

УДК 004.021

DOI: [10.26102/2310-6018/2020.28.1.012](https://doi.org/10.26102/2310-6018/2020.28.1.012)

Разработка программы оценки времени размещения сообщения в онлайн-социальной сети ВКонтакте

Е.Б. Грибанова, А.С. Савицкий

*Томский государственный университет систем управления и радиоэлектроники,
Томск, Российская Федерация*

Резюме: Актуальность исследования обусловлена высокой популярностью социальных сетей для передачи информации, а также влиянием времени размещения сообщения на число его просмотров и соответственно степень его распространения. В статье представлено описание программы для оценки времени размещения сообщения в группах социальной сети ВКонтакте. Представлены алгоритмы сбора и обработки используемых данных. Приводится описание используемых запросов к API ВКонтакте с помощью стандартных методов. Разработанная программа позволяет учитывать показатели активности участников и их индивидуальные характеристики (число друзей, групп, записей на стене и т. д.), а также число сообщений, публикуемых другими участниками социальной сети. Для получения интегральной характеристики была применена линейная свёртка критериев. Реализация программы выполнена с помощью языка C#, для хранения собранных данных и результатов их обработки был использован Табличный процессор Microsoft Excel. Рассмотрен пример оценки времени размещения сообщения в выбранной группе социальной сети ВКонтакте. Разработанная программа может быть использована администраторами сообществ социальной сети для оценки времени размещения в группах онлайн-социальной сети и выбора наилучшего момента публикации.

Ключевые слова: социальные сети, время размещения сообщений, новостная лента, активность пользователей.

Для цитирования: Грибанова Е.Б., Савицкий А.С. Разработка программы оценки времени размещения сообщения в онлайн-социальной сети ВКонтакте. *Моделирование, оптимизация и информационные технологии*. 2020;8(1). Доступно по: https://moit.vivt.ru/wp-content/uploads/2020/02/GribanovaSavitsky_1_20_1.pdf DOI: 10.26102/2310-6018/2020.28.1.012

Development of a program to estimate the time of posting a message in the online social network VKontakte

E.B. Gribanova, A.S. Savitsky

Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russian Federation

Abstract: The relevance of the study is due to the high popularity of social networks for the transmission of information, as well as the influence of the time of posting a message on the number of views and, accordingly, the degree of its distribution. The article presents a description of the program for estimating the time of posting messages in groups of the social network Vkontakte. Algorithms of data collection and processing are presented. The description of used requests to VK API by means of standard methods is given. The developed program takes into account the activity indicators of participants and their individual characteristics (the number of friends, groups, posts on the wall, etc.), as well as the number of messages published by other members of the social network. Linear convolution was applied to obtain the integral characteristic. The program is implemented using the C# language. Microsoft Excel spreadsheet processor was used to store data about subscribers and results. An example of estimating the time of posting a message in the selected group of the social network Vkontakte is

considered. The developed program can be used by administrators of social network communities to estimate the time of posting in groups of the online social network and to choose the best moment of publication.

Keywords: social networks, time of post, news feed, user activity.

For citation: Griбанова Е.В., Savitsky A.S. Development of a program to estimate the time of posting a message in the online social network VKontakte. *Modeling, Optimization and Information Technology*. 2020;8(1). Available from: https://moit.vivt.ru/wp-content/uploads/2020/02/GriбановаSavitsky_1_20_1.pdf DOI: 10.26102/2310-6018/2020.28.1.012 (In Russ).

Введение

Онлайновые социальные сети сегодня включают большое число участников и являются популярными площадками для распространения информации социального, экономического и политического характера [1–2]. Степень распространения информации определяется многими факторами: структурой сети, влиянием участников, публикующих сообщение [3–6], характеристиками самого поста (его интерес для аудитории) и т. д. Одним из важнейших факторов, определяющих число просмотров сообщения, является время его размещения. Наилучшим временем размещения сообщения считается тот момент, когда большее число пользователей находится в статусе «онлайн», так как в этом случае они с высокой вероятностью могут увидеть эту информацию. Использование показателей активности участников для определения наилучшего момента размещения сообщения рассматривается в существующих работах.

