Научная статья
УДК_004.624
DOI: 10.17072/1993-0550-2022-4-82-88

Выявление признаков информационных операций на основе анализа начальной частоты публикации дубликатов

Андрей Николаевич Рабчевский[1,2], Михаил Юрьевич Карпов[1], Евгений Георгиевич Ашихмин[2]
[1]Пермский государственный национальный исследовательский университет, Пермь, Россия
[2]ООО "СЕУСЛАБ", Пермь, Россия
[1]andrey@ranat.ru, AuthorID 1089772
[2]karpov.mu@psu.ru
[3]e.ashikhmin@seuslab.ru, https://orcid.org/0000-0001-9193-4535

Аннотация. В данной работе авторы проанализировали информационные потоки деструктивного контента в социальной сети ВКонтакте и выявили высокую начальную частоту публикации четких дубликатов как некое общее свойство, присущее информационным операциям и предложили использовать это свойство как признак информационных операций. Последующие аналитические исследования показали, что методика, использующая этот признак, позволяет выявлять информационные операции на самых ранних стадиях, что приводит к сокращению времени на принятие решения о мерах противодействия выявленным информационным операциям и повысить эффективность информационного противодействия.

Ключевые слова: *информационные войны; информационные операции; выявление признаков; дубликаты; частота публикаций*

Для цитирования: *Рабчевский А. Н., Карпов М. Ю., Ашихмин Е. Г.* Выявление признаков информационных операций на основе анализа начальной частоты публикации дубликатов // Вестник Пермского университета. Математика. Механика. Информатика. 2022. Вып. 4(59). С. 82–88. DOI: 10.17072/1993-0550-2022-4-82-88.

Статья поступила в редакцию 21.10.2022; одобрена после рецензирования 11.11.2022; принята к публикации 14.11.2022.

Research article

Identification of Information Operations Signs Based on the Initial Frequency of Duplicate Publications Analysis

Andrey N. Rabchevskiy[1,2], Mikhail Yu. Karpov[1], Evgeniy G. Ashikhmin[2]
[1]Perm State University, Perm, Russia
[2]"SEUSLAB" LLC, Perm, Russia
[1] andrey@ranat.ru, AuthorID 1089772
[2] karpov.mu@psu.ru
[3] e.ashikhmin@seuslab.ru, https://orcid.org/0000-0001-9193-4535

Abstract. Analyzed the information flows of destructive content in the social network VKontakte and revealed a high initial frequency of publication of clear duplicates as a certain common property inherent in information operations and proposed to use this property as a sign of information operations.

Keywords: *information wars; information operations; sign detection; duplicates; frequency of publications*

For citation: *Rabchevskiy A. N., Karpov M. Yu., Ashikhmin E. G.* Identification of Information Operations Signs Based on the Initial Frequency of Duplicate Publications Analysis // Bulletin of Perm University. Mathematics. Mechanics. Computer Science. 2022;4(59):82–88. (In Russ.). DOI: 10.17072/1993-0550-2022-4-82-88.

The article was submitted 21.10.2022; approved after reviewing 11.11.2022; accepted for publication 14.11.2022.

Введение

Развитие современных информационных и телекоммуникационных технологий привело к тому, что противостояние между противоборствующими сторонами на международной арене перешло на уровень информационных войн. Согласно проекту Конвенции об обеспечении информационной безопасности ООН [1], *информационная война,* это "межгосударственное противоборство в информационном пространстве с целью нанесения ущерба информационным системам, процессам и ресурсам, критически важным и другим структурам; для подрыва политической, экономической и социальной систем; массированной психологической обработки населения для дестабилизации общества и государства, а также принуждения государства к принятию решений в интересах противоборствующей стороны".

Основным средством ведения информационных войн являются *информационные операции*, которые по определению, данному в работе [2], есть "действия, предпринимаемые для достижения информационного превосходства в обеспечении национальной военной стратегии путем воздействия на информацию и информационные системы противника с одновременным укреплением и защитой собственной информации и информационных систем и инфраструктуры".

Одним из *объектов воздействия* информационных операций являются общество в целом (как гражданское население, так и личный состав вооруженных сил), его государственные, экономические и социальные институты. Согласно определению, представленному выше [1], такое воздействие производится в виде "*массированной психологической обработки населения для дестабилизации общества и государства*".

Противодействие информационным операциям является одной из основных задач в информационной войне. Основным элементом противодействия является выявление информационных операций.

Таким образом, актуальной является задача разработки методов, алгоритмов и программных средств, предназначенных для выявления информационных операций. В настоящей работе исследуются методы выявления признаков, позволяющих распознавать информационные операции на ранних стадиях.

