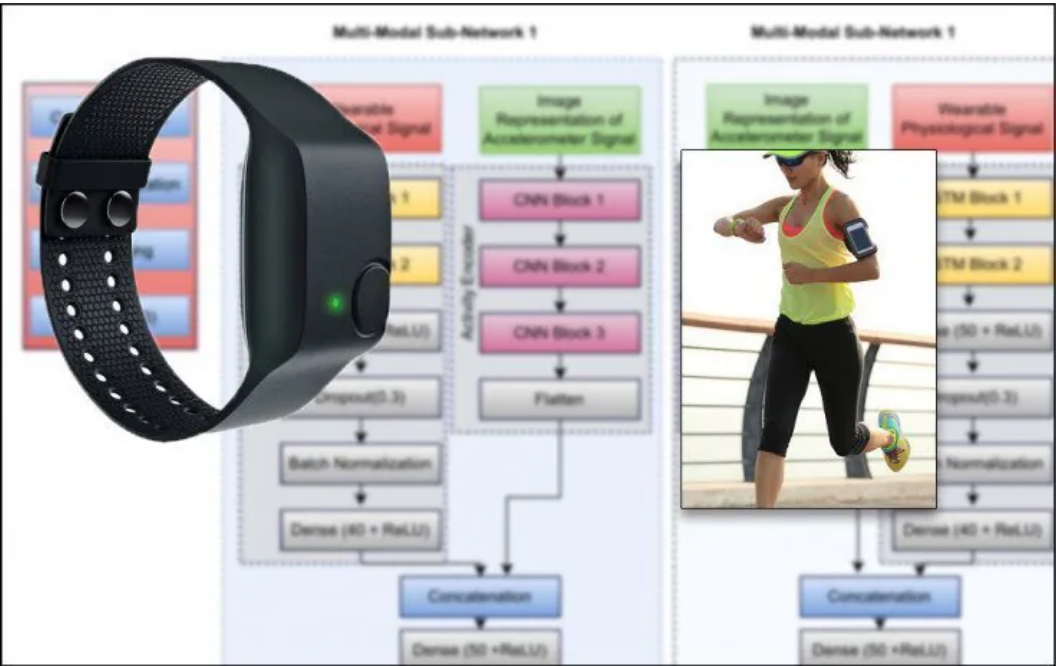


Re-Identifying People Through Wearable Health Data and Machine Learning

Originally published at <https://www.unite.ai/pri-attack-wearable-health-data-massachusetts/>



A new type of privacy attack based on wearable health data has been identified by researchers from the University of Massachusetts Lowell. [Person Re-identification Attack](#) (PRI-Attack) uses HIPAA-compliant, publicly available data from health wearables to establish the identity of individuals from heart rate, breathing and hand gesticulation data, among others.

The vulnerability is made possible in the US by the fact that the Health Insurance Portability and Accountability Act (HIPAA), while requiring that medical data remains anonymous, does not consider raw sensor data (such as skin temperature and accelerometer (ACC) data) as being privacy-sensitive, and therefore does not require that publicly-shared data of this type be encrypted or subject to the same general protections it affords to traditional forms of patient data, such as health records.

From Vector To Visual

A PRI-Attack uses interpreted image data to discern common patterns that correlate to other types of health data. A person's skin response, for instance, can be [evaluated from video](#) (photoplethysmography), and correlated to what ought to be completely anonymous vector information from health-monitoring devices such as wearable watches, and other kinds of monitoring apparatus. Photoplethysmography yields heart-rate data, which can be paired up with non-identified wearable cardiac data.

Gesture recognition is another 'key' that can be trivially translated from vector data into a visual matrix that, again, allows interpreted image/video data to be correlated to apparently anonymous accelerometer information in health data.

Up	Down	Left	Right
CW	CCW	Z	AZ
S	AS	Push	Pull

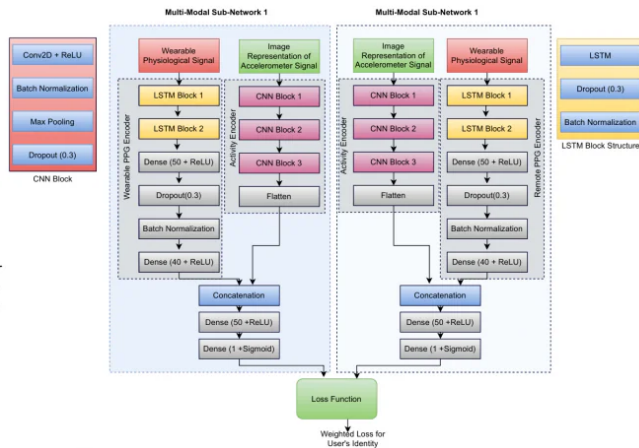
Hand gesture information from wearable data. Source: <https://arxiv.org/pdf/2106.11900.pdf>

Sensor Data As PII

The research, from UML Assistant Professor Mohammad Arif UI Alam, contends that physiological sensing data can indeed constitute PII, and is in effect a biological analogue of the browser fingerprinting techniques currently [believed](#) to undermine new initiatives to protect user privacy on the web.

