

HEATHER SHAW

THE IDENTITY IN EVERYONE'S POCKET

When people interact with their smartphones, the digital traces left behind can be used to infer their identity.

Around a quarter of an adult's daily behaviour is spent on their smartphone (Ellis et al., 2019; Shaw et al., 2020). As such, smartphone usage data can reveal important insights into a person's daily habits and can be used to infer their identity.

“ It was possible to find within a top-10 list, the person whom the application usage data belonged to 75% of the time.

In our study of 28,692 days of smartphone data usage from 780 people, we ranked each application from the most to least used per day, for each person. We found that people were consistent in their application

usage patterns on a day-to-day basis (e.g., consistently used Facebook the most and the calculator application the least). When we examined two randomly selected days from the same person, we found greater similarity in application use patterns than when we randomly selected two days that belonged to two different people.

To explore if application use could identify a single person, we fed 4,680 days of application usage data (equating to 6 days per person) into machine learning models. The models learned people's usage habits from the 6 days of application data and then tried to predict a person's identity when presented with an anonymous seventh day of data. The model was able to identify the correct person one-third of the time. Daily smartphone use can therefore act as a digital fingerprint.

The results further showed that it was possible to find within a top-10 list, the person to whom the application usage data belonged 75% of the time. In practical terms, this means that an investigation seeking to find a criminal's new phone from knowledge of their historic phone usage could reduce a pool of ~1,000 people's phones to 10 phones, with a 25% risk of missing them.

Our results suggest that access to smartphone application use data allows for a reasonable prediction about a person's identity even when they are logged-out of their account. This identification is possible with no monitoring of the conversations or behaviours within the applications themselves. Therefore, it is important to acknowledge that application usage data alone could risk our privacy if it is misused. It also questions whether usage data should be protected in the same manner as other personal identifiers.

Dr Heather Shaw is a lecturer in Psychology at Lancaster University.

