Бахчевников Валентин Владимирович – Южный федеральный университет; e-mail: bahchevnikov@sfedu.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: +79518289271; кафедра радиотехнических и телекоммуникационных систем; ассистент.

Деркачев Владимир Александрович – e-mail: vderkachev@sfedu.ru; 347922, г. Таганрог, ул. Шевченко, 2; тел.: +79614154733; Научно-конструкторское бюро цифровой обработки сигналов; конструктор.

Бакуменко Алексей Николаевич – e-mail: baku@sfedu.ru; 347900, г. Таганрог, ул. Петровская, 81; тел,: +79886031853; Инжиниринговый центр приборостроения радио и микроэлектроники; инженер.

Bakhchevnikov Valentin Vladimirovich – Southern Federal University; e-mail: bahchevnikov@sfedu.ru; 44, Nekrasovsky, Taganrog, 347928**,** Russia; phone: +79518289271; the department of radio engineering & telecommunication systems; associate professor.

Derkachev Vladimir Aleksandrovich – e-mail: vderkachev@sfedu.ru; 347922, 2, Shevchenko street, Taganrog, Russia; phone: +79614154733; Research and Design Bureau of Digital Signal Processing; constructor.

Bakumenko Alexey Nikolaevich – e-mail: baku@sfedu.ru; 81, Petrovskaya street, Taganrog, 347900, Russia; phone: +79886031853; Engineering Center of Instrument Making, Radio- and Microelectronics; engineer.

УДК 004.89 DOI 10.18522/2311-3103-2020-3-156-172

В.В. Бова, Д.Ю. Запорожец, Ю.А. Кравченко, Э.В. Кулиев, В.В. Курейчик, Н.А. Лызь

ИДЕНТИФИКАЦИЯ НЕЯВНЫХ УГРОЗ НА ОСНОВЕ АНАЛИЗА АКТИВНОСТИ ПОЛЬЗОВАТЕЛЯ В ИНТЕРНЕТ-ПРОСТРАНСТВЕ[*]

Статья посвящена проблеме идентификации неявных информационных угроз поисковой деятельности пользователя в Интернет-пространстве на основе анализа его активности в процессе данного взаимодействия. Применение знаний, хранящихся в интернет-пространстве, для реализации преступных намерений несет в себе угрозу для всего общества. Выявление злого умысла в действиях пользователей глобальной информационной сети не всегда является тривиальной задачей. Отработанные технологии анализа контекста интересов пользователя дают сбой в случае осторожных грамотных действий злоумышленников, которые в явном виде не демонстрируют преследуемой ими цели. В работе проведен анализ угроз, связанных с определенными сценариями реализации поисковых процедур, проявляющихся в поисковой деятельности. Описаны критерии оценки неэффективных и эффективного сценариев поиска. Среди признаков, указывающих на возможность наличия угрозы, выделены следующие основные: уход от решения задачи в бесцельную навигацию или к привлекательным ресурсам, поверхностный поиск, отсутствие смыслового погружения в решение поисковой задачи, хаотичные действия при поиске. Для определения наличия неблагоприятных признаков построена система показателей. Сформулированы признаки эффективного сценария организации поиска в Интернет-пространстве, описаны варианты наличия неявных угроз для подобной ситуации. Представлен подход идентификации описанных угроз с учетом заданных критериев оценки различных сценариев поведения пользователя в глобальном информационном пространстве. Разработан алгоритм машинного обучения для идентификации проблемных сценариев путем сравнения с ключевыми паттернами поведения. Создана программная реализация подсистемы идентификации информационных угроз, проведены экспериментальные исследования для подтвержде-

[*] Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 18-29-22019.

ния эффективности подсистемы. Экспериментальные исследования проводились на основе обработки открытых данных из социальных сетей, а также с применением анализа поисковой деятельности пользователей в университетской корпоративной информационной среде.

Информационный поиск; неявные угрозы; анализ активности пользователя; вектор признаков; методы машинного обучения; оптимизация; интеллектуальные системы.

V.V. Bova, D.Yu. Zaporozhets, Yu.A. Kravchenko, E.V. Kuliev, V.V. Kureyichik, N.A. Lyz

IMPLICIT THREATS IDENTIFICATION BASED ON ANALYSIS OF USER ACTIVITY ON THE INTERNET SPACE

The article is devoted to the problem of identifying implicit information threats of a user's search activity in the internet space based on an analysis of his activity in the course of this interaction. The use of knowledge stored in the Internet space for the implementation of criminal intentions poses a threat to the whole society. Identifying malicious intent in the users' actions of the global information network is not always a trivial task. The proven technologies for analyzing the context of user interests fail in the case of cautious and competent actions of attackers who do not explicitly demonstrate the goal they are pursuing. The paper analyzes the threats associated with certain scenarios for the implementation of search procedures that manifest themselves in search activities. Criteria of inefficient and effective search scenarios estimation are described. Among the signs indicating the possibility of a threat, the following main ones are highlighted: avoiding solving the problem in aimless navigation or attractive resources, superficial search, lack of meaningful immersion in solving the search problem, and chaotic actions during the search. To determine the presence of adverse signs, a system of indicators is built. The features of an effective scenario for organizing a search in the Internet space are formulated, options for the presence of implicit threats for a similar situation are described.An approach for identification the described threats is presented taking into account the specified criteria for evaluating various scenarios of user behavior in the global information space. A machine learning algorithm has been developed to identify problem scenarios by comparing with key behavioral patterns. The software implementation of the subsystem for identifying information threats has been created, experimental studies have been conducted to confirm the effectiveness of the subsystem. Experimental studies were carried out on the basis of processing open data from social networks, as well as using analysis of user search activity in the university corporate information environment.

Information search; implicit threats; analysis of user activity; feature vector; machine learning methods; optimization; intelligent systems.

Введение. Онлайн-обучение как восходящий тренд в развитии образовательных систем требует моделирования и разработки способов управления деятельностью обучающихся в интернет-пространстве. Широкое понимание онлайн-обучения позволяет говорить о многообразии видов информационно-образовательной деятельности, реализуемых в сети Интернет: учебно-познавательной, поисково-познавательной, коммуникативно-познавательной, информационно-созидательной, развлекательно-познавательной [1]. Одним из наиболее распространенных видов является поисково-познавательная деятельность, связанная с поиском информации в сети Интернет, включающая использование технологий поиска и работы с информацией, анализ и отбор релевантной информации, ее верификацию и применение [1]. Поиск и использование информации очень важны для обучающихся как в формальном, так и в неформальном и информальном образовании. Этот вид деятельности способствует накоплению не только «знаниевого» опыта, но и опыта работы с информацией, освоению технологий поиска и анализа, а также когнитивному и личностному развитию за счет расширения интересов и повышения мотивации познавательной деятельности обучающихся [2].