Авторами [7] предложена методика, согласно которой каждому пользователю присваивается некоторый балл в зависимости от числа его подписчиков, времени, прошедшего после последнего размещения сообщения и взаимодействия с другими пользователями, и определяется наилучшее время с точки зрения максимального суммарного балла.

В статье [8] приводятся данные о количестве пользователей социальных сетей Orkut, MySpace, Hi5 и LinkedIn, находящихся в статусе «онлайн» в различные моменты времени. Сбор данных осуществлялся в течение 12 дней, пиковые значения числа пользователей в статусе «онлайн» наблюдались в 15 часов (рассматривались участники из Бразилии).

Также различными компаниями, осуществляющими коммерческую деятельность в области социальных медиа (Oberlo, SproutSocial, Later, Hotsuite и т. д.), проводятся исследования и разрабатываются программы для размещения сообщений в сети [9] и определения наилучшего момента публикации. Среди существующих систем автоматического размещения информации в социальных сетях можно отметить NovaPress, AMPLIFR, SMMplanner и т. д., которые осуществляют публикацию сообщений согласно заданному пользователем расписанию. Некоторые разработки также обладают функциями определения наилучшего времени размещения. Например, Best Time to Post (Later) позволяет определить наилучшее время размещения в Instagram на основе анализа активности подписчиков.

Таким образом, существующие работы основаны на использовании информации об активностях пользователей для определения наилучшего времени публикации (при этом в существующих программных продуктах принцип построения оценки является скрытым от пользователя). Однако при определении момента времени размещения сообщения необходимо учитывать и другие факторы. Так, результаты проведенного

опроса [10] свидетельствуют о том, что большая часть пользователей (48,82%) просматривает только часть новостной ленты, содержащей опубликованные сообщения сообществ и участников, на которые подписан пользователь. Вследствие этого важную роль играет скорость обновления новостной ленты, которая определяется числом публикуемых сообщений участниками. Если скорость новостной ленты очень высока, то опубликованное участником сообщение может быть смещено другими постами и не прочитано.

Поскольку активность пользователей зависит от их места расположения и образа жизни для различных сообществ наилучшее время размещения сообщения будет отличаться. В связи с этим возникает необходимость разработать программу, позволяющую выполнять обработку данных выбранного сообщества. Кроме того, необходимость разработки программы обусловлена большим объемом обрабатываемых данных.

Цель данной работы заключается в разработке программы, позволяющей оценивать моменты времени размещения сообщений в социальной сети с использованием как характеристик пользователей (в том числе периодов их индивидуальной активности), так и информации о публикации сообщений другими участниками. При этом будут использованы данные социальной сети ВКонтакте, имеющей наибольшую аудиторию среди социальных сетей, используемых в России.

Разработка программы

Реализация программы выполнена на языке C#, при этом были использованы стандартные методы API ВКонтакте: `groups.getMembers`, `users.get`, `wall.get`, `newsfeed.get` и `execute`. Результаты сбора и обработки данных хранятся в Excel документе.

Работа с программой включает сбор сведений о числе сообщений, публикуемых в новостной ленте, который осуществляется через заданный промежуток времени. В результате определяется число публикуемых сообщений за установленный интервал. При формировании оценки могут быть использованы как значения, собранные ранее и заложенные в программу (базовая оценка, вычисленная на основе данных 300 групп, собранных в течение недели), так и выполнен новый сбор данных (пользовательская оценка).

После этого выполняется построение интегральной оценки с помощью следующих этапов:

Этап 1. Сбор данных о статусе пользователей с заданным интервалом (программа автоматически собирает сведения: `id` участника и статус).

Этап 2. Сбор данных о характеристиках участников: `id`, число друзей, число лайков, репостов и комментариев за последний месяц на странице пользователя, число сообществ участника.

Данные этапа 1 и 2 выгружаются в Excel и далее выполняется их последующая обработка.

