 1). Как следствие, вероятность того, что защищенная информационная система готова к безопасной эксплуатации в отношении атаки P_{0a}, последовательно использующей *R* угроз, *r* = 1,…,*R*, определяется следующим образом:

$$P_{0a} = 1 - \prod_{r=1}^{R} (1 - P_{0r}).$$

Аналогичным образом (параллельно включенный резерв) может рассматриваться и средство защиты, направленное на нивелирование угрозы, используемой атакой, которое может быть охарактеризовано аналогичными параметрами и характеристиками.

Под *готовностью защищенной информационной системы к безопасной эксплуатации в отношении атаки* понимается вероятность того, что защищенная ИС готова к безопасной эксплуатации – в ней отсутствуют обнаруженные либо внесенные в нее и не устраненные «каналы», как собственно создающие угрозы, используемые атакой, так и в средстве защиты информации, призванном нивелировать угрозу (угрозы), эксплуатируемую атакой, позволяющие осуществить несанкционированный доступ к информации.

Соответственно под *уровнем эксплуатационной информационной безопасности защищенной информационной системы в отношении атаки* понимается количественная оценка готовности защищенной ИС к безопасной эксплуатации – значение вероятности того, что защищенная в отношении атаки информационная система готова к безопасной эксплуатации.

Все сказанное можно отнести и к ИС в целом. Атаки в этом случае могут рассматриваться, в терминологии теории надежности [11], в качестве последовательно включенного в систему резерва (для осуществления атаки на систему в целом злоумышленнику достаточно реализовать одну из атак). Как следствие, вероятность того, что информационная система готова к безопасной эксплуатации в целом, $P_{0\text{исц}}$, при возможной совокупности из *D* атак, *d* = 1,…,*D*, каждая из которых соответственно характеризуется P_{0ad}, определяется следующим образом:

$$P_{0\text{исц}} = \prod_{d=1}^{D} P_{0ad}.$$

С учетом того, что каждая угроза, используемая при осуществлении атаки, может рассматриваться в терминологии теории надежности в качестве параллельно включенного в систему резерва, а каждая из совокупности возможных атак на ИС может рассматриваться в качестве последовательно включенного в систему резерва, в общем случае, применительно к ИС в целом, можем говорить о последовательно-параллельном резервировании. Таким образом, мы можем количественно оценить готовность ИС к безопасной эксплуатации (вероятность того, что информационная система готова к безопасной эксплуатации) как в отношении отдельной атаки, так и в целом.

Под *средним временем наработки на нарушение эксплуатационной информационной безопасности информационной системы* (среднее время наработки на отказ в теории надежности) понимается средний интервал времени, в течение которого ИС готова к безопасной эксплуатации (между моментами нарушения эксплуатационной информационной безопасности ИС). Данная характеристика определяется следующим образом:

$$T_{\text{о}} = \frac{P_0}{\mu\,(1 - P_0)}.$$

Это – параметры и характеристики непосредственно ИС, определяющие риск осуществления успешной атаки на ИС.

Под *риском нарушения безопасности информационной системы* будем понимать вероятность наступления условия нарушения эксплуатационной информационной безопасности ИС, позволяющего злоумышленнику осуществить успешную атаку на ИС с целью получения несанкционированного доступа к обрабатываемой в ней информации.

.

Риск нарушения безопасности ИС определяется значением вероятности P_0 (соответственно применительно к атаке P_{0a}, к ИС в целом $P_{0исц}$). Это эксплуатационные характеристики непосредственно ИС, определяющие риск осуществления успешной атаки на ИС.

Рассмотрим еще один важный эксплуатационный параметр.

Под *интенсивностью осуществления злоумышленником атак на защищенную информационную систему* будем понимать количество совершаемых им атак на ИС в единицу времени λ_a. Это уже эксплуатационный параметр нарушителя, характеризующий его заинтересованность в получении несанкционированного доступа к информации, обрабатываемой в системе (это характеристика ценности обрабатываемой информации для нарушителя).

Отметим, что значение параметра λ_a должно задаваться при проектировании системы защиты ИС с целью формирования требований к уровню эксплуатационной информационной безопасности проектируемой защищенной ИС (проектируемой системы защиты ИС). Требуемый для этого параметр λ_a может быть определен либо на основании опыта эксплуатации подобных (обрабатывающих аналогичную информацию) защищенных информационных систем, либо экспертным путем. Важным является уточнение данного параметра в процессе всего времени эксплуатации ИС с применением средств аудита – протоколирования событий, в данном случае – осуществления атак (и их отражения) на защищенную ИС, что, в свою очередь, определяет соответствующие требования к средствам аудита событий, реализуемым в ИС. Определение значения параметра λ_a как эксплуатационной характеристики (в процессе эксплуатации ИС) позволяет оценить корректность сделанных исходных предположений о ценности для злоумышленников обрабатываемой в ИС информации, а также учесть изменение данного параметра в процессе эксплуатации системы, в результате чего, при необходимости, доработать исходные требования к уровню эксплуатационной информационной безопасности ИС.

