

# EXTRACTING INFORMATION FROM WI-FI TRAFFIC ON PUBLIC TRANSPORT

ANDRÁS BÁNHALMI<sup>1</sup>, VILMOS BILICKI<sup>1</sup>, ISTVÁN MEGYERI<sup>1</sup>, ZOLTÁN MAJÓ-PETRI<sup>2</sup> & JÁNOS CSIRIK<sup>1</sup>

<sup>1</sup>Faculty of Science and Informatics, University of Szeged, Szeged, Hungary

<sup>2</sup>Institute of Business Studies, Faculty of Economics and Business Administration, University of Szeged, Szeged, Hungary

## ABSTRACT

The utilization and the quality of public transport are important for the customers, maintainers and service providers. Passive measurement techniques, when humans are not involved are the cheapest way for collecting large amounts of long-term data from multiple public transport lines. Useful data can be collected from various sources, such as from cameras, infrared sensors and Wi-Fi routers. We addressed the problems of estimating passenger counts in two different ways, and also to get travel statistics like the number of passengers getting on or off a vehicle at a bus stop; and even to compute an origin–destination matrix from Wi-Fi monitoring data. In this study, we focus on Wi-Fi data, which can be still useful for extracting relevant data after many years. Here we describe Wi-Fi data collection methods, and then prove the usefulness of applying simple artificial intelligence-based methods to extract information from the huge amount of Wi-Fi data. We will also show that ‘lower-level re-estimation’ can be useful for further optimization, which means that globally modelled data may have to be re-modelled on partially selected groups to get better results. Namely, after building linear models and estimating absolute and relative errors, we found that the relative error of the Wi-Fi-based estimation can be markedly reduced if data are processed and analysed in more detail. When a daily Wi-Fi analysis is split into between-stops parts, an additive linear correction can be computed and applied to these parts, and as a result, the relative error of estimates can be reduced.

*Keywords: axle load-based estimation, public transport, Wi-Fi frame monitoring.*

## 1 INTRODUCTION

The purpose of a public transport system can be viewed from different aspects. The most important ones stated are that it should help people change stops quickly, reduce car traffic and reduce pollutant emissions. Going deeper, the quality of public transport can be improved, if overcrowding on some lines can be reduced or when new direct lines are put in service. But if the maintainer is trying to cut costs, it may be necessary to reduce the frequency of some lines.

Hence, when the maintainer of the public transport system addresses these key questions, the answers can be found by solving lower-level tasks. Many of these tasks are related to the statistics that contains information about how many passengers travel from one point to another, which line (or lines) of public transport they use and when they travel.

A traditional way of acquiring this information was via a questionnaire, and then the completed forms were statistically analysed. Another common way for passenger counting is when it is done on the vehicles by humans. The problems with the traditional methods are that these methods are expensive and so can be used only for short time periods.

Over the past few years, many different methods have been elaborated in different kinds of public transport studies (some of them will be briefly described in the next section). We also sought to develop and describe a methodology that can be readily used for analysing the actual state of a public transport system. We will mention the limitations and also the errors of the proposed solutions. Our two basic methods are based on the Wi-Fi traffic and

on telemetry. These are not new ideas. However, the novelty is that we describe our analysis on these data and data processing in a systematic way and propose some lower-level analysis methods to reduce the estimation error further.

In the following sections, first we will summarize the state-of-the-art of the problem counting passengers who travel from one point to another. Then, our data collection methods will be described in more detail (data collections of Wi-Fi, telemetry and handcrafted data). After we will describe our actual results on the analysis of Wi-Fi packages, compared to telemetry-based estimates, from different aspects. The key idea will be, besides extracting global relations of the data involved, that if datasets are divided up in some way (daily data into between-stops data) and these parts are also analysed, then better results can be obtained. Our proposed method gives a reduction in relative error values from 22% to 8%.

## 2 RELATED WORK

Modelling the movement of people in a public transport system, measuring a crowd in a public square or a road, or measuring the traffic in a city were discussed earlier, and several models have been created for handling these tasks. Here, we focus on counting passengers, and the methods can be divided into two main parts.

