Determining whether Two Devices are on the Same Person using Accelerometers

Zachary Stence and Torry Johnson

Example

Suppose two people, Alice and Bob, buy health monitoring sensors that pair to their smartphones through cryptographic means. This works off of the assumption that the owner of the device is the only one to ever wear it. Suppose one day Alice accidentally wears Bob's sensor, and Bob wears Alice's. Throughout the day, Alice's sensor will be recording Bob's data and vice versa, mixing up their health records. Obviously this is a problem because if one of them has a serious health issue, the data would appear on the other person, causing both a false positive and false negative. The likelihood of a situation like this occurring will only increases as such sensors become more prevalent in society.

To fix this, we propose authentication of the user of the device using accelerometers to detect whether the owner is currently wearing the device.

Objective

- Wearable devices often have no authentication techniques to verify user
- Use various machine learning algorithms from `scikit-learn` [5] to analyze data from wearable device accelerometers
- Binary classification regarding whether or not two streams of accelerometer data are from devices being worn by the same person [4]

Potential Challenges

1. How to effectively use `scikit-learn`, what functionality from `scikit-learn` we need to use
2. Using `.mat` files from the UniMiB-SHAR [3] dataset in python
3. Manipulating UniMiB-SHAR data useful format
    - Extracting the data we need
    - Extracting features
    - Splitting into training and testing data
4. Understanding the complicated math behind feature matrices, coherence, and other signal processing techniques
5. Collecting our own data (if time permits)

Potential Solutions

1. Follow tutorials, read documentation, look for similar projects
2. Find a python package that can read `.mat` files into a useful data structure (scipy.io)
3. Conceptually figure out desired format, then figure out how to get the data to look like that (NumPy)
4. Research signal processing, starting with the fundamentals and building up to what we need to understand
5. Use apps that log accelerometer values into a spreadsheet, then read in python similar to other dataset

Project Plan

Meetings

Individual Meetings:

Monday at 9:00am

(rescheduled if missed)

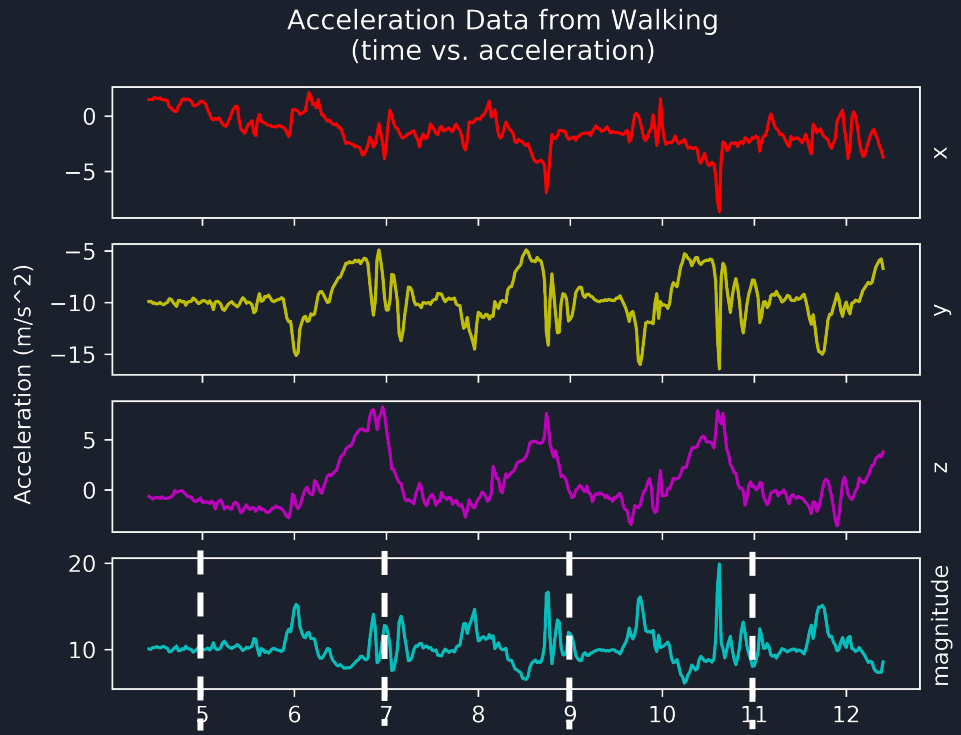
REU Group Meetings:

Wednesday at 11:00am

Resources

- UniMiB-SHAR dataset
- Smartphone
- Smartwatch
- Sensor Kinetics iPhone App
- Power Sense iPhone App
- Android App
- Manually collected data

Process of Data Analysis



Coherence of two feature matrices
Algorithm *classifies* as 'on same body' or 'not on same body'

$C(A, B)$ → Machine Learning Algorithm → 0 or 1

$\text{magnitude} = \sqrt{x^2 + y^2 + z^2}$

$F_1 = \{f_1, f_2, f_3, ...\}$ $F_2 = \{f_1, f_2, f_3, ...\}$ $F_3 = \{f_1, f_2, f_3, ...\}$

feature vectors

$\cdots$

feature matrix

$$A = \begin{bmatrix} \vdots & \vdots & \vdots & \\ F_1 & F_2 & F_3 & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

