Федулов Кирилл Евгеньевич
студент
Научный руководитель
Маяцкая Ольга Борисовна
канд. филос. наук, доцент

ФГБОУ ВО «Уфимский университет науки и технологий»
г. Уфа, Республика Башкортостан

DOI 10.31483/r-153095

ДИПФЕЙКИ В ПОЛИТИКЕ: НОВАЯ ЭРА ДЕЗИНФОРМАЦИИ И МАНИПУЛЯЦИЙ

Аннотация: *в статье представлен анализ проблемы дипфейков в политическом дискурсе современного информационного общества. Дипфейки представляют собой синтетический медиаконтент, созданный с использованием технологий искусственного интеллекта и глубокого обучения, что позволяет создавать правдоподобные видео и аудио записи политических деятелей. Автор рассматривает основные подходы к определению концепции дипфейков, классификацию их видов, а также анализирует угрозы, которые они представляют для политической стабильности, информационной безопасности и доверия к демократическим институтам. Особое внимание уделяется механизмам манипуляции общественным мнением посредством дипфейков, их роли в избирательных кампаниях и протестных движениях.*

Ключевые слова: *дипфейки, дезинформация, псевдореальность, манипуляция, политика, искусственный интеллект.*

В условиях цифровой трансформации общества и развития технологий искусственного интеллекта возникла принципиально новая угроза информационной безопасности – дипфейки. Термин «deepfake» впервые появился в 2017 году, обозначая синтетический видеоконтент, создаваемый с помощью технологий глубокого обучения, однако, данная технология корнями своими уходит весьма

глубоко, проведя ретроспективный анализ, увидим, что «первым был Томас Эдисон, который решил применить их еще в 1890-х годах. Во время испано-американской войны, когда необходимо было поднять дух патриотизма солдат, камеры были очень тяжелыми, и компания Edison Manufacturing поставила себе цель расширить возможности кинематографа. В компании пошли на хитрость: пленка реальных событий с маршем американских солдат сменялась постановочными атаками американцев на их врага. Люди были не в состоянии понять, были ли эти военные действия на самом деле правдивы. В итоге цель была достигнута, патриотический дух возрос» [5].

Если изначально дипфейки использовались в развлекательных целях, то в настоящее время они стали инструментом политической манипуляции и дезинформации, поскольку политическая сфера оказалась особенно уязвимой перед угрозой, которую представляют дипфейки: создание правдоподобного видео с политическим деятелем, произносящим речь, которую он никогда не произносил или, например, политиком, демонстрирующим компрометирующее поведение, может оказать значительное влияние на избирательные кампании, на доверие общества к институтам власти и социальную стабильность. По данным многочисленных экспертов и исследователей, количество дипфейков в интернете растет экспоненциально, а средства их создания становятся все более доступными.

Концепция дипфейков требует многоаспектного анализа, так как дипфейк – это многоаспектный феномен, создаваемый с использованием нейронных сетей и технологий глубокого обучения, технологическая основа дипфейков была заложена еще в 2014 году с появлением генеративно-состязательных сетей Гудфеллоу, однако их применение в политических целях произошло позже, определим наиболее распространенные типы дипфейков, которые включают в себя следующие компоненты:

– замена лица одного человека на лицо другого человека в видеозаписи (наиболее простой и часто встречающийся вид дипфейков);

− синхронизация движений губ с аудиозаписью (оригинальное видео сохраняется, но звуковое сопровождение может быть полностью заменено или отредактировано);

− технологии, при которых все черты лица человека управляются через марионеточный интерфейс, позволяя создавать видео с манипулирование всеми элементами выражения лица;

− изменение отдельных атрибутов лица (возраст, эмоции, выражение) без замены всего лица;

− создание синтетического голоса, способного воспроизводить речь с использованием технологий или синтеза голоса по образцу.

Кроме того, в научной литературе выделяются так называемые «cheapfakes» (дешевые подделки) – видео, которые создаются путем простого редактирования или замедления оригинального контента без использования искусственного интеллекта [7]. Однако угроза, которую представляют эти более простые формы манипуляции, также требует серьезного внимания, поскольку они могут быть столь же эффективны в деле дезинформации: «Дипфейки могут использоваться в различных сферах… но также вызывают опасения по поводу злоупотреблений, таких как распространение дезинформации или создание компрометирующих материалов. Это тревожный факт. Научным журналом «iScience», был проведен эксперимент: испытуемым показали 16 видеороликов, половина из которых была сделана с помощью технологии «дипфейк». В результате большая часть таких роликов осталась неузнанной. При этом испытуемые были крайне уверены в свои силах и правоте и не сомневались, делая выбор» [4, с. 12].

