УДК 004.056

РАЗРАБОТКА МЕТОДА ДЕТЕКТИРОВАНИЯ ВМЕШАТЕЛЬСТВА В АУДИОПОТОК С ЦЕЛЬЮ НЕПРАВОМЕРНОГО ИСПОЛЬЗОВАНИЯ БИОМЕТРИЧЕСКИХ ДАННЫХ

Вернигорова А.А. (Университет ИТМО), **Талынкова Е.Н.** (Университет ИТМО)
Научный руководитель – Роговой В.
(Университет ИТМО)

Введение. В последнее время все большее распространение в качестве метода биометрической аутентификации получает распознавание личности человека по совокупности уникальных характеристик голоса. Хотя голосовая биометрия предлагает безопасный способ аутентификации пользователей, она не застрахована от аудио-спуфинга. Достижения в области машинного обучения, технологии записи и синтетической речи позволяют создавать высококачественную подделку голоса или голосовые дипфейки, при помощи которых злоумышленник может обмануть систему автоматической верификации говорящего (ASV). Успешное проведение подобных атак приводит к серьезным последствиям, вплоть до получения доступа к банковским счетам легитимного пользователя. Поскольку существующие решения по предотвращению реализаций спуфинг-атак не имеют широкого применения и не демонстрируют требуемых показателей качества, существует необходимость в разработке эффективной меры противодействия подмене голоса, которая может быть надежно использована для защиты систем ASV.

Основная часть. Для решения проблемы безопасности систем автоматической верификации предлагается использовать методы машинного обучения, а именно, разработку и использование нейронной сети архитектуры LSTM (Long-short term memory) для детектирования аудио-спуфинга. Именно такая архитектура предназначена для выявления долговременных зависимостей, что особенно важно для аудиозаписей, так как позволяет отследить зависимости, характерные для дипфейков, не на отдельном отрезке аудиозаписи, а на всей аудиозаписи в целом. На вход нейросеть получает аудиозапись, которая считывается в виде временного ряда в массив numpy. Из полученных отрезков аудиозаписи извлекаются различные акустические характеристики (спектральные и частотные характеристики, кепстральные коэффициенты и т.д.). Акустические характеристики выделяются из каждого из этих отрезков, а нейронная сеть выделяет зависимости в значениях этих характеристик, определяя их принадлежность подлинникам или дипфейкам. В ходе анализа датасета с аудиозаписями выделяются значимые для анализа акустические характеристики (т.е. те, которые помогают отличить дипфейки от подлинных записей), а малозначимые отсеиваются с помощью дропаута. Данное действие позволяет оптимизировать решение, а также сокращает значения ошибок первого и второго рода. На выходе нейросеть выдает класс аудиозаписи (дипфейк или подлинная).

Таким образом, разрабатываемый метод позволяет с высокой точностью оценить принадлежность аудиозаписи к классу подлинников или дипфейков, а также выделить характерные для дипфейков изменения в акустических характеристиках.

Выводы. В результате проделанной работы были рассмотрены существующие методы распознавания подделки голосовых шаблонов, используемых с целью обмана системы автоматической верификации говорящего. На основании недостатков, выявленных в ходе их анализа, разработан собственный метод детектирования вмешательства в аудиопоток с целью неправомерного использования биометрических данных, позволяющий повысить точность выявления аудио-спуфинга и снизить ошибки первого и второго рода за счет уменьшения количества компонент при анализе аудиофайла и использования временных рядов.

Результаты исследования могут найти широкое применение в любых сферах, в которых применяются системы ASV, в том числе связанными с проведением финансовых операций.

Список использованных источников:

1. Awais Khan, Khalid Mahmood Malik, James Ryan1, and Mikul Saravanan. Voice Spoofing Countermeasures: Taxonomy, State-of-the-art, experimental analysis of generalizability, open challenges, and the way forward. arXiv:2210.00417v2 [eess.AS] 21 Nov 2022.
2. Chadha A, Abdullah A, Angeline L, Sivanesan S (2021) A review on state-of-the-art Automatic Speaker verification system from spoofing and anti-spoofing perspective. Indian Journal of Science and Technology 14(40): 3026-3050.
3. Aakshi Mittal, Mohit Dua. Automatic speaker verification systems and spoof detection techniques: review and analysis. International Journal of Speech Technology (2022) 25:105–134.
4. Xuechen Liu, Xin Wang, Md Sahidullah, Jose Patino, Héctor Delgado, Tomi Kinnunen, Massimiliano Todisco, Junichi Yamagishi, Nicholas Evans, Andreas Nautsch, Kong Aik Lee. ASVspoof 2021: Towards Spoofed and Deepfake Speech Detection in the Wild. arXiv:2210.02437v1 [cs.SD] 5 Oct 2022.
5. Junichi Yamagishi, Xin Wang, Massimiliano Todisco, Md Sahidullah, Jose Patino, Andreas Nautsch, Xuechen Liu, Kong Aik Lee, Tomi Kinnunen, Nicholas Evans, H´ector Delgado. ASVspoof 2021: accelerating progress in spoofed and deepfake speech detection. arXiv:2109.00537v1 [eess.AS] 1 Sep 2021.

Вернигорова А.А. (автор) Подпись

Талынкова Е.Н. (автор) Подпись

Роговой В.В. (научный руководитель) Подпись