

Борисова Анна Сергеевна

кандидат экономических наук, доцент  
кафедры информационных систем  
и математического моделирования  
Волгоградского института управления –  
филиала Российской академии народного  
хозяйства и государственной службы  
при Президенте РФ

## ТРОЛЛИНГ КАК НОВОЕ СОЦИАЛЬНОЕ ЯВЛЕНИЕ В СИСТЕМЕ ИНТЕРНЕТ-КОММУНИКАЦИИ [1]

### Аннотация:

*В статье проведен анализ интернет-троллинга как нового социального явления, получившего широкое распространение в Сети. Исследованы его типовые черты, описаны отличия от других форм антисоциального поведения в интернете (киберзапугивания или распространения спама). Представлена классификация троллей в зависимости от причин, провоцирующих их деструктивные действия. Предложены меры по нейтрализации интернет-троллинга, включающие не только способы, основанные на человеческих или автоматических механизмах идентификации троллей и их сообщений (реактивные меры), но и методы обнаружения постов, которые могут стать потенциальными целями проявления онлайн-агрессии (превентивные меры). Названы целесообразными методы борьбы с троллингом, имеющие законодательный характер (введение уголовной ответственности за троллинг, запрет на анонимность в глобальной сети), и массовое просвещение интернет-пользователей в данной области.*

### Ключевые слова:

*троллинг, интернет-коммуникация, онлайн-агрессия, механизм троллинга, виды троллей, деструктивное поведение в Сети, нейтрализация интернет-троллинга.*

Borisova Anna Sergeevna

PhD in Economics, Associate  
Professor, Information Systems and  
Mathematical Modeling Department,  
Volgograd Institute of Management –  
branch of the Russian Presidential  
Academy of National Economy  
and Public Administration

## TROLLING AS A NEW SOCIAL PHENOMENON IN THE INTERNET COMMUNICATION SYSTEM [1]

### Summary:

*The study analyzes Internet trolling as a new social phenomenon that has become widespread in the online community. The research investigates its typical features and shows the difference with other forms of anti-social behavior on the Internet (cyberbullying or spamming). Internet trolls are classified according to the reasons for provoking their destructive actions. The measures to neutralize Internet trolling are proposed. They include not only human or automatic mechanisms for identifying trolls and their messages (reactive measures) but also methods for detecting user's posts that can be potential targets of online aggression (preventive measures). In the author's opinion, anti-trolling legislative methods (introduction of criminal responsibility for trolling, ban on Internet anonymity) and mass awareness of Internet users are the appropriate measures in this regard.*

### Keywords:

*trolling, Internet communication, online aggression, trolling mechanism, types of trolls, destructive behavior on the Internet, neutralization of Internet trolling.*

**Введение.** По данным Internet World Stats за 2019 г., более 4,4 млрд человек в мире являются пользователями сети Интернет, что составляет 57,3 % от общей численности населения планеты [2]. Глобальная сеть стала отражением реального мира, с помощью онлайн-инструментов граждане устанавливают личные и деловые контакты. Особую роль в этом играют социальные медиа, ставшие основным форумом для публичных обсуждений. Более 3 млрд человек – представителей различных культур и слоев общества используют социальные сети (например, Facebook), публикуют сообщения в системах микроблогов (например, Twitter), участвуют в дискуссиях на форумах (например, Reddit), где добавляют комментарии и обмениваются мнениями по интересующим их темам, создавая диалог в глобальном масштабе [3]. Наряду с позитивными социальными аспектами расширения виртуальных возможностей для общения в сети Интернет появились новые формы антисоциального поведения. Некоторые пользователи участвуют в дискуссиях без намерения внести конструктивный вклад в диалог, а скорее пытаются нарушить и подорвать коммуникацию. Их ассоциируют с «агентами хаоса» в интернете и обычно называют троллями.

**Характеристика интернет-троллинга.** Изначально термин «троллинг» определял способ рыбалки, когда люди перетаскивают приманку за лодкой, чтобы привлечь рыбу [4]. В сказках тролль изображался монстром, который прячется под мостами, чтобы напугать и напасть на неизвестных прохожих [5].