Использование дипфейков в политике опирается на несколько ключевых механизмов воздействия на общественное мнение, среди которых выделим:

− дипфейки эксплуатирующие когнитивные предубеждения людей, которые склонны верить информации, соответствующей их уже сложившимся взглядам, и отвергать противоречащую им информацию, данное явление хорошо известно как «предвзятость подтверждения»;

− видеоконтент, воспринимаемый большинством людей как более убедительный и правдоподобный, чем текстовая информация или даже неподвижные изображения, как отмечает Й. Пил, режиссер, создавший известное видео, в котором президент Барак Обама предупреждает о опасности дипфейков, визуальная информация имеет мощное воздействие на восприятие реальности;

− дипфейки, распространяемые через социальные медиа-платформы, которые использованы для целевой доставки контента специфическим демографическим группам, исследования показывают, что таргетирование дипфейков на определенные сегменты аудитории может оказывать реальное влияние на политические установки избирателей.

Примечательно, что в контексте кампании Кембридж Аналитика была выявлена связь между использованием целевой рекламы и влиянием на политические решения граждан, что убедительно показывает, насколько эффективной может быть комбинация технологий манипуляции и платформ социальных сетей. Дипфейки представляют многоуровневую угрозу для политических систем:

− они подрывают доверие к источникам информации, когда граждане начинают сомневаться в подлинности видеозаписей, это создает общую атмосферу недоверия, в которой даже истинная информация может быть воспринята со скептицизмом;

− дипфейки могут быть использованы для компрометации политических лидеров накануне выборов, исследования свидетельствуют, что с 2018 по 2025 год количество политических дипфейков заметно возросло, особенно в период предвыборных кампаний;

− дипфейки способны спровоцировать социальные и политические кризисы;

− дипфейки используются для целей информационной войны и геополитического противостояния, так, например, различные акторы, включая государственные и негосударственные структуры, могут использовать дипфейки для подрыва авторитета враждебных политических режимов или расширения своего влияния.

Социальные сети играют ключевую роль в распространении дипфейков, платформы, такие как VK, Telegram, Instagram (деятельность Instagram и

Facebook в России признана экстремистской и запрещена *https*://*www.rbc.ru*/*technology_and_media*/*21*/*03*/*2022*/*6238a5e89a79477e5dc0245f* (дата обращения: 12.12.2025)) обладают огромной аудиторией и алгоритмами, которые могут способствовать вирусному распространению контента: «Молодому поколению россиян сегодня предоставляется информация, содержащая уже готовые выводы, формирующая мотивы и установки, определяющие поведение аудитории в сети Интернет, сюда же вписывается конструирование альтернативных фактов, моделирующих «постправду» под воздействием которых уже не имеет значения подлинность, истинность и реальность любой новости, главное, чтобы она формировала определенный эмоциональный настрой молодежи и соответствовала реальным политическим целям коммуникатора, развивающего негативные эмоции, «нужное» восприятие информации, в конечном итоге ведущее к радикализации взглядов, деструктивному поведению, разрушению национального самосознания и потере исторической памяти» [6, с. 12].

Значительная часть дипфейков в интернете распространяется именно через эти каналы. Исследования показывают, что платформы часто обладают недостаточными механизмами для быстрого выявления и удаления синтетического контента. Более того, сокращение времени проверки контента и усиление конкуренции между платформами за внимание пользователей, создают условия для более быстрого распространения дезинформации. По данным различных аналитических центров, в 2025 году было выявлено более 100 тысяч дипфейков на различных платформах, однако эта цифра, вероятно, значительно ниже реального количества подделок, поскольку многие из них остаются необнаруженными.

Первые попытки законодательного регулирования дипфейков появились на уровне национальных государств и международных организаций. В 2019 году в Конгрессе США был внесен проект закона о дипфейках, предусматривающий уголовную ответственность за создание и распространение дипфейков, способных причинить вред, австрийское правительство в 2022 году приняло план действий по противодействию дипфейкам, предусматривающий как технические, так и законодательные меры, «Китай стал одним из первых государств,

внедривших специальные меры по регулированию дипфейк-контента. Обязательная маркировка таких материалов помогает пользователям легче распознавать подделки и снижает вероятность распространения дезинформации. Это важный шаг в борьбе с манипуляциями и фейковыми новостями, так как повышает уровень доверия к информации и способствует более осознанному потреблению контента. В ноябре 2022 г. были приняты «Положения об управлении глубоким синтезом информационных сервисов в Интернете», в положениях законов «О кибербезопасности», «О безопасности данных», «О защите личной информации», «Об управлении информационными службами Интернета» и других законов и нормативных актов КНР закреплены запреты на использование дипфейков в целях угрозы национальной безопасности и интересам, для нанесения ущерба имиджу нации, посягательства на общественные интересы, нарушения экономического и социального порядка или посягательства на законные права и интересы других лиц» [1, с.123].