Этап 3. Расчет характеристик индивидуальных периодов активности пользователей (на основе данных этапа 1). Для каждого пользователя в исследуемые моменты времени вычисляется величина, принимающая значение от 0 до 1 и характеризующая число периодов до момента, когда пользователь будет в статусе «онлайн» (чем ближе величина к единице, тем ближе к моменту, когда участник будет в статусе «онлайн»).

Этап 4. Вычисление суммарной оценки на основе характеристик участников (с использованием данных этапа 2 и 3). Для каждого участника выполняется линейная свертка трех показателей: число друзей, активность участника (сумма лайков,

репостов и комментариев) и количество групп [11]. Полученное значение корректируется с учетом характеристики индивидуальной активности и определяется суммарная величина для всего сообщества.

Этап 5. Выполнение линейной свертки критериев (на основе данных этапа 4 и характеристики обновления новостной ленты) и вычисление интегральной оценки момента размещения сообщения. Момент времени, для которого оценка будет максимальной считается наилучшим. При осуществлении линейной свертки может быть привлечена экспертная информация и установлены коэффициенты важности отдельных показателей.

Сбор данных осуществляется с помощью запросов к API ВКонтакте. Рассмотрим алгоритм сбора данных о статусе пользователей.

Шаг 1. Определение начальных параметров: id группы, время окончания сбора t_n , текущее время t_c , интервал сбора i .

Шаг 2. Проверка условия: если $t_c \Rightarrow t_n$, то работа алгоритма завершается.

Шаг 3. Запуск стандартного метода execute (в связи с ограничением на число запросов к API ВКонтакте в секунду и для ускорения работы программы запросы выполняются с использованием метода execute, который позволяет запускать последовательность других методов).

Получение количества подписчиков группы NumUsers с использованием запроса getMembers:

```
var R = API.groups.getMembers({ "group_id": id, "sort": "id_asc", "offset": 0, "count": 1000});  
NumUsers = R.count;
```

Определение числа итераций, в течение которых будут собраны данные пользователей (за один запрос может быть получена информация о 1000 участниках):

```
step = NumUsers / 1000;
```

Далее в цикле выполняется сбор данных об участниках. Переменные offset (сдвиг) и numbers (число запрашиваемых участников) определяются в связи с ограничением на размер максимального запроса и позволяют выполнять перебор всех участников:

```
var R = API.groups.getMembers({ "group_id": id, "sort": "id_asc", "offset":  
offset, "count": numbers });  
a = R.items;
```

```
users = users + API.users.get({ "user_ids": a, "fields": "online" });
```

Шаг 4. Запись в документ Excel полученных данных (id и статус участников).

Шаг 5. Остановка выполнения программы в течение заданного интервала i .
Возврат на шаг 2.

Для получения расширенной информации об участнике (число групп, друзей и т. д.) выполняется тот же запрос (в цикле перебора участников с переменной цикла j) с дополнительным параметром counters, при этом осуществляется передача методу get идентификатора одного пользователя:

```
b = R.items[j];
```

```
users = users + API.users.get({ "user_ids" : b, fields : "online, counters" });
```

Сбор данных о числе лайков, репостов и комментариев на странице пользователя осуществляется путем циклического перебора набора записей, полученного с помощью запроса wall.get.

```
var W = API.wall.get({ "owner_id": id, "count": 100, "offset": offset });
```

Алгоритм сбора сведений о числе сообщений в новостной ленте включает следующие шаги.

Шаг 1. Определение начальных параметров: время окончания сбора t_n ,

текущее время t_c , интервал сбора k .

Шаг 2. Проверка условия: если $t_c \Rightarrow t_n$, то работа алгоритма завершается.

Шаг 3. Получение сообщений из новостной ленты (переменная $time$ определяется как текущий момент минус исследуемый интервал, например, для пяти минут: $var\ time = API.getServerTime() - 300$):

```
var W = API.newsfeed.get({"filters": "post", "start_time": time, "count": 100, "start_from": next})
```

Число сообщений в этом запросе равно $W.items.length$.

Переменная $next$ позволяет перейти к следующей странице новостей если число записей равно максимальному значению (100): $next = W.next_from$.

