

МАТЕМАТИЧЕСКИЕ ОСНОВЫ КОМПЬЮТЕРНОЙ БЕЗОПАСНОСТИ

УДК 004.94

DOI 10.17223/20710410/62/5

МОДЕЛЬ И МЕТРИКИ ОСВЕДОМЛЁННОСТИ В КОНФИДЕНЦИАЛЬНОЙ ИНФОРМАЦИИ. ЧАСТЬ 2. ФАКТИЧЕСКАЯ ОСВЕДОМЛЁННОСТЬ

Н. А. Гайдамакин

*Уральский федеральный университет имени первого Президента России Б. Н. Ельцина,
г. Екатеринбург, Россия*

E-mail: n.a.gaidamakin@urfu.ru

Фактическая осведомлённость пользователей в конфиденциальной информации рассматривается как владение соответствующей информацией, характеризующееся степенью восприятия (усвоения) и возможности использования информации («извлечения» из памяти). В рамках субъектно-объектного класса моделей управления доступом в компьютерных системах произведена формализация понятия осведомлённости как результата доступов пользователей к объектам, содержащим конфиденциальную информацию. Доступ к объекту (по чтению), имеющий временные рамки (продолжительность), формирует осведомлённость пользователя в конфиденциальной информации соответствующего объекта, величина которой пропорциональна объёму конфиденциальной информации объекта, индексу сложности (читабельности) текста объекта, продолжительности доступа и, кроме того, зависит от индивидуальных способностей пользователя в восприятии (скорости чтения) и освоении (понимании, обработке) информации. При этом объём конфиденциальной информации объекта определяется как величина, пропорциональная количеству слов в тексте объекта и коэффициенту информативности объекта. С течением времени по кривой забывания Г. Эббингауза осведомлённость пользователя в конфиденциальной информации уменьшается. Степень снижения осведомлённости зависит от индивидуальных особенностей пользователя и уровня конфиденциальности информации. Последующие доступы к объекту могут восстанавливать степень осведомлённости в зависимости от продолжительности доступов и времени, прошедшего с предыдущего доступа. Рассматриваются вид и параметры функции уменьшения/восстановления осведомлённости с течением времени и в зависимости от истории доступов. Осведомлённость пользователя во всей конфиденциальной информации, содержащейся (обрабатываемой) в компьютерной системе, складывается из осведомлённости по всем объектам доступа с учётом синергетического эффекта, который может быть как положительным (знание о системе объектов больше суммы знаний об объектах), так и отрицательным. Характер и особенности фактической осведомлённости пользователей в конфиденциальной информации иллюстрируются на примерах при различных параметрах объектов доступа, истории доступов и индивидуальных особенностей пользователей.

Ключевые слова: *конфиденциальная информация, осведомлённость, фактическая осведомлённость, модель осведомлённости, метрики осведомлённости, информативность объекта доступа, сложность текста объекта доступа, понимание текста объекта доступа, продолжительность доступа, кривая забывания информации, история доступов, эффект синергии в осведомлённости.*

THE MODEL AND METRICS OF AWARENESS IN CONFIDENTIAL INFORMATION. PART 2. ACTUAL AWARENESS

N. A. Gaydamakin

Ural Federal University named after the First President of Russia B. N. Yeltsin, Ekaterburg, Russia

The actual awareness of users in confidential information is considered as the possession of relevant information, characterized by the degree of perception (assimilation) and the possibility of using information ("extraction" from memory). Within the subject-object class of access control models in computer systems, the concept of awareness is formalized as a result of user access to objects containing confidential information. Access to the object (by reading), having a time frame (duration), forms the user's awareness of the confidential information of the corresponding object, the value of which is proportional to the volume of confidential information of the object, the index of complexity (readability) of the text of the object, the duration of access and also depends on the user's individual ability to perceive (reading speed) and master (understanding, processing) information. At the same time, the volume of confidential information of an object is defined as a value proportional to the number of words in the text and the coefficient of informativeness of the object. Over time, according to the Ebbinghaus forgetting curve, the user's awareness of sensitive information decreases. The degree of decrease in awareness depends on the individual characteristics of the user and the level of confidentiality of the information. Subsequent accesses to the object can restore the degree of awareness depending on the duration of the accesses and the time elapsed since the previous access. The type and parameters of the function of reducing/restoring awareness over time and depending on the access history are considered. The user's awareness of all confidential information contained (processed) in a computer system consists of awareness of all access objects, taking into account the synergetic effect, which can be either positive (knowledge about the system of objects is greater than the sum of knowledge about objects) or negative. The nature and features of users' actual awareness in confidential information are illustrated by examples with various parameters of access objects, access history and individual characteristics of users.

Keywords: *confidential information, awareness, actual awareness, awareness model, awareness metrics, informativeness of the access object, complexity of the access object text, understanding of the access object text, duration of access, information forgetting curve, access history, synergy effect in awareness.*

Введение

Как отмечено в [1], *осведомлённость* пользователей в конфиденциальной информации рассматривается в качестве синонима *владения* информацией, т. е. как знание сведений, сообщений, данных, составляющих информацию. Соответственно *фактическая осведомлённость* (Actual Awareness) пользователей в конфиденциальной информации

является результатом осуществления доступов к объектам компьютерной системы, содержащим конфиденциальную информацию.

Вместе с тем известные субъектно-объектные модели управления доступом в компьютерных системах [2, 3] не затрагивают результативную сторону процессов доступа, поскольку неявно полагается, что субъект-пользователь, осуществив доступ к объекту (по чтению), априори обладает соответствующей информацией и впоследствии в любой момент времени может её использовать. Интуитивно ясно, что это не совсем так. В процессе доступа можно успеть прочесть только часть текста (файла, объекта) и в результате обладать только частью информации, содержащейся в объекте доступа. Можно прочесть весь текст объекта (текстовый файл), но ничего не понять или понять только часть информации.

Таким образом, требует формализации процедурно-результативная и темпорально-целевая сторона процесса доступа к информации в компьютерных системах, включая доступ к конфиденциальной информации.

Целевая сторона понятия «осведомлённость» исследовалась в сфере феноменологической социологии [4, 5], где один из подходов трактует «осведомлённость» как *состояние, в котором субъект осведомлён о некоторой информации (владеет некоторой информацией) и эта информация непосредственно доступна для использования в направлении широкого спектра поведенческих действий*. Иначе говоря, осведомлённость, разделяемая в некоторых подходах по феноменологической социологии на «явную» и «неявную» («скрытую»), рассматривается в контексте способности индивидуума к эффективному применению имеющихся у него знаний и сведений, что предполагает, в первую очередь, способность «извлечь из памяти» информацию в последующие моменты времени.

