# Biometria vs deepfake głosowy

## Jak chronić Twój głos?

Od lat rośnie liczba zagrożeń w cyberprzestrzeni, w tym oszustw. Oszuści najczęściej skupiają się na wyłudzaniu danych, środków finansowych albo infekowaniu komputerów szkodliwym oprogramowaniem. Deepfake głosowy jest nowym rodzajem ataku, wykorzystującym głos. Bazując na archiwalnych nagraniach, oszuści tworzą syntezowane głosy kluczowych osób w organizacji. Następnie podszywając się pod te osoby, usiłują wymusić na pracownikach określone działanie. Szeroko komentowanym oszustwem głosowym był opisany na łamach „The Wall Street Journal" przypadek, w którym dokonano przelewu na kwotę 243 000 dolarów.


Tego typu ataki są wciąż mało znane i trudne do wykrycia. Jeszcze do niedawna uważano, że stworzenie syntezowanego głosu na podstawie dostępnej publicznie próbki nie jest łatwe, dlatego nie stanie się powszechnym zagrożeniem. Wbrew rachubom stało się inaczej, a deepfake głosowy jest coraz częstszym zjawiskiem w cyberprzestępczości.

Próba oszukania człowieka to jedno, ale czy można oszukać technologię? Biometria głosowa polega na uwierzytelnieniu głosem użytkownika. Wykorzystując głos jako czynnik biometryczny minimalizowane jest ryzyko oszustw i nadużyć, jest to jednocześnie wygodny sposób na weryfikację tożsamości. Czy deepfake może oszukać biometrię głosową? A jeśli tak, to w jaki sposób możemy lepiej zadbać o nasze bezpieczeństwo?

### Techmo i LARA

LARA to projekt realizowany przez firmę BiometrIQ. Jego owocem będzie system wieloaspektowego wsparcia biometrii. Techmo zostało zaproszone do wzięcia udziału w pracach nad projektem, gdzie będziemy mogli wykorzystać naszą wiedzę i doświadczenie w zakresie syntezy mowy.

W ramach prac stworzymy innowacyjny system budowania nowego głosu syntezy w oparciu o minimalną próbkę głosu, czyli tzw. konwersję głosu. Pozwoli to lepiej wykrywać deepfake głosowy i skutecznie mu przeciwdziałać. W efekcie podniesie to istotnie bezpieczeństwo użytkowników biometrii głosowej oraz zmniejszy liczbę bezprawnych użyć publicznie dostępnych nagrań.

BIOMETRIQ

# Kradzież głosu, ataki prezentacji i projekt LARA

**Andrzej Tymecki**
Dyrektor zarządzający, BiometrIQ

**Skąd pomysł na projekt*?**

**Andrzej Tymecki (AT):** BiometrIQ specjalizuje się w wykrywaniu i przeciwdziałaniu oszustwom biometrycznym. Pracując nad multimodalnym, biometrycznym systemem antyfraudowym, znaczną cześć czasu poświęciliśmy analizie tzw. ataków prezentacji. Przegląd dostępnych technologii uwydatnił łatwość realizacji takich ataków, szczególnie z wykorzystaniem syntezy lub konwersji głosu. Dotarliśmy także do ewidencji ataków zrealizowanych w międzynarodowych operacjach gospodarczych. Zebrane informacje stały się katalizatorem uruchomienia projektu LARA, którego celem jest realizacja systemu detekcji oraz zabezpieczenia przed atakami prezentacji w biometrii głosowej.

**Możemy doprecyzować, czym są ataki prezentacji?**

**AT:** Ataki prezentacji polegają na wykorzystaniu sztucznych danych biometrycznych do podszycia się pod określoną osobę w interakcji z systemem biometrycznym. Wykorzystywane dane mogą być wytworzone np. poprzez odtworzenie zarejestrowanych uprzednio wypowiedzi lub wygenerowanie strumieni głosowych, przy użyciu technologii syntezy lub konwersji głosu. W efekcie systemowi biometrycznemu prezentowane są dane biometryczne imitujące prawdziwego użytkownika – stąd też pochodzi nazwa atak prezentacji.