С точки зрения эксплуатационной информационной безопасности успешная атака будет реализована злоумышленником в том случае, если она будет проведена в тот интервал времени, когда информационная система не готова к безопасной эксплуатации. С учетом сказанного определим важнейшую характеристику эксплуатационной информационной безопасности – вероятность реализации успешной атаки на ИС.

Под *вероятностью реализации успешной атаки на информационную систему* будем понимать вероятность того, что атака будет осуществлена злоумышленником в промежуток времени, когда ИС не готова к безопасной эксплуатации.

Вероятность реализации успешной атаки на защищенную ИС P_a, определяемая вероятностью того, что атака (по крайней мере, одна из атак) придется на интервал времени $1/\mu$, может быть получена из закона Пуассона [10]:

$$P_a = 1 - e^{-\lambda_a/\mu}.$$

С учетом сказанного определим эксплуатационную характеристику риска.

Под *риском осуществления успешной атаки на информационную систему* будем понимать вероятность реализации успешной атаки на ИС с целью получения несанкционированного доступа к обрабатываемой в ней информации.

Риск осуществления успешной атаки на ИС определяется значением вероятности P_a.

Важным моментом является то, что мы рассматриваем восстанавливаемую систему, для которой можно вести понятие периода эксплуатации T, задаваемого следующим образом:

$$T = T_o + 1/\mu.$$

В этой формуле T_o – среднее время наработки на нарушение эксплуатационной информационной безопасности ИС, по истечении которого в течение времени $1/\mu$ безопасное состояние ИС восстанавливается – информационная система в течение этого промежутка времени уязвима для атак. Таким образом, если вероятность успешной атаки на ИС, P_a, определяется вероятностью того, что атака (по крайней мере, одна из атак) придется на интервал времени $1/\mu$, то с вероятностью, равной $1 - P_a$, в данном интервале времени успешная атака на ИС осуществлена не будет. Следующая подобная возможность злоумышленнику «будет предоставлена» в следующем периоде эксплуатации системы T – через промежуток времени Tо. При этом значение характеристики P_a определяется эксплуатационными параметрами ИС и нарушителя, как следствие, P_a одинаковы для различных периодов эксплуатации системы T_i, i=1,…,I.

Распределение характеристики P_a во времени (качественная оценка), определяемом периодами эксплуатации системы T_i, i=1,…,I, представлено на рис. 2.

Таким образом, если вероятность успешной атаки на ИС в периоде эксплуатации T_i в течение интервала времени $1/\mu$ на этом периоде, см. рис. 2 (в различных периодах это события независимые), равна P_a, то вероятность успешной атаки на ИС в i-ом периоде эксплуатации T_i, i = 1,…,I, равна

$$P_{ai} = \begin{cases} 0, \text{ если } i = 1; \\ P_a(1-P_a)^{i-2}, \text{ если } i \geq 2. \end{cases}$$

Замечание. Для удобства формального описания за первый период эксплуатации в этой формуле и далее (что необходимо учитывать при использовании приведенных ниже расчетных формул) принимаем T_{i-1}=Tо за последующие периоды эксплуатации:

$T_{i\geq 2} = T_{\text{о}} + 1/\mu$.

Данная эксплуатационная характеристика определяет крайне важное эксплуатационное свойство ИС, которое должно быть нами далее учтено. Поскольку информационная система – это система восстанавливаемая, т.е. восстанавливаются свойства ее информационной безопасности в процессе эксплуатации (с интенсивностью μ), несмотря на значение характеристики P_0, информационная система может с соответствующими вероятностями P_{ai}, i = 1,…,I, безопасно эксплуатироваться в различных периодах (на различных временных интервалах T_i). С учетом этого может быть введена следующая эксплуатационная характеристика ИС – вероятность того, что на ИС будет осуществлена успешная атака за время ее эксплуатации, характеризуемое I периодами эксплуатации системы T_i с номерами i = 1,…,I, P_{aI}. Эта характеристика рассчитывается по следующей формуле:

$$P_{aI} = \sum_{i=1}^{I} P_{ai} .$$

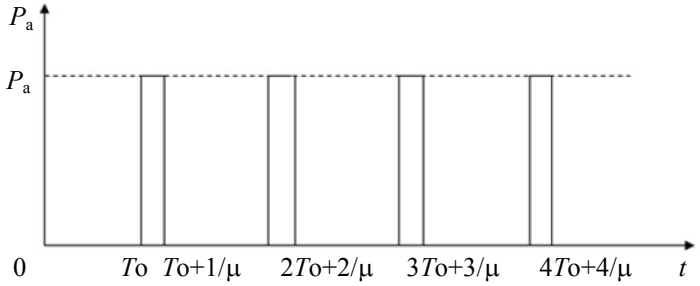

Рис. 2. Распределение характеристики P_a во времени

Для любого произвольного времени эксплуатации ИС t характеристика P_{aI} (в данном случае $P_a(t)$) может быть определена следующим образом:

$$P_a(t) = \begin{cases} 0, \text{ если } 0 < t \leq T_{\text{о}}; \\ \sum_{i=2}^{I} P_{ai}, \text{ если } (I-1)T_{\text{о}} + \dfrac{I-2}{\mu} < t \leq IT_{\text{о}} + \dfrac{I-1}{\mu}. \end{cases}$$

С учетом изменения во времени характеристики P_a и соответственно P_{ai}, (рис. 2), качественная зависимость изменения характеристики $P_a(t)$ во времени представлена на рис. 3. Таким образом, зная эксплуатационные параметры ИС $T_{\text{о}}$ и μ, для любого произвольного времени эксплуатации ИС t можно определить соответствующее ему число периодов эксплуатации системы I продолжительностью T_i, для которого уже определить искомое значение $P_a(t)$.