To test the hypothesis, the researcher developed a hand-gesture recognition and localization framework that interprets gesture data (recorded vector-based movement) from a wearable accelerometer, and translates the movements into a visual record that can be correlated to movements recorded by wearable health devices.

A Multi-Modal [Siamese Neural Network](#) (mm-SNN) was constructed to interpret gesture information classified via Support Vector Machine (SVM). One network deals with the vector information (interpreted as image information in a 3D space) and the second network treats of the physiological data recorded from sensor data.



Testing

The system was tested on various datasets, including a 'Gamer's Fatigue Dataset' obtained by collecting data on five volunteer students, aged 19-25, who played videogames for seven days while wearing the Empatica E4 [wristband](#). The watch features ACC, electrodermal context (EDA), skin temperature and photoplethysmography (PPG) sensors.

The E4 was also used in a novel 'restaurant data' dataset, wherein eight volunteers prepared and ate sandwiches for twenty minutes, and in an 'older adults' dataset, where 22 older subjects, aged 75-95, performed 13 scripted activities while wearing the watch.

Finally, the researchers used the [publicly available](#) 'Healthy Adults Fatigue Dataset', which monitored 28 healthy men and women with an average age of 42 over 1-219 consecutive days while wearing a multisensor wearable device broadly similar to the data-gathering capabilities of the E4, including a 3-axis ACC, galvanic skin response electrode, temperature and photo sensors, and a barometer.

The results indicate that heart rate and breathing rate are the surest means to reidentification, scoring an averaged >66%+ accuracy rate.

Data #Users	PPG	HR	BR	BVP	IBI	EDA	TC	PC	Temp
D1 (5)	69.65 ± 5.4	70.45 ± 2.4	71.58 ± 3.1	66.65 ± 4.6	67.38 ± 3.8	52.43 ± 9.4	53.42 ± 8.9	54.97 ± 7.6	49.23 ± 8.2
D2 (8)	67.12 ± 5.1	69.29 ± 3.5	70.54 ± 2.9	65.86 ± 3.9	65.43 ± 3.1	50.87 ± 9.1	50.38 ± 9.1	51.64 ± 8.1	48.43 ± 7.2
D3 (22)	66.44 ± 6.6	67.38 ± 4.7	67.16 ± 5.1	64.95 ± 4.5	64.88 ± 4.9	50.63 ± 9.6	51.75 ± 8.4	50.76 ± 7.5	45.5 ± 11.4
D4 (28)	65.65 ± 7.1	66.95 ± 4.3	66.85 ± 4.1	65.82 ± 6.1	66.75 ± 5.3				34.8 ± 12.5

Results from testing the PRI-Attack methodology. Crib: PPG: photoplethysmography; HR: heart rate; BR: breath rate; BVP: Blood Volume Pulse (obtained from PPG); IBI: Inter Beat Interval (obtained from PPG); TC: Tonic Component of EDA signal; Phasic Component of EDA data (Ibid); Temp: Temperature.

The research concludes:

'While modern computer vision technology can be easily utilized to learn hand gestures and corresponding physiological signal (heart rate, breathing rate) from public surveillance camera, these huge amount of recorded videos can be easily utilized by the attackers to learn user specific biometrics to reveal identity from HIPPA compliant serve stored wearable sensing data.'

HIPAA Considers PHR Data ‘Anonymized By Default’

The US government has acknowledged the growth of personal health records (PHR), and [classifies](#) such a record (including data from health wearables) as ‘an electronic record of an individual’s health information by which the individual controls access to the information and may have the ability to manage, track, and participate in his or her own health care’.

Nonetheless, since this is a phenomenon from the private sector, the government concedes no official oversight of such data, having established that it does not contain personally identifiable information (PII). A [report](#) in June 2016 on non-covered HIPAA entities from the U.S. Department of Health and Human Services states:

‘[Large] gaps in policies around access, security, and privacy continue, and confusion persists among both consumers and innovators. Wearable fitness trackers, health social media, and mobile health apps are premised on the idea of consumer engagement. However, our laws and regulations have not kept pace with these new technologies.’

Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG's Brahma.ai.

Portfolio site: martinanderson.ai

Contact: martin@martinanderson.ai



Discover More