Travelers-Tracing and Mobility Profiling Using Machine Learning in Railway Systems

Syed Muhammad Asad*†, Kia Dashtipour*, Sajjad Hussain*, Qammer Hussain Abbasi*, Muhammad Ali Imran*
* James Watt School of Engineering, University of Glasgow, Glasgow, G12 8QQ, UK
S.Asad.1@research.gla.ac.uk, {Kia.Dashtipour, Sajjad.Hussain, Qammer.Abbasi, Muhammad.Imran}@glasgow.ac.uk,
† Transport for London (TfL) Head-Quarter, 14 Pier Walk, Greenwich Peninsula, London SE10 0ES, UK

***Abstract*—With the advent of Coronavirus Disease 2019 (COVID-19) throughout the world, safe transportation becomes critical while maintaining reasonable social distancing that requires a strategy in the mobility of daily travelers. Crowded train carriages, stations, and platforms are highly susceptible to spreading the disease, especially when infected travelers intermix with healthy travelers. Travelers-profiling is one of the essential interventions that railway network professionals rely on managing the disease outbreak while providing safe commute to staff and the public. In this plethora, a Machine Learning (ML) driven intelligent approach is proposed to manage daily train travelers that are in the age-group 16-59 years and over 60 years (vulnerable age-group) with the recommendations of certain times and routes of traveling, designated train carriages, stations, platforms, and special services using the London Underground and Overground (LUO) Network. LUO dataset has been compared with various ML algorithms to classify different age-group travelers where Support Vector Machine (SVM) mobility prediction classification achieves up to 86.43% and 81.96% in age-group 16-59 years and over 60 years.**

***Index Terms*—Artificial Intelligence, COVID-19, Intelligent Transport Systems, Mobility Management, Travelers-Tracing.**

I. Introduction

The world-wide COVID-19 pandemic can be easily spread by nearby people especially in areas with high density where sufficient social distance is improbable between the mobile individuals (e.g., transport network, city centers, etc.). The increase in mobile devices present a new opportunity to combat the challenges encountered by traditional contact-tracing techniques in terms of monitoring, classifying, and advising different age-group travelers about the spread of COVID-19 in densely populated areas, such as overcrowded LUO network.

A globally accepted social distancing strategy to mitigate the spread of pandemic while avoiding crowded areas has gained crucial significance [1]. However, there is a need for devising mitigation strategies to further decelerate the disease spread by using Artificial Intelligence (AI) [2] that exhibits traits associated with the historical mobility of the different age-groups such as vulnerable age-group travelers. Our main strategy is to exploit mobility-aware travelers tracing sensors in existing railway systems including WiFi, Radio-frequency Identification (RFID), Bluetooth, Ultra-Wideband (UWB) [4] that collect different age-group travelers data anonymously and pseudonymized it for AI-driven decision analysis in real-time [3]. By these key enablers, we foresee improvements in advising travelers with designated safe routes, train carriages, stations and platforms.

In this study, we introduce a novel approach to manage daily train travelers that are in the range of ages 16-59 years and over 60 years (vulnerable age-group) by advising traveling in certain train carriages, stations and platforms using the LUO network. In addition, we outline travelers-tracing enabling technologies to manage their profiles and intelligent mobility. Furthermore, we structured the study with our travelers-tracing system model in Section II followed by a discussion on simulations results of our dataset, different age-groups ML classification, and key technology enablers in railway systems included in Section III. Discussion on results and recommendations for travelers is mentioned in Section IV. Finally, Section V concludes the paper.

II. System Model and Problem Formulation

We have assumed either a symptomatic or asymptomatic individuals are locomotive throughout the LUO network. When the age-groups, 16-59 years or over 60 years comes in contact with the LUO network, the probability of being exposed is β_u and β_o respectively. Generally, the rate of propagation is affected by contact intensity which is determined by the contact frequency and duration. Hence, it can be state that the rate of propagation is directly proportional to contact intensity. Human tracking in the railway systems is dependant upon travelers-tracing enablers such as; WiFi, RFID, Bluetooth, UWB where wireless cellular network signaling, Global Positioning System (GPS) are either not available or limited. The collected data is pseudonymized and aggregated by the railway systems portal for data analysis and decision-making. The main focus of our study is relies on the travelers-tracing dependant decision-making process to exploit different age-groups mobility management. The average contact intensity between the two age-group travelers i and j in the discrete time span τ can be denoted as [6]:

$$\delta_{i,j}(t) = \frac{\sum_{x=1}^{\alpha^c_{i,j}(t)} \tau(x)}{\alpha^c_{i,j}(t)} \tag{1}$$