Под *риском осуществления успешной атаки на информационную систему в течение времени ее эксплуатации t* будем понимать вероятность реализации успешной атаки на ИС с целью получения несанкционированного доступа к обрабатываемой в ней информации за время ее эксплуатации t.

Риск осуществления успешной атаки на ИС в течение времени ее эксплуатации t определяется значением вероятности $P_a(t)$ на соответствующем времени t числе периодов эксплуатации системы I продолжительностью T_i.

Акцентируем внимание читателя на особенностях введенной эксплуатационной характеристики риска. Как правило, на практике риск оценивается вероятностью взлома (либо вероятностью успешной атаки, которая, кстати, также не рассматривается в качестве эксплуатационной характеристики) ИС. При этом не рассматривается распределение характеристики риска во времени эксплуатации системы. Естественно, возникает вопрос, насколько адекватна подобная модель риска и что на самом деле ею описывается – для какого момента времени эксплуатации ИС определяется риск, с учетом того, что его значение (значение соответствующей вероятности) изменяется во времени? Насколько и как применима подобная характеристика, если непонятно, что ею определяется?

Если, по аналогии с теорией надежности [11], ранее нами рассматривалась характеристика $T_{\text{о}}$ – среднее время наработки на нарушение эксплуатационной информационной безопасности ИС, то с ис-

пользованием параметра λ_a можем ввести важнейшую временную эксплуатационную характеристику ИС – среднюю продолжительность безопасной эксплуатации ИС или среднее время наработки ИС до успешной атаки.

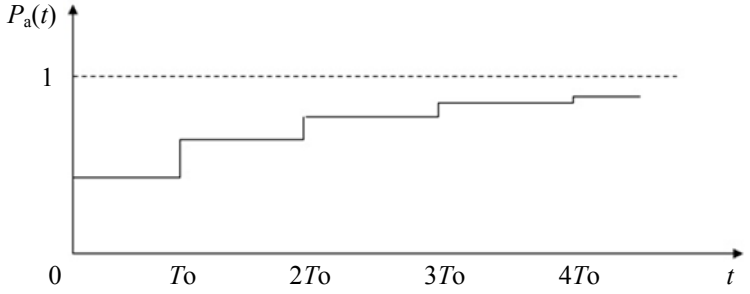

Рис. 3. Качественная зависимость изменения характеристики $P_a(t)$ во времени

Под *средней продолжительностью безопасной эксплуатации информационной системы* (или под *средним временем наработки информационной системы до успешной атаки*) будем понимать среднее время, в течение которого ИС будет находиться в безопасном состоянии с начала ее эксплуатации.

Данная характеристика $T_{\text{бэ}}$ может быть рассчитана по следующей формуле:

$$T_{\text{бэ}} = T_{\text{о}} + \sum_{i=2}^{\infty} (i-1) T P_{ai} .$$

Замечание. $T_i = T$, $i \geq 2$.

Итак, нами были введены важнейшие вероятностные и временные характеристики риска, позволяющие оценить безопасность эксплуатации ИС с учетом интенсивности осуществления злоумышленником атак на ИС и уровня эксплуатационной информационной безопасности ИС. В ряде случае подобных оценок риска будет достаточно для ответа на вопрос, следует ли реализовывать защиту ИС с целью повышения уровня ее эксплуатационной информационной безопасности.

Однако при высокой заинтересованности злоумышленника в получении несанкционированного доступа к обрабатываемой в ИС информации (характеризуемой большим значением параметра λ_a) и, как следствие, при принятии решения о необходимости реализации защиты ИС уже целесообразно рассматривать такие характеристики ИС (в том числе, защищенной ИС), как потери (экономическую характеристику эксплуатационной информационной безопасности). Ведь потери в защищенной ИС связаны как с потенциальной возможностью несанкционированного доступа злоумышленника к обрабатываемой информации (потенциальные потери, которые будут снижаться при реализации защиты), так и с реализацией и эксплуатацией системы защиты в ИС.

Прежде чем говорить о потерях, введем еще одну важнейшую эксплуатационную характеристику ИС – характеристику сложности реализации атаки на ИС для злоумышленника.

Как уже отмечалось, в рамках рассматриваемой теории эксплуатационной информационной безопасности каждая последующая угроза, используемая при осуществлении атаки, может рассматриваться в терминологии теории надежности в качестве параллельно включенного в систему резерва. При этом успешная реализация атаки возможна лишь при условии актуальности (выявленных и неустраненных «каналов» несанкционированного доступа к информации применительно к каждой угрозе) всех угроз, используемых атакой. Таким образом, для осуществления успешной атаки злоумышленник должен обладать определенным количеством информации – информацией о выявленных и неустраненных (существующих) «каналов» несанкционированного доступа к информации применительно к каждой угрозе, используемой атакой. Очевидно, что именно количество информации, которым должен обладать злоумышленник в отношении актуальности угроз, используемых атакой, и определяет сложность реализации успешной атаки на ИС.