The first part is when we do not wish to track people. This set of methods can be used for counting people on the vehicle, but they cannot be used for computing an origin–destination (OD) matrix. In this case, one can use telemetry data, if they are available. From telemetry data, the axle load can be used to estimate the load mass, and from this, using a linear regression model, an estimate can be computed on of how many people are on board [1, 2]. This method is more accurate than other estimation methods, such as infrared sensor-based methods [2]. When infrared sensors are employed, these sensors are typically used at the door for sensing the movement and counting the passengers getting on or off. Other state-of-the-art methods use cameras and object detection methods to count people [3, 4]. Each method has its own special problems, when it is utilized in a real environment. The mass-based passenger number estimation cannot handle widely varying values of weights, and when the vehicle is moving, the measured weight is noisy. When a camera is used, one problem is that covered or hidden people cannot be detected, but people outside a bus can be detected. When multiple cameras are used, the same person may be detected by several cameras.

The second part is when some kind of tracking can be carried out. In this case, the beginning and end points of a journey can be estimated, and after some aggregation, an OD matrix can also be computed. People can be tracked, if on-board camera systems are used [5], or something that belongs to some passengers and transmits a signal which can be tracked. If the transport system uses smart cards, these can be used for counting of passengers [6]. Another approach is the Wi-Fi- or Bluetooth-based tracking [7-11]. In the London Underground, an enormous amount of Wi-Fi data was collected to analyse the public transport [12]. Also in the London Underground in another project, the accelerometer data patterns collected by phones were investigated to identify journeys [13].

From these various approaches, here we will focus on Wi-Fi package measurement-based solutions, which may also have some drawbacks in the long term. It is expected that tracking devices with Wi-Fi capabilities will be harder or impossible in a couple of years because MAC address randomization will become commonplace [14, 15]. So in the next decade, many other solutions will be preferred for these kinds of tasks. In the SASMob project, which seeks to make the public transport system smarter and improve the quality, we will also do experiments by applying the above-mentioned alternative techniques (e.g. object detection and tracking-based ones).

Table 1: Three kinds of data were collected. Firstly, telemetry and manual passenger counts are compared (first row), then the relation between telemetry and Wi-Fi data are analysed (second row).

Data 1	Data 2	Data 3
Manual passenger counting, 16 hours	Telemetry data, 16 hours long	
	Telemetry data, 15 days long	Wi-Fi data, 15 days long

### 3 COLLECTING DATA

We have three different data sources:

The ‘Data 1’ data were collected by humans and they contain the passenger counts on a trolley bus on a specific date (from 18 April 2018 04:30:10) for 16 hours, including the numbers of those getting on or off at each station.

The ‘Data 2’ data were taken on the same trolley bus on the same day and the time interval of Data 1. The trolley has a telemetry system (type: SKODA electric control system, ATM data recorder, <http://www.atmavio.pl/en/rail-vehicles/user-list/>). The data of this telemetry system contain, for example, the three axle loads, velocity, GPS data and door state indicators (open-closed) measured per second.

The ‘Data 3’ data were collected for 15 days on the same trolley bus, using a Wi-Fi router (type: MikroTIK LtAP mini LTE) in monitoring mode. ‘Data 2’ are also available for this time period. These data were collected from different routes depending of the trolleys bus schedule.

‘Data 1’ and ‘Data 2’ need no further explanation, but the ‘Data 3’ Wi-Fi database needs to be described in more detail. Data were acquired by a hardware set-up which contained a mini-PC (Raspberry pi 3+), a Wi-Fi router mentioned earlier and a GSM module (for uploading the measured data from time to time). When the Raspberry booted, the Wi-Fi router was set to monitoring mode, and tshark was run to log all the Wi-Fi data with a type of Probe Request packets to a file. Only channel 1 (2,412 MHz) was monitored by the router because the Wi-Fi capable devices send probe requests on each channel [16]. The logs of course contained the timestamps, MAC addresses (they were hashed for privacy) and RSSI (signal strength).

### 4 PREPROCESSING THE DATA

Firstly we analysed the relation between the manual passenger counts (‘Data 1’), and the axle load data (from ‘Data 2’). As a linear relation is expected [1], a linear regression model was applied. From this, we got an estimation of the load, when the vehicle was empty, and after an optimal multiplier was calculated to get the number of passengers, when the total mass was known. However, when the trolley bus is moving, the measured axle weights are very noisy, for various reasons, such as the road surface and acceleration (see Fig. 1). So some data aggregation is needed for a more precise estimation.