В сфере обучения и психологии исследовались процессы восприятия (усвоения) информации, подаваемой в той или иной форме (пирамида Дейла) — устной (через слушание), текстовой (через чтение) и т. д. [6], и были получены эмпирические данные и соотношения по степени усвоения, сохранения и воспроизведения информации [7–13].

Тем не менее в рамках отмеченных формальных моделей управления доступом в компьютерных системах как в отношении результативности доступа к объектам, содержащим конфиденциальную информацию, так и в отношении способности к эффективному использованию информации в дальнейшем не рассматриваются временные аспекты, связанные с изменениями с течением времени актуальности, в том числе конфиденциальности информации, с изменениями способности индивидуума к эффективному «извлечению» и использованию информации, которой он обладает (забывание информации), а также то, что можно именовать временной диссипацией (деградацией) информации [14].

Наконец, понятие «осведомлённость» широко используется в сфере управления знаниями персонала организации (knowledge management), включая понятие осведомлённости персонала в вопросах информационной безопасности [15–18].

Целью данной работы является рассмотрение формализованной и процедурно-аналитической составляющей (фактической) осведомлённости пользователей в конфиденциальной информации, содержащейся в компьютерных системах.

1. Исходные положения

Будем основываться на субъектно-объектной формализации компьютерной системы, рассматривая её как совокупность *объектов* доступа $o \in O$, содержащих конфиденциальную информацию (файлы, каталоги, базы данных и/или их таблицы-

строки/записи-поля), и *субъектов* доступа $s \in S$ (выполняющиеся по командам *пользователей* $u \in U$ компьютерные программы).

Также будем предполагать, что в компьютерной системе действует дискретное время, в каждый момент t_k которого пользователи $u \in U$ осуществляют *доступы* субъектов $s \in S$ к объектам $o \in O$. Под доступом будем понимать имеющий временные рамки процесс воздействия субъекта $s \in S$ на объект $o \in O$, в результате которого формируется *поток информации* — однонаправленный, т. е. от объекта к субъекту или от субъекта (через субъект) к объекту, либо двунаправленный, т. е. одновременно от субъекта к объекту и от объекта к субъекту.

В контексте анализа осведомлённости, как и в [1], ограничимся рассмотрением только доступов вида «Чтение» (Read) к объектам, содержащим текстовую информацию, и будем использовать *детерминистскую* трактовку понятия *информационного потока* как процесса изменения слова, характеризующего объект-приемник информационного потока в зависимости от слова, характеризующего объект-источник информационного потока. Иначе говоря, объект доступа в теоретико-информационном смысле рассматривается как слово некоторого языка [19].

Под «*фактической осведомлённостью в конфиденциальной информации*» будем понимать величину, характеризующую в количественной (интервальной) шкале степень владения пользователем $u \in U$ конфиденциальной информацией, находящейся в компьютерной системе.

Под *владением* пользователем $u \in U$ конфиденциальной информацией будем понимать её знание, т. е. *ознакомленность* с ней, и её понимание, а также способность к её эффективному использованию, включая количественную (объём информации) и качественную (степень конфиденциальности и извлекаемости) стороны.

Таким образом, быть осведомлённым в конфиденциальной информации означает знать информацию, т. е. её воспринять и понять, переработать и интегрировать с ранее усвоенной информацией, и, кроме того, быть способным её извлечь из памяти и использовать в тех или иных целях.

Степень владения конфиденциальной информацией (величину осведомлённости) в интервальной шкале ограничим диапазоном $[0, 1]$, где 1 — максимальная, т. е. 100-процентная осведомлённость во всей конфиденциальной информации (абсолютно полное владение информацией), обрабатываемой в компьютерной системе.

2. Объём конфиденциальной информации, информативность, сложность и конфиденциальность информации объектов доступа

Как уже отмечалось, в теоретико-информационном смысле при детерминистском подходе объект доступа рассматривается как слово некоторого языка в определённом алфавите [19]. В случае конечного алфавита объём информации, реализуемый информационным потоком в результате доступа, характеризуется количеством слов, представляющих объект. Исходя из этого, понятие «*объём (конфиденциальной) информации объекта o_n* » можно определить как величину, пропорциональную количеству слов $Q(o_n, t_m)$, содержащихся в момент времени t_m в тексте объекта o_n :

$$V(o_n, t_m) = Q(o_n, t_m) \theta(o_n, t_m), \quad (1)$$

где $\theta(o_n, t_m)$ — изменяющийся в диапазоне $[0, 1]$ коэффициент информативности объекта o_n в момент времени t_m .

Введение коэффициента информативности $\theta(o_n, t_m)$ обусловлено тем, что языком письменной речи одну и ту же информацию можно выразить по-разному, с разной

ясностью, чёткостью, полнотой, «понятностью» и, следовательно, с различным словесным объёмом. Будем считать, что существует вещественно-значная функция, выражающая данное свойство объектов в числовом диапазоне $[0, 1]$.

В процессе восприятия (чтения) текстовой информации, её усвоения важным фактором является т. н. «*читабельность*» («*понятность*», «*трудность*») текста, во многих источниках отождествляемая с понятием *сложности* текста. Многочисленные исследования, обзор которых можно найти в [20], определили ряд критериев сложности текста, к которым относятся, прежде всего, количественно-статистические параметры (средняя буквенная или слоговая длина слов, среднее количество слов в предложениях, доля сложносочинённых предложений, количество специальных терминов), а также ряд других грамматических и семантико-тематических критериев (повествовательность изложения, конкретность/абстрактность слов и выражений, среднее значение словесного расстояния между связанными элементами предложений и др.). Предложены эмпирические соотношения для определения «индекса читабельности» (*индекса сложности*) текста, на основе которых разработаны специальные программные средства для автоматизированного вычисления индекса читабельности, в том числе для текстов на русском языке [20].

Будем считать, что существует вещественно-значная функция $\varphi(o_n, t_m)$, которая для каждого объекта доступа o_n в любой момент времени t_m ставит в соответствие количественное значение его индекса сложности. Диапазон значений индекса сложности объекта $\varphi(o_n, t_m)$ определим так, чтобы его среднее значение для группы однородно-тематических текстов (например, тексты финансово-экономической, кадровой, проектно-конструкторской и т. д. тематики) равнялось 1.

Конфиденциальность информации объектов доступа будем понимать как такое свойство соответствующей информации, когда может возникнуть какой-либо ущерб при свободном обороте этой информации, т. е. в результате доступа к ней (соответственно осведомлённости в ней) неопределённого круга лиц.

Отметим, что не все объекты доступа содержат конфиденциальную информацию и/или не вся информация объекта является конфиденциальной. Однако ввиду существенных трудностей формализации вопросов семантического анализа и разделения информации объекта доступа на конфиденциальную и неконфиденциальную части будем считать, что весь объём информации объекта доступа, определяемый по соотношению (1), в том числе с учётом информативности объекта доступа, характеризует конфиденциальную информацию, содержащуюся в объекте.

