RED-AID — REspectful and capability-centreD AI Device for Preventing Call Fraud

Prof. Max Van Kleek (University of Oxford) - Principal Investigator
Dr Peter Novitzky (University College London) - Co-investigator
PDRA - TBA

INTRODUCTION

Voice-based fraud ('vishing') accounts for one of the most significant crimes in the UK, in terms of frequency, impact and pervasiveness. The harms resulting from such attacks can be devastating, both financially and emotionally. Research suggests that, as AI systems are being applied by adversaries for automated open-source intelligence (OSINT) and execution, these attacks will become even more effective.

AIM

WatchDog will be a simple IoT Edge device that aims to be a first line of defence for making individuals more resilient to phone fraud. The anticipated proof of concept will consist of 2 embedded computers with a touchscreen (Intel NUC and Raspberry Pi) interfaced to a landline telephone via an audio tap adapter [2] or with VoIP integration to an IP-based PBX. The landline is our initial target as it is often the vector of phone fraud leveraged against older adults.

Our approach is novel and centres on automatic Conversational Analysis, the use of NLP techniques to analyse conversations. To do this, Watchdog will first perform real-time transcription of phone calls using a commercial high-accuracy speech-to-text API, and then apply a custom ensemble of deep NL classifiers on the resulting transcripts to identify features that indicate the call may be fraudulent, including analysis of the dialogue for evidence of coercion and deception. These classifiers will identify fragments of conversational turns that correspond to stages of social engineering fraud action, including pretext delivery and direct elicitation/call to action. This evidence will be fused with other information, including call origin and ID. The combined result will be provided immediately back to the user, in a representation appropriate to the needs and context. This might involve, for instance, highlighting aspects of a conversation that indicate coercion or deception; displaying an aggregate risk score; a stronger intervention for a more immediate risk might include immediate call termination.

PROJECT TIMELINE

Start date: 01-06-2022
Finish date: 30-11-2022

WHY

An estimated 4m incidents of phone fraud recorded annually by the National Crime Agency [1], amounting to an estimated £190bn of loss in 2017 UK, but this figure is thought to be much less than actual figures because it is a crime that tends to be grossly under-reported by individuals.

METHODOLOGY

The ethical design of systems such as WatchDog are particularly challenging because they aim to support vulnerable individuals situated in complex networks of care and support. We will investigate the normative underpinnings of such a widely applicable dynamic ethical account, by analysing and contributing to the co-creation of a state-of-the-art ML technology for a specific vulnerable population from the capabilities approach perspective.

EXPECTED IMPACT

The rapidly ageing population has led to an increased focus on Assistive Technologies (AT) like WatchDog, including Ambient Assisted Living (AAL) technologies [5]. Their goal is to enable their users to stay active and longer in their preferred dwellings and communities, and empower their activities of daily living [3,4,6].

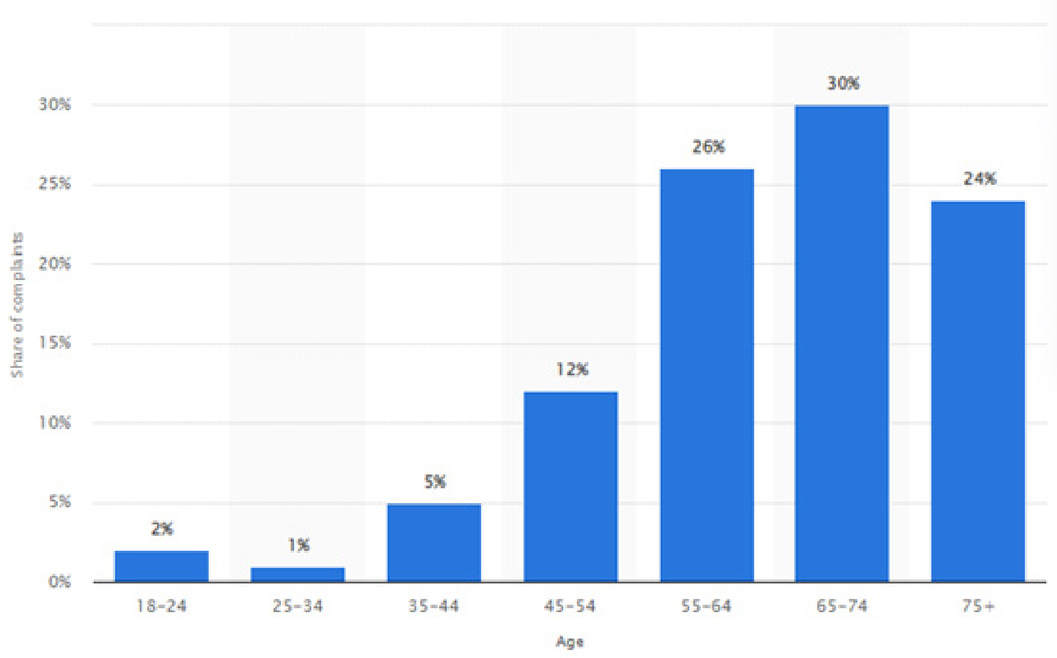

Figure 1: 'No-hang-up-fraud' complaints reported to the UK Ombudsman in 2015. Between mid-2012–2014, the ombudsman service resolved 185 complaints of financial fraud involving a no hang-up scam, brought by 173 individual consumers, 148 of which the age was known. 80% of affected individuals were over 55 [7].

Figure 2: Mock-up of Watchdog

KEY OUTCOMES

- Design and Development of a proof-of-concept Watchdog IoT Edge device for providing a 1st line of defence for older adults against phone fraud
- Increased resilience of individuals (mainly older adults) to phone fraud
- Ethical framework based on Capability Approach for Assistive Technology designed for vulnerable population.

USER PARTNERS

Howz

ACKNOWLEDGEMENTS

This work has been supported by the PETRAS National Centre of Excellence for IoT Systems Cybersecurity, which has been funded by the UK EPSRC under grant number EP/S035362/1

REFERENCES

[1] https://www.nationalcrimeagency.gov.uk/what-we-do/crime-threats/fraud-and-economic-crime
[2] Wooden case for Raspberry Pi: DIYProjects.io
[3] Marcelino, I., Laza, R., Domingues, P., Gómez-Meire, S., Fdez-Riverola, F., & Pereira, A. (2018). Active and Assisted Living Ecosystem for the Elderly. Sensors, 18(4), 1246. https://doi.org/10.3390/s18041246
[4] Sixsmith, A., & Sixsmith, J. (2008). Ageing in Place in the United Kingdom. Ageing International, 32(3), 219–235. https://doi.org/10.1007/s12126-008-9019-y
[5] Novitzky P, Chen C, Smeaton AF, Verbruggen R, Gordijn B: Informed Consent of Persons with Dementia and Ambient Assisted Living Technologies. In: Elger B, Wangmo T, Jotterand F, Ienca M (eds.): Intelligent Assistive Technologies for Dementia: Clinical, Ethical, Social, and Regulatory Implications, Oxford University Press 2019. Ch. 10., https://doi.org/10.1093/med/9780190459802.003.0010
[6] Hofmann, B. (2012). Ethical Challenges with Welfare Technology: A Review of the Literature. Science and Engineering Ethics, 19(2), 389–406. https://doi.org/10.1007/s11948-011-9348-1
[7] Distribution of complaints about "no hang-up fraud" reviewed by the Omsbudsman in the United Kingdom (UK) in 2015, by age group: https://www.statista.com/statistics/467409/age-of-victims-of-phone-financial-frauds-uk/

PETRAS · UK Research and Innovation · University of Oxford · UCL