УНИВЕРСИТЕТ ИТМО

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ
ноябрь–декабрь 2021 Том 21 № 6 http://ntv.ifmo.ru/
SCIENTIFIC AND TECHNICAL JOURNAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
November–December 2021 Vol. 21 No 6 http://ntv.ifmo.ru/en/
ISSN 2226-1494 (print) ISSN 2500-0373 (online)

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК УНИВЕРСИТЕТ ИТМО
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ

doi: 10.17586/2226-1494-2021-21-6-942-950
УДК 004.8 004.056.53

Идентификация аккаунтов пользователей социальных сетей при помощи сравнения графического контента

Анастасия Андреевна Корепанова[1], Максим Викторович Абрамов[2], Александр Львович Тулупьев[3]✉

[1,2,3] Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация
[1,2,3] Санкт-Петербургский государственный университет, Санкт-Петербург, 199034, Российская Федерация

[1] aak@mail.ru, https://orcid.org/0000-0003-2962-8670
[2] mva@dscs.pro, https://orcid.org/0000-0002-5476-3025
[3] alt@dscs.pro✉, https://orcid.org/0000-0003-1814-4646

Аннотация

Предмет исследования. В работе представлен новый подход к сопоставлению аккаунтов социальных сетей «ВКонтакте» и «Instagram» для определения принадлежности их одному пользователю. Подход основан на сопоставлении графического контента. Новизна подхода состоит в объединении нескольких методов для сопоставления графического контента. Впервые предложен метод сопоставления аккаунтов данных социальных сетей. **Метод.** Представленный метод заключается в объединении трех способов сопоставления графического контента с помощью: выделения из фотографий лиц владельцев аккаунта и их сопоставления; сопоставления всех лиц в обоих аккаунтах; попарного сравнения изображений для определения одинаковых изображений в обоих аккаунтах с помощью метода перцептивного хэша pHash. **Основные результаты.** Предложенный метод апробирован на наборе данных из более чем 8000 пар аккаунтов. По результатам эксперимента величина метрики F-мера достигла 0,87. **Практическая значимость.** Практическая значимость метода заключается в автоматизации сопоставления аккаунтов пользователей в различных социальных сетях через реализацию разработанного алгоритма в прототипе комплекса программ. Дальнейшее направление исследования направлено на расширение набора данных и атрибутов профилей, рассматриваемых для сравнения. Результаты работы могут быть внедрены в программный комплекс анализа защищенности пользователей информационных систем от социоинженерных атак. Перспективным направлением является объединение полученных результатов с методами сопоставления аккаунтов, основанными на структурном подобии социальных графов.

Ключевые слова

социальные сети, идентификация пользователя, обработка изображений, машинное обучение, социоинженерные атаки

Благодарности

Работа выполнена в рамках проекта по государственному заданию СПб ФИЦ РАН СПИИРАН № 0073-2019-0003 (формирование подхода сопоставления аккаунтов с помощью анализа графического контента); поддержана Санкт-Петербургским государственным университетом, проект № 73555239 (реализация базовых подходов сопоставления аккаунтов и их апробация); при финансовой поддержке РФФИ, проект № 20-07-00839 (реализация комбинированного метода сопоставления аккаунтов и апробация результатов).

Ссылка для цитирования: Корепанова А.А., Абрамов М.В., Тулупьев А.Л. Идентификация аккаунтов пользователей социальных сетей при помощи сравнения графического контента // Научно-технический вестник информационных технологий, механики и оптики. 2021. Т. 21, № 6. С. 942–950. doi: 10.17586/2226-1494-2021-21-6-942-950

Social media user identity linkage by graphic content comparison

Anastasia A. Korepanova[1], Maxim V. Abramov[2], Alexander L. Tulupyev[3]✉

[1,2,3] St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation

[1,2,3] Saint Petersburg State University, Saint Petersburg, 199034, Russian Federation

[1] aak@mail.ru, https://orcid.org/0000-0003-2962-8670

[2] mva@dscs.pro, https://orcid.org/0000-0002-5476-3025

[3] alt@dscs.pro✉, https://orcid.org/0000-0003-1814-4646

Abstract

The article proposes a new approach to comparing accounts of the social media "VKontakte" and "Instagram" to determine those accounts which belong to the same user. The approach is based on the comparison of graphic content; the novelty of the approach consists in combining several methods for matching graphic content, also for the first time a method is proposed for matching accounts of the mentioned social media. The proposed method combines three methods of matching graphic content: by extracting the faces of the account users from the photos in the account and matching them, by matching all faces in both accounts, and by pairwise comparison of images to determine the same images in both accounts using the perceptual pHash method. The described method was tested on a dataset of more than 8,000 pairs of accounts. According to the results of the experiment, the value of the F1-score metric reached 0.87. The practical significance lies in automating the comparison of user accounts in various social networks by implementing of the developed algorithm in the prototype of the software package. A further direction for research lies in expanding the set of data and attributes of profiles considered for comparison. The results can be introduced into a software package for the analysis of the security of a user of information systems against social engineering attacks. It seems promising to combine the obtained findings with account matching methods based on the structural similarity of social graphs.

Keywords

social media, user identity linkage, image processing, machine learning, social engineering attacks

Acknowledgements

This work was carried out within the framework of the project under the state assignment of SPC RAS SPIIRAS No. 0073-2019-0003 (approach formation); supported by Saint Petersburg State University, project No. 73555239 (implementation of the approach and its approbation); with the financial support of the RFBR, project No. 20-07-00839 (approbation of the results in the prototype of the software package).

For citation: Korepanova A.A., Abramov M.V., Tulupyev A.L. Social media user identity linkage by graphic content comparison. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2021, vol. 21, no. 6, pp. 942–950 (in Russian). doi: 10.17586/2226-1494-2021-21-6-942-950

Введение

В настоящее время анализ социальных сетей интегрирован во многие области науки и индустрии. Он применяется в различных контекстах, в таких сферах как информационная безопасность [1], таргетированная реклама [2, 3], банковский скоринг [4–6] и многих других. Список сфер применения постоянно расширяется. Например: анализ данных конкретного пользователя для определения его скрытых характеристик (личностных особенностей, особенностей взаимодействия с другими пользователями и др.); анализ некоторых групп пользователей, объединенных по какому-то признаку (например, жители одной страны, школьники одного возраста) для определения общих тенденций поведения внутри этих групп (склонность к гедонизму, депрессивным состояниям и др.) или реакций внутри этих групп на какие-то события. Сведения о пользователе, извлеченные посредством анализа его аккаунта в социальной сети, могут быть использованы для проведения социоинженерной атаки [7, 8]; предоставления релевантной рекламы в задачах маркетинга [2, 3] и т. п. Анализ групп пользователей позволяет, например, оценить отношение пользователей к некоторому продукту или событию [3], в последний год он получил особенную актуальность в связи с необходимостью оценки психического состояния пользователей во время пандемии и оценки распространения инфодемии по социальным сетям [9].