Также отметим, что из понятия конфиденциальности следует её различная степень — в зависимости от величины ущерба при свободном обороте соответствующей информации. В большинстве случаев степень конфиденциальности информации ввиду существенных сложностей в определении количественных значений ущерба выражается величиной в качественных (порядково-вербальных) шкалах («высокая», «средняя», «низкая»). Тем не менее будем считать, что существует вещественно-значная функция $f_{\text{conf}}(o_n, t_m)$, которая каждому объекту $o_n \in O$ компьютерной системы в каждый момент времени t_m ставит в соответствие некоторую величину конфиденциальности $K = f_{\text{conf}}(o_n, t_m)$ в диапазоне $[0, 1]$.

3. Осведомлённость в конфиденциальной информации пользователя в результате однократного доступа к объекту

Из рассмотренных выше определений и формализаций следует, что количество информации, которую получает субъект s (управляющий им пользователь u) в резуль-

тате доступа к объекту o определяется словесным объёмом соответствующего объекта (например, файла).

Однако здравый смысл говорит, что это не всегда так, поскольку, как уже отмечалось, в результате доступа можно ознакомиться только с частью информации или понять только часть информации объекта.

Разрешение этой проблемы представляется на основе введения темпорального аспекта в определение и характеристики доступа, т. е. временных рамок, и результативных аспектов последовательности доступов в понятие владения информацией.

Сделаем следующие предположения.

Положение 1. Пользователи $u_l \in U$, осуществляя в моменты времени t_m доступы к объектам $o_n \in O$ продолжительностью Δt_{mn} , формируют или повышают степень владения информацией, заключённой в соответствующих объектах.

Таким образом, атрибутами доступа, помимо идентификаторов субъекта и объекта, вида доступа (чтение, запись, редактирование), являются начало и окончание (длительность) доступа.

Положение 2. Приращение осведомлённости в конфиденциальной информации $\mathcal{A}_\Delta(u_l, o_n, t_m, \Delta t_{mn})$ пользователя u_l в результате доступа к объекту o_n в момент времени t_m продолжительностью Δt_{mn} определяется:

- объёмом $V(o_n, t_k)$ конфиденциальной информации, содержащейся в момент времени t_m в объекте o_n ;
- продолжительностью Δt_{mn} доступа пользователя u_l к объекту o_n в момент времени t_m ;
- величиной $f_{\text{conf}}(o_n, t_m)$ конфиденциальности информации, содержащейся в момент времени t_m в объекте o_n ;
- сложностью $\varphi(o_n, t_m)$ текста объекта доступа o_n в момент времени t_m ;
- индивидуальными особенностями (способностями) пользователя u_l в отношении скорости восприятия информации (скорости чтения);
- индивидуальными особенностями (способностями) пользователями u_l в отношении степени понимания воспринятой информации, т. е. её переработки и интегрирования с ранее усвоенной информацией.

Диапазон значений функции приращения осведомлённости $\mathcal{A}_\Delta(u_l, o_n, t_m, \Delta t_{mn})$ ограничим отрезком $[0, 1]$, полагая под значением 1 максимальную осведомлённость в конфиденциальной информации, имеющейся в объекте.

Исходя из положения 2, приращение осведомлённости $\mathcal{A}_\Delta(u_l, o_n, \Delta t_{mn})$ пользователя u_l в конфиденциальной информации объекта o_n в результате доступа в момент времени t_m продолжительностью Δt_{mn} будем определять на основе соотношения

$$\mathcal{A}_\Delta(u_l, o_n, t_m, \Delta t_{mn}) = \left(\frac{\bar{\vartheta} \beta_l \Delta t_{mn}}{Q(o_n, t_m)} \theta(o_n, t_m) \varphi(o_n, t_m) (1 - \chi_{mn}) \gamma_l \right)_{|1}, \quad (2)$$

где:

- $\bar{\vartheta}$ — средняя скорость чтения текста, равная по некоторым данным 200 слов в минуту [21];
- β_l — коэффициент индивидуальной способности пользователя u_l в отношении скорости чтения текста относительно среднего значения, принимающий по некоторым данным значения в диапазоне от 0,3 до 1,88;
- χ_{mn} — доля времени закрытия текста объекта o_n на экране программой типа «Заставка» от момента времени t_m до момента времени $t_m + \Delta t_{mn}$ (для многостранич-

- ных текстов, для которых требуется «пролистывание» текста на экране компьютера, в величину χ_{mn} можно добавлять долю времени «неактивности пользователя»);
- γ_l — коэффициент индивидуальной способности пользователя u_l в отношении понимания прочитанного текста в диапазоне от 0 до 1, среднее значение которого по некоторым данным равно 0,52 (52 %) (имеется в виду коэффициент понимания текста служебных документов. В отношении сложных учебных, технических, научных текстов можно использовать с известными оговорками (без учёта влияния иллюстраций, мультимедиа и т. д.) данные пирамиды Дейла [6], согласно которым объём усвоения однократно прочитанного текста обучаемыми в среднем составляет 10 %);
 - $(x)_{|1}$ обозначает функцию

$$(x)_{|1} = \begin{cases} x, & \text{если } x < 1, \\ 1, & \text{если } x \geq 1. \end{cases}$$

Значение $\mathcal{A}_\Delta(u_l, o_n, t_m, \Delta t_{mn}) = 1$ означает 100-процентную осведомлённость пользователя u_l в конфиденциальной информации, содержащейся в тексте объекта o_n .

Параметры β_l и γ_l , характеризующие индивидуальные особенности пользователей в отношении восприятия и усвоения воспринятой информации, могут оцениваться на основе аналитико-тестовых процедур, подобных процедурам определения профессиональной квалификации сотрудников организаций.

Выражение (2) задаёт приращение осведомлённости пользователя в результате однократного доступа к объекту.

4. Осведомлённость пользователя в конфиденциальной информации объекта в результате последовательности доступов

Сделаем следующие очевидные предположения.

Положение 3. Осведомлённость пользователя u_l в конфиденциальной информации объекта o_n с течением времени имеет тенденцию к снижению, при этом осведомлённость в более конфиденциальной информации снижается в меньшей степени.

Положение 4. Последующие доступы пользователя u_l к объекту o_n после доступа в момент времени t_m могут повышать (возобновлять, восстанавливать) осведомлённость пользователя в соответствующей информации.

Действительно, если осведомлённость пользователя в информации объекта o_n не достигла 1, то пользователь может её повысить в результате повторного и последующего доступов. При этом с течением времени соответствующая информация может забываться, воспроизводиться («извлекаться» из памяти) в меньшем объёме, с худшим качеством (с худшим пониманием) и в результате степень владения пользователем соответствующей информацией будет снижаться. Очевидно, что более важная, т. е. более конфиденциальная информация обрабатывается и хранится в сознании пользователя более тщательно и, следовательно, степень владения такой информацией снижается медленнее, чем менее конфиденциальной информацией.