1. Признаки информационных операций

Изучению наиболее характерных признаков информационных операций посвящено большое количество работ. По мнению [3], "основным признаком информационной операции является повторение близких по смыслу новостей и комментариев новостей, выполненных в соответствующем контексте. При этом совершенно неважно, соответствуют ли эти новости происходящим в эмпирическом мире событиям. Более того, чем меньше похожести, тем больше вероятность того, что новости являются частью спланированной операции". При этом, основными параметрами, которые являются значимыми для выявления информационных операций, по мнению авторов, являются *частота подачи материала*, относящегося к заданной теме, *значимость источника* и *охват населения*.

Основной особенностью информационных операций по утверждению [4] являются "экспрессивный и молниеносный характер действий, организованных с использованием технологий Интернет и создание гигантских "народных големов", охватывающих значительное число недовольных всех спектров", то есть, другими словами, достижение максимального охвата населения в минимальные сроки.

В работе [5] указывается, что основными особенностями информационных операций являются наличие большого количества нечетких дубликатов [6] и небольшой интервал времени их публикации.

Таким образом, по общему мнению, основными признаками информационных операций являются:

– большое количество публикаций близкого по смыслу целевого контента,

– высокая частота публикаций,

– небольшой интервал времени публикаций,

– максимальный охват аудитории.

2. Методы выявления информационных операций

Для выявления или распознавания информационных операций многие исследователи используют классический подход, когда анализируется число сообщений определенной направленности в равные промежутки времени. Если их количество больше порогового значения, то делается вывод о наличии признаков проведения информационной операции. Недостатком такого метода является эмпирический характер оценки этого порогового значения, которое может зависеть от многих факторов.

Наряду с классическим подходом, исследователями часто применяется анализ динамических свойств информационных потоков. Изучению этой проблемы также посвящено большое количество исследований [7–10]. Так, например, для анализа временных рядов, которые отображают зависимость объемов информационных потоков от времени, в работе [11] предлагаются анализ временных рядов как реализацию стохастического процесса, а также корреляционный анализ, анализ Фурье, вейвлет-анализ, корреляция с шаблоном, фрактальный и мультифрактальный анализ. Подобные подходы решают проблему эмпиричности выбора пороговых значений, однако их общий недостаток состоит в том, что для получения временного ряда необходим анализ сообщений за довольно продолжительный промежуток времени, необходимый для построения временного ряда, что не позволяет выявлять информационные операции на ранних стадиях.

В качестве признаков информационных операций в работе [5] используются некоторые метрики центральности и их соотношения для графов, состоящих из сообщений (вершин) и их взаимосвязей (ребер). Однако предложенный метод также не лишен недостатков. На выявление информационной операции требуется существенное время, которое включает время активной фазы информационной операции, а также время, необходимое на сбор и обработку информации.

Кроме того, использование в качестве признака информационной операции именно нечетких дубликатов, во-первых, накладывает повышенные требования к техническим средствам для их выявления, а во-вторых, снижает контрастность признаков.

Следует отметить, что с учетом времени, прошедшего с момента публикации данной работы, вероятно, исследуемая область несколько изменилась, и организаторы современных информационных операций не считают нужным маскироваться. Исследования актуальных информационных операций показывают, что наряду с нечеткими дубликатами в информационных операциях широко используются и "четкие" дубликаты (далее по тексту дубликаты). И наконец, потоки "нечетких" дубликатов суммарно формируют тренд в виде конкретного информационного повода, а потоки четких дубликатов прямо указывают на организованный целенаправленный характер деятельности по их публикации, а значит в большей мере являются признаком информационной операции.

В общем случае, большинство авторов выявляют информационные операции на основании исследования параметров, форм или содержания сообщений в сети, полученных на довольно продолжительном отрезке времени.

В то же время, для успешного противодействия информационным операциям, необходимо их выявление на как можно более ранней стадии.

Если исходить из того, что достижение целей информационной операции возможно только в случае обеспечения максимального охвата аудитории в минимальные сроки [3, 4], то логично предположить, что при проведении информационной операции будет наблюдаться одновременный массовый вброс целевого контента через множество каналов. При построении временного ряда, соответствующего такому массовому вбросу, логично ожидать аномально высокую начальную частоту публикации дубликатов, особенно в начальной стадии информационной операции.

Таким образом, измерение начальной частоты публикации дубликатов и выявление аномальных временных рядов позволяет существенно сократить период времени, необходимый для выявления информационных операций, так как анализу подлежит не весь временной ряд, а только его начальная фаза.