Одна из сфер применения анализа социальных сетей — область анализа защищенности пользователей информационных систем от социоинженерных атак. Данные атаки направлены не на технические уязвимости системы, а на уязвимости ее пользователей, связанные с их личностными особенностями [7, 8]. Существует подход к оценке защищенности пользователей информационных систем от социоинженерных атак, основанный на построении профиля уязвимостей пользователя [7, 8], который ассоциирован с его личностными особенностями. В качестве одного из источников информации для оценок личностных особенностей пользователя могут выступать данные, извлекаемые из социальных сетей [10–12]. Часто пользователи имеют несколько профилей в разных социальных сетях (термины «профиль пользователя в социальной сети» и «аккаунт пользователя» понимаются в настоящей работе как синонимы)[1], которые не всегда напрямую связаны взаимными ссылками, но каждый из которых может содержать некоторую часть полезной информации о пользователе. Актуальной является задача поиска аккаунтов пользователя в различных социальных сетях. На практике сложно найти все страницы пользователя в социальных сетях вручную, поэтому необходима автоматизация поиска. Один из

[1] Статистика социальных сетей в России на 2018 год [Электронный ресурс]. URL: https://hiconversion.ru/blog/statistika-socialnyh-setej-v-rossii-na-2018-god/ (дата обращения: 30.08.2021).

этапов решения данной задачи — автоматизация сопоставления двух профилей для оценки вероятности того, что они принадлежат одному пользователю. Подход состоит в сопоставлении двух профилей пользователей социальных сетей «ВКонтакте» и «Instagram» и определении, принадлежат ли они одному человеку. Обозначенная задача не является новой: существуют исследования, посвященные идентификации аккаунтов пользователей в различных социальных сетях, например, были рассмотрены социальные сети «ВКонтакте» и «Одноклассники» [13, 14], однако комбинация «ВКонтакте» — «Instagram» в данных исследованиях не рассматривалась. Социальные сети могут отличаться друг от друга в представлении данных и методов взаимодействия пользователей через них, и поэтому методы сопоставления аккаунтов пользователей разных социальных сетей должны разрабатываться индивидуально. Таким образом, актуальность настоящей работы подтверждается отсутствием устоявшегося подхода к сопоставлению профилей пользователей «ВКонтакте» и «Instagram», а также широким прикладным потенциалом.

Цель работы — повышение оперативности сопоставления профилей пользователей в различных социальных сетях: через разработку модели сопоставления, соответствующего ей алгоритма сопоставления и их автоматизации. Теоретическая значимость работы представлена разработанными новой моделью и алгоритмом сопоставления аккаунтов пользователей в социальных сетях.

Обзор литературы

Задача идентификации аккаунтов одного пользователя в различных социальных сетях становится все более актуальной в связи с ростом количества социальных сетей и их популярности. Эти факторы обусловливают рост числа научных исследований по данной теме. Существуют три стандартных подхода к решению задачи идентификации пользователя в социальных сетях: с помощью сопоставления анкетных данных, публикуемого контента и свойств социальных графов аккаунтов (графов, составленных из друзей пользователя в социальной сети).

Первый подход рассмотрен в работах [15–24]. В [15, 16] предложены методы, основанные на сопоставлении имен и никнеймов пользователей, точность таких методов оказалась невысока. В работах [17, 18] набор сопоставляемых атрибутов расширен до никнейма, города проживания и аватара. Метод поатрибутного сравнения применим только в случае, если публичная анкета пользователя заполнена достаточно подробно. Подходы, основанные на сопоставлении публикуемого в аккаунтах контента, представлены в работах [19, 20]. Исследование [19] сфокусировано на выделении тем интересов пользователя с помощью тематического моделирования его текстового контента, авторы [20] извлекают локации из постов, которые используют для составления уникального следа пользовательской активности. Предложенный в [21] метод сопоставления аккаунтов в «Instagram» и «Twitter» основан на анализе активности пользователя через частоту и периодичность постинга. К последнему подходу можно отнести работы [22–24]. В [22] изучена проблема сопоставления аккаунтов пользователей при ограниченном объеме известной информации (мало анкетных данных и/или контента). Предложено сопоставлять структурные свойства социальных графов с использованием метода деанонимизации графа на основе перколяции. В [23] изучены 40 характеристик социальных графов и их влияние на идентификацию пользователей. В [24] предложен подход, основанный на сопоставлении атрибутов пользователя и его социального окружения. Среди социального окружения пользователя определяются его близкие друзья, и сопоставляются с близкими друзьями аккаунта потенциального этого же пользователя в другой социальной сети.

Данные настоящей работы, несмотря на широкое освещение вопроса идентификации аккаунтов пользователя в различных социальных сетях, имеют существенные особенности.

Рассмотренные научные работы концентрируются на социальных сетях, популярных в англоязычном интернет-сегменте: «Forsquare», «Facebook», «Twitter» и др., но в этих работах не учтена специфика социальных сетей «ВКонтакте» и «Instagram». Подходы, основанные на сопоставлении значений атрибутов публичной анкеты, не могут быть применены для решения поставленной задачи настоящей работы, так как «Instagram» не предлагает пользователю анкеты для заполнения. Анализ социального графа имеет следующие сложности: так как «Instagram» сильно ограничивает возможность выгрузки данных, то сбор большого набора данных для обучения, содержащего информацию о социальных связях, затруднителен.

Таким образом, задача разработки нового метода сопоставления аккаунтов «ВКонтакте» и «Instagram» на данный момент не решена.