В результате, основываясь на положениях 3 и 4, величину $\mathcal{A}(u_l, o_n, t_k)$ осведомлённости пользователя u_l в конфиденциальной информации, содержащейся в объекте o_n в момент времени t_k после последнего доступа к этому объекту в момент времени t_m , можно определять на основе следующего итеративного соотношения:

$$\mathcal{A}(u_l, o_n, t_k) = (\mathcal{A}(u_l, o_n, t_m) f_{\text{Forg}}(u_l, o_n, t_k, t_m) + \mathcal{A}_\Delta(u_l, o_n, t_k, \Delta t_{nk}))_{|1}, \quad (3)$$

где $f_{\text{Forg}}(u_l, o_n, t_k, t_m)$ — функция «забывания» информации (доля остаточной в памяти информации от первоначальной) и, следовательно, снижения к моменту времени t_k степени владения пользователем u_l информацией объекта o_n с момента предыдущего ознакомления с ней в момент времени t_m , $m < k$.

Начало исследованиям процессов сохранения информации в памяти положили работы Г. Эббингауза (см. [7, 13]), который определил «кривую забывания» (*Forgetting Curve*), выражающуюся в разных интерпретациях эмпирически определённой функцией вида

$$f_{\text{Forg}}(u_l, o_n, t_k, t_m) = \frac{a}{a + b \log(t_k - t_m)},$$

где a и b — параметры «забывания», эмпирические средние значения которых при усреднении для различных индивидуумов (пользователей), т.е. без учёта их индивидуальных особенностей, равны 1,84 и 1,25 соответственно.

Следует отметить, что значения параметров a и b , экспериментально найденные Эббингаузом, относятся к способности обучаемых через определённое время воспроизводить слогословесную информацию (трёхбуквенные слоги, лишённые ассоциативной связи). Многочисленные последующие исследования подтвердили результаты Эббингауза, при этом были исследованы вопросы влияния на запоминание, сохранение и воспроизведение информации ассоциативных связей в элементах информации, структурных и ритмичных характеристик, эмоциональной окраски информации. Были предложены и другие виды функций кривых забывания [7–13, 22], характер поведения которых в общем виде сходен с кривой Г. Эббингауза.

В контексте анализа осведомлённости в конфиденциальной информации необходимо анализировать способность пользователей воспроизводить не словесно-речевую форму прочитанного текста, а усвоенное в результате предыдущего чтения смысловое содержание, т.е. основные смысловые составляющие текста и их параметры. Скорость забывания смысловых составляющих информации существенно медленнее и по данным многих исследователей [10, 13, 22] может характеризоваться функцией экспоненциального вида

$$f_{\text{Forg}}(u_l, o_n, t_k, t_m) \approx e^{-\lambda_l(t_k - t_m)}, \quad (4)$$

где λ_l — параметр скорости забывания информации l -м пользователем, среднее значение которого в отношении забывания (объёма воспроизведения) смысловых элементов текста характеризуется величиной, которую на основе аппроксимации функции (4) по методу наименьших квадратов экспериментальными данными, приведёнными в [23, 24], можно определить равной $5,36 \cdot 10^{-06}$ мин $^{-1}$.

Индивидуальное значение параметра скорости забывания информации l -м пользователем λ_l , как и параметры β_l и γ_l , может определяться на основе специальных аналитико-тестовых процедур.

В соответствии с положением 3, функция (4) должна учитывать влияние на скорость забывания уровня конфиденциальности информации. Этого можно достичь, введя в соотношение (4) функцию $f_{\text{conf_v}}(o_n, t_m)$ как некий коэффициент пропорциональности степени забывания информации от уровня конфиденциальности:

$$f_{\text{Forg}}(u_l, o_n, t_k, t_m) = f_{\text{conf_v}}(o_n, t_m) e^{-\lambda_l(t_k - t_m)}. \quad (5)$$

Здесь $f_{\text{conf_v}}(o_n, t_m)$ — функция со значениями в диапазоне от 0 до 1, определяющая степень влияния на скорость забывания уровня конфиденциальности информации.

При этом, ввиду неопределённости факторов, определяющих вид функции $f_{\text{conf_v}}(o_n, t_m)$, будем считать её тождественной функции $f_{\text{conf}}(o_n, t_m)$. Иначе говоря,

при самом высоком уровне конфиденциальности информации объекта o_n значение $f_{\text{conf_v}}(o_n, t_{i-1}) = 1$, т. е. максимально медленное забывание соответствующей информации; при уровнях конфиденциальности, меньших максимального, $f_{\text{conf_v}}(o_n, t_{i-1}) < 1$ и забывание происходит быстрее.

Пример 1. На рис. 1 для иллюстрации характера функции (3) приведены расчёты величины осведомлённости пользователя в конфиденциальной информации $\mathcal{A}(u_i, o_1, t_k)$, содержащейся в объекте o_1 , объём которого составляет 3000 слов (порядка 10 страниц текста) с неизменяющимися по времени уровнем конфиденциальности «Средний» (0,809 от максимального), уровне информативности $\theta = 0,5$, индексе сложности текста $\varphi = 1$, при полном задействовании времени доступов ($\chi_{m1} = 0$). Об эвристиках перевода показаний порядково-вербальной шкалы конфиденциальности в количественную см. [1].

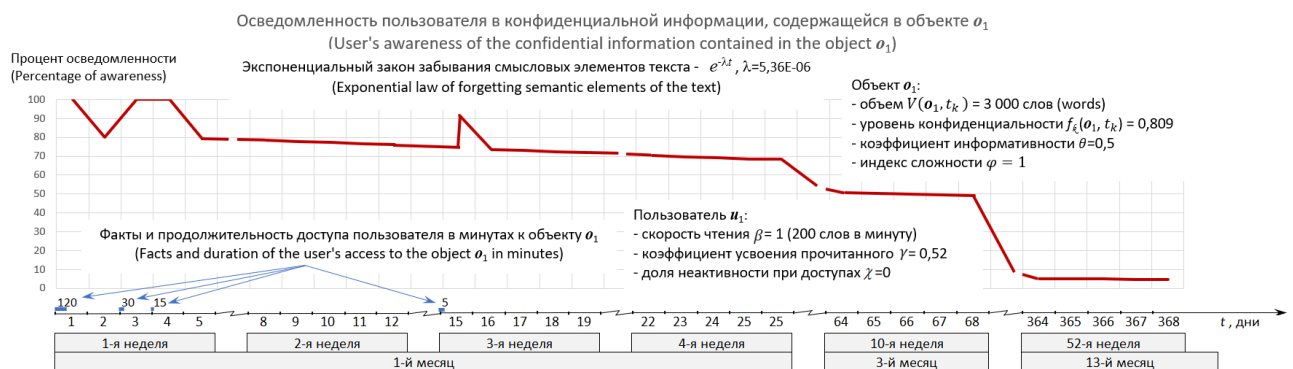


Рис. 1. Пример для иллюстрации осведомлённости пользователя в конфиденциальной информации, содержащейся в объекте, в результате определённой истории доступов

