

# МЕТОДЫ АНАЛИЗА ПИСЬМЕННОЙ КОММУНИКАЦИИ СКУЛШУТЕРОВ ДЛЯ ПРОГНОЗИРОВАНИЯ ПРЕДУПРЕЖДАЮЩЕГО ПОВЕДЕНИЯ<sup>1</sup>

Савельев А.О., Карпова А.Ю., Третьяков Д.А., Максимова Н.Г.

Национальный исследовательский Томский политехнический университет, Томск, Россия  
sava@tpu.ru, belts@tpu.ru, dat32@tpu.ru, nata8720@tpu.ru

*Аннотация.* В докладе изложены первичные результаты автоматизированного выявления значимых признаков, содержащихся в письменной коммуникации, для создания диагностической методики по выявлению предупреждающих сигналов насильственного поведения потенциальных скулшутеров. Выполнен анализ открытых наборов данных для определения характерных когнитивных, эмоциональных, психологических и поведенческих признаков.

*Ключевые слова:* обработка естественного языка, скулшутинг, письменная коммуникация, предупреждающее поведение.

## Введение

Феномен скулшутинга в настоящее время считается серьезной и растущей угрозой во всем мире. Исследования, посвященные скулшутингу (феномен насильственных инцидентов с расстрелами школьников или стрельбой в школах; англ. school shooting) в активной форме ведутся с начала 2000-х г.г. Всплеск исследований в мировой науке связан с волной инцидентов скулшутинга, последовавших после преступления, совершенного в старшей школе Колумбайн (США) [1]. Феномен экспоненциального роста скулшутинга получил название «эффект Колумбайн», суть которого в том, что в результате широкого освещения в СМИ, массовые расстрелы в учебных учреждениях имеют ярко выраженный эффект инцидентов подражания. За прошедшие 20 лет, в мировой науке накоплен значительный исследовательский бэкграунд в области изучения скулшутинга. Но, универсального ответа на вопрос, какие процессы/мотивы могут побудить человека принять образ мышления, при котором совершение насильственного инцидента оценивается как вероятно приемлемый способ решения проблемы, и чем такое мышление отличается от того, которое может привести к фактическому планированию, подготовке и осуществлению инцидента скулшутинга, – до сих пор нет. Сходство таких событий, как скулшутинг, происходящих по всему миру, дает основание для изучения его как особого феномена, существенно отличающегося от других форм массового насилия, таких как серийные убийства, бандитизм, разборки банд, захват школы террористами, насильственный экстремизм, терроризм одиночек.

На настоящий момент времени существует несколько исследовательских фокусов изучения скулшутинга, в которых условно можно выделить пять направлений:

- анализ поведенческих и социально-психологических признаков скулшутеров (идейно-смысловые основы, мотивы насильственного поведения, причины, условия, черты личности [2, 3]);
- изучение степени влияния социальных сетей на рост случаев имитации инцидентов и деструктивная роль СМИ (социальные связи в онлайн-среде, тематические сообщества, тематический контент; освещение в СМИ, провоцирующее вероятность имитационных действий [4, 5]);
- изучение типологии подражания (например, сходство механизмов подражания скулшутеров и синдрома подражающих самоубийств [6]);
- обоснование комплексной теории скулшутинга (например, пятифакторная теория 2004 г. [7]);
- многоуровневый подход – комплексная оценка риска скулшутинга через анализ и интерпретацию сигналов в письменной коммуникации (например, предупреждающие знаки агрессии, средствах, прямой или косвенной угрозе, намерении, как индикаторах риска целенаправленного насилия [8, 9]).

Последнее направление является наиболее перспективным, но методологически сложным из-за отсутствия доступных, надежных, эффективных исследовательских инструментов и сервисов в виде программного обеспечения, без которого невозможен процесс сбора, обработки и анализа больших данных в онлайн среде. На настоящий момент, пока в этом направлении только экспериментируют с комбинацией методов, методик для выявления наблюдаемых сигналов в социальных сетях. Наиболее эффективный результат применения такого подхода представлен в работе научной группы из Швеции [10]. Авторы предлагают методику прогнозирования насильственного поведения в социальных сетях.