**Już kilka lat temu, podczas wydarzenia „Techmo Day" poruszaliśmy temat vishingu (phishingu głosowego). Czy możemy mówić o rosnącej skali tego typu oszustw?**

**AT:** Zdecydowanie tak. Postęp w zakresie technologii cyfrowych umożliwił wykorzystanie głosu w procesie autoryzacji czy autentykacji użytkownika. Głos ludzki posiada przeszło 100 cech indywidualnych danego mówcy. Zaawansowane systemy biometrii głosowej są w stanie zidentyfikować mówcę z wysokim poziomem trafności, co ma ogromne znaczenie użytkowe. W sferze usług od wielu lat obsługa klienta jest realizowana w kanale telefonicznym. Wprowadzenie systemów biometrii głosowej pozwala przyspieszyć proces obsługi klienta przy jednoczesnym utrzymaniu, a często nawet zwiększeniu poziomu bezpieczeństwa. Dodatkowo, poprzez wykorzystanie technologii, takich jak Alexa, Siri, Cortana czy asystent Google, komunikacja głosowa wkracza masowo na urządzenia powszechnego użytku.

**Dlaczego deepfake głosowy jest coraz bardziej powszechny?**

**AT:** Popularyzacja głosowej obsługi urządzeń stanowi nieodpartą pokusę dla przestępców. O ile wykrycie osoby próbującej modulować swój głos tak, aby podszyć się pod inną osobę jest stosunkowo prostym zadaniem, to potwierdzenie, czy głos pochodzi z syntezatora dźwięku jest nie lada wyzwaniem. Świadczą o tym publikowane niemalże codziennie strumienie multimedialne, w których wykorzystano nowoczesne technologie do imitowania innej, często znanej, osoby. Wejście do budynku, wypłata pieniędzy z bankomatu, komendy wydawane asystentowi głosowemu – wszystkie te zdarzenia narażone są na ataki typu deepfake, w których głos autoryzowanego użytkownika może być podrobiony przez zaawansowany syntezator.

**W jaki sposób projekt LARA zwiększy bezpieczeństwo użytkowników?**

**AT:** W ramach projektu LARA stworzymy pakiet rozwiązań, który będzie wsparciem systemów biometrycznych w obronie przed atakami prezentacji. Zwiększenie bezpieczeństwa użytkowników rozpocznie się już na etapie tworzenia voiceprintów, dzięki wykorzystaniu technologii detekcji i analizy artefaktów charakterystycznych dla mówcy. Opracowany w ramach projektu autorski system oceny prawdziwości toru głosowego umożliwi nie tylko ocenę, czy emisja głosu miała miejsce w oczekiwanym źródle, lecz pozwoli również określić, czy otrzymany strumień głosowy faktycznie przeszedł określone stopnie toru głosowego. To w znacznym stopniu poprawi ocenę prawdopodobieństwa, czy strumień głosowy został wygenerowany sztucznie. Ostatnim elementem pakietu jest technologia przeciwdziałająca wykorzystywaniu strumieni akustycznych dostępnych w domenie publicznej. W dobie powszechnej internetowej transmisji strumieni multimedialnych, realnym zagrożeniem staje się synteza głosu na podstawie próbek pozyskanych z domeny publicznej. Nasze rozwiązanie zabezpieczy strumienie multimedialne przed możliwością użycia ich do syntezy lub konwersji głosu, jednocześnie nie pogarszając subiektywnych odczuć słuchaczy.

* Nazwa projektu LARA została zainspirowana fikcyjną postacią z gier komputerowych Larą Croft. Wspomniana Lara łączy inteligencję, pasję i determinację do realizowania trudnych misji a to są cechy, których nie może zabraknąć w projekcie.

**Wydawca:**
Techmo Sp. z o.o.
ul. Torfowa 1/5, 30-384 Kraków

Kraków, październik 2022 r.