Как видно из рис.1, пользователь u_1 в первый рабочий день в течение двух часов имел доступ к объекту o_1 , что позволило ему несколько раз его прочесть (напомним, средняя скорость чтения текста составляет 200 слов в минуту), как говорят, «полностью отработать», и в результате с учётом в среднем 52-процентной усвояемости прочитанного один раз текста получить полное представление о содержании (о смысле и деталях) информации объекта. В следующий рабочий день у пользователя u_1 доступа к объекту o_1 не было, и в результате процессов забывания информации, в том числе с учётом влияния на скорость забывания уровня конфиденциальности, степень владения конфиденциальной информацией соответствующего объекта у пользователя снизилась до 80 %. В третий и четвёртый рабочие дни пользователь u_1 вновь осуществлял доступы к объекту o_1 в течение 30 и 15 мин соответственно, что позволило ему повысить уровень владения соответствующей информацией вновь до 100 %.

В последующие два рабочих дня первой недели и всю вторую неделю доступа пользователя u_1 к объекту o_1 не было, шёл процесс забывания информации по экспоненциальному закону (сначала очень быстро, затем медленнее), в результате чего уровень владения конфиденциальной информацией объекта к началу третьей рабочей недели снизился до 75 % (график на рис. 1 отражает величину осведомлённости с шагом по времени в один день, поэтому имеет вид ломаной линии). В первый рабочий день третьей недели пользователь u_1 имел непродолжительный (5-минутный) доступ к объекту o_1 (видимо, чтобы вспомнить или уточнить какие-либо аспекты информации), что позволило повысить уровень владения до 91,6 %, т. е. из-за кратковременности доступа не до 100 %. Далее ввиду отсутствия у пользователя u_1 в последующее время

доступов к объекту o_1 шёл процесс экспоненциального снижения уровня владения соответствующей информацией и к концу года степень осведомлённости пользователя в конфиденциальной информации объекта o_1 составила всего 4,8 %.

5. Осведомлённость пользователя во всей конфиденциальной информации в результате последовательности доступов к объектам компьютерной системы

Сделаем следующие предположения.

Положение 5. Осведомлённость пользователя в конфиденциальной информации по некоторой системе объектов складывается из осведомлённости в конфиденциальной информации по отдельным объектам, её составляющим.

Положение 6. Осведомлённость пользователя в конфиденциальной информации некоторой системы объектов в целом характеризуется синергетическим эффектом, который в зависимости от особенностей пользователя и конфиденциальности информации по соответствующим объектам, сложности системы и характера взаимосвязей объектов, её составляющих, может быть как положительным, так и отрицательным.

Положительный эффект синергии в осведомлённости заключается в формировании нового знания о системе объектов на основе обработки, агрегирования и обобщения знаний по отдельным объектам. Иначе говоря, знание о системе больше суммы знаний о составляющих её объектах.

Отрицательный эффект синергии заключается в том, что знание о системе ухудшается (уменьшается, «запутывается») в результате неспособности индивидуума переработать, обобщить, интегрировать знание по отдельным объектам системы в целое, т. е. знание о системе в целом меньше суммы знаний о составляющих её объектах.

Исходя из положений 5 и 6, осведомлённость $\mathcal{A}(u_l, t_k)$ пользователя u_l во всей конфиденциальной информации, содержащейся в компьютерной системе в момент времени t_k , будем определять как аддитивно-мультипликативную функцию осведомлённости по объектам, содержащим конфиденциальную информацию, по отношению ко всему объёму конфиденциальной информации:

$$\mathcal{A}(u_l, t_k) = \left(\mathcal{E}(u_l, t_k, O) \sum_{n=1}^{N_{t_k}} \mathcal{A}_{\mathbf{N}}(u_l, o_n, t_k) \right)_{|1}. \quad (6)$$

Здесь:

- $\mathcal{E}(u_l, t_k, O)$ — функция эффекта синергии в осведомлённости по системе объектов в целом в зависимости от осведомлённости по каждому отдельному объекту, значениями которой являются положительные числа как больше 1 (положительный эффект синергии), так и меньше 1 (отрицательный эффект синергии);
- O — множество объектов доступа компьютерной системы с отношениями (с их взаимосвязями, содержанием и т. д.);
- N_{t_k} — количество объектов с конфиденциальной информацией в момент времени t_k ;
- $\mathcal{A}_{\mathbf{N}}(u_l, o_n, t_k)$ — величина осведомлённости пользователя u_l в момент времени t_k в конфиденциальной информации объекта o_n , вычисляемая с нормированием по отношению к объёму всей конфиденциальной информации, обрабатываемой в компьютерной системе в соответствующие моменты времени, с учётом степени конфиденциальности информации объектов доступа:

$$\mathcal{A}_{\mathbf{N}}(u_l, o_n, t_k) = (\mathcal{A}_{\mathbf{N}}(u_l, o_n, t_m) f_{\text{Forg}}(u_l, o_n, t_k, t_m) + \mathcal{A}_{\mathbf{N}_{\Delta}}(u_l, o_n, t_k, \Delta t_{nk}))_{|1},$$

где t_m — момент времени последнего доступа пользователя u_l к объекту o_n ;

- $\mathcal{A}_{N_{\Delta}}(u_l, o_n, t_k, \Delta t_{nk})$ — нормированное приращение осведомлённости пользователя u_l в конфиденциальной информации объекта o_n в результате доступа в момент времени t_k продолжительностью Δt_{nk} , которое, в свою очередь, определяется следующим соотношением:

$$\mathcal{A}_{N_{\Delta}}(u_l, o_n, t_k, \Delta t_{nk}) = \left(f_{\mathcal{K}}(o_n, t_k) \frac{\bar{\vartheta} \beta_l \Delta t_{nk}}{\sum_{j=1}^{N_{t_k}} V(o_j, t_k)} \theta(o_n, t_k) \varphi(o_n, t_k) (1 - \chi_{kn}) \gamma_l \right)_{|1} \quad (7)$$

Ввиду сложности природы эффекта синергии определение вида функция $\mathcal{E}(u_l, t_k, O)$ представляет отдельное направление исследований. В практических приложениях можно принять $\mathcal{E}(u_l, t_k, O) = 1$.

Пример 2. На рис. 2 для иллюстрации характера поведения величины $\mathcal{A}(u_l, t_k)$ приведены результаты расчёта осведомлённости пользователя в конфиденциальной информации, обрабатываемой в компьютерной системе, включающей три объекта, параметры которых (объём, конфиденциальность, информативность) представлены на рисунке.