Постановка задачи

Пусть имеется пара аккаунтов пользователей из социальных сетей «ВКонтакте» и «Instagram». В каждом из этих аккаунтов размещено некоторое множество изображений. Задача состоит в том, чтобы разработать метод, позволяющий через сравнение графического контента аккаунтов определять, принадлежит ли пара профилей одному пользователю или нет. Данную задачу можно формализовать как задачу классификации следующим образом: пусть X — множество пар профилей пользователей социальных сетей «ВКонтакте» и «Instagram», а Y — множество классов 0; 1, где 0 означает, что пара профилей не принадлежит одному пользователю, а 1 — принадлежит. Необходимо построить алгоритм $a: X \to Y$, который будет способен классифицировать любой $x \in X$, основываясь на анализе контента профилей.

Описание методов

В социальной сети «ВКонтакте» пользователь может указать о себе много информации: есть подробная анкета, в которой помимо социодемографических

атрибутов (пол, возраст, образование, семейное положение и др.) предусмотрена возможность для указания своих жизненных взглядов и интересов; можно добавить в друзья других пользователей; есть возможность подписываться на сообщества; можно добавлять себе на страницу музыку, видео и фотографии; можно как писать на своей странице текстовые посты, так и размещать чужие. В то же время социальная сеть «Instagram» ориентирована в первую очередь на обмен медиаконтентом в форме фотографий и видеозаписей с текстовыми описаниями ограниченного объема. Анкету о себе заменяет атрибут «биография» с ограничением в количестве символов; аудиозаписи добавить себе на страницу нельзя, сообществ в том же виде, что и в «ВКонтакте» в «Instagram» нет; для взаимодействия с другими пользователями можно «подписаться» на обновления на их странице. Таким образом, в аккаунтах одного и того же пользователя в данных социальных сетях может содержаться разная информация. Большая часть данных, доступных в «ВКонтакте», не имеют аналогов в «Instagram». Так как «Instagram» в большей степени содержит графический контент, в настоящей работе предложены методы сопоставления двух аккаунтов, основанные на анализе данного контента. Также «Instagram» сильно ограничивает возможность выгрузки данных, так что было решено на данном этапе использовать те данные, выгрузить которые наиболее просто, т. е. фотографии из постов. Были предложены и протестированы следующие методы сопоставления аккаунтов:

— с помощью распознавания лица пользователя аккаунта на фотографии поста — выделение на фотографиях лиц предположительных владельцев аккаунтов и определении, насколько эти лица в паре аккаунтов похожи;
— с помощью распознавания всех лиц в аккаунтах — поиск лиц одних и тех же людей в обоих аккаунтах в паре;
— с помощью сопоставления хэшей фотографий — поиск совпадающих изображений в двух аккаунтах;
— объединение методов в одной модели.

Рассмотрим перечисленные методы подробнее.

Сопоставление аккаунтов с помощью распознавания лица пользователя аккаунта

Для задач, связанных с распознаванием лиц на фотографиях, используем библиотеку, написанную на языке Python, face_recognition[1].

Метод основан на предположении, что пользователи социальных сетей «ВКонтакте» и «Instagram» размещают в основном свои фотографии. В связи с этим для определения, принадлежат ли аккаунты одному пользователю, можно выделить и сравнить на фотографиях лица владельцев аккаунтов. Данный подход заключается в следующей последовательности шагов.

Шаг 1. Выделение всех лиц на фотографиях с помощью библиотеки face_recognition. В данной библиотеке реализовано два подхода к выделению лиц на фотографиях: с помощью метода гистограмм направленных градиентов [25] и предобученной нейронной сети. Нейронная сеть точнее, но требует значительно больше вычислительных ресурсов, поэтому для скорости вычислений был использован метод гистограмм направленных градиентов. Выделенные лица кодируются в виде 128-элементного вектора чисел, т. е. строится эмбеддинг лица с помощью функции face_encodings библиотеки face_recognition.

Шаг 2. Проведение кластеризации выделенных лиц для определения лиц, принадлежащих разным людям. Выделенные лица кластеризуются с помощью алгоритма DBSCAN [26] реализованного в библиотеке sklearn[2]. Кластеризация позволяет выделить среди всех лиц на фотографиях в аккаунте лица, принадлежащие одному человеку. Алгоритм DBSCAN в качестве кластеров выделяет в множестве точек в векторном пространстве группы точек, которые расположены наиболее плотно. В качестве расстояния между двумя эмбеддингами лиц использовано значение евклидовой метрики. Принято, что две точки находятся достаточно близко друг к другу, если расстояние между ними меньше 0,5, минимальное число лиц, составляющих кластер — 5.

Шаг 3. Выполнение действий с кластером:

— если удалось получить хотя бы один кластер — выбирается кластер, содержащий наибольшее число лиц, и рассчитывается центроид этого кластера (в рассматриваемом случае – среднее арифметическое всех эмбеддингов, входящих в кластер);
— если не удалось выделить ни одного кластера — рассматриваются все лица в профиле как один кластер и рассчитывается его центроид.

Шаг 4. Сопоставление полученных центроидов. В качестве меры расстояния между центроидами использовано косинусное расстояние, т. е. значение косинуса угла между двумя векторами:

$$\cos(\mathbf{a}, \mathbf{b}) = \frac{\sum_{i=0}^{n} a_i b_i}{\sqrt{\sum_{i=1}^{n} (a_i)^2} \cdot \sqrt{\sum_{i=1}^{n} (b_i)^2}},$$

где **a**, **b** — векторы, представляющие центроиды кластеров; a_i — обозначает i-й элемент вектора **a**; b_i — i-й элемент вектора **b**.

Шаг 5. Применение значения косинусного расстояния с помощью логистической регрессии в качестве признака для классификации пар на классы «принадлежит одному пользователю» и «не принадлежит одному пользователю». Логистическая регрессия — одна из моделей машинного обучения, используемая для прогнозирования вероятности принадлежности объекта к одному из двух классов с помощью логистической функции [27]. Строятся следующие предположения о вероятности принадлежности объекта к классам 1 и 0:

$$P(y = 1|x) = h_w(x),$$

[1] Face Recognition [Электронный ресурс]. URL: https://github.com/ageitgey/face_recognition (дата обращения: 17.10.2021).