When the Wi-Fi frames are processed, there is much cleaning work and pre-processing to do. Many packets or frames may have come from the surrounding areas and not from the devices in the trolley bus, and these have to be filtered out. In the section ‘Data cleaning’, we describe what is needed to filter the data efficiently.

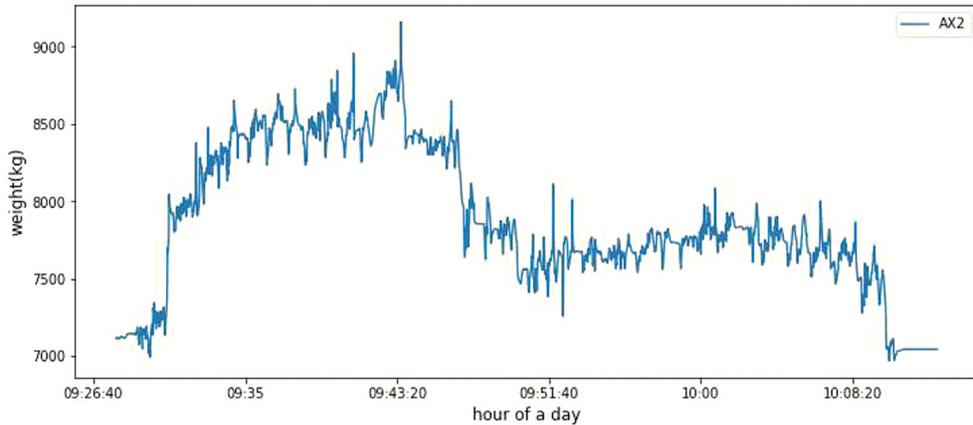


Figure 1: The raw signal of an axle load, in time.

#### 4.1 Data cleaning and pre-processing

We use different cleaning and pre-processing methods for the telemetry data and for the Wi-Fi collection.

From the telemetry data available, we just use the axle loads, timestamp, velocity, GPS and door state data. All these measurements are available every second. The first step in the pre-processing is that we need to know when the vehicle stops at a stop. For this, we use the velocity and the door state. The following algorithm is used: the bus is at a stop from time  $t_0$  to  $t_1$ , if between these times the velocity is 0, and when at least one door is opened. The state of the door is useful for deciding whether the vehicle is standing at a stop and not at a traffic light. Between two stops the number of passengers will not change, but the axle loads can vary greatly. Some strategies were tried to estimate the real weight of the trolley between stops. First, one can use the weight measured at the beginning or at the end of a stop. Another way is to aggregate in some way the measurements between stops, like computing a median value. From our experiments, we found that when the median value was used as an estimator, the correlation was the largest between the estimated passenger count and the gold standard manual count. So after the first phase, we have the median values for estimation. In the next part, we define the line number and the direction of the trolley bus and the stop identifiers for each stop. For this, for each time interval at a stop, a median GPS latitude and longitude are calculated, and from an available list, the identifiers (stop names) are defined. A list of lines with terminal stops is also available, so combining the information the concrete line and direction can be assigned.

For the Wi-Fi data cleaning, the whole dataset, which contains the timestamp, MAC address and RSSI is sorted by MAC. Then, iterating over all the MACs, ‘MAC trips’ are determined by connected timestamps. It is required because any device with a given MAC may be on a vehicle several times a day. The ‘MAC trips’ determination method finishes one trip and begins a new one, if the time difference between consecutive timestamps is larger than a threshold (for this we used 30 min, which is greater than the maximum travel time of the trolley on one line in one direction). Then, ‘MAC trips’ are filtered. If the distance between the beginning and the end point of a trip (computed from GPS coordinates) is below a threshold (here 200 m, which is greater than the range of Wi-Fi), then this trip is deleted because we think that this was a signal from the neighbourhood. If the time duration of a trip

is larger than a threshold (here 60 min), then this trip is also eliminated because these Wi-Fi signals were collected from a device that is permanently on the trolley bus. As the stops are also determined beforehand, another filter is applied too: if a trip does not contain at least two stops, then the trip is eliminated. We have to notice here that the signals from the randomized WIFI devices will be lost during this procedure.

#### 4.2 Telemetry data statistics

The following part describes the correlation analysis and the linear regression analysis between gold standard passenger counts and those estimated from axle loads. The method is based on [1]. We use here Data 1 and Data 2 with the same time period of measurements. The telemetry dataset was cleaned, pre-processed and filtered just as described earlier.