Однако исследования показывают, что студенты не готовы к эффективной поисковой деятельности, они весьма незначительно используют научные базы данных и электронные версии научных журналов. Так, для более чем 80 % студен-

тов вузов, наиболее важным, надежным и часто используемым источником академической информации являются поисковые системы общего назначения [3]. Большинство студентов не ищут новые подходы, а имеют тенденцию повторять стратегии, которые принесли успех в их предыдущем опыте [4]. Их поисковое поведение предвзято, они предпочитают результаты поиска с более высоким рейтингом, даже если это не вполне соответствует их потребностям [3]. Кроме того, поиск и обработка информации – это сложный процесс, включающий в себя постановку задачи, поиск соответствующих источников информации, извлечение и организацию соответствующей информации из каждого источника, синтез информации из различных источников и т.п. [5–7]. Эта деятельность востребует регулятивные, когнитивные и метакогнитивные способности студентов, которые не всегда развиты на должном уровне, следствием чего могут являться некачественные результаты поиска: плагиат, нерелевантная информация, некорректные выводы [5, 7, 8]. Клиповость мышления современной молодежи повышает вероятность данных следствий [9].

Дополнительные проблемы возникают в связи с тем, что поисковая деятельность имеет «затягивающий» характер и осуществляется не столько в специальных электронно-образовательных средах, сколько в сегментах интернет-пространства, где отсутствует контроль над потоком академической информации. В связи с этим к рискам потери эффективности деятельности добавляются кибер-риски, коммуникационные, контентные и другие риски, включая интернет-зависимость. Работая с информацией в Интернете, студенты могут испытывать дисфункциональные состояния, близкие к депрессии, связанные с ситуацией неопределенности и повышенной тревожностью [8]. У многих возникает «феномен потери в гиперпространстве», связанный с переживанием дезориентации из-за информационной перегрузки и бесцельного следования гиперссылкам [10]. Доказано, что такие интернет-активности, как поиск информации и общий серфинг, связаны с более высокими показателями проблематичного использования Интернета (интернет-зависимости) [11]. Веб-серфинг считается самым распространенным видом зависимости, в то же время он является и самым малоизученным [12]. Проявление и формирование склонности к интернет-зависимости может происходить в поисковой деятельности, когда она переходит в навязчивый веб-серфинг – просмотр многообразных интернет-ресурсов, хаотичную гипертекстовую навигацию, длительно и бесцельно осуществляемую субъектом.

Таким образом, важная роль поисково-познавательной деятельности в образовании, с одной стороны, и высокие риски для эффективности деятельности и психологического благополучия, с другой, требуют разработки как способов подготовки обучающихся, так и методов и инструментов внешнего сопровождения такой деятельности, в т.ч. интеллектуальных цифровых ассистентов. Решение второй задачи предполагает моделирование поисковой деятельности и угроз, возникающих в ней.

1. Описание проблемы и выбор критериев оценки угроз. Существуют различные по направленности исследования, в которых представлено моделирование поисковой деятельности человека в интернет-пространстве. Наиболее ранние психологические модели Б. Дэрвина, Д. Эллиса, К. Кухлтау, описывают только действия по поиску информации; расширенная модель Т. Уилсона включает пользователя информации, пользовательские потребности, запрос информации, обмен, использование и удовлетворенность (неудовлетворенность) пользователя результатами поискового запроса [13]. К настоящему времени создано множество поэтапных моделей поиска, при этом все многообразие выделяемых этапов можно свести к трем стадиям: начало и цепочка, мониторинг и дифференцирование, оценка и

извлечение [14]. Для изучения деятельности по поиску информации также используются модели поисковых стратегий, где акцент делается не столько на действиях, сколько на субъекте, ищущем информацию. Компонентами стратегий являются: планирование поиска, выбор ресурсов, реализация, извлечение знаний, способы преодоления трудностей, оценка, контроль и др. [5–7, 15, 16].

Поисковая деятельность характеризуется не только используемой стратегией, но и конкретным сценарием поведения, в т.ч. «механикой» использования поисковых систем. С этих позиций выделены различные способы поиска: сопоставление (соответствие) или исследование [6]; логический (булевый), наилучшего соответствия или комбинированный [16]; описана типовая модель поведения (тактика) пользователей при тематическом поиске информации [17]; построены формализованные вероятностные модели, обучающиеся из данных о пользовательских действиях (click logs) [18]; разработаны рекомендательные системы, поддерживающие направленный поиск в образовательных контекстах [19, 20].

Для документирования информации используются: метод навигационной карты (navigation flow map, NFM), который графически отображает многослойные взаимосвязи между веб-навигацией и поиском информации [6].

Как правило, моделируя действия пользователей, исследователи ставят задачи повышения интеллектуальности поисковых систем с целью удовлетворения потребностей и уменьшения информационной нагрузки на пользователя [17]. Ими не рассматривается контекст поисковой деятельности, несущие угрозу информационные события, возможности отвлечения и ухода пользователей в бесцельный серфинг. Без решения остаются важнейшие проблемы проявления и формирования склонности к интернет-зависимости в поисковой деятельности и другие вопросы обеспечения безопасности поискового поведения.

С позиции обеспечения безопасности поискового поведения обучающегося в части превенции веб-серфинга как интернет-зависимости можно выделить три угрозы и связанные с ними сценарии поведения: 1) уход от решения задачи в бесцельную навигацию или к привлекательным ресурсам (сценарий «*отдаление*»); 2) поверхностный поиск, отсутствие смыслового погружения в решение поисковой задачи (сценарий «*сканирование*»); 3) хаотичные действия при поиске (*сценарий «спонтанность»*).

Представим критерии для идентификации признаков наличия или отсутствия описанных угроз.

1. Для оценки признаков наличия сценария «Отдаление» выделим следующие критерии:

1.1) y_{11} – стабильность ключевых слов на страницах (релевантные слова);

1.2) y_{12} – переход по большому количеству гиперссылок (большое количество посещенных страниц);

1.3) y_{13} – постепенное увеличение времени пребывания на нерелевантных страницах;

1.4) y_{14} – длинные прямые цепочки без возвратов;

1.5) y_{15} – переход на рекламные баннеры;

1.6) y_{16} – переход от первой–второй к другим страницам поисковых результатов;

1.7) y_{17} – удержание «линии поиска».

2. Для оценки признаков наличия сценария «Сканирование» выделим следующие критерии:

2.1) y_{21} – большое количество просмотренных страниц с малым временем на просмотр отдельных страниц;

2.2) y_{22} – предпочтение простых и лаконичных источников;

2.3) y_{23} – использование легко доступных источников;

2.4) y_{24} – перепроверка найденной информации (возврат к посещенным страницам);

2.5) y_{25} – дочитывание страниц до конца, долгое нахождение на странице, существенная разница между временем просмотра отдельных страниц

2.6) $y_{26} = y_{16}$ – переход от первой–второй к другим страницам поисковых результатов;

3. Для оценки признаков наличия сценария «Спонтанность» выделим следующие критерии:

3.1) $\boldsymbol{y_{31}}$ – некорректная формулировка запроса;

3.2) $\boldsymbol{y_{32}}$ – использование результатов поиска, находящихся в первых строках списка;

3.3) $\boldsymbol{y_{33}}$ – частая смена запроса или отсутствие смены запроса;

3.4) $\boldsymbol{y_{34}}$ – фильтрация информации на ранней стадии цикла поиска

3.5) $\boldsymbol{y_{35}}$ – использование наиболее релевантных источников.

Противоположностью описанным нежелательным ситуациям в создаваемой модели является сценарий «Эффективность», который на начальном этапе оценки стал идентификатором отсутствия угроз. Для оценки признаков наличия сценария «Эффективность» выделим следующие критерии:

1) $\boldsymbol{x_1}$ – просмотр релевантных страниц;

2) $\boldsymbol{x_2}$ – существенное различие между временем нахождения на страницах;

3) $\boldsymbol{x_3}$ – присутствие качественного результата поиска.