---

<sup>1</sup> Исследование выполнено при финансовой поддержке ГЗ «Наука», в рамках проекта FSWW-2024-0023

На основе автоматизированного анализа текста, авторы извлекают психологические индикаторы и вычисляют обобщенный индекс риска насильственного поведения. Прогностическая валидность авторами оценивается по способности классифицировать преступников-одиночек в общей популяции пользователей. Утверждается, что точность классификации достигает 90-96%. Авторы используют данные о 76 ранее выявленных преступниках-одиночках, совершивших насильственные инциденты, включая скулшутеров, идеологически мотивированных преступников и массовых убийц. Представлены данные различного характера: самооправдательные манифесты, письма и сообщения официальным лицам, членам семей, друзьям. Также используются данные из 31 онлайн-среды (более 200 тысяч пользователей), в том числе включающие примеры подстрекательства к насилию, языка вражды, токсичных высказываний, пропаганды насильственного экстремизма. Выделено более 14 лингвистических маркеров, например, проявление гнева, тревоги; выражение отвращения; негативные и положительные эмоции; использование токсичных выражений, угроз и другие. Ключевой вывод авторов заключается в том, что разработанная методика и индекс могут служить диагностическим инструментом как для выявления потенциальных злоумышленников, так и для выявления онлайн-сред с высоким количеством пользователей, подверженных риску совершения насильственных действий.

Цель данной работы – разработать диагностическую методику по выявлению предупреждающих сигналов насильственного поведения потенциальных скулшутеров на основе комбинации методов обработки естественного языка и машинного обучения. Конкретной задачей текущего этапа работ является определение значимых для методики признаков.

Мы определяем скулшутинг как проблемно-мотивированное насилие, и в своих публикациях обозначили отличие скулшутинга от идеологически-мотивированного насилия, поддерживая основную идею, высказанную одним из ключевых исследователей скулшутинга, Питером Лангманом «there is no one cause of school shootings, despite efforts to pin the blame on a single cause» [11]. Доктор Лангман, на результатах сравнительного исследования 48 случаев скулшутинга и 55 случаев массовых убийств, совершенных одиночками, выявил и систематизировал различные типы влияния, подражания и вдохновения, и представил доказательства совпадений ролевых моделей поведения массовых убийц и скулшутеров [12]. Типология влияния включает героическую идеализацию, богоподобное поклонение, личное сходство, безрассудную страсть. В типологии подражания обозначены следующие маркеры: быстрое подражание, подражание как чествование, подражание как паломничество, подражание как копирование (речи, внешнего вида, поведения). Типология вдохновения содержит такие признаки, как общее вдохновение, сочувствие мотиву, выполнение чужой воли. Мы включили в набор анализируемых данных также найденные письменные тексты нескольких массовых убийц, чтобы проверить, насколько они совпадают по психологическим, поведенческим и когнитивным характеристикам.

## 1. Методология исследования

Для выявления потенциально значимых признаков для автоматизированной диагностики риска скулшутинга были обработаны тексты манифестов известных скулшутеров и массовых убийц, находящиеся в открытом доступе на ресурсе SchoolShooters.info. Первоначально было отобраны данные о 48 стрелках. После исключения данных, представленных в виде нераспознанных скан-образов текстов, написанных вручную, осталось 37 наборов текстов, сгруппированных по исполнителям. Затем были исключены манифесты и иные тексты за авторством стрелков, которые насчитывали в совокупности менее 10 предложений. Таким образом, итоговая анализируемая выборка данных включает 33 набора текстов, объемом от 16 до 29 617 предложений.

Отобранные наборы текстов были проанализированы для:

- Определения признаков наличия или отсутствия когнитивных искажений с помощью открытой (<https://huggingface.co/amedvedev/bert-tiny-cognitive-bias>) обученной BERT-модели с заявленным значением *precision* от 0,8 до 0,88.
- Классификации предложений по эмоциональной окраске на базе открытой обученной модели DistilRoBERTa (<https://huggingface.co/j-hartmann/emotion-english-distilroberta-base>), основанной на модели базовых эмоций Пола Экмана.
- Определения признаков токсичности в тексте на базе модели BERT-модели (<https://huggingface.co/unitary/toxic-bert>).
- Определения намерений, содержащихся в тексте, с помощью T5-модели [13].
- Выделения значимых именованных сущностей (персоналий) с использованием модели WikiNEuRal [14].