3. Постановка задачи

Предметом исследования в данной работе являются информационные потоки в виде публикаций пользователей в социальной сети ВКонтакте.

Далее по тексту будут использоваться следующие термины:

инфоповод – поток информационных сообщений, публикуемых пользователями социальной сети, объединенных одной тематикой.

информационный трек – поток сообщений, опубликованных пользователями социальной сети, состоящий из дубликатов, отсортированных по времени в единый временной ряд.

Целью данной работы было исследовать процессы распространения деструктивной информации в социальной сети ВКонтакте на предмет *выявления признаков* целенаправленной организованной деятельности в виде *информационных операций*. Для достижения указанной цели необходимо было решить следующие задачи:

– разработать программу анализа информационных треков (временных рядов публикации дубликатов);

– проанализировать с помощью программы коллекции материалов по нескольким инфоповодам;

– выявить треки с аномальными характеристиками;

– интерпретировать выявленные аномалии;

– предложить методику выявления признаков информационных операций на основе выявления аномалий информационных треков.

4. Метод

В соответствии с поставленными задачами была разработана программа[1], которая выявляет в общем информационном потоке четкие дубликаты, сортирует их по времени публикации, объединяет в отдельные информационные треки, анализирует начальную частоту публикации дубликатов в каждом треке и предоставляет информацию об информационных треках с аномально высокой начальной частотой публикации дубликатов. Если быть точным, выходные данные программы для каждого информационного трека содержат информацию о том, в течение какого времени были опубликованы первые 10 дубликатов и общее количество дубликатов.

Выходные данные программы для каждого инфоповода представляются в виде графика, сводной таблицы и набора таблиц для каждого информационного трека, включающих содержание дубликатов и ссылки на страницы, на которых они были опубликованы.

С помощью программы "Информационный трек-детектор" были проанализированы наиболее известные инфоповоды, среди которых самыми яркими являются инфоповод "Жыве Беларусь", связанный с протестными настроениями на выборах президента республики Беларусь, а также инфоповод "Дворец Путина", название которого говорит само за себя.

На рис. 1 представлен график информационных треков по инфоповоду "Жыве Беларусь". Цифрами обозначены номера информационных треков, которые представляют интерес для интерпретации в качестве признаков информационной операции.

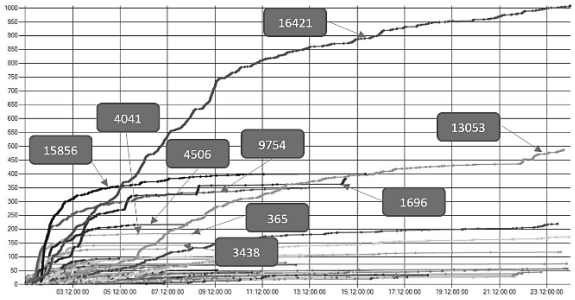

Рис. 1. *Сводный график информационных треков по инфоповоду "Жыве Беларусь"*

На данном графике присутствуют треки с высокой начальной частотой публикации дубликатов (все кроме 16421). Эти же треки имеют значительный охват по сравнению с другими, за исключением 16421, который представляет из себя набор дубликатов определенного хештега, под которым на страницах пользователей публиковались картинки с различным содержанием.

Перечень информационных треков по инфоповоду "Жыве Беларусь" и их параметры показаны в выдержке из сводной таблицы для этого информационного повода (см. табл. 1).

Таблица 1. *Выдержка из сводной таблицы по инфоповоду "Жыве Беларусь"*

Номер трека	Общее количество дубликатов	Период публикации первых 10 дубликатов
16421	1009	10:47:21
13053	487	21:41:03
15856	400	00:16:16
1696	382	00:12:37
9754	350	00:23:21
4506	217	00:10:35
365	184	00:19:33
3438	142	00:08:38
4041	129	00:06:55

[1] Программа "Информационный трек-детектор", свидетельство о регистрации программы для ЭВМ регистрационный № 2022668598 от 10.10.2022.

На рис. 2 представлен график информационных треков по инфоповоду "Дворец Путина".

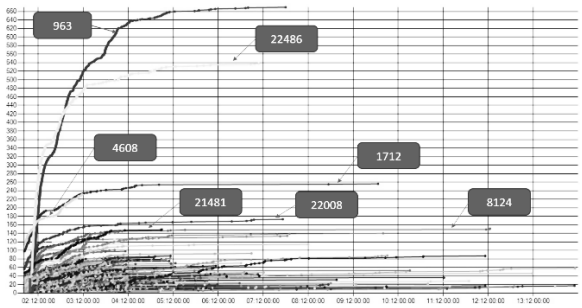

Рис. 2. *Сводный график информационных треков по инфоповоду "Дворец Путина"*

Данные расшифровки графика представлены в табл. 2.