Под *характеристикой сложности реализации атаки на информационную систему для злоумышленника* будем понимать количество информации, которым должен обладать злоумышленник в отношении актуальности угроз, используемых атакой.

Обратимся к основам теории информации. Следуя К. Шеннону [12], вероятностная мера количества информации определяется следующим образом. Пусть можно получить n сообщений по результатам некоторого опыта (т.е. у опыта есть n исходов), причем известны вероятности получения каждого сообщения (исхода) – p_i . Тогда в соответствии с идеей Шеннона, количество информации I в сообщении i определяется по формуле $I = -\log_2 p_i$, где p_i – вероятность i -го сообщения (исхода).

В нашем случае в отношении атаки возможны два исхода, определяющие то количество информации, которым должен обладать злоумышленник для реализации успешной атаки – готова или нет информационная система к безопасной эксплуатации.

С учетом же того, что вероятность готовности защищенной информационной системы к безопасной эксплуатации в отношении атаки, последовательно использующей R угроз, $r = 1,\ldots,R$, определяется следующим образом:

$$P_{0a} = 1 - \prod_{r=1}^{R}(1 - P_{0r}),$$

количество информации (как вероятностная мера этого исхода), необходимое злоумышленнику для осуществления успешной атаки – характеристика сложности реализации атаки на ИС для злоумышленника, обозначим ее через Sa – может быть определена по следующей формуле:

$$S\text{a} = -\log_2 \prod_{r=1}^{R}(1 - P_{0r}).$$

Это важнейшая эксплуатационная характеристика ИС, позволяющая количественно оценить и достаточно информативно представить сложность реализации злоумышленником успешной атаки. Например, рассмотрим, как изменится сложность реализации атаки, в предположении, что исходное для ИС значение P_{0a} составляло 0,7, а в результате использования средства защиты стало равно 0,99. Видим, что в первом случае Sa = 1,74, во втором случае Sa = 6,64, т.е. использование данного средства защиты в ИС усложнило злоумышленнику реализацию успешной атаки на ИС в 3,82 раза.

Замечание. За единицу сложности реализации атаки Sa = 1 при этом принимается условие P_{0a}=0,5 (система с равной вероятностью защищена или нет).

Введенную характеристику сложности реализации атаки на ИС для злоумышленника можно использовать и в процессе эксплуатации ИС для количественной оценки того, насколько упростится атака на ИС в случае актуализации (обнаружения «канала» несанкционированного доступа) одной или нескольких угроз (для актуальной угрозы r справедливо соотношение P_{0r}=0.

Данная важнейшая характеристика применима не только в отношении конкретной атаки, но и в отношении ИС в целом, обозначим ее Sис. С учетом вероятности того, что информационная система готова к безопасной эксплуатации в целом при возможной совокупности из D атак, $d = 1,\ldots,D$, каждая из которых соответственно характеризуется $P_{0\text{a}d}$, определяемой как

$$P_{0\text{исц}} = \prod_{d=1}^{D} P_{0\text{a}d},$$

характеристика Sис может быть рассчитана по следующей формуле:

$$S\text{ис} = -\log_2\left(1 - \prod_{d=1}^{D} P_{0\text{a}d}\right).$$

Теперь о потерях. Потери для незащищенной ИС связаны исключительно с успешной атакой на ИС, в результате которой злоумышленник получает несанкционированный доступ к обрабатываемой в ней информации (в данном случае речь идет о потенциальных потерях).

Пусть потери от несанкционированного доступа к информации (в результате ее хищения, удаления или модификации) составляют $C_{\text{инф}}$.

Под *риском потерь от осуществления успешной атаки на информационную систему в течение времени ее эксплуатации t* будем понимать экономическую эксплуатационную характеристику ИС, определяемую как потенциальные потери от успешной атаки на ИС в результате реализации злоумышленником несанкционированного доступа к обрабатываемой в ней информации за время ее эксплуатации t.

С учетом сказанного ранее (с учетом полученной расчетной формулы для $P_a(t)$), для любого произвольного времени эксплуатации ИС t риск потерь от осуществления успешной атаки на ИС, $R_{C_{\text{инф}}}(t)$, может быть рассчитан следующим образом:

$$R_{C_{\text{инф}}}(t) = \begin{cases} 0, \text{если } 0 < t \le T_0; \\ C_{\text{инф}} \sum_{i=2}^{I} P_{\text{a}i}, \text{если } (I-1)T_0 + \dfrac{(I-2)}{\mu} < t \le IT_0 + \dfrac{(I-1)}{\mu}. \end{cases}$$

Очевидно, что функция распределения риска потерь от осуществления успешной атаки на ИС во времени имеет тот же вид, что и соответствующая функция распределения вероятности успешной атаки на ИС (см. рис. 2), причем в пределе риск потерь стремится к значению $C_{\text{инф}}$.

