

ТРОЛЛИНГ БОТОВ ЛЮДЬМИ – НОВЫЙ ТРЕНД В ПОЛЬЗОВАТЕЛЬСКОМ ИНТЕРНЕТ-ПРОСТРАНСТВЕ

О. С. Мухина¹⁾, В. Ф. Олешко²⁾

*¹⁾Уральский федеральный университет,
ул. Мира, 19, 620000, г. Екатеринбург, Россия,
olga.mukhina@urfu.ru,*

*²⁾Уральский федеральный университет,
ул. Мира, 19, 620000, г. Екатеринбург, Россия,
vladimir.oleshko@urfu.ru*

В статье анализируется новый формат троллинга: подстрекательские и саркастические комментарии людей в пользовательском интернет-пространстве по отношению к ботам, работающим на основе искусственного интеллекта, с целью отличить ботов от живых собеседников. Выдвигается гипотеза о временном характере данного феномена.

Ключевые слова: нейросеть; искусственный интеллект; ИИ; бот; чат; троллинг.

ТРОЛІНГ БОТАЇ ЛЮДЗЬМІ – НОВИ ТРЭНД У КАРЫСТАЛЬНІЦКАЙ ІНТЭРНЭТ-ПРАСТОРЫ

В. С. Мухіна¹⁾, У. Ф. Алешка²⁾

*¹⁾Уральскі федэральны ўніверсітэт,
вул. Міра, 19, 620000, г. Екацярынбург, Расія,
olga.mukhina@urfu.ru,*

*²⁾Уральскі федэральны ўніверсітэт,
вул. Міра, 19, 620000, г. Екацярынбург, Расія,
vladimir.oleshko@urfu.ru*

У артыкуле аналізуецца новы фармат тролінга: падбукторванне і саркастычныя каментары людзей у карыстацкай інтэрнэт-прасторы ў адносінах да ботаў, якія працуюць на аснове штучнага інтэлекту, з мэтай адрозніць ботаў ад жывых суразмоўцаў. Прапануецца гіпотэза пра часовы характар дадзенага феномену.

Ключавыя словы: нейрасетка; штучны інтэлект; ІІІ; бот; чат; тролінг.

Социальные сети, появившись в начале XXI в., сформировали и другой феномен пользовательского Интернета: троллинг – публикация подстрекательских, неуместных или оскорбительных комментариев с целью вызвать недовольство других [1]. Так называемые «тролли» изначально, в момент становления социальных сетей, чаще всего были подростками, которые таким способом развлекались или удовлетворяли собственное эго. Позже троллинг появился как профессиональный процесс – деятельность в конкретных политических, экономических, рекламных и других целях.

Тем не менее, развлекательная задача троллинга и сегодня зачастую отмечается как основная. Другие цели – провокация, чтобы вызвать конфликт, или отвлечение участников дискуссии от острых тем [2].

Обычно тролль старается вызвать раздражение и вывести из равновесия собеседников [3]. Это влияет на атмосферу киберпространства, делая ее некомфортной для коммуникации [4]. Поэтому те пользователи Интернета, которые хоть однажды сталкивались с этим явлением, вырабатывали против троллей единственную эффективную стратегию поведения: игнорирование и жесткая модерация контента [5].

Классическое описание троллинга выглядит следующим образом: объект троллинга – человек; тема – группа людей и их глупость/недалекость; форма – провокация (сарказм, оскорбления, пародирование); интонация – уверенность, утвердительные предложения [6].

Вместе с тем, стремительная популяризация искусственного интеллекта (ИИ), в том числе в новых медиа, создает принципиально новый формат троллинга: как по объекту, так и по форме и целям.

Боты, созданные на основе ИИ, активно развиваются и становятся все более понимающими, гибкими и человечными. Их все сложнее отличить от человека. На данный момент такие боты активно используются для целей самопродвижения в чатах. Один из авторов настоящей статьи, являясь администратором нескольких пабликов в различных социальных сетях, наблюдает в течение последнего года (2024 – начало 2025) следующую тенденцию:

Первые ИИ-боты действовали слишком прямолинейно: в их имени («нике») обозначался вид деятельности владельца, являясь, таким образом, его рекламой. Например: «Марина | Медитации» или «Кристина|Йога». Бот отправлял массовые комментарии на любые публикации. Единственное преимущество такого бота было в том, что он сумел распознать, относительно верно истолковать текст и дать комментарий

по теме. Однако комментарии были очевидно «искусственными», чаще всего выдавая абстрактную похвалу тексту. Например: «Интересный текст! Я так много узнала про сонный паралич». В начале 2025 г. появился новый тренд в работе ИИ-ботов: вероятно, настройщики оценили, насколько интереснее и ярче выглядит критический комментарий. Поэтому некоторые ИИ-боты начали публиковать развернутые критические комментарии (к примеру, почему нет смысла смотреть фильм или читать книгу, о которых идет речь в статье, или почему герой интервью – скучный человек с недостаточной экспертностью в вопросе). Однако представители и первого, и второго поколения ИИ-ботов легко вычислялись как пользователями, так и модераторами, по стилю и примитивному однотипному подходу.