[2] Scikit-learn [Электронный ресурс]. URL: https://scikit-learn.org/stable/ (дата обращения: 17.10.2021).

$$P(y = 0|x) = 1 - h_w(x),$$
$$h_w(x) = \frac{1}{1 + e^{-(w_0 + w_1 x_1 + \ldots + w_n w_n)}},$$

где $P(y = 1|x)$ — вероятность принадлежности объекта x к классу 1; $P(y = 0|x)$ — вероятность принадлежности объекта x к классу 0; $h_w(x)$ — сигмоидная функция; $x_i, i \in 1\ldots n$ в сигмоидной функции — признаки объекта x; $w_i : i \in 0\ldots n$ — коэффициенты регрессии. В качестве объектов выступают пары аккаунтов, в качестве признака объекта — значение расстояния между центроидами, если два аккаунта принадлежат одному пользователю, то они относятся к классу 1, иначе пара аккаунтов относится к классу 0.

Сопоставление аккаунтов с помощью распознавания всех лиц на фотографиях

Данный подход основан на предположении, что, так как пользователи в своих аккаунтах могут выставлять не только свои фотографии, но и фотографии людей из своего социального круга: родных, друзей и знакомых; множества людей, чьи лица встречаются в разных аккаунтах одного и того же пользователя, будут в значительной степени пересекаться. Исходя из этого предположения, не обязательно выделять на фотографиях лицо владельца аккаунта, достаточно сопоставить все лица, встречающиеся в обоих аккаунтах. Недостаток этого подхода в том, что люди, чей социальный круг в значительной степени пересекается (это могут быть члены семьи, близкие друзья и др.), могут оказаться неотличимы с точки зрения сопоставления. Предположительно, этот метод не очень хорош в качестве самостоятельного, но может быть использован в комбинации с другими. Предложенный метод сопоставления аккаунтов состоит из следующих шагов:

— поиск лиц на фотографиях. Данный шаг аналогичен шагу 1 в разделе «Сопоставление аккаунтов с помощью распознавания лица пользователя аккаунта»;
— сопоставление лиц. Сопоставление лиц производится следующим образом: все лица из аккаунта «ВКонтакте» сопоставляются со всеми лицами из аккаунта «Instagram» с помощью косинусного расстояния. Если косинусное расстояние между лицами меньше порога 0,5, они считаются совпавшими, т. е. если для лица из аккаунта «ВКонтакте» нашлось лицо из аккаунта «Instagram», с которым оно совпало, лицо считается распознанным. Далее вычисляется доля распознанных лиц из аккаунта «ВКонтакте»;
— доля распознанных лиц используется в качестве признака в логистической регрессии для классификации аккаунтов.

Сопоставление аккаунтов с помощью сопоставления хэшей изображений

Сопоставление аккаунтов исключительно через сопоставление лиц обладает недостатком неуниверсальности, который может быть обусловлен следующими причинами: пользователи делятся в социальных сетях не только контентом со своим лицом, но и другими изображениями, например, природы, лиц родственников и др.; распознавание лиц может плохо сработать, если человек сфотографирован в очках, с ярким макияжем или при плохом освещении. При этом, согласно гипотезе, люди часто используют одинаковые фотографии в разных социальных сетях, поэтому полезным может быть сравнение не только через идентификацию лиц на фотографиях, но и через сравнение самих фотографий. В настоящей работе использован хэш — перцептивный хэш (pHash), реализованный в библиотеке ImageHash[1].

Изображение в цветовой модели RGB представлено в виде [28]:

$$(f(i,j))_{i,j=0}^{\substack{i=n-1\\j=m-1}} = \begin{pmatrix} (r,g,b)_{0,0} & (r,g,b)_{0,1} & \ldots & (r,g,b)_{0,n-1} \\ (r,g,b)_{1,0} & (r,g,b)_{1,1} & \ldots & (r,g,b)_{1,n-1} \\ \ldots & \ldots & \ldots & \ldots \\ (r,g,b)_{m-1,0} & (r,g,b)_{m-1,1} & \ldots & (r,g,b)_{m-1,n-1} \end{pmatrix}, \quad (1)$$

где $i \in [0, n-1]$, $j \in [0, m-1]$ — пространственные координаты пикселов изображения (индексы функции); значения функции — три трехзначных числа, которые определяют интенсивность (0 — минимальная, 255 — максимальная интенсивность) для каждого из цветовых каналов модели RGB (r — красного, g — зеленого, b — синего) [28].

pHash использует дискретное косинусное преобразование, в основе которого лежит быстрое преобразованием Фурье [29], позволяющее отбросить детали, которые могут возникнуть в результате незначительных преобразований изображения (изменение размера, цветокоррекция, соотношение сторон и др.), и сравнить основную структуру изображений.

$$F(i,j) = \sum_{k=0}^{m-1} \sum_{l=0}^{n-1} f(i,j) \cos\left[i\left(k + \frac{1}{2}\right)\frac{\pi}{m}\right] \cos\left[j\left(l + \frac{1}{2}\right)\frac{\pi}{n}\right],$$

где $F(i,j)$ — дискретное косинусное преобразование для матрицы $(f(i,j))_{i,j=0}^{\substack{i=n-1\\j=m-1}}$ [28]; $l \in [0, n-1]$, $k \in [0, m-1]$ — пространственные координаты пикселов изображения. Одно и то же изображение даже после изменений будет иметь близкий или одинаковый хэш.

Сравнение аккаунтов происходит следующим образом:

1) для каждой фотографии в аккаунте рассчитывается свой хэш;
2) происходит сравнение хэшей фотографий по схеме «каждый с каждым» с помощью расстояния Хэмминга (число позиций, в которых две строки одинаковой длины различаются) [30]. Если расстояние между двумя хэшами меньше порога i, фотографии считаются совпавшими;
3) если количество совпадений хэшей больше некоторого порога j, аккаунты считаются принадлежащими одному человеку [28].

[1] ImageHash 4.2.1 [Электронный ресурс]. URL: https://pypi.org/project/ImageHash/ (дата обращения: 17.10.2021).