Tasks

Tasks	Deliverable	Completion Date		Problems If Missed
		Expected	Actual	
In-depth machine learning research and familiarization	Good understanding of machine learning goals/techniques	5/21	5/21	Trouble understanding breadth/depth of project
Decide how UniMiB SHAR dataset will be used or if we will collect and use our own data	Know which dataset(s) we will use and why	5/23	5/21	Need to decide ASAP
Brush up on Python using online tutorials and coding our own examples	Progress in SoloLearn tutorials	5/23	5/22	Working with data may be difficult
Figure out how to import raw data for use in Python (full_data.mat)	Example program importing a small accelerometer dataset	5/23	5/21	Delay processing, finish ASAP
Determine features of dataset and how data will be used with machine learning algorithms	Outline of data and how it will be used (including specified features)	5/24	5/24	Proposal presentation may lack important info, need to finish ASAP
Project Proposal Submission	Completed Project Proposal	5/25	5/25	Finish ASAP
Determine what the data should look like • how will training/testing take place? • classify data as on same body/not	Conceptual idea of format of data and how it will interact with scikit-learn	5/28	5/28	Proposal presentation may lack important info, need to finish ASAP
Proposal Presentation	Completed Presentation	5/30	5/30	Complete ASAP

Tasks (continued)

Tasks	Deliverable	Completion Date		Problems If Missed
		Expected	Actual	
Get data into correct format (features) ready for use with scikit-learn	Dataset in python in the correct format for scikit-learn	6/8	6/6	Finish before train/test deadline
Outline what our Python code will accomplish	Generic outline of code	6/11	6/8	Difficulty coding program
Code preliminary program in Python for testing one machine learning algorithm	Complete python program	6/14	6/11	First training/testing will be delayed
Train and test our data with at least one algorithm	Results of training/testing (confusion matrix)	6/15	6/11	Finish ASAP. Needed by midterm presentation
Midterm Presentation	Completed presentation	6/20		Finish ASAP
Midterm Progress Report	Submitted midterm progress report	6/22		Finish ASAP
Midterm Report	Completed midterm report	6/22		Finish ASAP
Determine which classification algorithm will work the best with our dataset	List of classification algorithms ranked by train-time, test-time, and score	6/25	6/13	Finish soon, may run out of time
Tweaking dataset/algorithms used to solve any unexpected issues that may arise	Fully functional algorithm operating on data	7/7	6/13	Finish ASAP to allow for more in depth analysis
Find best classification algorithm(s)	Confidence level of authentication based on length of data	7/9		Finish ASAP

Tasks (gray cells are additional goals to be achieved if time permits)

Tasks	Deliverable	Completion Date		Problems If Missed
		Expected	Actual	
Obtain statistics about data and algorithm (visualizations) • Confusion matrix • How much time is needed for authentication? • How reliable is authentication? • How do different algorithms perform?	Graphs/plots and information/ statistics about data, algorithms, and overall project	7/11		Finish ASAP for final report
Collect and use our own data	Full dataset of our data collected	N/A		Finish if time permits
Further optimize analysis by refining data and/or algorithms used (hyperparameters)	Machine learning model with improved functionality	N/A		Finish if time permits
Obtain statistics after optimization and observe the increase in performance	Graphs/plots and information/ statistics about data, algorithms, and overall project	N/A		Finish if time permits
Create an example authentication app based off of data collected and analyzed in real-time [7]	Working app prototype	N/A		Finish if time permits
Final Report/Presentation	Completed final report/presentation	7/18		Finish ASAP
Poster Presentation	Completed poster	7/19		Finish ASAP

References

[1] Bianchi, Andrea and Ian Oakley. "Wearable authentication: Trends and opportunities" Usable privacy and security / Alexander De Luca, Emanuel von Zezschwitz. it - Information Technology, 58.5 (2016): 255-262. Retrieved 25 May. 2018, from doi:10.1515/itit-2016-0010

[2] Cornelius, Cory T, and David F Kotz. “Recognizing Whether Sensors Are on the Same Body.” Pervasive and Mobile Computing, vol. 8, no. 6, Dec. 2012, pp. 822–836. ScienceDirect, doi:10.1016/j.pmcj.2012.06.005.

[3] Micucci, Daniela, et al. “UniMiB SHAR: A Dataset for Human Activity Recognition Using Acceleration Data from Smartphones.” Applied Sciences, vol. 7, no. 10, 24 Oct. 2017. Applied Sciences, doi:10.3390/app7101101.

[4] Prettenhofer, Peter. “Classification with Scikit-Learn.” DataRobot Blog, 3 Mar. 2014, blog.datarobot.com/classification-with-scikit-learn.

[5] “Scikit-Learn: Machine Learning in Python.” Scikit-Learn, Oct. 2017, scikit-learn.org/stable/index.html.

[6] Taspinar, Ahmet. “Classification with Scikit-Learn.” Ahmet Taspinar, 1 Mar. 2018, ataspinar.com/2017/05/26/classification-with-scikit-learn/.

[7] Xu, Weitao. “Gait-based authentication system on smart wearable devices.” Online video clip. YouTube. YouTube, 22 Aug. 2016. Web. 21 May 2018.