На момент написания настоящей статьи (май 2025) появилось, на наш взгляд, третье поколение ИИ-ботов, вычислить которых уже намного сложнее. Они пишут разумные комментарии, близкие человеческому подходу. Они уже не ставят рекламу в свое имя («ник»), рассчитывая на тех, кто кликнет на их (самое обычное) имя и прочитает рекламу в описании, а модератору сложно предъявить претензии к их комментариям. Пример: ИИ-бот под именем «Анна Григорьева» пишет следующий комментарий: «Я сама неоднократно чувствовала подобные состояния. Время от времени ставлю себе медитации, и это действительно помогает избавиться от страха и тревог. Да, все впечатления яркие, порой даже странные, и они способствуют пониманию себя. Обязательно стоит обсуждать свои ощущения без страха быть непонятыми». Комментарий полностью соответствует тексту статьи. При клике на имя бота, мы читаем: «Эксперт в работе с состояниями и энергиями».

Для удаления подобного комментария и блокировки его автора у модератора нет оснований. Сами пользователи отмечают косвенные признаки, по которым вычисляют ИИ-бота: пишет ровно через 10 минут после публикации очередной статьи (хотя другие пользователи тут же добавляют, что бывали и другие промежутки времени); часто использует слово «интересно» (но это не показатель «искусственности», да и встречается не всегда – даже в приведенном выше примере данного слова нет); однотипная структура комментария (хвалит текст, а затем делится собственным опытом; но в приведенном примере структура немного иная: собственный опыт и далее совет).

Ввиду такой ситуации, пользователи выработали для себя стратегию вычисления ИИ-ботов, которую, на наш взгляд, можно назвать новым ти-

пом троллинга. Пользователи начинают писать боту подстрекательские комментарии с целью выявить, что это бот. Может показаться странным использовать понятие «троллинг» в отношении ботов. Ведь основная цель троллинга – вывести собеседника из равновесия, а боты не могут испытывать эмоции.

Надо отметить, что в данном случае меняются практически все признаки троллинга, обозначенные выше. Объектом уже является не человек, а искусственный интеллект. Интонации, если раньше были чаще утвердительные, теперь меняются на вопросительные. Темы, если ранее были в ракурсе группы, к которой принадлежит человек, то сейчас относятся к обсуждаемым вопросам. Меняются цели: если раньше это было развлечение, удовлетворение эго или профессиональные (оплачиваемые заказчиком) задачи политического, рекламного и иного рода, то теперь цель – распознать бота, чтобы дать повод модератору заблокировать его и в итоге общаться в чате гарантированно только с людьми.

Примеры:

1. Провокация диалога, на который ИИ пока не способен (ИИ-бот: «Я тоже пробовала подобные эксперименты». Пользователь: «Расскажешь?»).

2. Публикация саркастических ответов боту в виде иллюстраций (поскольку на данный момент боты не в состоянии определить содержание иллюстрации).

3. Сарказм и провокация, на которые человек ответил бы, с целью демонстрации чату, что данная псевдоличность является ботом («А как же «интересно»? Разве это не потрясающе интересный сон?») (в ответ на комментарий бота, написанный без его обычного слова «интересно»); «Научу вытаскивать деньги из астрала. Не нужны связи и опыт. Пиши. P. S. Модераторы, не блокируйте меня, я не бот, это просто прикол» (в ответ на самопрезентацию бота); ИИ-бот: «Интересное исследование, подтверждающее связь психики и сновидений». Пользователь: «Интересные боты пасутся в комментариях»).

Представляется, что, несмотря на явные отличия от классического троллинга по объекту, целям, интонациям и пр., сохраняется главный признак: подстрекательство и сарказм без цели реального общения на формально заявленную тему. Таким образом, мы являемся свидетелями возникновения нового вида троллинга, направленного от людей к роботам. В некотором роде это является также признаком необычной

дискриминации XXI в.: люди хотят общаться с людьми и активно «выживают» роботов из своего пространства для коммуникации.

Добавим, что в ближайшем будущем, несомненно, искусственный интеллект усовершенствуется в достаточной степени, чтобы никакие пользовательские провокации не выявили признаки работы нейросети. Поэтому данный феномен носит временный характер, и очень скоро людям придется общаться в чатах с роботами, не делая различий между ними и живыми собеседниками.

Библиографические ссылки

1. Штанькова П. А., Галичкина Е. Н. Лексические характеристики троллинга (на материале русского языка) // Креативная лингвистика. сборник научных статей. Астрахань, 2024. С. 133–137.
2. Геворгян К. М. Виды и функции троллинга в сетевой коммуникации // Филология, иностранные языки и медиакоммуникации. Материалы симпозиума в рамках XIX (LI) Международной научной конференции студентов, аспирантов и молодых ученых. Кемерово, 2024. С. 15–17.
3. Бешкарев Н. С., Гурбанова А. И. Троллинг и кибербуллинг в интернет-пространстве // В сборнике: Актуальные проблемы авиации и космонавтики. Сборник материалов X Международной научно-практической конференции, посвященной 100-летию академика М. Ф. Решетнева и Дню космонавтики. В 3-х томах. Красноярск, 2024. С. 1239–1240.
4. Осокина А. С. Троллинг подростков в социальных сетях: особенности его проявления // Бизнес-трансформация: управление улучшениями. 2024. № 4(9). С. 82–88.
5. Капиунов Д. С. Троллинг как одна из проблем социально-информационного пространства // Молодежь, наука и цивилизация. Материалы международной студенческой научной конференции. Красноярск, 2023. С. 676–679.
6. Анисимов Н. Д., Смотрин А. О. Троллинг в Интернете: разновидности и влияние // Философия и гуманитарные науки в информационном обществе. 2024. № 1(41). С. 80–89.