Рассмотрим возможные способы учета ценности информации. Для этого нами были введены два параметра – λ_a и $C_{\text{инф}}$. Отметим, что данные параметры характеризуют альтернативные подходы к определению ценности обрабатываемой информации: λ_a – с точки зрения злоумышленника, $C_{\text{инф}}$ – с точки зрения владельца информации (или оператора, обрабатывающего информацию). В общем случае – это альтернативные подходы к определению ценности информации, которые должны использоваться совместно, поскольку, как бы ни была ценна обрабатываемая информация для ее владельца, на деле она может не представлять никакой ценности для потенциальных злоумышленников, и наоборот.

Эксплуатационные характеристики риска для защищенной ИС. Экономическая оценка эффективности системы защиты

Как говорилось ранее, реализация в ИС системы защиты может, в терминологии теории надежности, рассматриваться в качестве параллельно включенного в систему резерва. Это обусловливается тем, что для осуществления успешной атаки на ИС злоумышленнику потребуется использовать как возникшие и неустраненные угрозы, эксплуатируемые атакой, собственно в ИС (например, выявленные уязвимости операционной системы и приложений), так и соответствующие угрозы, выявляемые непосредственно в системе защиты.

Несложно проиллюстрировать и то, что высокий уровень эксплуатационной информационной безопасности собственно системы защиты (значение вероятности готовности системы защиты к безопасной эксплуатации 0,95 и выше) вполне достижим на практике. Для этого оценим потенциальные возможности средств защиты в части повышения уровня эксплуатационной информационной безопасности информационных систем. С этой целью определим граничную оценку (при худших условиях уровень эксплуатационной информационной безопасности ИС полностью определяется уровнем эксплуатационной безопасности средства защиты). Расчетные значения приведены в таблице.

Интенсивность устранения «каналов» несанкционированного доступа к информации в средстве защиты	Готовность к безопасной эксплуатации при различных интенсивностях обнаружения «каналов» несанкционированного доступа к информации в средстве защиты			
	1/ 3 месяца	1/ 6 месяцев	1/ 12 месяцев	1/ 18 месяцев
1/3 дня	0,97	0,98	0,99	0,99
1/7 дней	0,93	0,96	0,98	0,99
1/14 дней	0,87	0,93	0,96	0,98

Таблица. Оценка уровня эксплуатационной информационной безопасности защищенной ИС

Поскольку приведенные в таблице эксплуатационные характеристики средств защиты достижимы на практике, за счет их использования в качестве параллельно включенного резерва могут быть существенно улучшены эксплуатационные характеристики ИС – в данном случае уже защищенной.

Некоторые технические решения, иллюстрирующие возможности средств защиты, применение которых направлено на повышение уровня эксплуатационной информационной безопасности информационных систем, описаны, например, в [13–15] (заметим, что данные технические решения на сегодняшний день авторами патентуются).

Для обозначения соответствующих эксплуатационных параметров и характеристик защищенной ИС далее будем использовать индекс «з». При этом исходные значения эксплуатационных параметров непосредственно ИС λ и μ не изменятся, но добавятся соответствующие эксплуатационные параметры системы защиты информации, $\lambda_{\text{сзи}}$, $\mu_{\text{сзи}}$, соответственно имеем $P_{0\text{з}}=f(\lambda, \mu, \lambda_{\text{сзи}}, \mu_{\text{сзи}})$. Отметим, что при расчете характеристик защищенной ИС, $T_{0\text{з}}$, $T_{\text{з}}$, $P_{a\text{з}}$ и соответственно $P_{ai\text{з}}$, в качестве параметра μ при этом уже используется $\mu_{\text{сзи}}$ (при этом будем определять граничные, худшие значения соответствующих характеристик).

Не изменится и исходное (при реализации защиты) значение параметра λ_a, поскольку им характеризуется ценность для злоумышленника обрабатываемой в ИС информации, вне зависимости от того, защищена информационная система или нет.

С учетом сказанного приведенные ранее формулы для расчета основных введенных характеристик эксплуатационной информационной безопасности, применительно к защищенной ИС, примут следующий вид:

$$P_{ai\text{з}} = \begin{cases} 0, \text{если } i = 1; \\ P_{a\text{з}}(1 - P_{a\text{з}})^{i-2}, \text{если } i \geq 2; \end{cases}$$

$$P_{a\text{з}}(t) = \begin{cases} 0, \text{если } 0 < t \leq T_{0\text{з}}; \\ \sum_{i=2}^{I} P_{ai\text{з}}, \text{если } (I-1)T_{0\text{з}} + \dfrac{(I-2)}{\mu_{\text{сзи}}} < t \leq IT_{0\text{з}} + \dfrac{(I-1)}{\mu_{\text{сзи}}}; \end{cases}$$

$$T_{\text{бэз}} = T_{0\text{з}} + \sum_{i=2}^{\infty} (i-1)T_{\text{з}} P_{ai\text{з}}.$$

Распределения характеристик $P_{\text{а}}$ и $P_{аз}$ во времени (качественная оценка), определяемом периодами эксплуатации системы T_i, $i = 1,\ldots,I$, иллюстрирующие причины положительного эффекта от внедрения системы защиты, представлены на рис. 4.