Шаг 4. Запись документ Excel количества постов.

Шаг 5. Остановка выполнения программы на k минут. Возврат на шаг 2.

На Рисунке 1 представлен интерфейс программы.

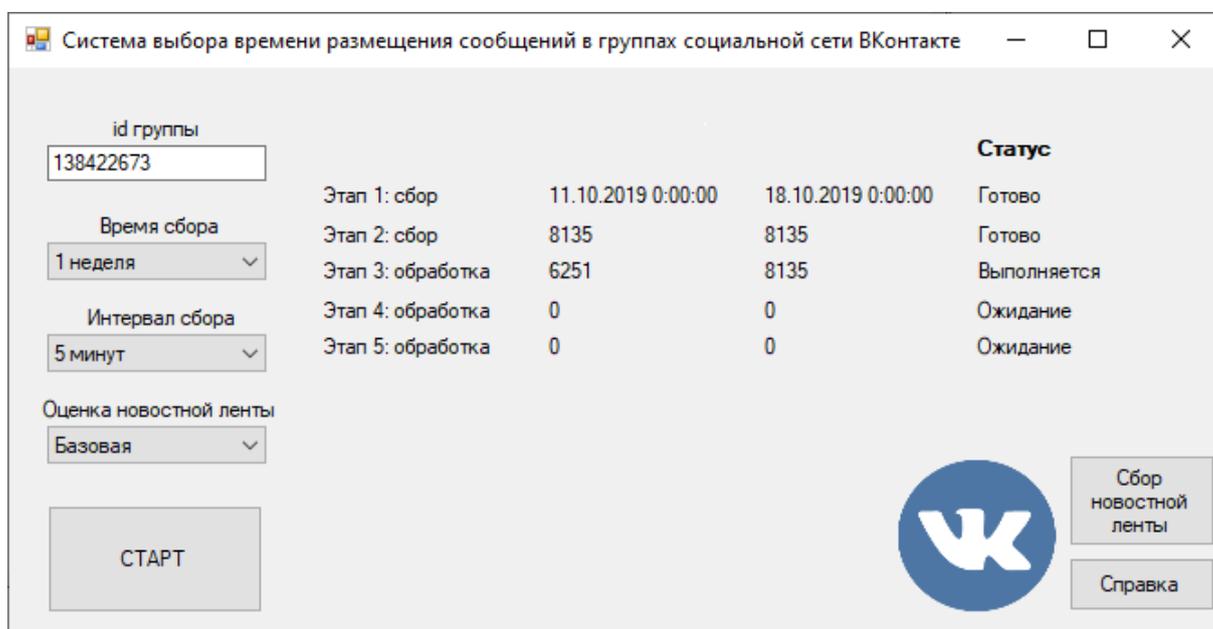


Рисунок 1 – Интерфейс программы
 Figure 1 – Software interface

Результаты определения наилучшего времени размещения сообщения

Для проведения анализа сообщества, была использована информация об участниках группы социальной сети ВКонтакте «Golden time Anime». Программа выполняла автоматический сбор данных об участниках группы через каждые 5 минут.

В Таблице 1 представлены значения средней доли участников, находящихся в статусе «онлайн», для различных дней недели (с использованием данных, собранных в течение двух недель). Наибольшее значение средней доли участников, находящихся в статусе «онлайн» наблюдается в период с 17 до 3 часов.

На Рисунке 2 представлены значения среднего числа публикуемых сообщений в новостной ленте (интервал сбора равен 5 минутам), при этом рассмотрен период с 21:00 до 02:00 часов. Из Рисунка 2 можно сделать вывод, что на начало каждого часа приходятся пиковые значения числа публикуемых сообщений, таким образом, это время является наиболее популярным для размещения сообществами постов.

В Таблице 2 представлены шесть лучших значений интегрального показателя (на момент вычисления оценки в группе насчитывалось 8135 участников, данные о периодах активностях участников были собраны в течение 12 дней). Таким образом, наилучшим моментом размещения сообщения в рассматриваемой группе будет 20:35.