Показателями, определяющими наличие описанных критериев, являются следующие значения:

1. Формулировка запросов (абсолютное количество, относительное количество релевантных).
2. Работа с выданной поисковиком информацией (к каким по счету результатам поиска (строкам) происходит обращение, есть ли переход ко второй и последующим страницам результатов, подряд или выборочно).
3. Цепочки: длина, количество возвратов.
4. Количество просмотренных страниц: всего, релевантных, нерелевантных, относительное количество релевантных.
5. Время просмотра страницы: максимальное, минимальное, общее среднее, среднее по релевантным, среднее по нерелевантным, «разброс» во времени просмотра отдельных страниц (или частотное распределение).
6. На какой минуте поиска появляются нерелевантные страницы.
7. По каждому показателю в соотнесении со средним по выборке: наличие завышенных (превышающих среднее + среднеквадратичное отклонение) или заниженных показателей.

На основе представленного описания проблемы и построенных сценариев наличия или отсутствия угроз разработаем подход идентификации неявных угроз на основе анализа активности пользователя в интернет-пространстве.

2. Подход идентификации неявных угроз на основе анализа активности пользователя в интернет-пространстве. Рассмотрим описанные в предыдущем пункте сценарии «отдаление», «сканирование» и «спонтанность» как слагаемые обобщенного критерия, однозначно указывающего на недопустимую форму организации работы пользователя (обучающегося) в интернет-пространстве, назовем эту форму организации условно «неэффективным сценарием». Опасность здесь представляет случайный выход пользователя на вредоносный контент. Для построения вектора признаков идентификации «неэффективного сценария» рассмотрим в табл. 1 описанные ранее критерии. Наличие признака обозначим *«1»*, отсутствие – *«0»*.

Таблица 1

Построение вектора признаков для угрозы «неэффективный сценарий»

	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{21}	y_{22}	y_{23}	y_{24}	y_{25}	y_{31}	y_{32}	y_{33}	y_{34}	y_{35}
Отдаление	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Сканирование	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0
Спонтанность	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0
	Общие векторы признаков																
Max угроза	**0**	**1**	**1**	**1**	**1**	**0**	**0**	**1**	**1**	**1**	**0**	**0**	**1**	**1**	**1**	**0**	**0**
Min угроза	**1**	**0**	**0**	**0**	**0**	**1**	**1**	**0**	**0**	**0**	**1**	**1**	**0**	**0**	**0**	**1**	**1**

В таком виде вектор признаков для идентификации возможной угрозы нежелательной формы организации активности пользователя в интернет-пространстве является суммой векторов наборов критериев трех рассмотренных неблагоприятных сценариев (рис. 1)

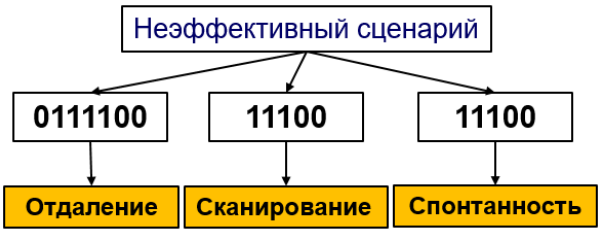

Рис. 1. Слагаемые вектора признаков «неэффективного сценария»

Обозначим сценарии «отдаление», «сканирование» и «спонтанность» через $\boldsymbol{Q_1, Q_2, Q_3}$ соответственно. Таким образом, обобщенный критерий имеет следующий вид:

$$\boldsymbol{Q_{int} = \tau_1 Q_1 + \tau_2 Q_2 + \tau_3 Q_3 \rightarrow max,} \quad (1)$$

где $\boldsymbol{\tau_i}$ – вес каждого из рассмотренных сценариев, заданный на основе экспертных оценок.

Рассмотрим теперь возможный «эффективный сценарий», как форму организации активности пользователя (обучающегося) в интернет-пространстве. За основу данной формы организации активности взят описанный в предыдущем пункте сценарий «эффективность», основными критериями которого приняты: $\boldsymbol{x_1}$ – просмотр релевантных страниц; $\boldsymbol{x_2}$ – существенное различие между временем нахождения на страницах; $\boldsymbol{x_3}$ – присутствие качественного результата поиска.

Будем утверждать, что даже в случае максимально эффективного сценария активности пользователя в интернет-пространстве, чему соответствует расширенный вектор признаков (табл. 2), остается возможность появления неявных угроз системного характера, имеющих значительный уровень и масштаб опасности для всего общества.

Таблица 2

Максимально эффективный сценарий активности пользователя в Интернет-пространстве

x_1	x_2	x_3	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{21}	y_{22}	y_{23}	y_{24}	y_{25}	y_{31}	y_{32}	y_{33}	y_{34}	y_{35}
1	**1**	**1**	**1**	**0**	**0**	**0**	**0**	**1**	**1**	**0**	**0**	**0**	**1**	**1**	**0**	**0**	**0**	**1**	**1**

Применение знаний, хранящихся в интернет-пространстве, для реализации преступных намерений несет в себе значительную угрозу. Выявление злого умысла в действиях пользователей глобальной информационной сети не всегда является тривиальной задачей. Отработанные технологии анализа контекста интересов пользователя дают сбой в случае осторожных грамотных действий злоумышленников, которые в явном виде не демонстрируют преследуемой ими цели.

Например, миллионы людей интересуются историей огнестрельного оружия, но только несколько из них могут интересоваться этой темой в преступных целях. Именно анализ семантики поисковой деятельности данных субъектов позволит выявить и классифицировать детали их интересов, которые дадут основания дифференцировать этих лиц от остального множества людей, интересующихся данной темой без злого умысла. Также возможен вариант присутствия пока еще несформировавшейся преступной личности, информационные интересы которой в будущем могут привести к печальным последствиям потери обществом законопослушного гражданина.

Оба эти случая простой анализ контекста интересов поиска информации может вовремя не идентифицировать, так как будут отсутствовать явные признаки преступных намерений. Единственным отличительным системно значимым признаком таких видов деятельности в среде Интернет будет наличие сценария высокой эффективности действий пользователя. Данные субъекты будут отличаться стабильной тематической привязкой (релевантностью) изучаемых материалов, длительным временем нахождения на определенных страницах поиска, что позволит им получить качественный результат проанализированного материала.

Этот признак, конечно, не указывает на преступные намерения или нежелательную траекторию интересов, но критерием угрозы в данном случае будет являться время, которое данные субъекты уделяют изучению тем, имеющих, с одной стороны, признаки двойного назначения, а, с другой – напрямую не связаны с профессиональной или образовательной деятельностью субъекта. Таким образом, авторы предлагают ввести в рассмотрение еще два следующих критерия: $\boldsymbol{x_4}$ – изучение контента с информацией двойного назначения; $\boldsymbol{x_5}$ – изучение контента с информацией, не имеющей прямого отношения к профессиональной или образовательной деятельности субъекта (табл. 3).