Для пунктов 1-4 приведенного списка было классифицировано каждое из предложений в наборах текстов. С учетом значительной разницы в объемах различных наборов текстов, полученные результаты были нормированы на общее количество предложений в наборе.

## 2. Результаты анализа письменной коммуникации скулшутеров

### 2.1. Наличие признаков когнитивных искажений

Использованная модель распознавания признаков когнитивных искажений предусматривают наличие следующих классов (исключая класс «отсутствие признаков когнитивных искажений (КИ)»):

1. Навешивание ярлыков (labeling).
2. Персонализация (personalization).
3. Долженствование (should statement).
4. Эмоциональное рассуждение (emotional reasoning).
5. Сверхобобщение (overgeneralizing).
6. Катастрофизация (catastrophizing).
7. Заблуждение о вознаграждении (reward fallacy).

Наиболее значимыми видами выявленных когнитивных искажений являются «навешивание ярлыков» и «персонализация». В таблице 1 приведены значения долей предложений относительно общего объема каждого из наборов данных, по указанным видам когнитивных искажений. Цветовым градиентом построчно (от зеленого к красному) обозначены наибольшие и наименьшие значения.

Таблица 1. Результаты выявления признаков когнитивных искажений в наборах текстов скулшутеров и массовых убийц

Стрелок	Отсутствие признаков КИ	Класс когнитивных искажений						
		1	2	3	4	5	6	7
Алваро Кастилло	0,280	0,187	0,160	0,107	0,093	0,093	0,080	0,000
Себастьян Боссе	0,534	0,080	0,093	0,042	0,081	0,110	0,052	0,009
Джеффри Уиз	0,125	0,063	0,063	0,063	0,313	0,250	0,125	0,000
Марк Лепин	0,475	0,164	0,115	0,000	0,131	0,033	0,082	0,000
Пекка-Эрик Аувинен	0,377	0,240	0,077	0,071	0,060	0,049	0,115	0,011
Аарон Ибарра	0,274	0,194	0,194	0,081	0,081	0,097	0,081	0,000
Адам Ланца	0,379	0,146	0,098	0,062	0,077	0,111	0,121	0,007
Алекс Грибал	0,071	0,286	0,107	0,000	0,107	0,107	0,321	0,000
Брайан Оливер	0,265	0,339	0,094	0,074	0,060	0,131	0,034	0,003
Брендан О'Рурк	0,307	0,218	0,178	0,050	0,042	0,094	0,106	0,005
Ган Лу	0,425	0,161	0,093	0,010	0,140	0,062	0,078	0,031
Джон Романо	0,459	0,071	0,112	0,112	0,094	0,071	0,059	0,024
Дуэйн Моррисон	0,286	0,156	0,163	0,109	0,048	0,122	0,102	0,014
Итан Крамбли	0,467	0,067	0,200	0,067	0,000	0,133	0,067	0,000
Карл Пирсон	0,265	0,061	0,153	0,122	0,112	0,224	0,020	0,041
Кимвир Джилл	0,389	0,170	0,125	0,065	0,038	0,098	0,105	0,010
Кип Кинкель	0,249	0,193	0,133	0,149	0,077	0,133	0,044	0,022
Крис Харпер-Мерсер	0,367	0,225	0,058	0,054	0,104	0,113	0,054	0,025
Люк Вудхэм	0,354	0,177	0,192	0,023	0,046	0,131	0,077	0,000
Майкл Слободян	0,300	0,300	0,300	0,100	0,000	0,000	0,000	0,000
Майрон Мэй	0,336	0,294	0,120	0,012	0,057	0,089	0,067	0,025
Николас Круз	0,319	0,213	0,170	0,021	0,000	0,213	0,043	0,021
Роберт Флорес	0,320	0,163	0,224	0,042	0,059	0,150	0,037	0,005
Скай Буш	0,514	0,096	0,085	0,107	0,113	0,068	0,017	0,000
Сын Хи Чо	0,401	0,184	0,105	0,059	0,072	0,039	0,118	0,020
Ти Джей Лейн	0,208	0,250	0,167	0,083	0,042	0,229	0,021	0,000