Таблица 2. *Выдержка из сводной таблицы по инфоповоду "Дворец Путина"*

Номер трека	Общее количество дубликатов	Период публикации первых 10 дубликатов
963	670	00:06:26
22486	538	00:06:00
1712	256	00:04:38
4608	231	00:09:54
22008	173	00:42:27
8124	151	00:08:28
1973	142	00:07:06

Интерес представляют треки, в которых количество дубликатов не менее 100 и первые 10 публикаций были сделаны в течение не более 30 минут. Совокупность высокой начальной частоты публикаций и большого охвата можно интерпретировать как признак информационной операции. Очевидно, что вероятность стихийной публикации полностью идентичного контента десятью различными пользователями в течение первых 4 минут (трек 1712) стремится к нулю, то есть наблюдается организованный одновременный массовый вброс целевого контента, который мы можем однозначно интерпретировать как признак информационной операции.

То же касается и других представленных информационных треков с аномальными характеристиками.

Как показали исследования, информационные треки с высокой начальной частотой публикации как правило имеют гораздо большее количество дубликатов, чем те, которые не идентифицируются как признаки информационной операции.

Таким образом, если анализировать только начальную частоту публикации дубликатов, можно существенно сократить время выявления признаков информационной операции.

Таким образом, методика выявления признаков информационных операций может выглядеть следующим образом:

– сбор коллекции публикаций по заданному информационному поводу;

– выявление информационных треков с аномально высокой начальной частотой, характеризующих организованность действий;

– проверка треков на предмет релевантности;

– формирование окончательного списка треков с признаками информационных операций.

Заключение

В данной работе авторы проанализировали информационные потоки деструктивного контента, содержащие в общей сложности 129 тысяч информационных сообщений, и выявили высокую начальную частоту публикации четких дубликатов как некое общее свойство, присущее информационным операциям. Это свойство было положено в основу методики выявления признаков информационных операций. Последующие аналитические исследования показали, что данная методика позволяет выявлять информационные операции на самых ранних стадиях, что приводит к сокращению времени на принятие решения о мерах противодействия выявленным информационным операциям и повысить эффективность информационного противодействия.

В дальнейшем необходимо провести дополнительные исследования для выявления количества не связанных между собой каналов массового вброса, а также использования в информационных операциях пользователей, обладающих наибольшим потенциальным уровнем влияния.

Список литературы

1. *Концепция* Конвенции ООН об обеспечении международной информационной безопасности. URL: http://www.scrf.gov.ru/security/information/document112/. (дата обращения: 04.10.2022).
2. *Макаренко С.И.* Информационное противоборство и радиоэлектронная борьба в сетецентрических войнах начала XXI века: монография. СПб.: Наукоемкие технологии, 2017. 546 с.

3. *Расторгуев С.П., Литвиненко М.В.* Информационные операции в сети Интернет / под. общ. ред. А.Б. Михайловского. М.: АНО "Центр стратегических оценок и прогнозов", 2014. 128 с.
4. *Еременко В.Т., Рязанцев П.Н.* Информационное противоборство в социотехнических системах. Орел: ОГУ им. И.С. Тургенева, 2016. 209 с.
5. *Потемкин А.В.* Распознавание информационных операций средств массовой информации сети Интернет // Интернет-журнал "Науковедение". 2015. № 7.
6. *Загорулько Ю.А., Саломатина Н.В., Серый А.С., Сидорова Е.А., Шестаков В.К.* Выявление нечетких дубликатов при автоматическом формировании тематических коллекций документов на основе web-публикаций // Вестник НГУ. Серия: Информационные технологии. 2013. № 4.
7. *Kleinberg J.* Temporal Dynamics of On-Line Information Streams. 2016. P. 221–238.
8. *del corso G., Gulli A., Romani F.* Ranking a stream of news. 2005. P. 97–106.
9. *Rakesh V., Singh D., Vinzamuri B., Reddy Chandan*. Personalized recommendation of twitter lists using content and network information // Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014. 2014. P. 416–425.
10. *Ландэ Д.В., Фурашев В.Н., Брайчевский С.М., Григорьев А.Н.* Основы моделирования и оценки электронных информационных потоков: монография. Киев: Инжиниринг, 2006. 176 с.
11. *Додонов А.Г., Ландэ Д.В., Цыганок В.В., Андрейчук О.В., Каденко С.В., Грайворонская А.Н.* Распознавание информационных операций. Киев: ООО "Инжиниринг", 2017. 282 с.