Таблица 3

Идентификация угроз при эффективном сценарии

№		x_1	x_2	x_3	x_4	x_5
1	Отсутствие угроз	**1**	**1**	**1**	**0**	**0**
2		**1**	**1**	**1**	**0**	**1**
3		**1**	**1**	**1**	**1**	**0**
4	Угроза	**1**	**1**	**1**	**1**	**1**

Первые три варианта этого сегмента вектора признаков имеют незначительный вес при идентификации возможных угроз, так как в первом варианте дополнительные признаки x_4 и x_5 вообще отсутствуют, во втором варианте нет сведений об изучении информации двойного назначения, а в третьем – интерес может быть связан с непосредственной профессиональной или образовательной деятельностью субъекта.

Наибольшее опасение должны вызвать максимальные значения критериев-признаков $\boldsymbol{x_1,x_2,x_3,x_4,x_5}$. Максимальные значения данного критерия указывают на

одержимость и должны стать причиной более детального рассмотрения перспектив применения полученной субъектом информации в результате проведенного им поиска. В дальнейшем решение задачи классификации указанной информации, проведенной на основе методов оценки семантической близости (рис. 2), позволит исключить наличие угроз, например, в случае если речь идет о хобби человека, либо подтвердить опасения, что потребует определенного вмешательства.

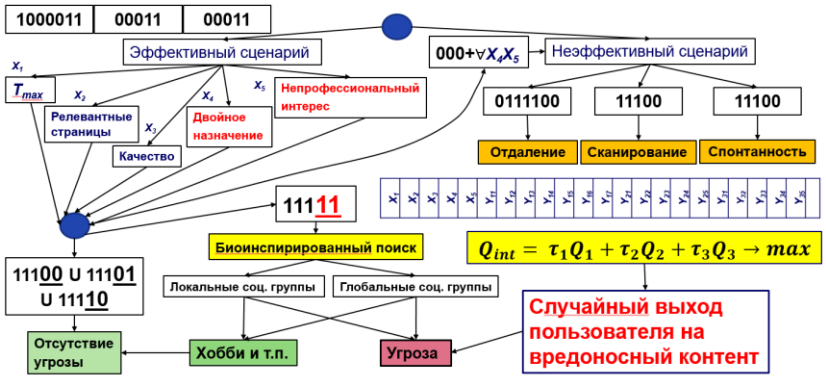

Рис. 2. Схема подхода идентификации неявных угроз на основе анализа активности пользователя

Идентификацию угроз в описанных сценариях предлагается проводить на основе комбинированных биологически правдоподобных методов машинного обучения. На первом этапе поиска необходимо исследовать локальные и глобальные социальные группы (рис. 2). Такой поиск эффективно выполняется на основе роевых методов с децентрализованной схемой управления, позволяющих провести максимальное количество локальных сравнений векторов признаков с достаточным количеством – глобальных (например, CS), на втором уровне для поиска в информационных пространствах, отобранных по описанным сценариям и результатам простого контекстного поиска субъектов, используются роевые методы с наиболее последовательными механизмами поиска, необходимыми для подтверждения или опровержения наличия угроз в рассматриваемом контенте (например, BFO).

3. Разработка биоинспирированного алгоритма идентификации неявных угроз. Метод расширенного интеллекта (AUI) в контексте интеллектуального анализа использует концепцию, в которой традиционный искусственный интеллект (ИИ) поддерживает принятие решений, обучение или планирование. Вместо группы из двух или более учащихся, есть конечный пользователь и машина, которые совместно строят решение для конкретной проблемы контекста, общаясь через среду машинного обучения. Роль конечного пользователя заключается в предоставлении набора данных для алгоритма машинного обучения, который строит модель из набора данных. Роль конечного пользователя активна, и целью является не просто эксперимент с такими параметрами, как количество кластеров или скорость обучения, которые влияют на то, как данный алгоритм строит модель. Вместо этого цель состоит в том, чтобы настроить саму модель.

Основываясь на информационных изменениях конечного пользователя, компьютер перестраивает новую модель, используя полученные знания и данные, которые он приобрел. Процесс повторяется итерационно, и на каждом этапе цикла вырабатываются новые знания о контексте и изучаемом явлении.

Процесс AUI приводит к циклическому процессу обнаружения знаний как конечного пользователя, так и компьютера. Вместе эти результаты углубляют предыдущие знания конечного пользователя о контексте, в котором были собраны наборы данных.

Чтобы расширить область применения метода AUI из легко интерпретируемых деревьев решений, необходимо также разработать новые вычислительные методы для обслуживания различных задач интеллектуального анализа данных в современных интеллектуальных средах обучения. Алгоритм нейронные ***N-Tree*** для образовательного кластерного анализа представляет собой сбалансированное двоичное дерево, каждый узел которого содержит точечный вектор. Длина точечного вектора равна длине входных векторов. Нейронное ***N***-дерево строится рекурсивно путем инициализации каждого точечного вектора случайными числами в диапазоне от минимального значения до максимального значения, встречающегося в анализируемом наборе данных. Сначала алгоритм берет число кластеров ***n*** и создает вектор, длина которого $n \times 2 - 1$ со случайными точечными векторами. Во-вторых, из созданного вектора строится сбалансированное двоичное дерево.

После построения нейронного ***N***-дерева со случайными точечными векторами каждый конечный узел индексируется путем обхода нейронного ***N***-дерева с обходом по порядку и назначения терминальных узлов с другим индексом. Обучение модифицированного алгоритма делится на три фазы. На первом этапе обучения нейронного ***N-Tree*** производится на основе предложенного авторами биоинспирированного алгоритма.

Для анализа пользовательской активности в сети Интернет был разработан биоинспирированный алгоритм сбора и анализа данных о посещённых пользователями информационных ресурсах. В качестве биоинспирированного алгоритма авторами предложен модифицированный алгоритм капли воды. Опишем более подробно предложенный алгоритм на примере поведенческой модели в живой природе.

В данном алгоритме моделируется несколько искусственных капель воды, которые зависят друг от друга, способны менять свое окружение таким образом, что находят оптимальный путь по пути наименьшего сопротивления. Данный алгоритм является конструктивным популяционно-ориентированным алгоритмом оптимизации.

Для каждой капли воды ***k*** применимы следующие показатели: количество грунта $soil_k$; скорость передвижения vel_k. Среда нахождения капли воды дискретна. Каждая ***k*** представляет собой граф ***(N, U)*** с набором узлов ***N*** и множеством ребер ***U***. Данный граф является средой для капель воды и их потока по ребрам графа. Каждая капля воды начинает строить свое решение постепенно, передвигаясь между узлами графа до тех пор, пока не завершится ее решение. Каждая итерация завершается, когда все капли воды завершили свой проход по ребрам графа. После каждой итерации вычисляется лучшее решение ***F*** в текущей итерации. После определения лучших решений на каждой итерации ***F***, определяется набор лучших решений ***Z*** с начала работы алгоритма капли воды. В процессе перемещения капли собираются в поток (рис. 3).