Обсудим теперь дополнительные потери. Риск потерь для незащищенной ИС связан исключительно с успешной атакой на ИС, в результате которой злоумышленник получает несанкционированный доступ к обрабатываемой в ней информации. Внедрение в ИС системы защиты связано с дополнительными затратами (потерями), определяемыми собственно стоимостью внедряемой системы защиты $C_{\text{сзи}}$ и удельной стоимостью (стоимостью в единицу времени) ее эксплуатации $C_{\text{у эсзи}}(t)$ (техническая поддержка от разработчика, увеличение штата сотрудников, эксплуатирующих средство защиты и т.п. (предполагаем, что она характеризуется линейной зависимостью), также являющейся эксплуатационной характеристикой уже средства защиты.

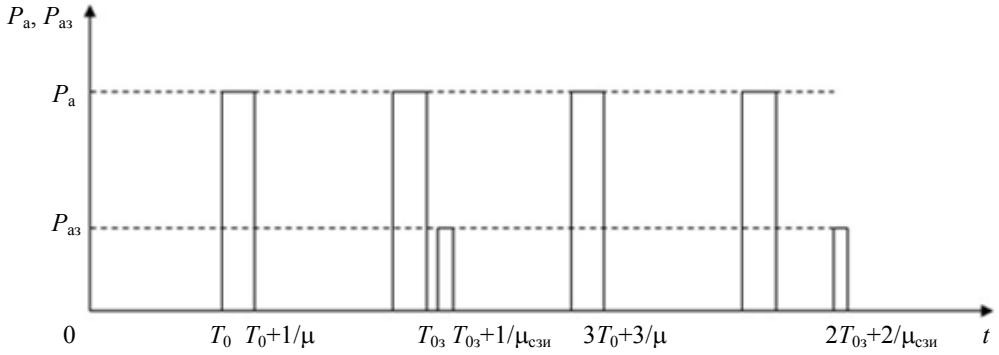

Рис. 4. Распределения характеристики P_a и $P_{аз}$ во времени

Под риском потерь от осуществления успешной атаки на защищенную информационную систему будем понимать экономическую эксплуатационную характеристику защищенной ИС, определяемую совокупностью потерь – затрат, определяемых как стоимостью системы защиты и ее эксплуатации в системе, так и потенциальными потерями от успешной атаки на защищенную ИС.

С учетом сказанного ранее (с учетом полученной расчетной формулы для $P_{аз}(t)$), для любого произвольного времени эксплуатации ИС *t* риск потерь от осуществления успешной атаки на защищенную ИС, $R_{C_{\text{инфз}}}(t)$, может быть рассчитан следующим образом:

$$R_{C_{\text{инфз}}}(t) = \begin{cases} C_{\text{сзи}} + tC_{\text{у эсзи}}(t), \\ \text{если } 0 < t \le T_{0\text{з}}. \\ C_{\text{сзи}} + tC_{\text{у эсзи}}(t) + C_{\text{инф}} \sum_{i=2}^{I} P_{ai\text{з}}, \\ \text{если } (I-1)T_{0\text{з}} + \dfrac{(I-2)}{\mu_{\text{сзи}}} < t \le IT_{0\text{з}} + \dfrac{(I-1)}{\mu_{\text{сзи}}}. \end{cases}$$

Используя полученные результаты, соответствующим образом определим оценки эффективности системы защиты как количественные меры снижения риска осуществления успешной атаки на ИС в течение времени ее эксплуатации *t*, $\Delta P_{аз}(t)$, и снижения риска потерь от осуществления успешной атаки на ИС в течение времени ее эксплуатации *t*, $\Delta R_{C_{\text{инфз}}}(t)$, которые могут быть рассчитаны по следующим формулам:

$$\Delta P_{аз}(t) = P_a(t) - P_{аз}(t),$$
$$\Delta R_{C_{\text{инфз}}}(t) = R_{C_{\text{инф}}}(t) - R_{C_{\text{инфз}}}(t).$$

Отметим, что, если значение $R_{C_{\text{инф}}}(t)$ в пределе стремится к $C_{\text{инф}}$, то значение $R_{C_{\text{инфз}}}(t)$ в пределе стремится к ∞. Это позволяет утверждать, что величина выигрыша в потерях от реализации в ИС защиты информации будет сильно изменяться в процессе эксплуатации системы, причем по мере увеличения продолжительности эксплуатации системы в общем случае будет иметь две фазы распределения – фазу увеличения и фазу уменьшения (начальную фазу эксплуатации, которая, естественно, проигрышна, так как связана с дополнительными затратами на реализацию в ИС системы защиты информации, здесь не рассматриваем).

Качественная оценка изменения (распределения) характеристики $\Delta R_{C_{\text{инфз}}}(t)$ во времени проиллюстрирована на рис. 5.