References

1. *Koncepciya* Konvencii OON ob obespechenii mezhdunarodnoj informacionnoj bezopasnosti. URL: http://www.scrf.gov.ru/security/information/document112/. (In Russ.). (Access date: 04.10.2022).
2. *Makarenko S.I.* Informacionnoe protivoborstvo i radioelektronnaya bor'ba v setecentricheskih vojnah nachala XXI veka: monografiya. SPb.: Naukoemkie tekhnologii; 2017. 546 p. (In Russ.).
3. *Rastorguev S.P., Litvinenko M.V.* Informacionnye operacii v seti Internet / pod. obshch. red. A.B. Mihajlovskogo. M.: ANO "Centr strategicheskih ocenok i prognozov", 2014. 128 p. (In Russ.).
4. *Eremenko V.T., Ryazancev P.N.* Informacionnoe protivoborstvo v sociotekhnicheskih sistemah. Orel: OGU imeni I.S. Turgeneva; 2016. 209 p. (In Russ.).
5. *Potemkin A.V.* Raspoznavanie informacionnyh operacij sredstv massovoj informacii seti Internet. Internet-zhurnal "NAUKOVEDENIE". 2015;(7). (In Russ.).
6. *Zagorul'ko Yu.A., Salomatina N.V., Seryj A.S., Sidorova E.A., SHestakov V.K.* Vyyavlenie nechetkih dublikatov pri avtomaticheskom formirovanii tematicheskih kollekcij dokumentov na osnove web-publikacij. Vestnik NGU. Seriya: Informacionnye tekhnologii. 2013;(4). (In Russ.).
7. *Kleinberg J.* Temporal Dynamics of On-Line Information Streams. 2016:221–238.
8. *del corso G., Gulli A., Romani F.* Ranking a stream of news. 2005:97–106.
9. *Rakesh V., Singh D., Vinzamuri B., Reddy Chandan*. Personalized recommendation of twitter lists using content and network information. Proceedings of the 8th International Conference on Weblogs and Social Media, ICWSM 2014; 2014:416–425.
10. *Lande D.V., Furashev, V.N., Brajchevskij, S.M., Grigor'ev, A.N.* Osnovy modelirovaniya i ocenki elektronnyh informacionnyh potokov: monografiya. Kiev: Inzhiniring; 2006. 176 p. (In Russ.).
11. *Dodonov A.G., Lande D.V., Cyganok V.V., Andrejchuk O.V., Kadenko S.V., Grajvoronskaya A.N.* Raspoznavanie informacionnyh operacij. Kiev: OOO "Inzhiniring"; 2017. 282 p. (In Russ.).

Информация об авторах:

А. Н. Рабчевский – кандидат технических наук, старший преподаватель кафедры информационной безопасности и систем связи Пермского государственного национального исследовательского университета (614068, г. Пермь, ул. Букирева, 15), генеральный директор ООО "СЕУСЛАБ" (614087, г. Пермь, ш. Космонавтов, д.111, корп.3), andrey@ranat.ru, AuthorID 1089772;

М. Ю. Карпов – старший преподаватель кафедры информационной безопасности и систем связи Пермского государственного национального исследовательского университета (614068, г. Пермь, ул. Букирева, 15), karpov.mu@psu.ru;

Е. Г. Ашихмин – начальник учебно-аналитического центра ООО "СЕУСЛАБ" (614087, г. Пермь, ш. Космонавтов, д.111, корп.3), e.ashikhmin@seuslab.ru, https://orcid.org/0000-0001-9193-4535.

Information about the authors:

A. N. Rabchevskiy – Candidate of Technical Sciences, Senior Lecturer of the Department of Information Security and Communication Systems of Perm State University (15, Bukireva street, Perm, 614068), General Director of SEUSLAB LLC (111 Kosmonautov st., bld. 3, Perm, 614087), andrey@ranat.ru, AuthorID 1089772;

M. Yu. Karpov – Senior Lecturer of the Department of Information Security and Communication Systems of Perm State University (15 Bukireva St., Perm, 614068), karpov.mu@psu.ru;

E. G. Ashikhmin – Head of the Training and Analytical Center of SEUSLAB LLC (111/3, Kosmonavtov highway, Perm, Russia, 614087), e.ashikhmin@seuslab.ru, https://orcid.org/0000-0001-9193-4535.