The linear relation between the two datasets is shown in Fig. 2. The Pearson correlation coefficient between the two datasets is 0.98. The linear regression model provides a net weight estimation of the trolley bus as 20875.0 kg and an average passenger weight of 70.0 kg. These values have been included in eqn (1). Namely,

$$\begin{aligned} weight_{total} &= weight_{net} + weight_{averageperson} \cdot count_{passenger} + error \\ &= 20875.0 + 70.0 \cdot count_{passenger} + error. \end{aligned} \quad (1)$$

We will use this linear relation to estimate the number of passengers based on the total weight, and then estimate the absolute and relative mean errors.

The difference between the estimate and the ‘gold standard’ number is defined in eqn (2):

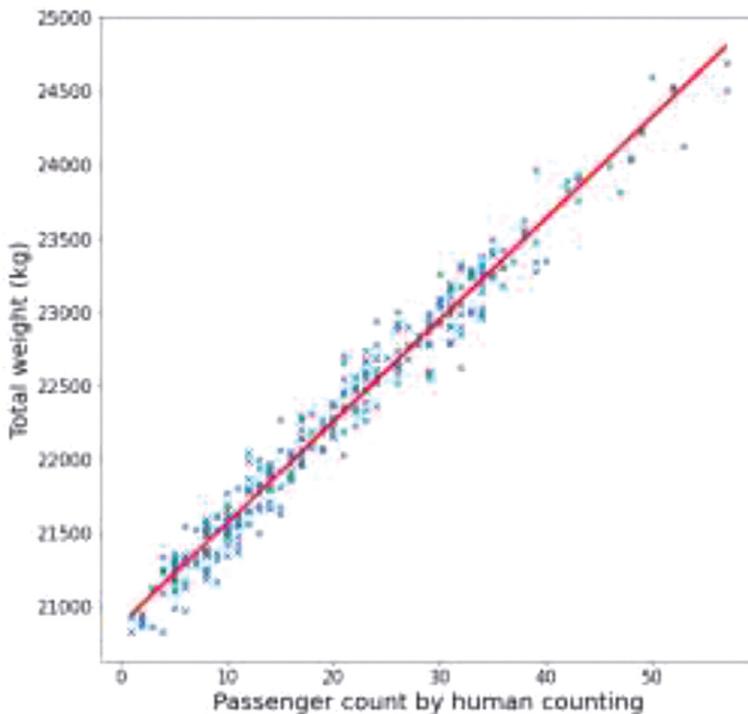


Figure 2: Linear relation between the gold standard passenger count and the total axle load.

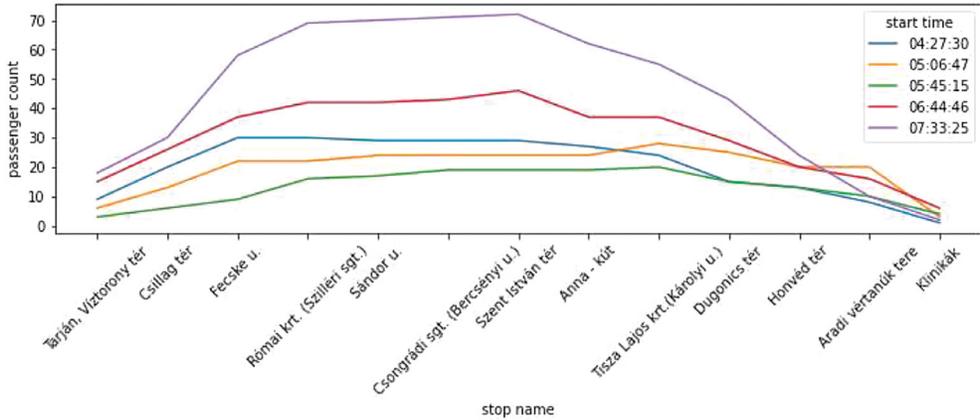


Figure 3: Passenger count computed for the same line and direction at different times. The horizontal axis shows the names of trolley bus stops.

$$diff = count_{passenger} - \max\left(0, \text{round}\left(\frac{weight_{total} - 20875}{70}\right)\right). \quad (2)$$

From this difference, various error measures can be calculated. The standard deviation of this difference variable is 2.17, and the mean of this variable was  $-0.01$ , so we can neglect the mean error bias. Thus, if we suppose that the difference samples are drawn from a normal distribution (the Shapiro–Wilk test  $P$ -value is 0.66 here), then a 95% confidence interval for the error is about , that is,  $0.01 \pm 2.17 \cdot 1.96, \pm 4$  passengers.