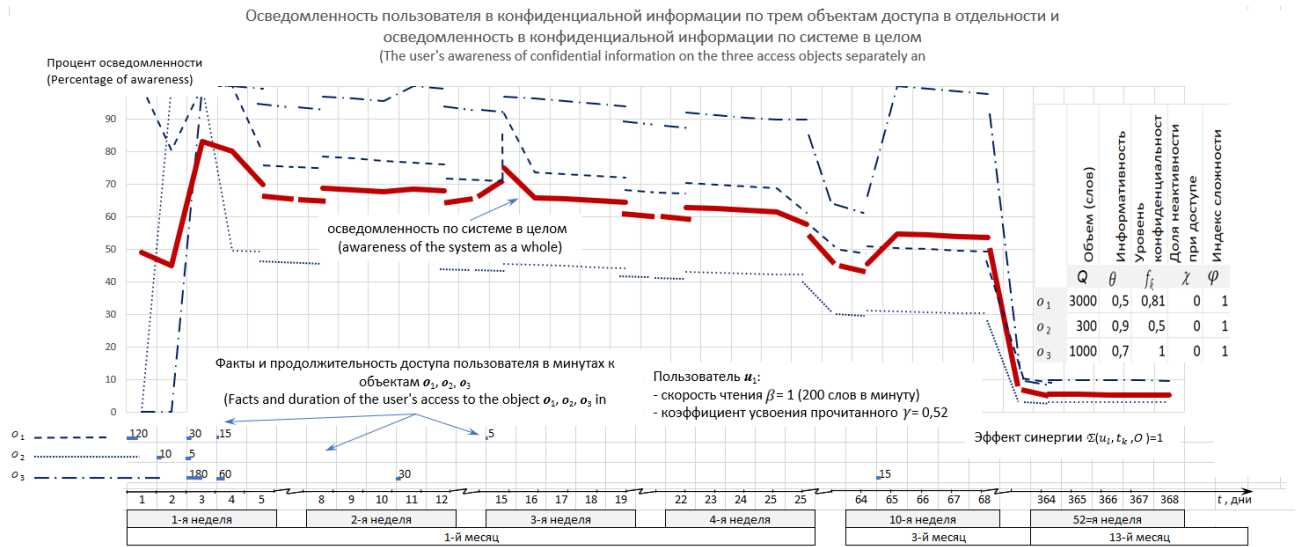


Рис. 2. Иллюстрация осведомлённости пользователя о конфиденциальной информации в компьютерной системе, включающей три объекта, в результате определённой последовательности (истории) доступов

Как видно из рис. 2, пользователь u_1 осуществлял серию доступов к 10-страничному объекту o_1 со средним уровнем конфиденциальности, коэффициентами информативности и сложности по фактам и продолжительности доступов аналогично данным рис. 1.

К небольшому (одностраничному) объекту o_2 с низким уровнем конфиденциальности, но с высоким уровнем информативности, пользователь u_1 во второй и третий рабочий день осуществил 10-минутный и 5-минутный доступы соответственно, что с учётом небольшого объёма информации объекта обеспечило ему 100-процентную осведомлённость в соответствующей конфиденциальной информации. Больше доступов к объекту o_2 не было и шёл процесс экспоненциального снижения осведомлённости.

При этом вклад осведомлённости в соответствующей информации в «общую копилку» осведомлённости ввиду небольшого объёма и невысокого уровня конфиденциальности объекта o_2 являлся незначительным.

К объекту o_3 , самому конфиденциальному ($f_{\text{conf}}(o_3, t_k) = 1$) и довольно объёмному ($Q(o_3, t_k) = 1000$ слов — 3–4 страницы текста) с высокой степенью информативности ($\theta(o_3, t_k) = 0,7$), пользователь u_1 осуществлял продолжительные доступы (180 и 60 мин) в третий и четвёртый рабочие дни, что обеспечило ему 100-процентную осведомлённость в соответствующей конфиденциальной информации. При этом снижение степени владения (осведомлённости) до следующего доступа на 11-й рабочий день ввиду высокой степени конфиденциальности соответствующей информации происходило существенно медленнее, чем по объектам o_1 и o_2 . Последующие доступы пользователя к объекту o_3 продолжительностью 30 минут (в 11-й рабочий день) и 15 минут (через два месяца в 65-й рабочий день) ввиду высоких остаточного уровня осведомлённости и величины продолжительности доступа по отношению к объёму объекта позволяли восстанавливать осведомлённость до 100 %.

Как видно из рис. 2, характер изменения осведомлённости пользователя u_1 по всей конфиденциальной информации системы с учётом времени доступа, объёма, информативности и конфиденциальности объектов определялся в основном доступами к объектам o_1 и o_3 . При этом вклад в общую осведомлённость доступов к объекту o_3 был примерно равным вкладу от доступов к существенно более объёмному, но менее конфиденциальному и менее информативному объекту o_1 .

Рассмотренные примеры расчётов и анализа осведомлённости в конфиденциальной информации демонстрируют результаты, отражающие интуитивные качественные представления о содержании понятия «осведомлённость» и характере её изменений.

Следует отметить, что как и в случае с анализом потенциальной осведомлённости, дополнительные возможности анализа фактической осведомлённости предоставляет тематико-иерархическое управление доступом [25, 26]. В таких системах конфиденциальная информация объектов доступа помечается (индексируется) мультирубриками [27], которые являются совокупностью рубрик-тематик $\mathcal{T}_i^{(m)} = \{\tau_{i_1}, \tau_{i_2}, \dots, \tau_{i_l}\}$, организованных в иерархический тематический рубрикатор $T_{\text{TH}} = \{r, \tau_1, \tau_2, \dots, \tau_K\}$, где r — корень дерева (вся тематика), отражающего иерархическую структуру рубрикатора. Управление доступом осуществляется на основе сравнения мультирубрик объектов доступа и тематических полномочий субъектов доступа (пользователей), которые выражаются также совокупностью разрешённых рубрик-тематик (разрешённой мультирубрикой). В результате появляется возможность анализировать фактическую осведомлённость пользователей не только в целом по всей конфиденциальной информации системы, но и осведомлённость в конфиденциальной информации по определённой рубрике-тематике или по некоторой тематической мультирубрике $\mathcal{T}_i^{(m)} — \mathcal{A}(u_l, \mathcal{T}_i^{(m)}, t_k)$. С этой целью в соотношениях (2), (7) необходимо ввести в качестве дополнительного сомножителя двоичную функцию $\delta(\mathcal{T}^{(m)}(o_n, t_k), \mathcal{T}_i^{(m)})$, равную 1, если тематика (мультирубрика) объекта доступа $\mathcal{T}^{(m)}(o_n, t_k)$ шире (охватывает) анализируемую тематику (мультирубрику) конфиденциальной информации $\mathcal{T}_i^{(m)}$, и 0 в противном случае.