Однако единого международного подхода к регулированию этого явления пока не выработано, сложность законодательного регулирования состоит в необходимости баланса между защитой от дезинформации и соблюдением прав человека на свободу выражения, кроме того, технологии совершенствуются быстрее, чем появляются законодательные инициативы, что требует постоянной адаптации правовых норм. Параллельно с развитием технологий создания дипфейков развиваются и методы их детектирования, к основным подходам относятся:

− компьютерное зрение и анализ артефактов -- выявление несоответствий и странных элементов в видео, которые указывают на синтетическое происхождение контента;

− анализ видеосжатия и искажений – дипфейки часто содержат специфические паттерны, возникающие из-за процесса обучения нейронных сетей;

− анализ лица в реальном времени – использование биометрических параметров для выявления несоответствий в выражении лица, мимике и движениях глаз;

− фактологическая проверка и проверка источников – верификация контекста и источника видео, проверка соответствия события известным фактам.

Несмотря на прогресс в разработке методов детектирования, текущие технологии все еще не обеспечивают 100% точность, особенно по мере совершенствования технологий создания дипфейков, одним из наиболее эффективных способов противодействия дипфейкам является повышение уровня медиаграмотности населения: люди, обладающие критическим мышлением и навыками анализа информации, менее подвержены влиянию синтетического контента, образовательные инициативы, направленные на обучение молодежи критическому анализу медиаконтента, должны быть интегрированы в школьные программы и программы высшего образования.

Дипфейки представляют собой принципиально новую и сложную угрозу в информационной экосистеме современного общества, особенно в политической сфере синтетический медиаконтент, созданный с помощью искусственного интеллекта, может быть использован для манипуляции общественным сознанием, компрометации политических лидеров, провоцирования социальных конфликтов и ведения информационных войн.

Решение проблемы дипфейков требует координации действий всех заинтересованных сторон – государства, технологических компаний, средств массовой информации, образовательных учреждений и гражданского общества, только таким образом можно минимизировать угрозы, которые несут дипфейки для политической стабильности, информационной безопасности и функционирования демократических институтов.

Список литературы

1. Епифанова Т.В. Проблемы законодательного регулирования объектов, созданных с использованием дипфейк-технологии в России и за рубежом / Т.В. Епифанова, К. И. Копейкин // Северо-Кавказский юридический вестник. – 2024. – №3. – С. 121–128. DOI 10.22394/2074-7306-2024-1-3-121-128. EDN KMWGMQ

2. Иванов В.Г. Deepfakes: перспективы применения в политике и угрозы для личности и национальной безопасности / В.Г. Иванов, Я.Р. Игнатовский // Вестник Российского университета дружбы народов. Серия: Государственное и муниципальное управление. – 2020. – Т. 7. №4. – С. 379–386. DOI 10.22363/2312-8313-2020-7-4-379-386. EDN YJJUWH

3. Касперский Е.В. Deepfake, дипфейк / Е.В. Касперский // Лаборатория Касперского [Электронный ресурс]. – Режим доступа: https://encyclopedia.kaspersky.ru/glossary/deepfake/ (дата обращения: 23.11.2025).

4. Кёбис Н.К. Одураченные дважды: люди не могут обнаружить дипфейки, но думают, что могут / Н.К. Кёбис, Б. Долежалова, И. Сораперра // iScience. – 2021. – Т. 24. Вып. 11. – С. 1–17.

5. Лукина Ю.В. Использование дипфейков в общественно-политической жизни/ Ю.В. Лукина // КиберЛенинка [Электронный ресурс]. – Режим доступа: https://cyberleninka.ru/article/n/ispolzovanie-dipfeykov-v-obschestvenno-politicheskoy-zhizni (дата обращения: 23.11.2025).

6. Маяцкая О.Б. Информационный экстремизм и его деструктивное влияние на поведение молодого поколения россиян / О.Б. Маяцкая // Бюллетень науки и практики. – 2022. – Т. 8. №12. – С. 512–515. DOI 10.33619/2414-2948/85/66. EDN GQAQCM

7. Galloway M. Deepfakes may use new technology but they are based on an idea / M. Galloway // Popular Science [Electronic resource]. – Access mode: https://www.popsci.com/technology/deepfakes-history-museum-exhibit/ (date of application: 23.11.2025).