INTERMEDIATE PROJECT REPORT
IMU DATA CLASSIFICATION

AUTHOR: JAN CHODOROWSKI
INSTRUCTOR: WITOLD PALUSZYŃSKI, PhD

DEPARTMENT OF ELECTRONICS
WROCŁAW UNIVERSITY OF TECHNOLOGY

Wrocław, 29.01.2017

ABSTRACT

The aim of this project was to build a system capable of recognizing which mean of transport an individual is using. The data should be gathered with inertial measurement unit of a smartphone.
The smartphone application should be capable of sending gathered data to the remote server, so that multiple users can provide their data and no precious data samples are lost. With the data stored on the server classification could be performed giving as precise information about the transportation mean as possible.

INTRODUCTION

The data was gathered with help of a mobile application. It has three buttons with names of transportation means. If any of these three buttons is pressed the data collecting begins. Data being collected comes from Inertial Measurement Unit and involves accelerometer and gyroscope readings.

Data is being saved into file in following format:

class,timestamp, accX, accY, accZ, gyrR, gyrP, gyrY

where:

class – is the name of transportation mean from the button

timestamp – time in ms measured when the sample was taken, it is measured from beginning of the acquisition

accX, accY, accZ – accelerometer readings along axis X, Y and Z

gyrR, gyrP, gyrY – angular displacement in three dimensions

The samples are taken in 20 ms intervals. Knowing the exact sampling rate is important while investigating frequency domain.

Having collected data the acquisition is stopped by pressing the active transportation mean button again. Then the user may send the collected data to remote server by pressing SEND DATA button.
The user may also remove old collected data by pressing CLEAR DATA button.

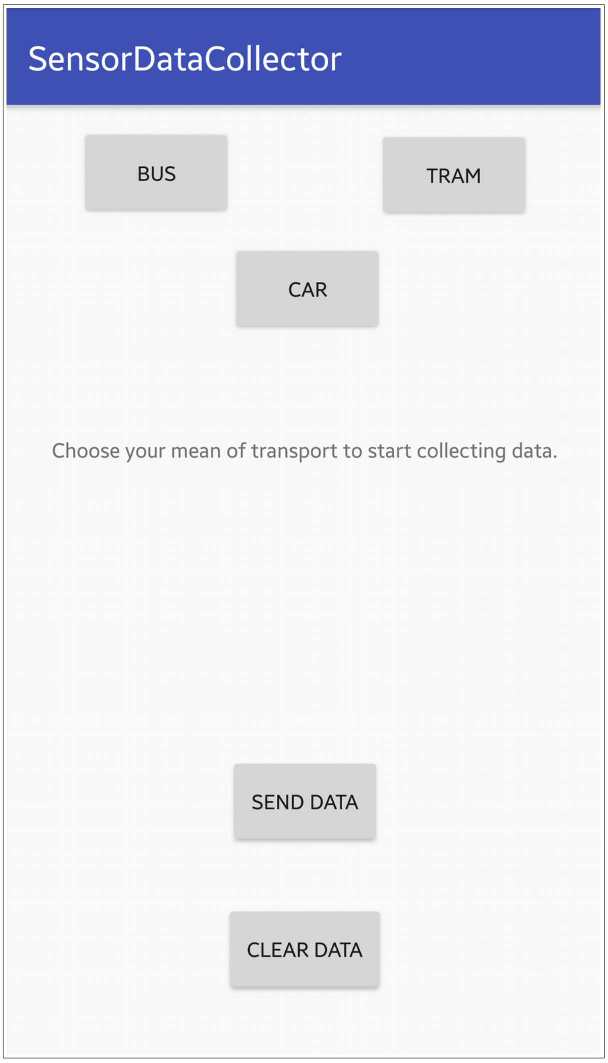

Fig.1 Applications user interface.

SERVER SIDE

The remote server is a ReadyNAS® 104 Network Attached Storage (NAS). The device physically is located in authors home. Because authors internet provider uses non-static IP, additional effort was made to assure constant server access.
The above was done by installing NO-IP Dynamic DNS software. The software using cron job checks for IP changes. If it detects such a change it send information to NO-IP DNS server.

The software web server used was NGINX with PUMA application server on top of it.

The web application was developed with Ruby on Rails 5 framework. The application provided RESTful API for storing collected IMU data.

DATA PREPROCESSING

When reasonable amount of data was collected, before classification, additional transformations had to be performed.

The following steps were taken in order to adjust data to classification purposes:

1. The intervals of 10 seconds were considered, therefore 500 samples were processed at once.
2. Norm of three accelerometer vectors was calculated
3. Norm of three gyroscope vectors was calculated
4. Accelerometer data norm with removed DC component[1] was calculated
5. Gyroscope data norm with removed DC component was calculated

FEATURE EXTRACTION

The following features were extracted for each accelerometer and gyroscope time window:

1. Signal energy
2. Signal standard deviation
3. Signal variance
4. Signal skewness
5. Signal kurtosis
6. Three main frequencies of the signal

All data processing and feature extraction was done within a python script.

CLASSIFICATION

Obtained in feature extraction process data samples were used in classification process in WEKA environment.

The most effective algorithm as Random Forest which achieved accuracy of 85.39%. The algorithm had unconstrained depth of a single tree and 1000 iterations. The experiments were conducted with 10 fold cross-validation.

1 Term originated in electronics, is the mean value of the waveform.

The ultimate goal of the project was to build system capable of determining which mean of transport an individual is using. Assuming an individual is uploading to server data from only one mean of transport at a time additional classification process had to be done to determine accuracy.

Dataset	car1	car3	bus3	tram1	tram3
Correctly Classified Instances Percentage [%]	78 83.871 %	204 94.4444 %	65 95.5882 %	7 11.1111 %	13 12.381 %
Incorrectly Classified Instances Percentage [%]	15 16.129 %	12 5.5556 %	3 4.4118 %	56 88.8889 %	92 87.619 %
Mean absolute error	0.2675	0.1656	0.1584	0.5254	0.5441
Root mean squared error	0.3346	0.2287	0.2188	0.5933	0.6129
Relative absolute error	82.1186 %	48.5689 %	35.7429 %	87.7804 %	89.5292 %
Root relative squared error	93.2733 %	60.9059 %	43.7716 %	92.8205 %	94.4578 %
Total Number of Instances	93	216	68	63	105

The table above shows that classification for class „car” and „bus” are more or less accurate when the „tram” class has accuracy much worse than a flip of a coin.

The inaccuracy of classification „tram” class may be caused by much smaller amount of this class samples in whole dataset but the problem demands further investigation.

SUMMARY

Great majority of the project goals were achieved. The acquisition, preprocessing, and classification system was successfully built and tested. The classification of uploaded data samples was performed successfully for two of three classes. The “tram” class came out to be impossible to classify. That may be caused by many reasons; bad preprocessing, not enough features extracted, too few data samples. In future further investigation will be conducted to find the reason of such a bad “tram” classification.

REFERENCES

1. Motion Mode Recognition and Step Detection Algorithms for Mobile Phone Users, Melania Susi *et al.*

2. Accelerometer Signal Processing for User Activity Detection, Jonghun Baek *et al.*