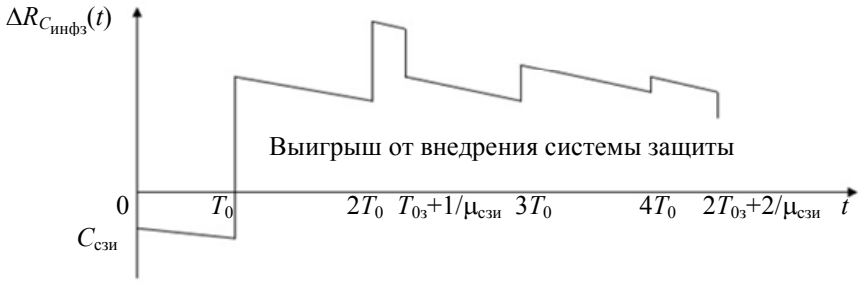

Рис. 5. Качественная оценка изменения характеристики $\Delta R_{C_{\text{инфз}}}(t)$ во времени

Как видим из рис. 5, характеристика $\Delta R_{C_{\text{инфз}}}(t)$ имеет пиковые значения в процессе эксплуатации системы. Это обусловливается различием в общем случае характеристик T и $T_з$ (соответственно для незащищенной и защищенной информационных систем).

Таким образом, используя введенные эксплуатационные параметры и характеристики информационной (в том числе защищенной) системы, эксплуатационные характеристики риска, применяя представленные расчетные формулы, можно количественно оценить выигрыш (проигрыш) от внедрения системы защиты для различных временных интервалов эксплуатации защищенной ИС, что позволяет получить экономическое обоснование целесообразности и эффективности реализации в ИС системы защиты и соответствующим образом, исходя из подобной оценки, сформировать требования к эксплуатационным параметра и характеристикам системы защиты для конкретной ИС.

Заключение

Предложенный в работе подход к оцениванию риска, в том числе риска потерь, реализуемый в рамках излагаемой авторами теории эксплуатационной информационной безопасности, предоставляет возможность оценить распределение потерь в процессе эксплуатации информационной (в том числе защищенной) системы. Использование введенных эксплуатационных характеристик риска позволяет построить адекватную модель ИС с учетом нарушений и восстанавливаемости свойств ее информационной безопасности в процессе эксплуатации и соответственно получить адекватные требования к параметрам и характеристикам (эксплуатационным) системы защиты, предполагаемой к использованию в конкретной ИС.

References

1. Korolev V.Yu., Bening V.E., Shorgin S.Ya. *Matematicheskie osnovy teorii riska* [Mathematical foundations of the theory of risk]. Moscow, FIZMATLIT Publ., 2011, 591 p.
2. Shapkin A.S., Shapkin V.A. *Teoriya riska i modelirovanie riskovykh situatsii* [Theory of risk and risk situations model-ing]. Moscow, Dashkov i K Publ., 2005, 880 p.
3. Mel'nikov V.P. *Zashchita informatsii v komp'yuternykh sistemakh* [Information protection in computer systems]. Moscow, Finansy i statistika Publ., 1997, 368 p.
4. Pavlukhin D.V. *Teoriya informatsionnoi bezopasnosti i metodologiya zashchity informatsii* [Theory of information security and methodology of information protection]. Tambov, Tambov State University Publ., 2005, 104 p.
5. Malyuk A.A. *Informatsionnaya bezopasnost': kontseptual'nye i metodologicheskie osnovy zashchity informatsii* [Information security: conceptual and methodological bases of protection of the information]. Moscow, Goryachaya liniya – Telekom Publ., 2004, 280 p.
6. Petrenko S.A., Simonov S.V. *Upravlenie informatsionnymi riskami. Ekonomicheski opravdannaya bezopasnost'* [Information risk management. Reasonable security]. Moscow, Kompaniya AiTi; DMK Press, 2004, 384 p.
7. Korchenko A.G. *Postroenie sistem zashchity informatsii na nechetkikh mnozhestvakh. Teoriya i prakticheskie resheniya* [Building security systems on fuzzy sets. Theory and practical solutions]. Kiev, MK-Press Publ., 2006, 320 p.
8. Miznov A.S., Shevyakhov M.Yu. *Nekotorye podkhody k otsenke informatsionnykh riskov s ispol'zovaniem nechetkikh mnozhestv* [Some approaches to evaluation of information risks based on fuzzy sets]. Sistemnyi analiz v nauke i obrazovanii, 2010, no. 1, pp. 54–60.
9. Shcheglov K.A., Shcheglov A.Yu. Zashchita ot atak na uyazvimosti prilozhenii. Modeli kontrolya dostupa [Defending against application exploits model of access control]. *Voprosy zashchity informatsii,* 2013, no. 2 (101), pp. 36–43.
10. Saaty T.L. *Elements of Queueing Theory with Applications.* McGraw-Hill, 1961, 423 p. (Russ. ed.: Saati T.L. *Elementy teorii massovogo obsluzhivaniya i ee prilozheniya.* Moscow, Sov. radio Publ., 1965, 511 p.)
11. Polovko A.M., Gurov S.V. *Osnovy teorii nadezhnosti* [Fundamentals of the theory of reliability]. St. Petersburg, BKhV-Peterburg Publ., 2006, 704 p.
12. Shannon C.E. A Mathematical Theory of Communication. *Bell System Technical Journal,* 1948, vol. 27, pp. 379–423, 623–656. (Russ. ed.: Shennon K.E. Matematicheskaya teoriya svyazi. *Raboty po teorii informatsii i kibernetike.* Moscow, Izdatel'stvo inostrannoi literatury Publ., 1963, pp. 243–332.)
13. Shcheglov K.A., Shcheglov A.Yu. Prakticheskaya realizatsiya diskretsionnogo metoda kontrolya dostupa k sozdavaemym failovym ob"ektam [Practical realization of discretionary access control method for newly created objects].*Vestnik komp'yuternykh i informatsionnykh tekhnologii*, 2013, no. 4, pp. 43–49.