Уильям Атчисон	0,277	0,222	0,122	0,041	0,042	0,148	0,139	0,009
Хосе Рейес	0,294	0,353	0,176	0,059	0,059	0,000	0,059	0,000
Чарльз Уитмен	0,438	0,180	0,079	0,101	0,090	0,034	0,067	0,011
Эван Рэмси	0,286	0,036	0,107	0,071	0,107	0,143	0,214	0,036
Эллиот Роджер	0,310	0,144	0,140	0,074	0,079	0,167	0,066	0,019
Эрик Харрис	0,489	0,123	0,073	0,076	0,075	0,082	0,072	0,010
Эрик Хейнсток	0,338	0,162	0,183	0,070	0,077	0,063	0,077	0,028
Среднее значение	0,339	0,179	0,135	0,066	0,078	0,109	0,082	0,012

## 2.2. Эмоциональная окраска предложений

В таблице 2 приведены результаты классификации предложений каждого из рассматриваемых наборов данных по эмоциям в соответствии с моделью базовых эмоций Экмана. Доминирующими эмоциями из выявленных являются «гнев» и «печаль», со средними долями относительно всего объема анализируемых текстов равными 0,16 и 0,14 соответственно. Характерным также является полное отсутствие предложений, классифицированных моделью как «страх».

Таблица 2. Результаты выявления эмоциональной окраски в наборах текстов скулишутеров и массовых убийц

Стрелок	Гнев	Отвращение	Страх	Радость	Печаль	Удивление	Нейтрально
Алvaro Кастилло	0,027	0,067	0,000	0,067	0,253	0,027	0,507
Себастьян Боссе	0,147	0,067	0,000	0,055	0,050	0,054	0,592
Джеффри Уиз	0,250	0,063	0,000	0,000	0,188	0,063	0,438
Марк Лепин	0,164	0,148	0,000	0,033	0,098	0,016	0,393
Пекка-Эрик Аувинен	0,202	0,115	0,000	0,055	0,038	0,033	0,508
Аарон Ибарра	0,242	0,097	0,000	0,081	0,145	0,065	0,274
Адам Ланца	0,079	0,098	0,000	0,063	0,061	0,109	0,538
Алекс Грибал	0,143	0,214	0,000	0,071	0,071	0,000	0,500
Брайан Оливер	0,091	0,272	0,000	0,034	0,047	0,097	0,309
Брендан О'Рурк	0,071	0,058	0,000	0,060	0,101	0,069	0,540
Ган Лу	0,093	0,047	0,000	0,078	0,119	0,021	0,560
Джон Романо	0,094	0,059	0,000	0,200	0,112	0,059	0,353
Дуэйн Моррисон	0,048	0,048	0,000	0,041	0,150	0,034	0,524
Итан Крамбли	0,067	0,067	0,000	0,067	0,267	0,000	0,467
Карл Пирсон	0,122	0,112	0,000	0,102	0,041	0,020	0,551
Кимвир Джилл	0,137	0,148	0,000	0,110	0,094	0,111	0,365
Кип Кинкель	0,249	0,083	0,000	0,050	0,243	0,055	0,238
Крис Харпер-Мерсер	0,079	0,067	0,000	0,075	0,092	0,029	0,563
Люк Вудхэм	0,277	0,085	0,000	0,023	0,069	0,008	0,438
Майкл Слободян	0,100	0,100	0,000	0,000	0,100	0,100	0,500
Майрон Мэй	0,057	0,087	0,000	0,034	0,067	0,031	0,589
Николас Круз	0,064	0,000	0,000	0,106	0,128	0,000	0,553
Роберт Флорес	0,067	0,106	0,000	0,084	0,182	0,027	0,436
Скай Буш	0,045	0,034	0,000	0,023	0,045	0,158	0,605
Сын Хи Чо	0,342	0,092	0,000	0,020	0,053	0,039	0,382
Ти Джей Лейн	0,063	0,292	0,000	0,021	0,167	0,021	0,333
Уильям Атчисон	0,151	0,127	0,000	0,051	0,103	0,042	0,423
Хосе Рейес	0,294	0,059	0,000	0,000	0,235	0,000	0,235