Каждая капля воды принадлежит узлу ***i*** и должна из всех узлов ***N*** достичь конечного узла ***j***. Вероятность попадания в конечный узел ***j*** вычисляется следующим образом [21]:

$$P_i^k(j) = \frac{f(soil(i,j))}{\sum_{l \notin V_{visited}^k} f(soil(i,l))}, \quad (2)$$

где

$$f\left(soil(i.j)\right) = \frac{1}{\varepsilon + g(soil(i,j))}. \quad (3)$$

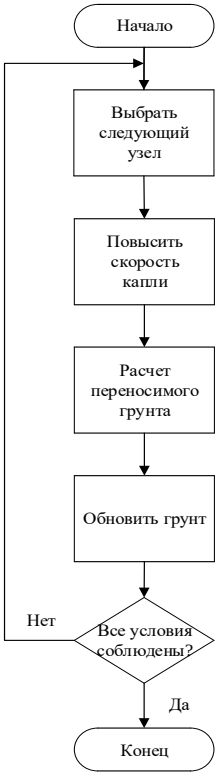

Рис. 3. Процесс обработки решения

Числовая постоянная $\boldsymbol{\varepsilon \geq 0}$. Данное ограничение вводится для исключения деления на $\boldsymbol{0}$. Функция $\boldsymbol{g(soil(i,j))}$ используется для обозначения наличия грунта $\boldsymbol{(i, j)}$, которая обозначает весь грунт на пути между узлами $\boldsymbol{i}$ и $\boldsymbol{j}$:

$$\boldsymbol{g\big(soil(i,j)\big) = \begin{cases} soil(i,j) if min_{\forall l \notin V^k_{visited}}(soil(i,l)) \geq 0 \\ soil(i,j) - min_{\forall l \notin V^k_{visited}}(soil(i,l) otherwise \end{cases}}. \quad (4)$$

По формуле 5 вычислим новую скорость:

$$\boldsymbol{vel^k(t+1) = vel^k(t) + \frac{a_v}{b_v + c_v soil(i,j)}}, \quad (5)$$

где $\boldsymbol{a_v}$, $\boldsymbol{b_v}$ и $\boldsymbol{c_v}$ являются наименьшими положительными значениями, чтобы избежать деления на нуль, а $\boldsymbol{soil(i,j)}$ определяет количество грунта от узла $\boldsymbol{i}$ к узлу $\boldsymbol{j}$.

Отметим основные принципы поведения капли воды [22]:

- предпочитает путь с меньшим количеством почвы, чем пути с большим количеством почты;
- предпочитает более легкий путь, когда приходится выбирать между несколькими маршрутами, которые существуют на пути от источника к месту назначения;
- легкость или твердость пути определяется количеством почвы на этом пути. Путь с большим уровнем почвы считается трудным путем, тогда как путь с меньшим уровнем почвы считается легким путем.

В данной работе предлагается увеличение скорости расчетов по формуле 4, где показатель $\boldsymbol{soil}$ экспоненциален $\boldsymbol{2}$ [21].

1: Input: Постановка задачи.
2: Output: Оптимальное решение.
3: Формулировка проблемы оптимизации как связного графа.
4: Инициализация постоянных параметров.
5: Repeat
6: Инициализация динамических параметров.
7: Распределите капли воды в случайном порядке по построенному графу.
8: Обновление списка посещенных узлов ($V_{visited}^{k}$), включая исходный узел
9: Repeat
10: For k = 1 to i do
11:i = исходный узел для капли k.
12:j = следующий выбранный узел, который не противоречит ограничениям.
13:Движение капли k от узла i к узлу j.
14:Обновление следующих параметров:
a) Скорость капли k.
b) Количество грунта в капле k.
c) Количество грунта на выходе e_{ij}.
15:End for
16:Until Не удовлетворены условия прекращения.
17:Выбор лучшего решения в этом цикле (T^{IB}).
18:Обновление значения содержания грунта на выходе для всех показателей (T^{IB}).
19:Обновление лучшего решения среди всех циклов (T^{TB}).
20:If (качество T^{TB}< качестваT^{IB}).
21:$T^{TB} = T^{IB}$.
22: Until не найдено решение, удовлетворяющее всем требованиям.
23: Return (T^{TB}).

На втором этапе обучения нейронного ***N-Tree*** происходит поиск наилучшего совпадающего блока осуществляется из терминальных узлов путем сравнения вектора произвольной выборки с каждым терминальным узлом.

На третьем этапе обучение начинается с корневого узла, и следующий узел входного вектора является либо левым, либо правым дочерним элементом текущего узла, в зависимости от расстояния входного вектора и точек. Правого и левого потомка (выбран узел с наименьшим). Таким образом, входной вектор проходит через нейронное ***N***-дерево от корневого узла к одному из терминальных узлов. После каждого шага вычисляется порядок уровней, начиная с текущего узла (кроме корневого узла), и обновляются точки поддерева порядка уровней, начиная с текущего узла. Целью процесса прямого обучения является корректировка путей для входных векторов для процесса кластерного анализа.

4. Программное приложение и экспериментальные исследования. С целью апробации разработанных методов и алгоритмов анализа пользовательской активности в сети Интернет была разработана подсистема сбора и анализа данных о посещённых пользователей информационных ресурсов. Подсистема включает в себя два компонента: клиентский модуль, реализованный в виде расширения для браузера и серверного модуля, который реализует функции аутентификации пользователей, сбора информации, переданной с клиентского модуля, первичной обработки поступившей информации, ее хранения, а также идентификации угроз в случае эффективного сценария. На данном этапе реализации системы любой сценарий считается эффективным.

Функции клиентского модуля достаточно просты и заключаются исключительно в сборе информации и метаданных каждой загруженной пользователем страницы и передачи данной информации на серверный модуль.

Серверный модуль реализует более сложные функции. Работа модуля заключается в приеме информации от клиентского модуля, разбор контента страницы на термы, фильтрация через табу-лист, нормализация термов, подсчет повторений данного терма на странице, сохранение результатов в базу данных. На данном этапе реализации системы позволяет работать только с существительными. Табу-лист представляет собой список термов, которые потенциально не несут никакой смысловой нагрузки, но часто встречаются в контенте страниц. Исследования показали, что, например, как минимум один их термов, таких как: «комментарий», «предпросмотр», «запрос», входят в 10 наиболее часто встречающихся термов для 100% пользователей. Это затрудняет процесс распознавания угроз при расчетах. Поэтому использование фильтра по табу-списку на ранних этапах обработки входящей информации позволяет сократить пространство будущего поиска, сократить ресурсы на хранение информации и повысить точность определения угрозы.

В качестве эталона для определения угроз был использован «Список товаров и технологий двойного назначения, которые могут быть использованы при создании вооружений и военной техники и в отношении которых осуществляется экспортный контроль» (далее «целевой список»). С помощью разработанной подсистемы из данного целевого списка были выделены термы. В ходе исследований было также определено, что термы из целевого списка часто используются со «словами сателлитами», которые не входят в целевой список, но могут также влиять на оценку угрозы. Поэтому было предложено разбить термы из целевого списка на категории, в соответствии с категориями целевого списка (перечислить) и включить в каждую категорию термы вместе с их сателлитами. Стоит отметить, что терм может быть использован в одной и более категориях. Начальное разбиение на категории выполнялось автоматически, на основе данных их из целевого списка. Дальнейшее наполнение категорий термами-сателлитами реализовывалось в виде опроса экспертов. Экспертам задавался вопрос: «К какой из представленных категории можно отнести следующий терм?». В качестве терма в общем случае выбирается произвольный терм из базы данных системы. Однако такой подход показал свою неэффективность из-за большого количества термов. Поэтому список термов, представляемых экспертам, задается заранее администратором системы в зависимости от текущих потребностей. Администратор имеет возможность изменять список термов для экспертизы в любое время.