where, $\alpha^c_{i,j}(t)$ is the total number of contacts before the time slot t, and $\tau(x)$ is the corresponding contact duration. Given the dataset with different age-groups, 16-59 years and over 60 years, the vertex ω representing daily travelers by using Eq. (1) and probabilities β_u and β_o would be:

$$\omega_{i,j}(t)(i|j \subseteq D) = \delta_{i,j}(t)\left[\sum_{i=1}^{\alpha_i^u(t)} \beta_u + \sum_{j=1}^{\alpha_j^o(t)} \beta_o\right] \quad (2)$$

where, $\alpha_i^u(t)$ and $\alpha_j^o(t)$ represent the 16-59 years and over 60 age-groups' neighbors. Let $\mathbb{U}(t)$, $\mathbb{O}(t)$, and $\mathbb{D}(t)$ denote travelers in age-group under 60 years, over 60 years, and available resources at LUO network to accommodate travelers accordingly. The optimization to advise certain safe routes to different age-groups is expressed in the following equation:

$$\sum_D \int_1^T \frac{\mathbb{U}(t)}{\mathbb{O}(t)}\, dt \quad (3)$$

where, $\{\mathbb{U}(t), \mathbb{O}(t)\} \leqslant \mathbb{D}(t)$ in the given time t.

III. Intelligence Enabled Railway Systems

We further exploit an existing LUO railway network functionalities intended to manage travelers of different age-groups, i.e, 16-59 years and over 60 years mobility and to ensure designated routing while using travelers-tracing method [6] against COVID-19 spread. We explain the key enablers that will play an instrumental role of travelers-tracing in existing railway systems as shown in Fig. 1. In the following subsections, we first present the simulation results for daily train travelers dataset to analyse the existing mobility of all age-groups as shown in Fig. 2 and finally we discuss and present results obtained from ML-driven classifiers as shown in Tables I and II. Since almost everyone carries cellular mobile devices which do not serve as always-on human trackers, therefore, the multiple technologies are exploited for travelers-tracking in railway systems. More specifically, the higher the number and mobility of user equipment (UEs) recorded by travelers-tracing enablers, the higher the number and mobility of people lying in different age-groups to be served and advised accordingly.

A. ML-driven Mobility Prediction Classification

We propose ML-driven mobility prediction classification for the interpretation of performance measures supported by accuracy, precision, recall, and F1 score results. These performance measures are used to predict daily train travelers according to the travelers age-groups, (i) 16-59 years, and (ii) over 60 years in LUO environment. ML is used in the assortment of pattern recognition and mobility traces to establish automation in intelligent decisions while learning from history and adapt to the testing environment [3], [7]. We have modelled our dataset to classify different age-group travelers in order to take necessary actions such as; monitoring potential contacts/proximity travelers, advising the travelers mobility to certain safe mobility pathways/routes to safeguard the vulnerable age-group travelers from the outspread disease. By using MATLAB libraries, we have modelled six classifiers to obtain accuracy, precision, recall, and F1 score results as shown in Tables I and II. As it can be seen that Support Vector Machine (SVM) classifier with default RBF kernel parameters settings and 200 kernel size outperformed in both age-groups compared to other classifiers.

TABLE I
Age-group 16-59 years Mobility Prediction Classification

Machine Learning Classifier	**Accuracy**	**Precision**	**Recall**	**F-Measure**
Logistic Regression (LR)	79.1	0.78	0.77	0.77
Multi-layer Perceptron (MLP)	80.91	0.80	0.79	0.79
Support Vector Machine (SVM)	86.43	0.86	0.84	0.84
Random Forest (RT)	83.36	0.83	0.81	0.81
K-Nearest Neighbour (KNN)	84.91	0.84	0.82	0.82
Decision Tree (DT)	85.63	0.85	0.83	0.83

TABLE II
Age-group Over 60 years Mobility Prediction Classification

Machine Learning Classifier	**Accuracy**	**Precision**	**Recall**	**F-Measure**
Logistic Regression (LR)	75.21	0.75	0.73	0.73
Multi-layer Perceptron (MLP)	76.59	0.76	0.74	0.74
Support Vector Machine (SVM)	81.96	0.81	0.79	0.79
Random Forest (RT)	77.86	0.77	0.75	0.75
K-Nearest Neighbour (KNN)	79.24	0.79	0.77	0.77
Decision Tree (DT)	79.6	0.79	0.79	0.79

B. Travelers-Tracing & Profiling Enablers

According to a recent study [8], severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) remains viable in aerosols for up to three hours exhaled by unhealthy people while speaking, coughing or even breathing, whether symptomatic or not [9]. We are particularly concerned with the scenario where vulnerable age-group (over 60s) travelers using the LUO environment would be protected against the disease spread where they will be recommended of certain times, safe train carriages, stations, and platforms. All areas of the LUO where human mobility is possible are considered as 'high-risk' as daily commuters use the network without knowing the contagious people traveling with them. The main objective is to detect high-risk age-group travelers by using travelers-tracing enablers, allowing prioritization for further monitoring and risk management.