Таблица 1 – Средняя доля участников, находящихся в статусе «онлайн»
 Table 1 – Average percentage of participants who are in the «online» status

Время	День недели						
	Пн.	Вт.	Ср.	Чт.	Пт.	Сб.	Вс.
0:00	0,2189	0,2220	0,2229	0,2218	0,2179	0,2222	0,2054
1:00	0,2143	0,2155	0,2220	0,2176	0,2117	0,2064	0,2173
2:00	0,2059	0,2131	0,2151	0,2188	0,2121	0,2025	0,2009
3:00	0,2001	0,1902	0,1954	0,1938	0,1888	0,1885	0,1885
4:00	0,1720	0,1525	0,1577	0,1579	0,1567	0,1548	0,1665
5:00	0,1311	0,1082	0,1126	0,1078	0,1090	0,1165	0,1401
6:00	0,0891	0,0734	0,0733	0,0722	0,0728	0,0841	0,0991
7:00	0,0603	0,0534	0,0532	0,0560	0,0547	0,0581	0,0649
8:00	0,0537	0,0505	0,0466	0,0469	0,0493	0,0487	0,0493
9:00	0,0474	0,0502	0,0520	0,0491	0,0497	0,0414	0,0442
10:00	0,0565	0,0622	0,0663	0,0610	0,0666	0,0521	0,0447
11:00	0,0713	0,0818	0,0897	0,0910	0,0889	0,0643	0,0558
12:00	0,0943	0,1202	0,1191	0,1183	0,1187	0,0897	0,0688
13:00	0,1163	0,1272	0,1221	0,1307	0,1351	0,1058	0,0966
14:00	0,1309	0,1375	0,1384	0,1453	0,1499	0,1290	0,1304
15:00	0,1603	0,1607	0,1587	0,1576	0,1640	0,1493	0,1595
16:00	0,1789	0,1718	0,1708	0,1736	0,1721	0,1710	0,1743
17:00	0,1792	0,1823	0,1779	0,1792	0,1805	0,1775	0,1921
18:00	0,1919	0,1915	0,1901	0,1797	0,1858	0,1867	0,1975
19:00	0,2007	0,2023	0,1961	0,1962	0,1965	0,1894	0,1965
20:00	0,2060	0,2065	0,1971	0,2045	0,2065	0,1963	0,2034
21:00	0,2118	0,2047	0,2096	0,2102	0,2016	0,1937	0,2078
22:00	0,2179	0,2142	0,2158	0,2208	0,1986	0,2038	0,2143
23:00	0,2193	0,2224	0,2221	0,2162	0,2159	0,2062	0,2203

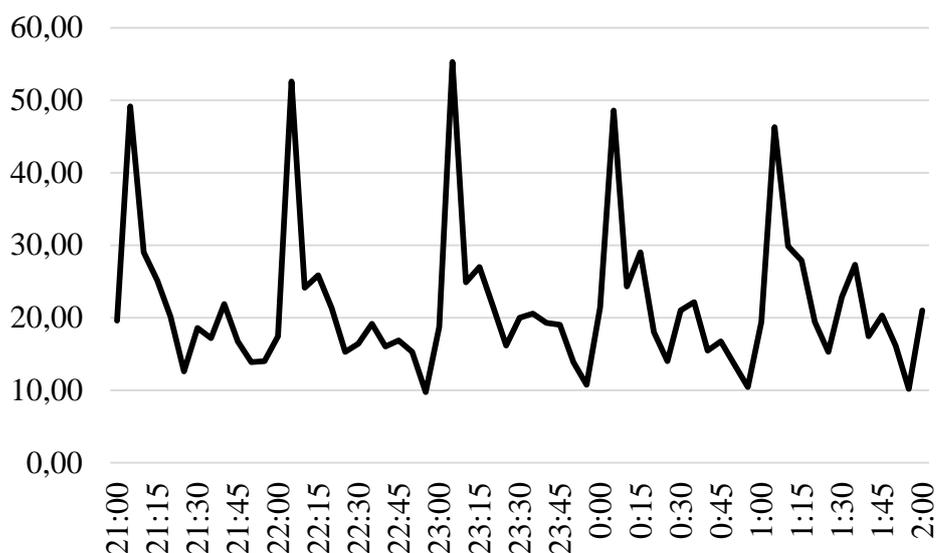


Рисунок 2 – Число публикуемых сообщений в новостной ленте
Figure 2 – Number of published messages in the news feed