Для расчета угрозы, исходящей от пользователя в системе применяются, методы AUI, в частности нейросетевой подход. Для заданного пользователя на входы нейронной сети подаются значения рисков каждой категории системы $\boldsymbol{R_k}, \boldsymbol{R_k} \epsilon [\boldsymbol{0}, \boldsymbol{1}], \boldsymbol{k} \in [\boldsymbol{1}, \boldsymbol{K}]$, где $\boldsymbol{K}$ – общее количество категорий.

Для расчета рисков каждой категории была применена следующая формула:

$$\boldsymbol{R_k} = \frac{\sum_{x=1}^{|X^k|} (arctg(x * \alpha(t_e - t_b)) * \pi * \beta)}{|X^k|}, \quad (6)$$

где $\boldsymbol{X^k}$ – множество термов, принадлежащих категории $\boldsymbol{k}$; $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ – поправочные коэффициенты. Эмпирически были выявлены следующие значения: $\boldsymbol{\alpha = 0.05}$; $\boldsymbol{\beta = 0.2}$.

Для начального обучения нейронной сети были сгенерированы наборы данных, имитирующие пользовательскую активность с заданными характеристиками. Для обучения сети, работающей с ***10*** категориями в общей сложности, было сгенерировано ***50000*** примеров для обучающей выборки и ***50000*** примеров для тестовой выборки. После обучения результаты тестов показали, что среднее отклонение значений результатов сети от эталонных результатов составило ***0.651%***. На рисунке 4 графически представлено распределение ошибок. Данное отклонение является приемлемым, так как ответ нейронной сети интерпретируется как бинарный сиг-

нал присутствия или отсутствия угрозы. Для подтверждения необходимости и достаточности количества обучающих примеров были проведены следующие обучающие и тестовые серии (табл. 4).

Таблица 4

Обучающие и тестовые выборки

Размер обучающей выборки (тыс.)	Размер тестовой выборки (тыс.)	Средняя ошибка (%)
20	50	21.38
40	50	2.19
50	50	0.66
60	50	0.65
80	50	0.64
100	50	0.63

Результаты экспериментальных серий приведены на рисунке 5. На данном рисунке можно видеть, что после ***50000*** обучающих примеров точность сети повышается незначительно, тогда так временные затраты на обучение имеют квадратичный характер.

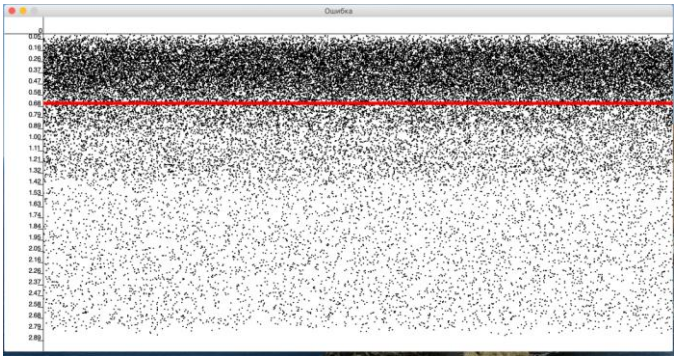

Рис. 4. Распределение ошибок

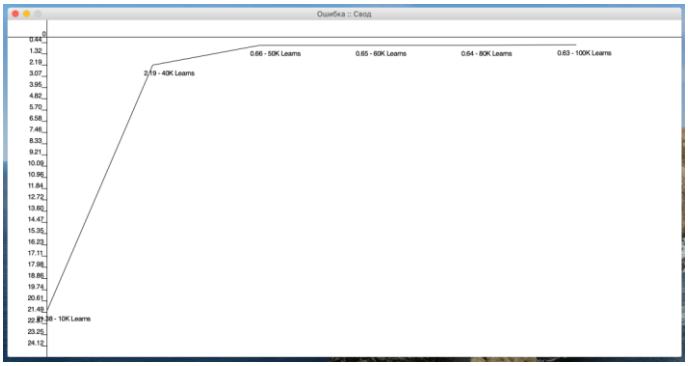

Рис. 5. Результаты тестирования обученной нейронной сети

Предложенный подход позволяет анализировать в реальном времени активность пользователя в сети Интернет и оценивать риски его поведения в контексте анализируемой проблемы на основе внедрения заранее обученной нейронной сети. За счет того, что нейронная сеть работает с категориями, а не с самими термами не

требуется переобучения нейронной сети при изменении состава категорий за счет обновления перераспределении экспертных ответов по назначению термов и их сателлитов. Недостатком предложенного подхода является невозможность расширения количество категорий без переобучения сети, а также невозможности получения объяснений результатов о наличии или отсутствии рисков. На данном этапе разработки все случаи, квалифицированные как опасные (с высоким уровнем риска) требуют анализа человеком-экспертом.

Заключение. Проведенные в представленной работе исследования направлены на решение проблемы идентификации неявных информационных угроз поисковой деятельности пользователя в Интернет-пространстве на основе анализа его активности в процессе данного взаимодействия. Полученные в ходе работы результаты направлены на повышение эффективности интеллектуальных систем-ассистентов, обеспечивающих безопасность и эффективность деятельности пользователя в Интернет-пространстве, на основе биологически правдоподобных методов машинного обучения.

При изучении возможных явных и неявных информационных угроз поисковой деятельности авторами были построены векторы признаков, необходимые для проведения процедур классификации сценариев поведения пользователя и следующих за ними событий. Определены критерии оценки неэффективных и эффективного сценариев поиска, разработана система показателей для определения значений данных критериев.

Предложен подход идентификации описанных угроз с учетом заданных критериев оценки различных сценариев поведения пользователя в глобальном информационном пространстве. Разработан модифицированный биоинспирированный алгоритм капли воды для сбора и анализа данных о посещённых пользователями информационных ресурсах. Данный алгоритм является конструктивным популяционно-ориентированным алгоритмом оптимизации.

Для проведения экспериментальных исследований была создана подсистема сбора и анализа данных о посещённых пользователей информационных ресурсов. Для разработки системы использован нейросетевой подход. Паттерном определения угроз использован «Список товаров и технологий двойного назначения, которые могут быть использованы при создании вооружений и военной техники и в отношении которых осуществляется экспортный контроль».

Результаты экспериментальных исследований подтвердили эффективность предложенного подхода идентификации информационных угроз, который позволяет анализировать в реальном времени активность пользователя в сети Интернет и оценивать риски его поведения в контексте анализируемой проблемы на основе внедрения заранее обученной нейронной сети.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Лызь Н.А., Истратова О.Н.* Информационно-образовательная деятельность в интернет-пространстве: виды, факторы, риски // Педагогика. – 2019. – № 4. – С. 16-26.
2. *Раицкая Л.К.* Влияние интернета на личность студента // Коммуникация в современном поликультурном мире: диалог культур: Ежегодный сборник научных трудов / отв. ред. Т.А. Барановская. – М., 2014. – С. 429-441.
3. *Salehi S., Du J. T., Ashman H.* Use of Web search engines and personalisation in information searching for educational purposes // Information Research. – 2018. – Vol. 23, No. 2. – Режим доступа: http://informationr.net/ir/23-2/paper788.html.
4. *Cen Y., Gan L., Bai C.* Reinforcement learning in information searching // Information Research. – 2013. – Vol, 18, Issue 1. – Режим доступа: http://informationr.net/ir/18-1/paper569.html#.Xl99magzaUk.