Сейчас термин «троллинг» используется для описания различных типов деструктивного поведения в глобальной сети, начиная от шуточных комментариев и заканчивая оскорбительным и

угрожающим поведением. Ярко выраженными мотивами троллинга являются скука, поиск внимания, чувство мести, желание получить удовольствие или нанести ущерб интернет-сообществу. Эти причины отличают троллинг от других форм антисоциального поведения в Сети, например киберзапугивания, где личности преступников обычно известны, а намерения более прямолинейны. В отличие от спамеров, тролли не стремятся к финансовой выгоде, создание беспорядка на самом деле является для них самоцелью. Типичные примеры троллинг-поведения включают в себя насмешки и дискредитацию участников дискуссии, подстрекательство и эскалацию споров, выдачу себя за опытных пользователей с целью распространения вредных советов и ложной информации.

Исследования психологического портрета тролля показали, что подобное деструктивное поведение может проявляться вне зависимости от возраста и гендерных различий, но чаще всего интернет-агрессия исходит от мужчин в молодом возрасте [6]. Кроме того, доказана связь антисоциального поведения с чертами личности, входящими в «темную триаду»: нарциссизмом (страстной фиксацией на себе), макиавеллизмом (получением удовольствия от манипулирования чьим-либо поведением) и психопатией (отсутствием эмпатии) [7]. Однако самым значимым предиктором проявления троллинга является садизм (коэффициент корреляции равен 0,5) [8]. Как и садисты, тролли беспощадны, эмоционально жестоки и получают удовольствие от боли и дискомфорта других людей. Можно сказать, что кибертроллинг является интернет-проявлением повседневного садизма. Психопатия также выступает предиктором троллинга, но с меньшей силой связи (коэффициент корреляции равен 0,47) [9].

**Механизм интернет-троллинга.** Общение в интернет-сообществе развивает у его членов чувство принадлежности к определенной группе, основанной не на тесной дружбе, а на скрытых связях. Подобные связи возникают при прочтении и написании комментариев и обладают высоким потенциалом для формирования отношений с людьми, с которыми пользователь лично не знаком, но взгляды которых разделяет. Появление интернет-тролля в подобных сообществах может иметь значимые негативные последствия. Согласно проведенным опросам [10], около 30 % подростков когда-либо становились жертвами интернет-троллей.

Причинами, по которым пользователи поддаются на провокации троллей, являются высокий уровень доверия к участникам интернет-сообщества, отсутствие невербальных сигналов (мика, жесты, тон) при передаче сообщения, вследствие чего невозможно оценить достоверность информации. Кроме того, большинство людей надеются, что каждый пользователь должен соблюдать нормы онлайн-поведения и будет наказан за их нарушение.

Тролли могут искусственно противопоставлять различные мнения, чтобы использовать склонность человека к бинарному выбору – видеть мир в черно-белом цвете, разделять всех на богатых и бедных, плохих и хороших. Для манипулирования общественным мнением тролли используют ложь, хитрость, альтернативные факты, маловероятные теории, искажение информации.

Развитию антисоциального поведения также способствуют такие характеристики сети Интернет, как анонимность пользователей, быстрое распространение информации, отсутствие строгого контроля публикуемого контента.

**Классификация интернет-троллей.** Термин «троллинг» широко используется для характеристики различных типов антисоциального и деструктивного поведения в сети Интернет. Некоторые из его проявлений могут иметь только развлекательный характер, но встречаются примеры угрожающего и запугивающего поведения, эквивалентного онлайн-преследованиям.

Наиболее достоверна классификация троллей в зависимости от причин, провоцирующих антисоциальное поведение. Так, Дж. Бишоп выделяет 12 типов троллей, разделенных на 4 группы [11].

Первую группу можно обозначить как «ненавистники». Это члены интернет-сообществ, которым нравится разжигать споры и проявлять онлайн-агрессию без какой-либо определенной цели. К этой группе относят пользователей, движимых силами мести, которые стремятся раскрыть истинные мотивы собеседника, поставив его в неудобную, искусственно созданную ситуацию. В эту группу входит «бунтарь», движимый деструктивными силами. Он принимает участие в троллинге с целью раскрытия публике известных фактов, способных спровоцировать состояния испуга или ужаса у аудитории. «Шутником» движут антисоциальные силы. Целью троллинга в этом случае является получение морального удовлетворения от негативной реакции членов интернет-сообщества.

Вторую группу составляют «провокаторы», нацеленные на привлечение внимания. Сюда можно отнести интернет-тролля, публикующего негативный контент, отвечающий ожиданиям интернет-сообщества. В данную группу также входит «плакса», движимый силами самобичевания, который хочет получить ложное чувство сопереживания от аудитории. Завоевать доверие других членов интернет-сообщества и впоследствии манипулировать ими в собственных целях – причины антисоциального поведения еще одного типа тролля из этой группы.