Таблица 2 – Наилучшие значения интегрального показателя
Table 2 – Best values of the integral indicator

Время	Интегральный показатель
20:35	0,794286
20:40	0,793982
20:45	0,791246
21:40	0,789898
21:20	0,789467
20:30	0,788878

Заключение

В статье приводится описание программы для оценки времени размещения сообщений в группах онлайн-социальной сети ВКонтакте. Представлены алгоритмы сбора и обработки данных, а также описание используемых запросов к API ВКонтакте с помощью стандартных методов. В отличие от существующих работ разработанная программа позволяет учитывать не только показатели активности участников, но и их индивидуальные характеристики (число друзей, групп, записей на стене и т. д.), а также число сообщений, публикуемых другими участниками социальной сети. Для получения интегральной характеристики была применена линейная свёртка. Кроме того, программа может быть использована для получения оценки на основе использования единичных характеристик (например, число пользователей, находящихся в статусе «онлайн»). Несмотря на то, что такой подход не учитывает все факторы формирования решения о размещении сообщения, однако он требует меньше вычислительных ресурсов за счёт меньшего объема обрабатываемых данных и, следовательно, обеспечивает меньшее время обработки.

Реализация программы выполнена с помощью языка C#, для хранения данных был использован Табличный процессор Microsoft Excel. Рассмотрен пример оценки

времени размещения сообщения в выбранной группе социальной сети ВКонтакте. Разработанная программа может быть использована администраторами групп для оценки времени размещения в группах онлайн-социальной сети и выбора наилучшего момента публикации.

ЛИТЕРАТУРА

1. Goyal S., Gagnon J. Social networks and the firm. *Revista de Administracao*. 2016;51(2):240-243.
2. Грибанова Е.Б., Катасонова А.В. Модель оценки групп социальной сети для реализации маркетинговых мероприятий. *Доклады Томского государственного университета систем управления и радиоэлектроники*. 2017;20(2):68-72.
3. Bakshy E., Hofman J., Mason W., Watts D. Everyone's an Influencer: Quantifying Influence on Twitter. *Proceedings of the 4-th International Conference on Web Search and Web Data Mining*. Hong Kong. 2011:1-10.
4. Bhagat S., Goyal A., Lakshmanan, L. Maximizing product adoption in social. *Proceedings of the 5-th ACM International Conference on Web Search and Data Mining*. Seattle. 2012: 603-612.
5. Chen W., Collins A., Cummings R., Ke T., Liu Z., Rincon D., Sun X., Wei W., Wang Y., Yuan Y. Influence maximization in social networks when negative opinions may emerge and propagate. *Proceedings of the 2011 SIAM International Conference on Data Mining*. Mesa. 2011:379-390.
6. Грибанова Е.Б., Логвин И.Н., Ширенков И.В. Алгоритм оценки маркетинговых мероприятий онлайн-социальной сети «ВКонтакте» на основе каскадной модели распространения информации. *Доклады Томского государственного университета систем управления и радиоэлектроники*. 2018;21(3):66-73.
7. Booth J.A., Howard J., Rankin A. System and Methods for Generating Optimal Post Times for Social Networking Sites. *United States Patent*. US 0275348 A1. 2013:1-8.
8. Benevenuto F., Rodrigues T., Cha M., Almeida V. Characterizing User Behavior in Online Social Networks. *Proceedings of the 9th ACM conference on Internet measurement*. Chicago. 2009:49-62.
9. Chumak A.A, Ukustov S.S, Kravets A.G., Voronin, J.F. Social Networks Message Posting Support Module. *World Applied Sciences Journal*. 2013;24(24):191-195.
10. Савицкий А.С. Модель оценки времени размещения сообщений в группах социальной сети с учетом скорости обновления новостной ленты. *Материалы международной научно-технической конференции студентов, аспирантов и молодых учёных «Научная сессия ТУСУР-2019»*. Часть 3. Томск. 2018:114-116.
11. Савицкий А.С. Модель выбора времени размещения сообщений в группах онлайн-социальной сети. *Сборник научных трудов научной конференции «Информационные технологии в науке, управлении, социальной сфере и медицине»*. Часть 2. Томск. 2018:90-92.