5. *Горюнова Л.Н., Круглова М.А., Провоторова Я.А., Цыган В.Н.* Стратегии информационного поиска и их взаимосвязь с личностными особенностями студентов // Петербургский психологический журнал. – 2013. – № 2. – С. 1-15.
6. *Lin C.-C., Tsai C.-C.* A navigation flow map method of representing students' searching behaviors and strategies on the web, with relation to searching outcomes // Cyberpsychology and Behavior. – 2007. – Vol. 10, Issue 5. – P. 689-695. – DOI: 10.1089/cpb.2007.9969.
7. *Walraven A., Brand-Gruwel S., Boshuizen H.P.* Information-problem solving: A review of problems students encounter and instructional solutions // Computers in Human Behavior. – 2008. – Vol. 24, Issue 3. – P. 623-648.
8. *Поршнев А.В.* Психологические аспекты эффективного использования интернета в образовательных целях // Культурно-историческая психология. – 2008. – № 3. – С. 43-50.
9. *Лозицкий В.Л.* Феномен клипового мышления и информационно-коммуникационные технологии в высшем профессиональном образовании // Научные труды Республиканского института высшей школы. – 2016. – № 16-2. – С. 375-380.
10. *Scholl P., Benz B.F., Böhnstedt D., Rensing C., Schmitz B., Steinmetz R.* Implementation and evaluation of a tool for setting goals in self-regulated learning with Web resources / In U. Cress & V. Dimitrova (Eds.), Lecture Notes in Computer Science. EC-TEL 2009, LNCS 5794. – Berlin: Springer-Verlag, 2009. – P. 521-534.
11. *Ioannidis K., Treder M.S., Chamberlain S.R., Kiraly F., Redden S.A., Stein D.J., Lochner C. Grant J.E.* Problematic internet use as an age-related multifaceted problem: Evidence from a two-site survey // Addictive Behaviors. – 2018. – Vol. 81. – P. 157-166.
12. *Янг К.С.* Диагноз – интернет-зависимость // Мир Интернет. – 2000. – № 2. – С. 24-29.
13. *Горюнова Л.Н.* Развитие моделей информационного поведения с позиции обобщенной психологической теории деятельности // Вестник Санкт-Петербургского университета. – 2008. – Сер. 12. – Вып. 3. – С. 439-444.
14. *Ho L.-A., Kuo T.-H., Lin B.* The mediating effect of website quality on Internet searching behavior // Computers in Human Behavior. – 2012. – Vol. 28, Issue 3. – P. 840-848. – DOI: 10.1016/j.chb.2011.11.024.
15. *Ek S.* Factors relating to problems experienced in information seeking and use: findings from a cross-sectional population study in Finland // Information Research. – 2017. – Vol. 22, No. 4. – P. 775. – Режим доступа: http://informationr.net/ir/22-4/paper775.html.
16. *Ford N., Miller D., Moss N.* Web search strategies and approaches to studying // Journal of the American Society for Information Science and Technology. – 2003. – Vol. 54, Issue 6. – P. 473-489. – DOI: 10.1002/asi.10233.
17. *Брумштейн Ю.М., Васьковский Е.Ю., Куаншкалиев Т.Х.* Поиск информации в Интернете: анализ влияющих факторов и моделей поведения пользователей // Известия Волгоградского государственного технического университета. – 2017. – № 1 (196). – С. 50-55.
18. *Николенко С.И., Фишков А.А.* Обзор моделей поведения пользователей для задачи ранжирования результатов поиска // Тр. СПИИРАН. – 2012. – № 3 (22). – С. 139-175.
19. *Liu C.-C., Chang C.-J., Tseng J.-M.* The effect of recommendation systems on internet-based learning for different learners: a data mining analysis // British Journal of Educational Technology. – 2013. – Vol. 44, No. 5. – P. 758-773. – https://doi.org/10.1111/j.1467-8535.2012.01376.x.
20. *Бова В.В., Кравченко Ю.А., Кулиев Э.В., Курейчик В.В.* Моделирование поведения субъекта в Интернет-сервисах на основе модифицированного алгоритма бактериальной оптимизации // Информационные технологии. – 2019. – Т. 25, № 7. – С. 397-404. – DOI: 10.17587/it.25.397-404.
21. *Пантелюк Е.А., Кравченко Ю.А., Цырульникова Э.С.* Решение задачи управления знаниями на основе алгоритма умной капли воды // Информатика, вычислительная техника и инженерное образование. – 2017. – № 1.– С. 59-67.
22. *Смирнова О.С., Богорадникова А.В., Блинов М.Ю.* Описание роевых алгоритмов, инспирированных неживой природой и бактериями, для использования в онтологической модели // International Journal of Open Information Technologies. – 2015. – Vol. 3, No. 12. – P. 28-37. – ISSN: 2307-8162.