Объединенная модель

Представленные методы сопоставления имеют как ограничения в применимости, так и преимущества. Исходя из этого, сделано объединение методов для получения наилучшего результата. В предложенном подходе результаты сопоставления графического контента разными методами выступают в качестве признаков единой модели классификации пар аккаунтов. В качестве модели классификации рассмотрена логистическая регрессия [27]. В данном случае в качестве объектов выступают пары аккаунтов, в качестве признаков объекта — значение расстояния между центроидами наибольших кластеров лиц, доля распознанных лиц из всех лиц в аккаунте «ВКонтакте» и результат сопоставления хэшей фотографий (1, если больше 5 совпало, и 0 иначе), в качестве классов — принадлежат ли два аккаунта одному пользователю (1, если принадлежат, и 0 иначе).

Эксперимент

Описание данных. Необходимо собрать датасет, состоящий из таких пар аккаунтов одних и тех же пользователей в различных социальных сетях, при условии, что аккаунты в одной паре принадлежат одному пользователю, а аккаунты в различных парах – разным. Набор данных сформирован в несколько шагов, на первом шаге отбирались подходящие аккаунты «ВКонтакте». Чтобы получить априорную информацию о принадлежности аккаунтов одному пользователю, выбраны аккаунты «ВКонтакте», у которых на странице указан соответствующий аккаунт пользователя в «Instagram». Датасет собран последовательным перебором аккаунтов «ВКонтакте» с ID в случайно выбранном промежутке 496 234–1 564 968. Пары аккаунтов в набор данных выбраны согласно требованиям:

— аккаунт не должен быть деактивирован;
— аккаунт должен быть открыт, т. е. из него должна быть возможность выгрузки фотографий;
— для сопоставления аккаунтов использованы фотографии. В аккаунтах пользователей количество фотографий может варьироваться от 0 до бесконечности, так как «Instagram» на данный момент не имеет ограничений на количество постов в профиле. Для формирования набора данных выставлены границы на количество фотографий у пользователя «ВКонтакте»: аккаунты, содержащие не более 200 фотографий, чтобы ускорить вычисления, а также отсечь аккаунты, не относящиеся к личным (например, онлайн-магазинов); аккаунты, содержащие менее 10 фотографий, чтобы отсечь аккаунты, не содержащие достаточно информации для сравнения. Обе границы были выставлены экспертно.

В результате из более миллиона просмотренных аккаунтов отобрано 27 720.

На следующем шаге из полученного набора данных отсекались те, чей аккаунт «Instagram» не соответствовал следующим требованиям:

— аккаунт должен быть открыт;
— количество фотографий в «Instagram» может варьироваться от 0 до бесконечности, по причинам, указанным выше; были отобраны аккаунты, количество фотографий в которых лежало в диапазоне 10–200;
— для скачивания данных использована библиотека Instaloader[1] которая позволяет выгружать данные из «Instagram» без авторизации. К сожалению, работа библиотеки не стабильна, не все посты скачиваются, поэтому аккаунты, для которых не удалось скачать больше 10 фотографий, были исключены.

В результате набор данных составил 4269 пар аккаунтов, аккаунты в паре считались принадлежащими одному человеку. Для каждой пары были скачены фотографии с датами размещения в социальной сети. 4269 пар аккаунтов, не принадлежащих одному пользователю, были составлены из имеющихся аккаунтов случайным образом. Можно отметить следующие недостатки набора данных, полученного описанным в этом разделе способом:

— можно выбрать только пользователей, которые добровольно указали ссылку на свою страницу в другой социальной сети. Применимость полученных в ходе данной работы результатов на пользователях, каким-либо образом скрывающих свою страницу, необходимо проверить дополнительно;
— сложно гарантировать, что пара состоит именно из личных аккаунтов пользователя. Например, пользователь может указать на своей странице «ВКонтакте» ссылку на свой интернет-магазин или публичную страницу своей компании, размещенные в «Instagram»;
— собранный набор данных не содержит разметки для изображений, т. е. нет априорной информации о том, какие изображения и лица совпадают.

Результаты эксперимента. Для каждого метода сопоставления построены отдельные модели, которые впоследствии были объединены в одну. Оценка качества классификации каждой модели выполнена с помощью следующих метрик: точность, чувствительность, F-мера и accuracy. Пусть TP — число пар, правильно отнесенных к классу аккаунтов, принадлежащих одному человеку, FP — число пар аккаунтов, принадлежащих разным людям, которые были ошибочно отнесены к классу аккаунтов, принадлежащих одному человеку, FN — число пар, принадлежащих одному человеку, ошибочно отнесенных к классу аккаунтов, принадлежащих разным людям. Тогда точность (precision), чувствительность (recall) и F-мера будут иметь следующий вид:

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}},$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$

$$\text{F-мера} = 2\frac{\text{precision}\cdot\text{recall}}{\text{precision} + \text{recall}}.$$

Используем процедуру скользящего контроля по двадцати блокам [31]. Для построения и обучения моделей применим библиотеку sklearn.

[1] Instaloader [Электронный ресурс]. URL: https://instaloader.github.io/# (дата обращения: 17.10.2021).

Таблица. Оценка качества разных моделей классификации
Table. Performance measures for the different classification models

Модель	Метрика			
	Точность	Чувствительность	F-мера	Accuracy
Сопоставление лиц владельца аккаунта	0,78	0,83	0,80	0,78
Сопоставление всех лиц в аккаунтах	0,88	0,81	0,84	0,84
Объединение первых двух	0,89	0,83	0,86	0,85
Сопоставление хэшей изображений ($i = 0, j = 5$)	0,72	0,51	0,52	0,59
Объединение всех методов	0,91	0,84	0,87	0,87

Для сопоставления аккаунтов с помощью хэшей картинок методом перебора выберем пороги $i = 0$ (расстояние Хэмминга, меньше которого картинки считаются совпавшими) и $j = 5$ (количество совпавших фотографий, необходимое для того, чтобы считать аккаунты совпавшими). Результаты эксперимента отражены в таблице.

В результате эксперимента модель, объединяющая несколько методов сопоставления аккаунтов, показала наилучшие значения метрик, в частности F-мера и accuracy достигли значения 0,87. Таким образом, предложенный подход показал высокое качество классификации на собранном наборе данных.