After the parameters of the linear relation are known, for any route of the trolley bus, the between-stops weights can be determined, the passenger counts can be estimated and the computed data can be visualized as like in the following chart showing the passenger counts on the same line and direction at different times (Fig. 3).

## 5 WI-FI DATA ANALYSIS

The tracking of Wi-Fi devices is one way of getting an estimate of an OD matrix. The number of passengers can also be estimated, even if there are no axle load data available. However, this kind of passenger count estimation does not appear to be so accurate, but it can still be used, if no further data are available.

In this section, we investigate the precision of measurements in our given environment. This description will also show how one can recalibrate the computation, if a key part of the system is modified. Here, we will calibrate the Wi-Fi-based passenger count estimation using the previously described axle load-based estimation as the gold standard, as we do not have manually counted passenger numbers.

After cleaning, we have MAC trips from the public transport vehicle and these MACs are not randomized ones. Using GPS coordinates and timestamp data of stops, we can easily estimate from and up till which stop the MAC address was logged. Then, we can calculate between two stops how many passengers were travelling. The number of passengers getting on or off can be computed in parallel. Moreover, a matrix can be filled, after the lines,

direction and stop names have been identified. This is the OD matrix for a line and direction and a time interval.

### 5.1 The sources of Wi-Fi-based estimation error

The main reasons for Wi-Fi-based estimation errors are:

- The time intervals between the consecutive frames may be big,
- It is more and more common for a MAC address to be randomized (for privacy),
- There is no known relation between the number of passengers and the number of active devices present.

### 5.2 Time durations between consecutive probe requests

Several studies [14] indicate that the probe requests are sent in various channels in a very short time, but between these ‘bursts’, a lot of time can pass. We compiled some statistics on the duration between these ‘bursts’ (here, we defined a burst such that the duration between the consecutive frames of probe requests is less than 1 second). Figure 4 shows an example of the time differences between consecutive timestamps of frames.

We computed some statistics of time differences between consecutive ‘bursts’. The average of time durations between ‘bursts’ is 82.8 sec. This value means that the estimates of getting on and off will be uncertain, and they can differ from the given place with one or more stops depending on the travelling time.

### 5.3 Randomized MAC addresses

From the MAC trips collected in a real scenario, one can determine how many MACs are assigned trips, but one cannot determine the ratio of devices which randomize their MAC address as these signals cannot be differentiated from the signals from the signals coming from the neighbourhood. In addition, the result of the linear regression in the sections later tells us that about 30% of the total passenger count can be detected by non-randomized Wi-Fi probe requests.

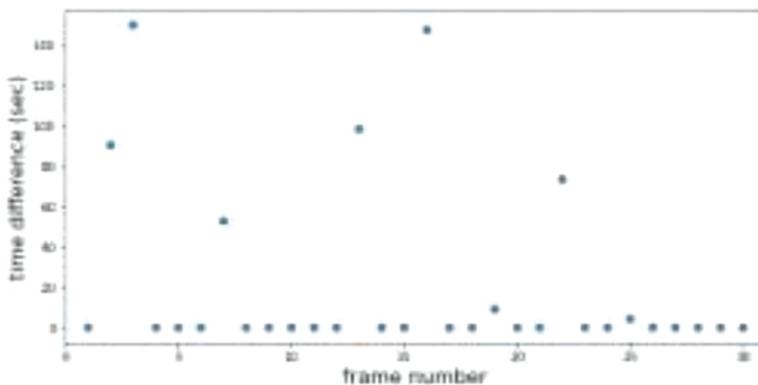


Figure 4: An example of the time durations between consecutive probe requests.

#### 5.4 Data used in our estimations

A set of 15 whole-day data from Data 3 is examined here. For each day, a linear regression model was applied between the Wi-Fi-based device count and the axle load-based passenger count estimation. The mean of the correlations was 0.8 with a standard deviation of 0.05. The mean slope of linear regression was 0.3 with a standard deviation of 0.03. The linear relation between the two data is shown in Fig. 5.