14. Shcheglov K.A., Shcheglov A.Yu. Sistema zashchity ot zapuska vredonosnykh program [Malware startup protection system]. *Vestnik komp'yuternykh i informatsionnykh tekhnologii*, 2013, no. 5, pp. 38–43.
15. Shcheglov K.A., Shcheglov A.Yu. Realizatsiya metoda mandatnogo dostupa k sozdavaemym failovym ob"ektam sistemy [Implementation of mandatory access control to newly created file objects method]. *Voprosy zashchity informatsii*, 2013, no. 4 (103), pp. 15–20.

Щеглов Константин Андреевич – студент, Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики; менеджер по развитию, ЗАО «НПП «Информационные технологии в бизнесе», Санкт-Петербург, Россия, schegl_70@mail.ru

Щеглов Андрей Юрьевич – доктор технических наук, профессор, Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики; генеральный директор, ЗАО «НПП «Информационные технологии в бизнесе», Санкт-Петербург, Россия, info@npp-itb.spb.ru

Konstantin Shcheglov Student, Saint Petersburg National Research University of Information Technologies, Mechanics and Optics; development manager, JSC "Information Technologies in Business", Saint Petersburg, Russia, Scheglov.konstantin@gmail.ru

Andrei Shcheglov D.Sc., Professor, Saint Petersburg National Research University of Information Technologies, Mechanics and Optics; General director, JSC "Information Technologies in Business", Saint Petersburg, Russia, info@npp-itb.spb.ru

УДК 004.89

ИДЕНТИФИКАЦИЯ АНОНИМНЫХ ПОЛЬЗОВАТЕЛЕЙ ИНТЕРНЕТ-ПОРТАЛОВ НА ОСНОВАНИИ ТЕХНИЧЕСКИХ И ЛИНГВИСТИЧЕСКИХ ХАРАКТЕРИСТИК ПОЛЬЗОВАТЕЛЯ[1]

А.А. Воробьева[a], А.В. Гвоздев[a]

[a] Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, Санкт-Петербург, Россия, Alice_w@mail.ru

Задача идентификации анонимных пользователей Интернет-порталов становится все более актуальной научной задачей, это обусловлено ростом числа интернет-пользователей, в том числе анонимных, ростом числа случаев совершения противоправных действий (например, анонимных угроз и экстремистских высказываний) и несовершенством существующих подходов и алгоритмов идентификации анонимных пользователей.

В контексте работы под идентификацией пользователя понимается распознавание анонимного пользователя в Интернете [1–5]. Распознавание производится путем соотнесения набора характеристик анонимного пользователя с характеристиками, собранными ранее и уже имеющимися в базе данных. К характеристикам пользователя относятся ряд технических (IP-адрес, версия операционной системы и пр.) и лингвистических (стиль письменной речи автора сообщения) характеристик. В работе рассматривается возможность идентификации пользователей по различным наборам таких характеристик (техническим, лингвистическим и комбинированным). Анализируется возможность применения различных методов классификации (метод опорных векторов, нейросети, логическая регрессия) для решения задачи по идентификации анонимных пользователей.

Проведенные эксперименты показали, что использование лингвистических характеристик совместно с техническими позволяет повысить точность идентификации анонимного пользователя Интернет-портала.

Ключевые слова: идентификация анонимных пользователей, атрибуция текстов, авторство сообщений, компьютерная лингвистика, информационная безопасность.

ANONYMOUS WEBSITE USER IDENTIFICATION BASED ON COMBINED FEATURE SET (WRITING STYLE AND TECHNICAL FEATURES)[2]

A. Vorob'yeva[b], A. Gvozdev[b]

[b] Saint Petersburg National Research University of Information Technologies, Mechanics and Optics, Saint Petersburg, Russia, Alice_w@mail.ru

The task of anonymous web users identification becomes more and more important research task. The number of users is increased dramatically and usage of the Internet for criminal purposes (such as anonymous threats and extremist statements) becomes more frequent. Existing approaches and algorithms for identifying anonymous users are not enough efficient. In the context of this work, user identification means recognizing of an anonymous user on the Internet. Identification is performed by correlating the set of anonymous user features with stored in the database features collected previously. Feature set of the user consists of technical features (IP- address, OS version, etc.) and writing-style features of the user (for short texts in the Russian language). We compared the discriminating power of three feature sets (technical, writing-style and combined) and

[1] Работа выполнена в рамках НИР «Идентификация пользователей порталов сети Интернет».
[2] Done in the framework of S&R work «Identification of Internet portals users»