Заключение

В работе представлен новый подход к сопоставлению аккаунтов пользователей из социальных сетей «ВКонтакте» и «Instagram» для определения тех, которые принадлежат одному человеку. Данный подход использует следующие предположения о содержания контента пользователей социальных сетей: пользователи выставляют преимущественно фотографии со своим лицом, пользователи выставляют фотографии одних и тех же людей из своего социального круга в разных социальных сетях, пользователи часто публикуют одни и те же изображения в разных социальных сетях. Предложенный метод показал высокую точность на сформированном наборе данных, состоящем из более чем 4000 пользователей. Значимость работы заключается в новых предложенных методах и подходах для идентификации аккаунтов пользователей в различных социальных сетях. Дальнейшее направление для исследования лежит в расширении набора данных и атрибутов профилей, рассматриваемых для сравнения, а также внедрение программной реализации в программный комплекс, посвященный анализу защищенности пользователей информационных систем от социоинженерных атак. Перспективным кажется объединение полученных результатов с методами сопоставления аккаунтов, основанными на структурном подобии социальных графов.

Полученный результат может быть применен в широком спектре задач, связанных с социальными сетями: в области защиты от социоинженерных атак, маркетинга, социологических исследований.

Литература

1. Camacho D., Panizo-LLedot Á., Bello-Orgaz G., Gonzalez-Pardo A., Cambria E. The four dimensions of social network analysis: An overview of research methods, applications, and software tools // Information Fusion. 2020. V. 63. P. 88–120. https://doi.org/10.1016/j.inffus.2020.05.009
2. Yamane D., Yamane P., Ivory S.L. Targeted advertising: Documenting the emergence of Gun Culture 2.0 in Guns magazine, 1955–2019 // Palgrave Communications. 2020. V. 6. N 1. P. 61. https://doi.org/10.1057/s41599-020-0437-0
3. Hinds J., Williams E.J., Joinson A.N. “It wouldn’t happen to me”: Privacy concerns and perspectives following the Cambridge Analytica scandal // International Journal of Human Computer Studies. 2020. V. 143. P. 102498. https://doi.org/10.1016/j.ijhcs.2020.102498
4. Yu X., Yang Q., Wang R., Fang R., Deng M. Data cleaning for personal credit scoring by utilizing social media data: An empirical study // IEEE Intelligent Systems. 2020. V. 35. N 2. P. 7–15. https://doi.org/10.1109/MIS.2020.2972214
5. Óskarsdóttir M., Bravo C., Sarraute C., Vanthienen J., Baesens B. The value of big data for credit scoring: Enhancing financial inclusion using mobile phone data and social network analytics // Applied Soft Computing Journal. 2019. V. 74. P. 26–39. https://doi.org/10.1016/j.asoc.2018.10.004
6. Guo G., Zhu F., Chen E., Liu Q., Wu L., Guan C. From footprint to evidence: An exploratory study of mining social data for credit scoring // ACM Transactions on the Web. 2016. V. 10. N 4. P. 1–38. https://doi.org/10.1145/2996465

References

1. Camacho D., Panizo-LLedot Á., Bello-Orgaz G., Gonzalez-Pardo A., Cambria E. The four dimensions of social network analysis: An overview of research methods, applications, and software tools. *Information Fusion*, 2020, vol. 63, pp. 88–120. https://doi.org/10.1016/j.inffus.2020.05.009
2. Yamane D., Yamane P., Ivory S.L. Targeted advertising: Documenting the emergence of Gun Culture 2.0 in Guns magazine, 1955–2019. *Palgrave Communications*, 2020, vol. 6, no. 1, pp. 61. https://doi.org/10.1057/s41599-020-0437-0
3. Hinds J., Williams E.J., Joinson A.N. “It wouldn’t happen to me”: Privacy concerns and perspectives following the Cambridge Analytica scandal. *International Journal of Human Computer Studies*, 2020, vol. 143, pp. 102498. https://doi.org/10.1016/j.ijhcs.2020.102498
4. Yu X., Yang Q., Wang R., Fang R., Deng M. Data cleaning for personal credit scoring by utilizing social media data: An empirical study. *IEEE Intelligent Systems*, 2020, vol. 35, no. 2, pp. 7–15. https://doi.org/10.1109/MIS.2020.2972214
5. Óskarsdóttir M., Bravo C., Sarraute C., Vanthienen J., Baesens B. The value of big data for credit scoring: Enhancing financial inclusion using mobile phone data and social network analytics. *Applied Soft Computing Journal*, 2019, vol. 74, pp. 26–39. https://doi.org/10.1016/j.asoc.2018.10.004
6. Guo G., Zhu F., Chen E., Liu Q., Wu L., Guan C. From footprint to evidence: An exploratory study of mining social data for credit scoring. *ACM Transactions on the Web*, 2016, vol. 10, no. 4, pp. 1–38. https://doi.org/10.1145/2996465