Заключение

Как и в случае с «потенциальной осведомлённостью» [1], представленные понятие и величины фактической осведомлённости пользователей в конфиденциальной ин-

формации $\mathcal{A}_\Delta(u_l, o_n, \Delta t_{mn})$, $\mathcal{A}(u_l, o_n, t_k)$ и $\mathcal{A}(u_l, t_k)$ могут составить базис специального программного обеспечения в процессах мониторинга информационной безопасности компьютерных систем. Анализируя значения осведомлённости пользователей в конфиденциальной информации отдельных объектов или по всей совокупности объектов, администраторы компьютерных систем могут формировать на этой основе обоснованные решения в контексте обеспечения информационной безопасности.

Также следует отметить, что, как и в отношении технических аспектов анализа потенциальной осведомлённости, основные параметры для вычисления величин $\mathcal{A}_\Delta(u_l, o_n, \Delta t_{mn})$, $\mathcal{A}(u_l, o_n, t_k)$ и $\mathcal{A}(u_l, t_k)$ либо автоматически определяются в современных офисных системах работы с документами и электронного документооборота (объём информации в словах, история и продолжительность доступов), либо могут определяться на основе регламентации процедур делопроизводства.

Отдельных специальных исследований требуют вид функций «кривых забывания» в контексте владения конфиденциальной информацией и связанные с ними параметры восприятия, усвоения конфиденциальных данных, в том числе вопросов «старения» информации [14], а также вопросов, связанных с эффектом синергии осведомлённости по системе объектов доступа. Направлениями развития представленного подхода может быть рассмотрение осведомлённости и соответствующих параметров $\theta(o_n, t_k)$, $\varphi(o_n, t_k)$, β_l , γ_l , λ_l , $\mathcal{E}(u_l, t_k, O)$ как случайных величин с распределениями вероятностей, требующими теоретических и экспериментальных исследований, или как величин, являющихся элементами нечётких множеств [28, 29].

Вместе с тем представленные данные по соответствующим функциям и параметрам можно рассматривать в качестве «первого приближения» для использования в практических системах.

ЛИТЕРАТУРА

1. Гайдамакин Н. А. Модель и метрики осведомленности в конфиденциальной информации. Часть 1. Потенциальная осведомленность // Прикладная дискретная математика. 2023. № 61. С. 86–103.
2. Девянин П. Н. Модели безопасности компьютерных систем. Управление доступом и информационными потоками. М.: Горячая линия — Телеком, 2020. 352 с.
3. Гайдамакин Н. А. Разграничение доступа к информации в компьютерных системах. Екатеринбург: Изд-во Урал. ун-та, 2003. 328 с.
4. Chalmers D. The Conscious Mind: In Search of a Fundamental Theory. Oxford: Oxford University Press, 1996. 225 p.
5. Шютц А. Смысловая структура повседневного мира: очерки по феноменологической социологии. М.: Институт Фонда «Общественное мнение», 2003. 336 с.
6. Lee S. J. and Reeves T. C. Edgar Dale: A significant contributor to the field of educational technology // Educational Technology. 2007. V. 47. No. 6. P. 56–59.
7. Хрестоматия по общей психологии. Психология памяти / под ред. Ю. Б. Гиппенрейтер, В. Я. Романова. М.: Изд-во МГУ, 1979. 272 с.
8. Аткинсон Р. Человеческая память и процесс обучения. М.: Прогресс, 1980. 527 с.
9. Зинченко Т. П. Память в экспериментальной и когнитивной психологии. СПб.: Питер, 2002. 320 с.
10. Ланге В. Г. О скорости забывания // Вопросы психологии. 1983. № 4. С. 142–145.
11. Wickelgren W. A. Trace resistance and the decay of longterm memory // J. Math. Psychology. 1972. No. 9. P. 418–455.

12. Wickelgren W. A. Single-trace fragility theory of memory dynamics // *Memory Cognition*. 1974. V. 2(4). P. 775–780.
13. <http://www.ideationizing.com/2009/06/brief-history-of-mathematical.html> — A Brief History of the Mathematical Definition of Forgetting Curves. 2023.
14. Ефимов А. Н. Информация: ценность, старение, рассеяние. М.: Знание, 1978. 64 с.
15. ГОСТ Р 53894-2016. Менеджмент знаний. Термины и определения. М.: Стандартинформ, 2020. 24 с.
16. ГОСТ Р ИСО/МЭК 27002. Информационная технология. Методы и средства обеспечения безопасности. Свод норм и правил менеджмента информационной безопасности. М.: Стандартинформ, 2021. 74 с.
17. Building an Information Technology Security Awareness and Training Program. NIST Special Publication 800-50, 2003.
18. Астахова Л. В., Ульянов Н. Л. Модель политики управления осведомленностью сотрудников организации в области информационной безопасности // Материалы 67-й науч. конф. ЮУрГУ, Челябинск, 14–17 апреля 2015 г. С. 678–682.
19. Грушо А. А., Применко Е. А., Тимонина Е. Е. Теоретические основы компьютерной безопасности. М.: Академия, 2009. 272 с.
20. Солнышкина С. И., Кисельников А. С. Сложность текста: этапы изучения в отечественном прикладном языкознании // Вестник Томского госуниверситета. Филология. 2015. № 6(38). С. 86–99.
21. Чмыхова Е. В., Давыдов Д. Г., Лаврова Т. П. Экспериментальное исследование факторов скорости чтения // Психология обучения. 2014. № 9. С. 26–36.
22. Буймов А. Г. Закономерности поведения кривых забывания // Доклады ТУСУРа. 2017. Т. 20. № 4. С. 138–141.
23. Thalheimer W. How Much Do People Forget? <http://www.work-learning.com/catalog.html>. 2023
24. Шардаков М. Н. Усвоение и сохранение в обучении. Учёные записки Ленингр. гос. пед. ин-та им. А. И. Герцена. 1940. Т. 36. 195 с.
25. Гайдамакин Н. А. Модель тематического разграничения доступа к информации при иерархической структуре классификатора в автоматизированных системах управления // Автоматика и телемеханика. 2003. № 3. С. 177–189.
26. Гайдамакин Н. А. Многоуровневое тематико-иерархическое управление доступом (MLTHS-система) // Прикладная дискретная математика. 2018. № 39. С. 42–57.
27. Гайдамакин Н. А., Баранский В. А. Алгебра мультирубрик на корневых деревьях иерархических тематических классификаторов // Сиб. электрон. матем. изв. 2017. Т. 14. С. 1030–1040.
28. Нечеткие множества в моделях управления и искусственного интеллекта / под ред. Д. А. Поспелова. М.: Наука, 1986. 312 с.
29. Пегат А. Нечеткое моделирование и управление: пер. с англ. 2-е изд. М.: БИНОМ. Лаборатория знаний, 2013. 798 с.