The aim here was to learn the behaviour of errors when estimations are averaged for many occurrences of the same route. For this, a route identification process is the first step. Stop identification information was assigned automatically (using GPS coordinates), with the name of that stop, the line number and direction. Then a line and direction ('Line 8' from 'Makkosház') was chosen for further analysis (over 15 days the trolley bus travelled on 8 different routes, and from these routes, the chosen one was the most frequent in the schedule of the trolley). It turned out that our trolley bus made 90 full trips on Line 8 in these 15 days. We analysed these 90 trips further and will present the results later.

#### 5.5 The relative error as a function of the passenger count

We investigated the relative error of the Wi-Fi-based count estimates after the linear regression was applied (see Fig. 6). This shows the decreasing relative error, if more passengers are on the vehicle, and the absolute error is not increasing linearly with the axle-based passenger count. The relative error computed here is defined in eqn (3). Here, we made bins of uniform intervals (number of passengers 5–9, 10–14 and so on) and did not take into account the first (0–4) interval for the relative error computation because it may be high and dividing by zero is also a problem.

$$E_R(bin) = \frac{\sum_{bin} \frac{|PC_{wifi} - PC_{axleload}|}{PC_{axleload}}}{|bin|}. \quad (3)$$

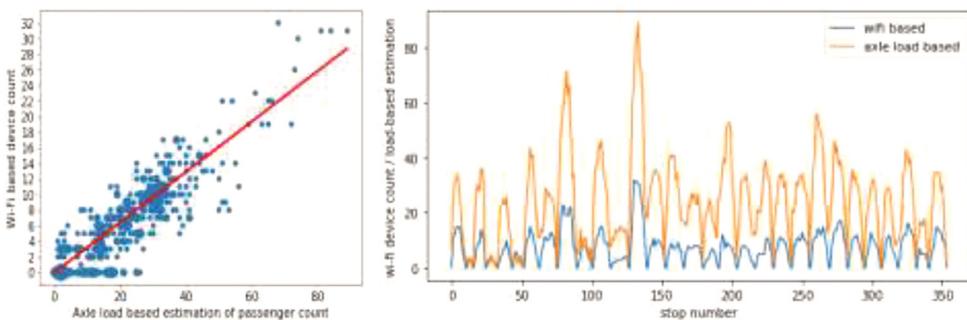


Figure 5: Left: The scatter plot of the axle load-based passenger count and Wi-Fi-based device count. Right: The parallel plot of a one-day Wi-Fi-based device count set and axle load-based passenger count set. On this particular day, the Pearson correlation coefficient was 0.86, the intercept was  $-0.48$  and the slope was 0.32.

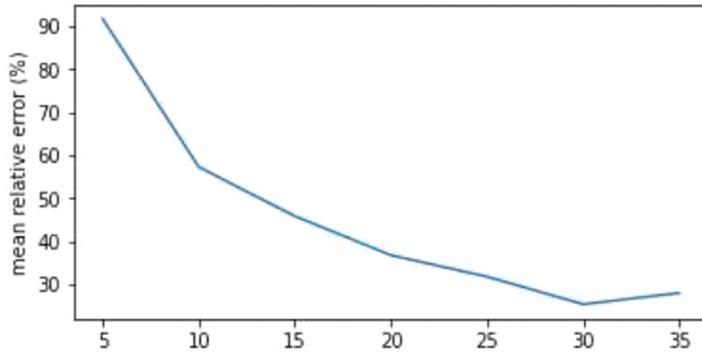


Figure 6: The relative percentage error of passenger count estimates as a function of the number of passengers.

### 5.6 A detailed analysis on between stops statistics

Error statistics were calculated on passenger count estimates between specific consecutive stops, and some surprising results were obtained. We examined for each stop-to-stop interval the deviation between the average of Wi-Fi-based passenger count estimates and the average of the axle load-based estimates. We used the 90 full trips defined earlier.

One might think that both of these averages should be around the average of the hypothetical true passenger counts travelling between two specific stops and the error would decrease if the size of the sample were larger. However, the results show something else: the difference between two estimation averages (the Wi-Fi-based and the axle load-based) is significant in our sample. To demonstrate this, first we show a scatter plot for two stop-to-stop intervals of a trolley bus line (Fig. 7).

