From media fears to research reality: how ready are countermeasures against speaker verification spoofing?

Jasmin Liu[1], Federico Alegre[1], Benoit Fauve[1], Anssi Kanervisto[2], Md Sahidullah[2], Tomi Kinnunen[2], Héctor Delgado[3], Massimiliano Todisco[3], Nicholas Evans[3], Mauro Falcone[4]

[1]ValidSoft, UK
[2]University of Eastern Finland, Finland
[3]EURECOM, France
[4]Fondazione Ugo Bordoni, Italy

{jasmin.liu, federico.alegre, benoit.fauve}@validsoft.com
anssi.kanervisto@uef.fi, {sahid, tkinnu}@cs.uef.fi,
{delgado, todisco, evans}@eurecom.fr, falcone@fub.it

Abstract

The subject of biometrics spoofing has received more attention over the last few years within the research community. For speech technology spoofing attacks [1] relate to replay attacks [2, 6] and a category sometimes referred to as synthetic voice attacks, including attacks with artificial signals such as voice conversion and speech synthesis [3, 4, 5]. Some significant initiatives came with the ASVspoof2015 evaluation focused on spoofing with synthetic voice [3] and more recently ASVspoof2017, an evaluation focusing on replay detection [6].

We present some results about the latest development on voice anti-spoofing, most of them derived from work undertaken during the H2020 OCTAVE project [7]. Results on synthetic voice detection [Figure 2] are presented with synthetic audio from systems recently mentioned in press coverage [Figure 1]. More results are presented about the recent development of replay detection systems within the Octave project and for the ASVspoof2017 challenge. Finally, we discuss the challenges ahead, with results showing the difficulty of building up representative databases and further results showing adverse effects of codec and noise.

Fig. 1: Some of the recent coverage about vulnerability of voice biometrics systems from mainstream media including BBC, the Telegraph and the Economist.

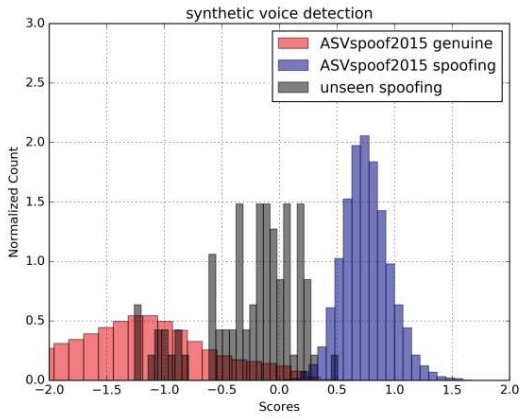

Fig. 2: Score distribution from a state of the art synthetic voice detection system on ASVSpoof2015 standard database (red and blue distributions) and on a set of 80 files (unseen set in grey distribution) from companies mentioned in recent press coverage.

References

[1] Z. Wu, N. Evans, T. Kinnunen, J. Yamagishi, F. Alegre, and H. Li, "Spoofing and countermeasures for speaker verification: a survey," *Speech Communication*, vol. 66, no. 0, pp. 130–153, 2015.

[2] F. Alegre, A. Janicki, and N. Evans, "Re-assessing the threat of replay spoofing attacks against automatic speaker verification," in BIOSIG 2014 - Proceedings of the 13th International Conference of the Biometrics Special Interest Group, 10.-12. September 2014, Darmstadt, Germany, 2014, pp. 157–168.

[3] Z. Wu, T. Kinnunen, N. Evans, J. Yamagishi, C. Hanilci, M. Sahidullah, and A. Sizov, "ASVspoof 2015: "The First Automatic Speaker Verification Spoofing and Countermeasures Challenge," in INTERSPEECH 2015, September 6-1, Dresden, Germany, 2015.

[4] M. Sahidullah, T. Kinnunen, and C. Hanilçi, "A Comparison of Features for Synthetic Speech Detection," in INTERSPEECH 2015, Annual Conference of the International Speech Communication Association, September 6-10, Dresden, Germany, 2015.

[5] M. Todisco, H. Delgado, and N. Evans, "A New Feature for Automatic Speaker Verification Anti-Spoofing: Constant Q Cepstral Coefficients," in ODYSSEY 2016, The Speaker and Language Recognition Workshop, June 21-24, Bilbao, Spain, 2016.

[6] T. Kinnunen, M. Sahidullah, H. Delgado, M. Todisco, N. Evans, J. Yamagishi, and K. A. Lee, "The ASVspoof 2017 challenge: Assessing the limits of replay spoofing attack detection," in INTERSPEECH 2017, August 20-24, 2017, Stockholm, Sweden, 2017.

[7] List of publications from the H2020 Octave project: https://www.octave-project.eu/documents/publications/