Чарльз Уитмен	0,202	0,034	0,000	0,090	0,101	0,034	0,517
Эван Рэмси	0,036	0,036	0,000	0,036	0,143	0,000	0,714
Эллиот Роджер	0,107	0,100	0,000	0,218	0,138	0,060	0,228
Эрик Харрис	0,223	0,068	0,000	0,058	0,049	0,090	0,484
Эрик Хейнсток	0,085	0,141	0,000	0,113	0,056	0,028	0,549
Среднее значение	0,161	0,096	0	0,057	0,141	0,041	0,429

### 2.3. Наличие признаков токсичности в тексте

Результаты анализа текстов на наличие признаков токсичности приведены в таблице 3. Отметим, что несмотря на заметную долю предложений в каждом из наборов текстов, которые отнесены к «токсичным» (в среднем 0,134), модель не позволила выявить какого-либо значимого специфического класса «токсичности», который был бы характерен для большинства стрелков.

Таблица 3. Результаты анализа текстов скулитутеров и массовых убийц на наличие признаков токсичности

Стрелок	Токсичный	Выражено токсичный	Нецензурный	Угроза	Оскорбление	Ненависть
Алvaro Кастилло	0,027	0,000	0,000	0,013	0,000	0,000
Себастьян Боссе	0,122	0,002	0,074	0,013	0,023	0,001
Джеффри Уиз	0,188	0,000	0,188	0,000	0,063	0,000
Марк Лепин	0,000	0,000	0,000	0,000	0,000	0,000
Пекка-Эрик Аувинен	0,109	0,000	0,033	0,000	0,049	0,005
Аарон Ибарра	0,194	0,000	0,048	0,081	0,016	0,000
Адам Ланца	0,063	0,000	0,036	0,002	0,014	0,002
Алекс Грибал	0,107	0,000	0,036	0,000	0,000	0,000
Брайан Оливер	0,044	0,000	0,007	0,000	0,007	0,000
Брендан О'Рурк	0,029	0,000	0,008	0,002	0,008	0,002
Ган Лу	0,005	0,000	0,000	0,000	0,000	0,000
Джон Романо	0,076	0,000	0,047	0,000	0,024	0,000
Дуэйн Моррисон	0,007	0,000	0,000	0,000	0,000	0,000
Итан Крамбли	0,000	0,000	0,000	0,000	0,000	0,000
Карл Пирсон	0,071	0,000	0,041	0,000	0,000	0,000
Кимвир Джилл	0,179	0,000	0,121	0,003	0,071	0,011
Кип Кинкель	0,166	0,006	0,050	0,066	0,022	0,000
Крис Харпер-Мерсер	0,113	0,000	0,013	0,008	0,017	0,025
Люк Вудхэм	0,146	0,000	0,015	0,015	0,000	0,000
Майкл Слободян	0,200	0,000	0,100	0,000	0,000	0,000
Майрон Мэй	0,006	0,000	0,000	0,002	0,000	0,000
Николас Круз	0,128	0,000	0,043	0,064	0,043	0,000
Роберт Флорес	0,000	0,000	0,000	0,000	0,000	0,000
Скай Буш	0,017	0,000	0,000	0,000	0,000	0,000
Сын Хи Чо	0,395	0,020	0,224	0,053	0,191	0,020
Ти Джей Лейн	0,125	0,000	0,000	0,021	0,000	0,000
Уильям Атчисон	0,240	0,001	0,142	0,017	0,076	0,015
Хосе Рейес	0,353	0,000	0,118	0,176	0,118	0,000
Чарльз Уитмен	0,022	0,000	0,000	0,011	0,000	0,000
Эван Рэмси	0,071	0,000	0,036	0,000	0,000	0,000



- преобладание в числе эмоционально окрашенных предложений признаков «гнева» и «печали», при полном отсутствии признаков «страха»;
- преобладание в текстах признаков намерений с явной положительной коннотацией (например, to be helpful; to be a good friend);
- высокой доли упоминания среди персоналий известных скулшутеров и массовых убийц.

