УДК 004.8:004.81

ДИПФЕЙК КАК ФАКТОР ТРАНСФОРМАЦИИ УГРОЗ СОЦИАЛЬНОЙ ИНЖЕНЕРИИ: АНАЛИЗ ТЕНДЕНЦИЙ И МЕТОДОВ ПРОТИВОДЕЙСТВИЯ

Бондаренко В.С. (ВКА), **Плаксеев Д.А.** (ВКА), **Тельбух В.В.** (ВКА)
Научный руководитель – кандидат технических наук, преподаватель Тельбух В. В.
(Военно-космическая академия имени А.Ф.Можайского)

Введение. Социальная инженерия долгое время оставалась искусством манипуляции, ограниченным текстом и голосом. Злоумышленники могли представиться сотрудником банка или написать письмо от имени руководства, но для этого требовалась лишь психологическая подготовка. Ситуация кардинально изменилась с появлением дипфейков – технологий синтеза медиа на основе глубокого обучения [1]. Сегодня атакующий способен не просто имитировать, а в реальном времени «клонировать» личность на видеозвонке. Это ставит под сомнение то, что раньше казалось незыблемым: доверие к аудиовизуальной информации перестало быть надёжным фундаментом коммуникации [2, с. 1760]. Цель работы – показать, что дипфейк выступает не просто новым вектором угроз, а фактором, меняющим саму природу социальной инженерии, что требует пересмотра устоявшихся подходов к защите.

Основная часть.

Причины распространения дипфейк-атак

Во-первых, технологии стали общедоступными: если раньше для создания подделки требовались глубокие знания в области нейросетей, то сегодня сервисы Deepfake-as-a-Service позволяют любому пользователю за небольшую плату сгенерировать реалистичное видео. Во-вторых, высокая эффективность, ведь дипфейки бьют по самому уязвимому звену – доверию человека к собственным глазам и ушам. Когда подделка технически безупречна, критическое мышление отключается, и жертва выполняет требуемые действия. В-третьих, низкие риски для злоумышленников, так как отследить источник синтетического контента крайне сложно, особенно при использовании анонимных сетей и зарубежных платформ.

Масштабы угрозы

Статистика последних лет подтверждает превращение дипфейка из технологической диковинки в системное оружие. Согласно данным Sensity AI, в 2024 году зафиксирован 231 уникальный политический дипфейк, а общее число синтетических файлов в соцсетях достигло 8 млн [3]. В российском сегменте за январь–сентябрь 2025 года выявлено 342 дипфейка – в 4,1 раза больше, чем за весь 2024-й. Примечательно, что 79% из них имитировали глав регионов и госслужащих, что указывает на политический заказ. Финансовый ущерб от крупных инцидентов только в первом квартале 2025 года превысил 200 млн долларов, а эксперты прогнозируют, что к концу 2025 года с дипфейк-атакой столкнётся каждый второй россиянин.

Методы борьбы с дипфейками.

1. Программно-аппаратные средства выявления подделок.

Современные детекторы, основанные на машинном обучении, анализируют артефакты сжатия, несоответствия освещения, аномалии в движении губ и спектральные характеристики голоса. Однако, как показало исследование CSIRO (2025), ни один из 16 популярных детекторов не демонстрирует надёжных результатов в реальных условиях из-за сжатия видео в социальных сетях и постоянного совершенствования генеративных моделей [4]. Перспективным направлением является использование ансамблевых методов, комбинирующих разные подходы, а также внедрение стандартов цифровой подписи контента (C2PA), позволяющих отслеживать происхождение файлов.

2. Развитие цифровой грамотности и критического мышления.

Формирование у населения навыков проверки достоверности видео- и аудиоматериалов становится важным барьером. Пользователи должны усвоить правило, что даже если они видят знакомое лицо и слышат родной голос в мессенджере, это не гарантирует подлинности. Необходимы образовательные программы, разъясняющие природу дипфейков и способы защиты, включая курсы медиаграмотности в учебных заведениях и информационные кампании в СМИ.

3. Деятельность фактчекинговых организаций и верификационных платформ.

Независимые экспертные сообщества и специализированные сервисы (например, Reality Defender) занимаются анализом сомнительного мультимедийного контента. Важно создавать открытые базы данных известных дипфейков и обеспечивать быстрое опровержение вирусных подделок через те же каналы, по которым они распространяются. Координация усилий таких организаций на международном уровне способствует оперативному выявлению и нейтрализации угроз.

4. Совершенствование нормативно-правовой базы.

В Европейском Союзе AI Act (2024) с августа 2026 года требует обязательной маркировки синтетического контента. В России в ноябре 2025 года внесён законопроект о маркировке видео, созданных с помощью ИИ, а при Минцифры создана рабочая группа, рассматривающая введение уголовной ответственности за использование дипфейков в мошеннических целях. Важно не только вводить запреты, но и стимулировать разработку средств верификации – например, через налоговые льготы для компаний, внедряющих стандарты C2PA, а также гармонизировать национальные законодательства с международными подходами.

Выводы. Проведённый анализ позволяет утверждать, что дипфейк – это не просто новый вид мошенничества, а фактор, трансформирующий парадигму кибербезопасности. Социальная инженерия перешла из сферы текста в сферу биометрического копирования, и старая формула «доверяй, но проверяй» окончательно устарела. На смену ей приходит принцип «никогда не доверяй, всегда проверяй» (Zero Trust). Предложенные меры – от многослойных детекторов до законодательной маркировки – должны реализовываться комплексно. Только объединив усилия разработчиков, государства и граждан, можно сдержать волну дипфейк-атак и сохранить доверие к информации в цифровой среде.

Список использованных источников:

1. Goodfellow I., Pouget-Abadie J., Mirza M. et al. Generative Adversarial Nets // Advances in Neural Information Processing Systems. 2014. Vol. 27. P. 2672–2680.

2. Chesney R., Citron D. Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security // California Law Review. 2019. Vol. 107. No. 6. P. 1753–1820.

3. Sensity AI. The State of Deepfakes 2024 [Электронный ресурс]. London: Sentinel Ltd, 2024. URL: https://sensity.ai/reports (дата обращения: 25.02.2026).

4. CSIRO Data61 & Sungkyunkwan University. Deepfake Detection Benchmarking Initiative: Results and Recommendations [Электронный ресурс]. Technical Report. Sydney, 2025. 94, p. URL: https://data61.csiro.au/en/News/News-releases/2025/March/Deepfake-detection-benchmarking (дата обращения: 25.02.2026).