REFERENCES

1. *Lyz' N.A., Istratova O.N.* Informatsionno-obrazovatel'naya deyatel'nost' v internet-prostranstve: vidy, faktory, riski [Information and educational activities in the Internet space: types, factors, risks], *Pedagogika* [Pedagogy], 2019, No. 4, pp. 16-26.
2. *Raitskaya L.K.* Vliyanie interneta na lichnost' studenta [Influence of the Internet on the student's personality], *Kommunikatsiya v sovremennom polikul'turnom mire: dialog kul'tur: Ezhegodnyy sbornik nauchnykh trudov* [Communication in the modern multicultural world: a dialogue of cultures: Annual collection of scientific papers], ed. by T.A. Baranovskaya. Moscow, 2014, pp. 429-441.
3. *Salehi S., Du J. T., Ashman H.* Use of Web search engines and personalisation in information searching for educational purposes, *Information Research*, 2018, Vol. 23, No. 2. Available at: http://informationr.net/ir/23-2/paper788.html.
4. *Cen Y., Gan L., Bai C.* Reinforcement learning in information searching, *Information Research*, 2013, Vol, 18, Issue 1. Available at: http://informationr.net/ir/18-1/paper569.html#.Xl99magzaUk.
5. *Goryunova L.N., Kruglova M.A., Provotorova Ya.A., TSygan V.N.* Strategii informatsionnogo poiska i ikh vzaimosvyaz' s lichnostnymi osobennostyami studentov [Information search strategies and their relationship to students ' personal characteristics], *Peterburgskiy psikhologicheskiy zhurnal* [Petersburg psychological journal], 2013, No. 2, pp. 1-15.
6. *Lin C.-C., Tsai C.-C.* A navigation flow map method of representing students' searching behaviors and strategies on the web, with relation to searching outcomes, *Cyberpsychology and Behavior*, 2007, Vol. 10, Issue 5, pp. 689-695. DOI: 10.1089/cpb.2007.9969.
7. *Walraven A., Brand-Gruwel S., Boshuizen H.P.* Information-problem solving: A review of problems students encounter and instructional solutions, *Computers in Human Behavior*, 2008, Vol. 24, Issue 3, pp. 623-648.
8. *Porshnev A.V.* Psikhologicheskie aspekty effektivnogo ispol'zovaniya interneta v obrazovatel'nykh tselyakh [Psychological aspects of effective use of the Internet for educational purposes], *Kul'turno-istoricheskaya psikhologiya* [Cultural and historical psychology], 2008, No. 3, pp. 43-50.
9. *Lozitskiy V.L.* Fenomen klipovogo myshleniya i informatsionno-kommunikatsionnye tekhnologii v vysshem professional'nom obrazovanii [The phenomenon of clip thinking and information and communication technologies in higher professional education], *Nauchnye* trudy Respublikanskogo instituta vysshey shkoly [Scientific works of the Republican Institute of higher education], 2016, No. 16-2, pp. 375-380.
10. *Scholl P., Benz B.F., Böhnstedt D., Rensing C., Schmitz B., Steinmetz R.* Implementation and evaluation of a tool for setting goals in self-regulated learning with Web resources, In U. Cress & V. Dimitrova (Eds.), Lecture Notes in Computer Science. EC-TEL 2009, LNCS 5794. Berlin: Springer-Verlag, 2009, pp. 521-534.
11. *Ioannidis K., Treder M.S., Chamberlain S.R., Kiraly F., Redden S.A., Stein D.J., Lochner C. Grant J.E.* Problematic internet use as an age-related multifaceted problem: Evidence from a two-site survey, *Addictive Behaviors*, 2018, Vol. 81, pp. 157-166.
12. *Yang K.S.* Diagnoz – internet-zavisimost' [Diagnosis-Internet addiction], *Mir Internet* [Mir Internet], 2000, No. 2, pp. 24-29.
13. *Goryunova L.N.* Razvitie modeley informatsionnogo povedeniya s pozitsii obobshchennoy psikhologicheskoy teorii deyatel'nosti [Development of models of information behavior from the position of a generalized psychological theory of activity], *Vestnik Sankt-Peterburgskogo universiteta* [Vestnik of Saint Petersburg University], 2008, Ser. 12, Issue 3, pp. 439-444.
14. *Ho L.-A., Kuo T.-H., Lin B.* The mediating effect of website quality on Internet searching behavior, *Computers in Human Behavior*, 2012, Vol. 28, Issue 3, pp. 840-848. DOI: 10.1016/j.chb.2011.11.024.
15. *Ek S.* Factors relating to problems experienced in information seeking and use: findings from a cross-sectional population study in Finland, *Information Research*, 2017, Vol. 22, No. 4, pp. 775. – Режим доступа: http://informationr.net/ir/22-4/paper775.html.
16. *Ford N., Miller D., Moss N.* Web search strategies and approaches to studying, *Journal of the American Society for Information Science and Technology*, 2003,– Vol. 54, Issue 6, pp. 473-489. DOI: 10.1002/asi.10233.
17. *Brumshteyn Yu.M., Vas'kovskiy E.Yu., Kuanshkaliev T.Kh.* Poisk informatsii v Internete: analiz vliyayushchikh faktorov i modeley povedeniya pol'zovateley [Search for information on the Internet: analysis of influencing factors and user behavior patterns], *Izvestiya Volgogradskogo gosudarstvennogo tekhnicheskogo universiteta* [Izvestiya Volgograd state technical University], 2017, No. 1 (196), pp. 50-55.

18. *Nikolenko S.I., Fishkov A.A.* Obzor modeley povedeniya pol'zovateley dlya zadachi ranzhirovaniya rezul'tatov poiska [Overview of user behavior models for the task of ranking search results], *Tr. SPIIRAN* [SPIIRAS Proceedings], 2012, No. 3 (22), pp. 139-175.
19. *Liu C.-C., Chang C.-J., Tseng J.-M.* The effect of recommendation systems on internet-based learning for different learners: a data mining analysis, *British Journal of Educational Technology*, 2013, Vol. 44, No. 5, pp. 758-773. Available at: https://doi.org/10.1111/j.1467-8535.2012.01376.x.
20. *Bova V.V., Kravchenko Yu.A., Kuliev E.V., Kureychik V.V.* Modelirovanie povedeniya sub"ekta v Internet-servisakh na osnove modifitsirovannogo algoritma bakterial'noy optimizatsii [Modeling the behavior of a subject in Internet services based on a modified algorithm of bacterial optimization], *Informatsionnye tekhnologii* [Information technologies], 2019, Vol. 25, No. 7, pp. 397-404. DOI: 10.17587/it.25.397-404.
21. *Pantelyuk E.A., Kravchenko Yu.A., Tsyrul'nikova E.S.* Reshenie zadachi upravleniya znaniyami na osnove algoritma umnoy kapli vody [Solving the problem of knowledge management based on the smart water drop algorithm], *Informatika, vychislitel'naya tekhnika i inzhenernoe obrazovanie* [Informatics, computer engineering and engineering education], 2017, No. 1, pp. 59-67.
22. *Smirnova O.S., Bogoradnikova A.V., Blinov M.Yu.* Opisanie roevykh algoritmov, inspirirovannykh nezhivoy prirodoy i bakteriyami, dlya ispol'zovaniya v ontologicheskoy modeli [Description of swarm algorithms inspired by inanimate nature and bacteria for use in an ontological model], *International Journal of Open Information Technologies*, 2015, Vol. 3, No. 12, pp. 28-37. ISSN: 2307-8162.

Статью рекомендовал к опубликованию к.т.н. С.Г. Буланов.

Бова Виктория Викторовна – Южный федеральный университет; e-mail: vvbova@yandex.ru; 347928, г. Таганрог, Некрасовский, 44; тел.: 88634371651; кафедра систем автоматизированного проектирования; доцент.

Запорожец Дмитрий Юрьевич – e-mail: duzaporozhets@sfedu.ru; кафедра систем автоматизированного проектирования; доцент.

Кравченко Юрий Алексеевич – e-mail: yakravchenko@sfedu.ru; кафедра систем автоматизированного проектирования; доцент.

Кулиев Эльмар Валерьевич – e-mail: ekuliev@sfedu.ru; кафедра систем автоматизированного проектирования; доцент.

Курейчик Владимир Викторович – e-mail: vvkur@sfedu.ru; кафедра систем автоматизированного проектирования; зав. кафедрой; профессор.

Лызь Наталья Александровна – e-mail: nlyz@sfedu.ru; тел.: 88634361586; кафедра систем автоматизированного проектирования; зав. кафедрой; профессор.

Bova Victoria Victorovna – Southern Federal University; e-mail: vvbova@yandex.ru; 44, Nekrasovskiy, Taganrog, 347928, Russia; phone: +78634371651; the department of computer aided design; associate professor.

Zaporozhets Dmitrii Yurievich – e-mail: duzaporozhets@sfedu.ru; the department of computer aided design; associate professor.

Kravchenko Yury Alekseevich – e-mail: yakravchenko@sfedu.ru; the department of computer aided design; associate professor.

Kuliev Elmar Valerievich – e-mail: ekuliev@sfedu.ru; the department of computer aided design; associate professor.

Kureichik Vladimir Victorovich – e-mail: vvkur@sfedu.ru; the department of computer aided design; head of department; professor.

Lyz Nataliya Alexandrovna – e-mail: nlyz@sfedu.ru; phone: +78634361586; the department of psychology and life safety; head of department; professor.