Your motion sensors know more about you than you would think.

It's essential for information communicated through your smartphone to only reach its intended receiver. Using only easily available motion sensor data measured while typing, though, it is possible to extract more than half of keystrokes typed.

With the emergence of Artificial Intelligence and Machine Learning, systems conventionally believed to be secure may be more vulnerable than previously thought. Due to the immense potential of machine learning to quickly sift through and learn from large amounts of data, algorithms can increasingly use unexpected sources of information to uncover secrets seemingly unrelated to the information gathered, from electrical devices. This, using seemingly unrelated information to covertly infer something about a system, is known as a side-channel attack. Non machine learning examples of side-channel attacks go far back; in 1943 an American state of the art encrypted teletype was found to have a fatal flaw. The electromagnetic emissions could be picked up by a freestanding oscilloscope, and the information captured could then be translated into the exact text that was being encrypted. In the 80s, Soviet spies could plant a bug in an IBM typewriter to monitor the electrical noise generated by the machine, and as different keys had specific characteristics, the plain text written could easily be retrieved.

More modern examples include looking at the movement of the torso while typing on the computer keyboard in a video call to extract what is being written. Now, according to the author of the thesis "Keystroke Classification of Motion Sensor Data - An LSTM Approach", it may also be possible to extract text and passwords written on your smartphone keyboard from data measured by motion sensors. But that can't be that big of a problem, apps need permission to read from the phone's embedded sensors, right? Turns out this is true for many sensors, like the camera and microphone, but *not* for the motion sensors accelerometer and gyroscope, measuring the acceleration force and rate of rotation along all three axes, respectively. So, any app on your phone may use them for any purpose, without your knowledge.

The idea is simple. Imagine for example typing the letter "q" on your smartphone's soft keyboard, this could potentially cause you to slightly tilt your phone leftwards, which the gyroscope would pick up on. Advanced machine learning networks (the "LSTM", or "Long Short Term Memory network" in the title) could then use this information to understand exactly what is being typed.

And it works! At least to some extent; half the keystrokes get classified correctly. This is using text only seen once, the model would get better results on strings typed recurrently, like passwords and PIN codes. While on the topic of PIN codes, there are good reasons to believe the model to have better results in situations where there are fewer possible keys to classify between, like, as in the case of PIN codes, only numbers. And keep in mind that the models are only taking the motion sensor data into account, also using word and sentence legibility could improve the results in some areas.

So, is there need for worry? There might, but there are ways to deal with the problem. First, phone makers may need to reconsider whether the motion sensors' data is unintrusive enough to not require permission. Before that is done, we users need to be more careful with what we allow on our phones. We may also need to find other ways to stay safe online, like two factor authorization and behavioural biometric authentication methods. The latter is a somewhat new field that often uses machine learning techniques to make sure that only *you* have access to your accounts. In this way, machine learning can be used to help solve the problem it has created.