REFERENCES

1. Gaydamakin N. A. Model' i metriki osvedomlennosti v konfidentsial'noy informatsii. Chast' 1. Potentsial'naya osvedomlennost' [The model and metrics of awareness in confidential information. Part 1. Potential awareness]. *Prikladnaya Diskretnaya Matematika*, 2023, no. 61, pp. 86–103. (in Russian)

2. *Devyanin P. N.* Modeli bezopasnosti kompyuternykh sistem. Upravleniye dostupom i informatsionnymi potokami [Security models of computer systems. Access and information flow management]. Moscow, Goryachaya liniya — Telekom, 2020. 352 p. (in Russian)
3. *Gaydamakin N. A.* Razgranichenie dostupa k informatsii v komp'yuternykh sistemakh. [Differentiation of access to information in computer systems]. Ekaterinburg, UrFU Publ., 2003. 328 p. (in Russian)
4. *Chalmers D.* The Conscious Mind: In Search of a Fundamental Theory. Oxford, Oxford University Press, 1996. 225 p.
5. *Schutz A.* The Phenomenology of the Social World. Northwestern University Press, 1972. 255 p.
6. *Lee S. J. and Reeves T. C.* Edgar Dale: A significant contributor to the field of educational technology. Educational Technology, 2007, vol. 47, no. 6, pp. 56–59.
7. Khrestomatiya po obshchey psikhologii. Psikhologiya pamyati [Common Psychology reader. Memory Psychology]. Yu. B. Gippenreyter and V. Ya. Romanov (eds), Moscow, MSU Publ., 1979. 272 p. (in Russian)
8. *Atkinson R.* Chelovecheskaya pamyat' i protsess obucheniya [Human Memory and Learning Process]. Moscow, Progress Publ., 1980. 527 p. (in Russian)
9. *Zinchenko T. P.* Pamyat' v eksperimental'noy i kognitivnoy psikhologii [Memory in experimental and cognitive psychology]. Saint-Petersburg, Piter, 2002. 320 p. (in Russian)
10. *Lange V. G.* O skorosti zabyvaniya [About forgetting speed]. Voprosy Psikhologii, 1983, no. 4, pp. 142–145. (in Russian)
11. *Wickelgren W. A.* Trace resistance and the decay of longterm memory. J. Math. Psychology, 1972, no. 9, pp. 418–455.
12. *Wickelgren W. A.* Single-trace fragility theory of memory dynamics. Memory Cognition, 1974, vol. 2(4), pp. 775–780.
13. <http://www.ideationizing.com/2009/06/brief-history-of-mathematical.html> — A Brief History of the Mathematical Definition of Forgetting Curves, 2023.
14. *Efimov A. N.* Informatsiya: tsennost', starenie, rasseyaniye [Information: Price, Consenescence, Diffusion]. Moscow, Znanie, 1978. 64 p.
15. GOST R 53894-2016. Menedzhment znaniy. Terminy i opredeleniya [Knowledge Management. Terms and Definitions]. Moscow, Standartinform, 2020. 24 p. (in Russian).
16. GOST R ISO/MEK 27002. Informatsionnaya tekhnologiya. Metody i sredstva obespecheniya bezopasnosti. Svod norm i pravil menedzhmenta informatsionnoy bezopasnosti [Information Technology. Methods and Means of Ensuring Security. Information Security Management Norms and Rules]. Moscow, Standartinform, 2021. 74 p. (in Russian)
17. Building an Information Technology Security Awareness and Training Program. NIST Special Publication 800-50, 2003.
18. *Astakhova L. V. and Ul'yanov N. L.* Model' politiki upravleniya osvedomlennost'yu sotrudnikov organizatsii v oblasti informatsionnoy bezopasnosti [Model of control policy of awareness of the organization staff in the field of information security]. Proc. 67th Conf. SUSU, Chelyabinsk, April 14–17, 2015, p. 678–682. (in Russian)
19. *Grusho A. A., Primenko E. A., and Timonina E. E.* Teoreticheskie osnovy komp'yuternoy bezopasnosti [Theoretical foundations of computer security]. Moscow, Akademiya Publ., 2009. 272 p. (in Russian)
20. *Solnyshkina S. I. and Kisel'nikov A. S.* Slozhnost' teksta: etapy izucheniya v otechestvennom prikladnom yazykoznanii [Text difficulty: learning stages in practical homeland glossology]. Vestnik TSU. Filologiya, 2015, no. 6(38), pp. 86–99. (in Russian)

21. *Chmykhova E. V., Davydov D. G., and Lavrova T. P.* Eksperimental'noe issledovanie faktorov skorosti chteniya [Reading speed factors experimental study]. *Psikhologiya Obucheniya*, 2014, no. 9, pp. 26–36. (in Russian)
22. *Buymov A. G.* Zakonomernosti povedeniya krivykh zabyvaniya [Forgetting curves action patterns]. *Doklady TUSURa*, 2017, vol. 20, no. 4, pp. 138–141. (in Russian)
23. How Much Do People Forget? <http://www.work-learning.com/catalog.html>, 2023.
24. *Shardakov M. N.* Usvoenie i sokhranenie v obuchenii [Adoption and Saving in Learning]. *Scientific Notes of the Leningrad State Pedagogical University named after Herzen*, 1940, vol. 36. 195 s. (in Russian)
25. *Gaydamakin N. A.* A model of thematic differentiation of access to information for the hierarchical classifier in automatic control systems. *Autom. Remote Control*, 2003, vol. 64, no. 3, pp. 505–516.
26. *Gaydamakin N. A.* Mnogourovnevoe tematiko-ierarkhicheskoe upravlenie dostupom (MLTHS-sistema) [Multilevel thematic-hierarchical access control (MLTHS-system)]. *Prikladnaya Diskretnaya Matematika*, 2018, no. 39, pp. 42–57. (in Russian)
27. *Gaydamakin N. A. and Baranskiy V. A.* Algebra mul'tirubrik na kornevykh derev'yakh ierarkhicheskikh tematicheskikh klassifikatorov [Algebra of multirubric on root trees of hierarchical thematic classifiers]. *Siberian Electronic Math. Reports*, 2017, vol. 14, pp. 1030–1040. (in Russian)
28. Nechetkie mnozhestva v modelyakh upravleniya i iskusstvennogo intellekta [Illegible sets in artificial intelligence handling models]. D. A. Pospelov (ed.), Moscow, Nauka, 1986. 312 p. (in Russian)
29. *Piegat A.* Fuzzy Modeling and Control. Berlin; Heidelberg, Springer Verlag, 2001. 728 p.