Следующим этапом работы, мы рассматриваем возможность сравнения полученных результатов в эксперименте с другими обученными моделями, на тех же данных.

Необходимо также упомянуть ограничения полученных результатов. Их применимость для создания методики выявления предупреждающих сигналов требует дополнительного уточнения. Необходимо соотнести полученные результаты обработки текстовой коммуникации скулшутеров с аналогичными результатами обработки нейтральных текстов авторов тех же возрастных категорий. Также следует учитывать ограничения самих использованных открытых обученных моделей, использованных при анализе данных.

## Литература

1. *Kostinsky S., Bixler E.O., Kettl P.A.* Threats of School Violence in Pennsylvania After Media Coverage of the Columbine High School Massacre: Examining the Role of Imitation // *Arch Pediatr Adolesc Med.* 2001. N 155(9). – P. 994–1001.
2. *Rocque M.* Exploring school rampage shootings: Research, theory, and policy // *The Social Science Journal.* 2012. N 49. – P. 304–313.
3. *Seguin M., Chawky N., Lesage A., Boyer R., Guay S., Bleau P., Miquelon P., Szkrumelak N., Steiner W., Roy D.* Evaluation of the Dawson College Shooting Psychological Intervention: Moving Toward a Multimodal Extensive Plan // *Psychological Trauma-Theory Research Practice and Policy.* 2013. N 5. – P. 268–276.
4. *Meindl J., Ivy J.* Mass shootings: The role of the media in promoting generalized imitation // *American Journal of Public Health.* 2017. N 107. – P. 368–370.
5. *Cohen K., Johansson F., Kaati L., Mork J.* Detecting linguistic markers for radical violence in social media // *Terrorism and Political Violence.* 2013. P. 246–256.
6. *Malkki L.* Political Elements in Post-Columbine School Shootings in Europe and North America // *Terrorism and Political Violence.* 2014. N 26(1). – P. 185–210.
7. *Newman S., Fox C., Roth W., Mehta J., Harding D.* Rampage: The Social Roots of School Shootings. – NY: Basic Books, 2004. – 352 p.
8. *Kaati L., Shrestha F., Cohen K.* Linguistic analysis of lone offender manifestos // 2016 IEEE International Conference on Cybercrime and Computer Forensic (ICCCF). – Vancouver, 2016. P. 1–8.
9. *Kaati L., Shrestha A., Akrami N.* Predicting Targeted Violence from Social Media Communication // 2022 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). – Istanbul, 2022. P. 383–390.
10. *Kaati L., Shrestha A., Akrami N.* General Risk Index: A Measure for Predicting Violent Behavior Through Written Communication // 2023 IEEE International Conference on Big Data. – Sorrento, 2023. P. 4065-4070.
11. *Langman P.* A Bio-Psycho-Social Model of School Shooters // *The Journal of Campus Behavioral Intervention.* 2017. – (5). – P. 27.
12. *Langman P.* Different Types of Role Model Influence and Fame Seeking Among Mass Killers and Copycat Offenders // *American Behavioral Scientist.* 2018. – N 2 (62). – P. 210–228.
13. *Raffel C., Shazeer N., Roberts A., Lee K., Narang S., Matena M., Zhou Y., Li W., Liu P.* Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer // *Journal of Machine Learning Research.* 2020. – N 21. – P. 1–67.
14. *Tedeschi S., Maiorca V., Campolungo N., Cecconi F., Navigli R.* WikiNEuRal: Combined Neural and Knowledge-based Silver Data Creation for Multilingual NER // *Findings of the Association for Computational Linguistics: EMNLP 2021.* – Punta Cana, 2021. P. 2521–2533.