For the t-test for significance, we have the null hypothesis that the mean of the difference between the Wi-Fi-based estimate and the axle load-based one is zero. For the two stop-to-stop intervals on the scatter, we find that the  $P$ -values are  $3e-4$  and  $4e-10$  for ‘Ortutay utca’ and ‘Dugonics t r’ begin stops. This means that these differences are significant.

Based on the previous observation, we propose a correction for between stop estimates. First, on a part (a half) of the full data, a linear regression is applied, then this regression can be applied on the rest of the data to reduce the estimation error.

To evaluate the proposed correction method, we used a modified cross-validation technique because the results using standard  $N$ -fold cross-validation technique might be statistically unstable since the number of samples is very low. For example, if we use 10-fold cross-validation, then each test set would contain just 9 examples. The steps of evaluation used here are the following:

- Repeat the following 10 times:
  - Shuffle the examples randomly, and halve it to train and test sets
  - On the train set, compute the parameters of the linear regression
  - On the test set, apply the linear regression and compute error measures before and after applying this regression
- After collecting the error measures in the loop, compute error statistics

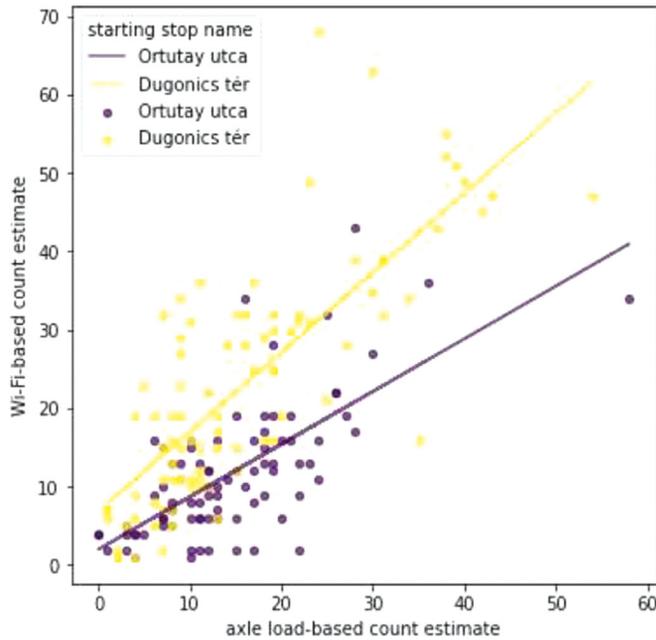


Figure 7: The scatter plot of estimates for specific stop-to-stop intervals. The linear regression lines have also been plotted.

The averaged relative error measures were collected for nine different stop-to-stop intervals and are listed in Table 2. These relative errors are computed on the average estimated passenger counts.

The results shown in Table 2 indicate that the relative error measures can be reduced if a linear correction is applied on the estimated value.

### 5.7 Comparison of results

In the literature, we did not find any article that contained relative error computations for the Wi-Fi-based estimates; hence, we cannot compare our results with those of others. Our proposed method cuts the relative error computed on-average passenger counts from an average of 22% to an average of 8.2%. We also have results on the correlations between axle load-based estimates and valid passenger counts, and Wi-Fi-based estimates and a gold standard passenger count (see Sections 4.2 and 5.4). In section 4.2 there is the correlation value 0.98 for the weight data and passenger count, in 5.4 a correlation value of 0.8 for the wifi and pass. count. Table 3 provides a comparison.

## 6 CONCLUSIONS AND FUTURE WORK

In this study, we addressed the problems of estimating passenger counts in two different ways. After constructing linear models and estimating absolute and relative errors, we found that the relative error of the Wi-Fi-based estimation can be markedly reduced if data sets are processed properly. When a daily Wi-Fi analysis is split into between-stops parts, a linear correction method can be applied to these parts, and as a result, the relative errors can be reduced.

Table 2: Average relative errors in different stop-to-stop intervals before and after correction.

Begin stop name	Diófa Vendéglő	Tisza Lajos krt.	Honvéd tér	Tündér utca	Ortutay utca	Dugonics tér	Rózsa utca	Anna - kút	Szent István tér
Relative error without correction	7.6%	16.1%	20.1%	36.6%	21.8%	16.2%	16.8%	42.3%	21.9%
Relative error after correction	7.7%	6.5%	9.5%	9.9%	7.6%	8.3%	11.4%	12.5%	5.3%

Table 3: Correlations between passenger count estimates.