Третью группу составляют «болтуны», которым нравится принимать участие в коммуникации вне зависимости от правильности и полезности публикуемых комментариев. В данную группу входят соперечающие тролли, движимые силами всеобщего прощения. Они формируют у собеседников ложное ощущение поддержки в решении их проблем. К этой группе можно отнести тролля, движимого творческими силами, который создает и публикует контент, обладающий исключительно юмористическим оттенком. В данную группу входят тролли, движимые социальными силами, цель которых – стимулировать общение в интернет-сообществе разными методами, в том числе с помощью подстрекательства других участников коммуникации.

Четвертую группу интернет-троллей можно обозначить как «наблюдатели». Они изучают диалог членов интернет-сообщества с целью поиска подходящего момента для публикации провокационного контента. В данную группу входит тролль, открыто не проявляющий онлайн-агрессию. Его цель – скрытый троллинг через «лайки» и «репосты» негативно окрашенной информации, публикацию результатов опросов общественного мнения с измененными данными. К этой группе также относится «истинный тролль», цель которого – внести в групповую коммуникацию хаос и разжечь конфликт. Последним видом антисоциального поведения по данной классификации является троллинг новых участников интернет-сообщества с целью проверки их на соответствие целям группы.

**Методы нейтрализации интернет-троллинга.** Среди потенциальных способов борьбы с онлайн-агрессией можно выделить несколько направлений.

Очевидным методом управления деструктивным поведением в сети Интернет является работа модераторов, которые отслеживают весь публикуемый контент и определяют его соответствие нормам онлайн-сообщества. Несмотря на высокую эффективность и популярность данного метода, он очень трудоемок в ситуациях, когда большое количество людей начинают одновременно размещать комментарии и каждый из них требует рассмотрения перед опубликованием.

Решение данной проблемы заключается в распределении обязанностей по контролю контента между всеми участниками интернет-коммуникации. Так, пользователи сообщают администратору о провокационной информации, после проверки он удаляет контент, отключает функцию чата или удаляет нарушителей из сообщества.

Еще одним направлением нейтрализации интернет-троллинга является автоматизация обнаружения вредоносных постов. Возможно использование программных алгоритмов, способных отслеживать информацию, размещенную в интернете, и автоматически отфильтровывать нежелательную онлайн-активность.

Перечисленные подходы к борьбе с троллингом основаны на человеческих или автоматических механизмах идентификации троллей и их постов. Даже в случае успеха обнаружение троллей не полностью решает проблему. Тролли хорошо работают с системой, например обходят запреты, используя разные имена пользователей или маскируя смысл своих сообщений. Важно, что эти меры являются реактивными: они обычно применяются после того, как опубликованы клеветнические, угрожающие или вводящие в заблуждение комментарии, и ущерб уже нанесен.

Другим направлением контроля за деструктивным поведением является использование нейтронной сети для обнаружения постов, которые могут стать потенциальными целями проявления онлайн-агрессии [12]. Т. е. вместо обнаружения и удаления последствий интернет-троллинга целесообразно предвидеть активность троллей и принимать превентивные меры.

Еще один способ, позволяющий решить проблему троллинга, носит законодательный характер. Предполагается, что примеров антисоциального поведения в сети Интернет станет значительно меньше, когда киберпреступники будут нести уголовную ответственность за свои действия. Поскольку сделан вывод, что анонимность в глобальной сети способствует распространению троллинга, одной из превентивных мер является запрет на использование псевдонимов.

Кроме того, чтобы свести к минимуму вредные психологические последствия троллинга, необходимо проводить эффективную просветительскую деятельность в данной отрасли. Важно обучать пользователей методам идентификации различных типов интернет-троллей и способам уменьшения их негативного влияния. Один из самых известных подходов в этом направлении – игнорирование спорных и провокационных комментариев в социальных медиа.

**Заключение.** Интернет-троллинг – это серьезная проблема, которая подрывает работу всех каналов интернет-коммуникации. Борьба с ним должна иметь комплексный характер и проходить при поддержке всех участников глобального диалога: со стороны государства возможно принятие превентивных мер запрета на анонимность в сети Интернет и ужесточение наказания за подобные преступления; со стороны экономических институтов – разработка программных продуктов по автоматизации обнаружения вредоносной информации и постов, которые могут

стать потенциальными целями проявления агрессии; со стороны гражданского общества – модерирование онлайн-диалога пользователями или администраторами, просветительская деятельность в данной области.