7. Азаров А.А., Тулупьева Т.В., Суворова А.В., Тулупьев А.Л., Абрамов М.В., Юсупов Р.М. Социоинженерные атаки. Проблемы анализа. СПб.: Наука, 2016. 349 с.
8. Абрамов М.В., Тулупьева Т.В., Тулупьев А.Л. Социоинженерные атаки: социальные сети и оценки защищенности пользователей. СПб.: ГУАП, 2018. 266 с.
9. Cinelli M., Quattrociocchi W., Galeazzi A., Valensise C.M., Brugnoli E., Schmidt A.L., Zola P., Zollo F., Scala A. The COVID-19 social media infodemic // Scientific Reports. 2020. V. 10. P. 16598. https://doi.org/10.1038/s41598-020-73510-5
10. Khlobystova A.O., Abramov M.V., Tulupyev A.L. Soft estimates for social engineering attack propagation probabilities depending on interaction rates among instagram users // Studies in Computational Intelligence. 2020. V. 868. P. 272–277. https://doi.org/10.1007/978-3-030-32258-8_32
11. Oliseenko V., Korepanova A. How old users are? Community analysis // CEUR Workshop Proceedings. 2020. V. 2782. P. 246–251.
12. Хлобыстова А.О., Абрамов М.В., Тулупьев А.Л., Золотин А.А. Поиск кратчайшей траектории социоинженерной атаки между парой пользователей в графе с вероятностями переходов // Информационно-управляющие системы. 2018. № 6. С. 74–81. https://doi.org/10.31799/1684-8853-2018-6-74-81
13. Корепанова А.А., Абрамов М.В., Тулупьева Т.В. Идентификация аккаунтов пользователей в социальных сетях «ВКонтакте» и «Одноклассники» // Семнадцатая Национальная конференция по искусственному интеллекту с международным участием. КИИ-2019: сборник научных трудов. в 2-х томах. Т. 2. 2019. С. 153–163.
14. Корепанова А.А., Тулупьева Т.В. Идентификация аккаунтов пользователя в различных социальных сетях по социальному окружению // Информационная безопасность регионов России (ИБРР-2019): материалы конференции. СПб., 2019. С. 442–443.
15. Liu J., Zhang F., Song X., Song Y.-I., Lin C.-Y., Hon H.-W. What's in a name? An unsupervised approach to link users across communities // Proc. of the 6th ACM International Conference on Web Search and Data Mining (WSDM). 2013. P. 495–504. https://doi.org/10.1145/2433396.2433457
16. Zafarani R., Liu H. Connecting users across social media sites: a behavioral-modeling approach // Proc. of the 19th ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD). 2013. P. 41–49. https://doi.org/10.1145/2487575.2487648
17. Zhang H., Kan M., Liu Y., Ma S. Online social network profile linkage // Lecture Notes in Computer Science. 2014. V. 8870. P. 197–208. https://doi.org/10.1007/978-3-319-12844-3_17
18. Mu X., Zhu F., Lim E., Xiao J., Wang J., Zhou Z. User identity linkage by latent user space modelling // Proceedings of the 22nd ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD). 2016. P. 1775–1784. https://doi.org/10.1145/2939672.2939849
19. Nie Y., Jia Y., Li S., Zhu X., Li A., Zhou B. Identifying users across social networks based on dynamic core interests // Neurocomputing. 2016. V. 210. P. 107–115. https://doi.org/10.1016/j.neucom.2015.10.147
20. Riederer C.J., Kim Y., Chaintreau A., Korula N., Lattanzi S. Linking users across domains with location data: Theory and validation // Proc. of the 25th International Conference on World Wide Web (WWW). 2016. P. 707–719. ttps://doi.org/10.1145/2872427.2883002
21. Chen X., Song X., Cui S., Gan T., Cheng Z., Nie L. User identity linkage across social media via attentive time-aware user modeling // IEEE Transactions on Multimedia. 2020. in press. https://doi.org/10.1109/TMM.2020.3034540
22. Nurgaliev I., Qu Q., Bamakan S.M.H., Muzammal M. Matching user identities across social networks with limited profile data // Frontiers of Computer Science. 2020. V. 14. N 6. P. 146809. https://doi.org/10.1007/s11704-019-8235-9
23. Li Y., Su Z., Yang J., Gao C. Exploiting similarities of user friendship networks across social networks for user identification // Information Sciences. 2020. V. 506. P. 78–98. https://doi.org/10.1016/j.ins.2019.08.022
24. Ma T., Guo L., Wang X., Qian Y., Tian Y., Al-Nabhan N. Friend closeness based user matching cross social networks // Mathematical Biosciences and Engineering. 2021. V. 18. N 4. P. 4264–4292. https://doi.org/10.3934/mbe.2021214
25. Dalal N., Triggs B. Histograms of oriented gradients for human detection // Proc. of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). 2005. V. 1. P. 886–893. https://doi.org/10.1109/CVPR.2005.177

7. Azarov A.A., Tulupeva T.V., Suvorova A.V., Tulupev A.L., Abramov M.V., Iusupov R.M. *Social Engineering Attacks: The Problems of Analysis*. St. Petersburg, Nauka Publ., 2016, 349 p. (in Russian)
8. Abramov M.V., Tulupeva T.V., Tulupev A.L. *Social Engineering Attacks: Social Media and Users Security Estimates*. St. Petersburg, SUAI Publ., 2018, 266 p. (in Russian)
9. Cinelli M., Quattrociocchi W., Galeazzi A., Valensise C.M., Brugnoli E., Schmidt A.L., Zola P., Zollo F., Scala A. The COVID-19 social media infodemic. *Scientific Reports*, 2020, vol. 10, pp. 16598. https://doi.org/10.1038/s41598-020-73510-5
10. Khlobystova A.O., Abramov M.V., Tulupyev A.L. Soft estimates for social engineering attack propagation probabilities depending on interaction rates among instagram users. *Studies in Computational Intelligence*, 2020, vol. 868, pp. 272–277. https://doi.org/10.1007/978-3-030-32258-8_32
11. Oliseenko V., Korepanova A. How old users are? Community analysis. *CEUR Workshop Proceedings*, 2020, vol. 2782, pp. 246–251.
12. Khlobystova A.O., Abramov M.V., Tulupyev A.L., Zolotin A.A. Search for the shortest trajectory of a social engeneering attack between a pair of users in a graph with transition probabilities. *Informatsionno-Upravliaiushchie Sistemy*, 2018, no. 6, pp. 74–81. (in Russian). https://doi.org/10.31799/1684-8853-2018-6-74-81
13. Korepanova A.A., Abramov M.V., Tulupeva T.V. Identification of user accounts on the social networks Vkontakte and Odnoklassniki. *Seventeenth National Conference on Artificial Intelligence with international participation (KII-2019). Proceedings in 2 volumes.* Vol. 2. 2019, pp. 153–163. (in Russian)
14. Korepanova A.A., Tulupieva T.V. User identification across different social networks through social circles. *Proceedings of the Conference Information Security of Russian Regions (ISRR-2019)*, St. Petersburg, 2019, pp. 442–443. (in Russian)
15. Liu J., Zhang F., Song X., Song Y.-I., Lin C.-Y., Hon H.-W. What's in a name? An unsupervised approach to link users across communities. *Proc. of the 6th ACM International Conference on Web Search and Data Mining (WSDM)*, 2013, pp. 495–504. https://doi.org/10.1145/2433396.2433457
16. Zafarani R., Liu H. Connecting users across social media sites: a behavioral-modeling approach. *Proc. of the 19th ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD*), 2013, pp. 41–49. https://doi.org/10.1145/2487575.2487648
17. Zhang H., Kan M., Liu Y., Ma S. Online social network profile linkage. *Lecture Notes in Computer Science*, 2014, vol. 8870, pp. 197–208. https://doi.org/10.1007/978-3-319-12844-3_17
18. Mu X., Zhu F., Lim E., Xiao J., Wang J., Zhou Z. User identity linkage by latent user space modelling. *Proceedings of the 22nd ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD)*, 2016, pp. 1775–1784. https://doi.org/10.1145/2939672.2939849
19. Nie Y., Jia Y., Li S., Zhu X., Li A., Zhou B. Identifying users across social networks based on dynamic core interests. *Neurocomputing*, 2016, vol. 210, pp. 107–115. https://doi.org/10.1016/j.neucom.2015.10.147
20. Riederer C.J., Kim Y., Chaintreau A., Korula N., Lattanzi S. Linking users across domains with location data: Theory and validation. *Proc. of the 25th International Conference on World Wide Web (WWW)*, 2016, pp. 707–719. ttps://doi.org/10.1145/2872427.2883002
21. Chen X., Song X., Cui S., Gan T., Cheng Z., Nie L. User identity linkage across social media via attentive time-aware user modeling. *IEEE Transactions on Multimedia*, 2020, in press. https://doi.org/10.1109/TMM.2020.3034540
22. Nurgaliev I., Qu Q., Bamakan S.M.H., Muzammal M. Matching user identities across social networks with limited profile data. *Frontiers of Computer Science*, 2020, vol. 14, no. 6, pp. 146809. https://doi.org/10.1007/s11704-019-8235-9
23. Li Y., Su Z., Yang J., Gao C. Exploiting similarities of user friendship networks across social networks for user identification. *Information Sciences*, 2020, vol. 506, pp. 78–98. https://doi.org/10.1016/j.ins.2019.08.022
24. Ma T., Guo L., Wang X., Qian Y., Tian Y., Al-Nabhan N. Friend closeness based user matching cross social networks. *Mathematical Biosciences and Engineering*, 2021, vol. 18, no. 4, pp. 4264–4292. https://doi.org/10.3934/mbe.2021214
25. Dalal N., Triggs B. Histograms of oriented gradients for human detection. *Proc. of the 2005 IEEE Computer Society Conference on*