Correlation, mass-based vs. human counted	0.97 [1]	0.99 [2]	0.98 (ours)
Correlation, Wi-Fi based vs. human counted or load based (*)	0.62 [10]	0.76 [11]	0.8 (ours*)

In the future, we plan to investigate several issues. Firstly, the reason for the shifts between the two kinds of means (measuring the same thing) when the data set is restricted to a stop-to-stop set should be ascertained. Secondly, we need to examine other ways of how Wi-Fi MAC addresses can be divided into two categories, that is, whether a device with a MAC was on a trolley bus (or a vehicle) or not.

#### ACKNOWLEDGEMENTS

The SASMob ‘Smart Alliance for Sustainable Mobility’ project (UIA02-239) is based on the initiative of Urban Innovative Actions, funded by the European Regional Development Fund.

#### REFERENCES

- [1] Nielsen, B. F., Fr olich, L., Nielsen, O. A., & Filges, D. (2014). Estimating passenger numbers in trains using existing weighing capabilities. *Transportmetrica A: Transport Science*, 10(6), 502-517.
- [2] Kotz, A. J., Kittelson, D. B., & Northrop, W. F. (2015). Novel vehicle mass-based automated passenger counter for transit applications. *Transportation Research Record*, 2563(1), 37-43.
- [3] Hsu, Y. W., Wang, T. Y., & Perng, J. W. (2020). Passenger flow counting in buses based on deep learning using surveillance video. *Optik*, 202, 163675.
- [4] Erlik Nowruzzi, F., El Ahmar, W. A., Laganieri, R., & Ghods, A. H. (2019). In-vehicle occupancy detection with convolutional networks on thermal images. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 0-0).
- [5] Balasubramanian, Y., Chandrasekaran, N., Asokan, S., & Subramanian, S. S. (2019). Deep-facial feature-based person reidentification for authentication in surveillance applications. *Visual Object Tracking in the Deep Neural Networks Era*. IntechOpen.
- [6] Li, T., Sun, D., Jing, P., & Yang, K. (2018). Smart card data mining of public transport destination: a literature review. *Information*, 9(1), 18.
- [7] Kurkcu, A., & Ozbay, K. (2017). Estimating pedestrian densities, wait times, and flows with Wi-Fi and Bluetooth sensors. *Transportation Research Record*, 2644(1), 72-82.
- [8] Chilipirea, C., Baratchi, M., Dobre, C., & Steen, M. V. (2018). Identifying stops and moves in WiFi tracking data. *Sensors*, 18(11), 4039.
- [9] Myrvoll, T. A., H akeg ard, J. E., Matsui, T., & Septier, F. (2017, October). Counting public transport passenger using WiFi signatures of mobile devices. *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)* (pp. 623-628). IEEE.
- [10] Oransirikul, T., Nishide, R., Piumarta, I., & Takada, H. (2014). Measuring bus passenger load by monitoring Wi-Fi transmissions from mobile devices. *Procedia Technology*, 18, 120-125.
- [11] Kang, L., Qi, B., & Banerjee, S. (2016, February). A wireless-based approach for transit analytics. *Proceedings of the 17th International Workshop on Mobile Computing Systems and Applications* (pp. 75-80).

- [12] Review of the TfL WiFi pilot, available online: <http://content.tfl.gov.uk/review-tfl-wifi-pilot.pdf>
- [13] Nguyen, K. A., Wang, Y., Li, G., Luo, Z., & Watkins, C. (2019). Realtime tracking of passengers on the London underground transport by matching smartphone accelerometer footprints. *Sensors*, 19(19), 4184.
- [14] Freudiger, J. (2015, June). How talkative is your mobile device? An experimental study of Wi-Fi probe requests. *Proceedings of the 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks* (pp. 1-6).
- [15] Martin, J., Mayberry, T., Donahue, C., Foppe, L., Brown, L., Riggins, C. & Brown, D. (2017). A study of MAC address randomization in mobile devices and when it fails. *Proceedings on Privacy Enhancing Technologies*, 2017(4), 365-383.
- [16] Castignani, G., Montavont, N., & Arcia-Moret, A. (2010). Analysis and evaluation of WiFi scanning strategies. *Proceeding of IV Cibelec*, 3-7.