#### Ссылки и примечания:

1. Исследование выполнено при финансовой поддержке РФФИ и администрации Волгоградской области в рамках научного проекта № 18-410-343006 «Повышение уровня доверия населения к региональным органам власти через интернет-коммуникации».
2. Internet World Stats: Internet Usage Statistics [Электронный ресурс]. URL: <https://www.internetworldstats.com/stats.htm> (дата обращения: 19.07.2019).
3. Digital 2018: Q4 Global Digital Statshot [Электронный ресурс]. 2018. 17 Oct. URL: <https://datareportal.com/reports/digital-2018-q4-global-digital-statshot> (дата обращения: 19.07.2019).
4. Fu-Yuan Hong, Kuang-Tsan Cheng. Correlation Between University Students' Online Trolling Behavior and Online Trolling Victimization Forms, Current Conditions, and Personality Traits // *Telematics and Informatics*. 2018. Vol. 35, iss. 2. P. 397–405. <https://doi.org/10.1016/j.tele.2017.12.016>.
5. Ibid.
6. Buckels E.E., Trapnell P.D., Paulhus D.L. Trolls Just Want to Have Fun // *Personality and Individual Differences*. 2014. Vol. 67. P. 97–102. <https://doi.org/10.1016/j.paid.2014.01.016>.
7. Which Personality Traits are Related to Traditional Bullying and Cyberbullying? A Study with the Big Five, Dark Triad and Sadism / M. van Geel, A. Goemans, F. Toprak, P. Vedder // *Personality and Individual Differences*. 2017. Vol. 106. P. 231–235. <https://doi.org/10.1016/j.paid.2016.10.063>.
8. March E. Psychopathy, Sadism, Empathy, and the Motivation to Cause Harm: New Evidence Confirms Malevolent Nature of the Internet Troll // *Personality and Individual Differences*. 2019. Vol. 141. P. 133–137. <https://doi.org/10.1016/j.paid.2019.01.001>.
9. Ibid.
10. Buckels E.E., Trapnell P.D., Paulhus D.L. Op. cit.
11. Bishop J. Representations of 'Trolls' in Mass Media Communication: A Review of Media-Texts and Moral Panics Relating to 'Internet Trolling' // *International Journal of Web Based Communities*. 2014. Vol. 10, no. 1. P. 7–24. <https://doi.org/10.1504/ijwbc.2014.058384>.
12. Tsantarliotis P., Pitoura E., Tsaparas P. Defining and Predicting Troll Vulnerability in Online Social Media // *Social Network Analysis and Mining*. 2017. Vol. 7, no. 1. 26 p. <https://doi.org/10.1007/s13278-017-0445-2>.

#### References:

Bishop, J 2014, 'Representations of 'Trolls' in Mass Media Communication: A Review of Media-Texts and Moral Panics Relating to 'Internet Trolling'', *International Journal of Web Based Communities*, vol. 10, no. 1, pp. 7-24, <https://doi.org/10.1504/ijwbc.2014.058384>.

Buckels, EE, Trapnell, PD & Paulhus, DL 2014, 'Trolls Just Want to Have Fun', *Personality and Individual Differences*, vol. 67, pp. 97-102, <https://doi.org/10.1016/j.paid.2014.01.016>.

Fu-Yuan, Hong & Kuang-Tsan, Cheng 2019, 'Correlation Between University Students' Online Trolling Behavior and Online Trolling Victimization Forms, Current Conditions, and Personality Traits', *Telematics and Informatics*, vol. 35, iss. 2, pp. 397-405, <https://doi.org/10.1016/j.tele.2017.12.016>.

Geel, M van, Goemans, A, Toprak, F & Vedder, P 2017, 'Which Personality Traits are Related to Traditional Bullying and Cyberbullying? A Study with the Big Five, Dark Triad and Sadism', *Personality and Individual Differences*, vol. 106, pp. 231-235, <https://doi.org/10.1016/j.paid.2016.10.063>.

March, E 2019, 'Psychopathy, Sadism, Empathy, and the Motivation to Cause Harm: New Evidence Confirms Malevolent Nature of the Internet Troll', *Personality and Individual Differences*, vol. 141, pp. 133-137, <https://doi.org/10.1016/j.paid.2019.01.001>.

Tsantarliotis, P, Pitoura, E & Tsaparas, P 2017, 'Defining and Predicting Troll Vulnerability in Online Social Media', *Social Network Analysis and Mining*, vol. 7, no. 1, 26 p., <https://doi.org/10.1007/s13278-017-0445-2>.