26. Schubert E., Sander J., Ester M., Kriegel H.-P., Xu X. DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN // ACM Transactions on Database Systems. 2017. V. 42. N 3. P. 19. https://doi.org/10.1145/3068335
27. Rymarczyk T., Kozłowski E., Kłosowski G., Niderla K. Logistic regression for machine learning in process tomography // Sensors. 2019. V. 19. N 15. P. 3400. https://doi.org/10.3390/s19153400
28. Олисеенко В.Д., Абрамов М.В., Тулупьев А.Л. Идентификация аккаунтов пользователей при помощи сравнения изображений: подход на основе phash // Научно-технический вестник информационных технологий, механики и оптики. 2021. Т. 21. № 4. С. 562–570. https://doi.org/10.17586/2226-1494-2021-21-4-562-570
29. Brigham E.O. The Fast Fourier Transform. New York, USA: Prentice-Hall, 2002.
30. MacKay D.J.C. Information Theory, Inference, and Learning Algorithms. Cambridge: Cambridge University Press, 2003. 628 p.
31. Воронцов К.В. Комбинаторный подход к оценке качества обучаемых алгоритмов // Математические вопросы кибернетики. 2004. Т. 13. С. 5–36.

Computer Vision and Pattern Recognition (CVPR'05), 2005, vol. 1, pp. 886–893. https://doi.org/10.1109/CVPR.2005.177
26. Schubert E., Sander J., Ester M., Kriegel H.-P., Xu X. DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN. *ACM Transactions on Database Systems*, 2017, vol. 42, no. 3, pp. 19. https://doi.org/10.1145/3068335
27. Rymarczyk T., Kozłowski E., Kłosowski G., Niderla K. Logistic regression for machine learning in process tomography. *Sensors*, 2019, vol. 19, no. 15, pp. 3400. https://doi.org/10.3390/s19153400
28. Oliseenko V.D., Abramov M.V., Tulupyev A.L. Identification of user accounts by image comparison: The phash-based approach. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2021, vol. 21, no. 4, pp. 562–570. (in Russian). https://doi.org/10.17586/2226-1494-2021-21-4-562-570
29. Brigham E.O. *The Fast Fourier Transform*. New York, USA, Prentice-Hall, 2002.
30. MacKay D.J.C. *Information Theory, Inference, and Learning Algorithms*. Cambridge, Cambridge University Press, 2003, 628 p.
31. Vorontcov K.V. A combinatorial approach to assessing the quality of training algorithms. *Matematicheskie voprosy kibernetiki*, 2004, vol. 13, pp. 5–36. (in Russian)

Авторы

Корепанова Анастасия Андреевна — младший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация; студент, Санкт-Петербургский государственный университет, Санкт-Петербург, 199034, Российская Федерация, SC 57218191916, https://orcid.org/0000-0003-2962-8670, aak@mail.ru

Абрамов Максим Викторович — кандидат технических наук, руководитель лаборатории, старший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация; доцент, Санкт-Петербургский государственный университет, Санкт-Петербург, 199034, Российская Федерация, SC 56938320500, https://orcid.org/0000-0002-5476-3025, mva@dscs.pro

Тулупьев Александр Львович — доктор физико-математических наук, профессор, профессор, Санкт-Петербургский государственный университет, Санкт-Петербург, 199034, Российская Федерация; главный научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация, SC 13608565400, https://orcid.org/0000-0003-1814-4646, alt@dscs.pro

Authors

Anastasia A. Korepanova — Junior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation; Student, Saint Petersburg State University, Saint Petersburg, 199034, Russian Federation, SC 57218191916, https://orcid.org/0000-0003-2962-8670, aak@mail.ru

Maxim V. Abramov — PhD, Head of Laboratory, Senior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation; Associate Professor, Saint Petersburg State University, Saint Petersburg, 199034, Russian Federation, SC 56938320500, https://orcid.org/0000-0002-5476-3025, mva@dscs.pro

Alexander L. Tulupyev — D.Sc., Full Professor, Saint Petersburg State University, Saint Petersburg, 199034, Russian Federation; Chief Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation, SC 13608565400, https://orcid.org/0000-0003-1814-4646, alt@dscs.pro

Статья поступила в редакцию 19.09.2021
Одобрена после рецензирования 27.10.2021
Принята к печати 29.11.2021

Received 19.09.2021
Approved after reviewing 27.10.2021
Accepted 29.11.2021