REFERENCES

1. Goyal S., Gagnon J. Social networks and the firm. *Revista de Administracao*. 2016;51(2):240-243.
2. Gribanova E.B., Katasonova A.V. System to evaluate social network groups for the implementation of marketing activities. *Proceedings of TUSUR*. 2017;20(2):68-72.

3. Bakshy E., Hofman J., Mason W., Watts D. Everyone's an Influencer: Quantifying Influence on Twitter. *Proceedings of the 4-th International Conference on Web Search and Web Data Mining*. Hong Kong. 2011:1-10.
4. Bhagat S., Goyal A., Lakshmanan, L. Maximizing product adoption in social. *Proceedings of the 5-th ACM International Conference on Web Search and Data Mining*. Seattle. 2012: 603-612.
5. Chen W., Collins A., Cummings R., Ke T., Liu Z., Rincon D., Sun X., Wei W., Wang Y., Yuan Y. Influence maximization in social networks when negative opinions may emerge and propagate. *Proceedings of the 2011 SIAM International Conference on Data Mining*. Mesa. 2011:379-390.
6. Griбанова Е.В., Логвин И.Н., Ширенков И.В. Algorithm for evaluating the marketing activities of the online social network Vkontakte based on the cascade model of information dissemination. *Proceedings of TUSUR*. 2018;21(3):66-73.
7. Booth J.A., Howard J., Rankin A. System and Methods for Generating Optimal Post Times for Social Networking Sites. *United States Patent*. US 0275348 A1. 2013:1-8.
8. Benevenuto F., Rodrigues T., Cha M., Almeida V. Characterizing User Behavior in Online Social Networks. *Proceedings of the 9th ACM conference on Internet measurement*. Chicago. 2009:49-62.
9. Chumak A.A, Ukustov S.S, Kravets A.G., Voronin, J.F. Social Networks Message Posting Support Module. *World Applied Sciences Journal*. 2013;24(24):191-195.
10. Savitskii A.S. Model' otsenki vremeni razmeshcheniya soobshchenii v gruppakh sotsial'noi seti s uchetom skorosti obnovleniya novostnoi lenty. *Materialy mezhdunarodnoi nauchno-tekhnicheskoi konferentsii studentov, aspirantov i molodykh uchenykh «Nauchnaya sessiya TUSUR-2019»*. Chast' 3. Tomsk. 2018:114-116.
11. Savitskii A.S. Model' vybora vremeni razmeshcheniya soobshchenii v gruppakh onlainovoi sotsial'noi seti. *Sbornik nauchnykh trudov nauchnoi konferentsii «Informatsionnye tekhnologii v nauke, upravlenii, sotsial'noi sfere i meditsine»*. Chast' 2. Tomsk. 2018:90-92.

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Грибанова Екатерина Борисовна, канд. техн. наук, доцент, кафедра автоматизированных систем управления, Томский государственный университет систем управления и радиоэлектроники, Томск, Российская Федерация.

e-mail: geb@asu.tusur.ru

ORCID: [0000-0001-6499-5893](https://orcid.org/0000-0001-6499-5893)

Савицкий Александр Сергеевич, магистрант, кафедра автоматизированных систем управления, Томский государственный университет систем управления и радиоэлектроники, Томск, Российская Федерация.

e-mail: bourbon7850@gmail.com

Ekaterina B. Griбанова, Candidate of Sciences in Engineering, Associate Professor, Department of Automated Control Systems, Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russian Federation

Alexandr S. Savitsky, Master student, Department of Automated Control Systems, Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russian Federation