1) WiFi: There is a wide availability of Wi-Fi currently being used within railway systems which in particular have been deployed in more than 97% of all LUO stations to facilitate users[1]. The ubiquity of WiFi access through the deployment of WiFi access points within the LUO network can be exploited to obtain traveler's mobility safe routes. The

[1]http://content.tfl.gov.uk/review-tfl-wifi-pilot.pdf

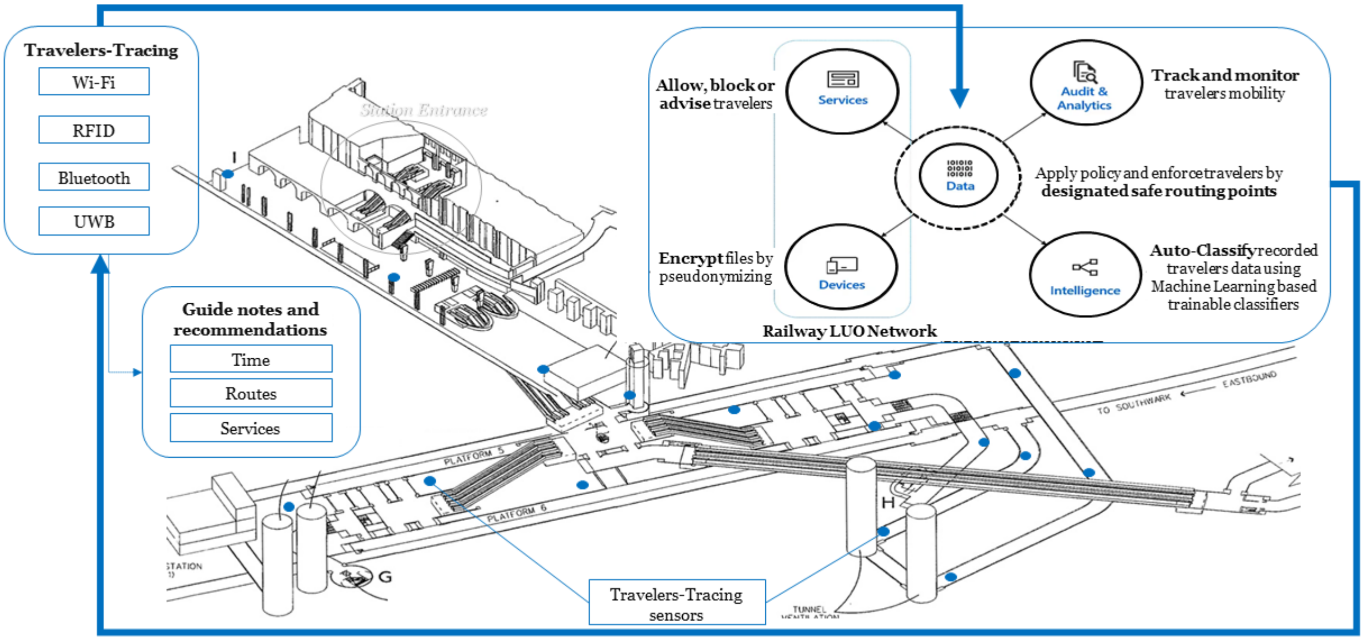


Fig. 1. Mobility-Aware travelers-tracing and profiling strategy using Real-Time Railway Tracker Network (RRTN).

uniqueness of mobility traces provides a traveler's movement from one station to another accurately when Wifi is switched on. The information of recently visited stations, platforms, and train carriages [10] is pseudonymized to prevent the identification of the travelers. This method plays an important role in determining designated safe routing to combat the spread of epidemics including COVID-19 [1].

2) Radio-frequency Identification (RFID): RFID is one of the major and widely-used technologies that are in operation in railway systems, covering almost all parts of train carriages, railway stations, and platforms. RFID is an electromagnetic waves-based identification method that has numerous applications including public transport, trains movement monitoring from depots to main lines and vice versa, building access, patient monitoring, inventory, assembly lines & supply chains, food tagging, security identification, localization, etc. [1]. UHF RFID uses passive tags connected to smartphones and objects in the LUO environment where tracking of different age-group travelers is possible using tap-in and tap-out touch system in Real-Time Railway Tracker Network (RRTN). Antenna in the RFID tag with beam steering capability detects the signal level (RSSI), bearings, and logarithmic points along an axis estimated by applying ML techniques.

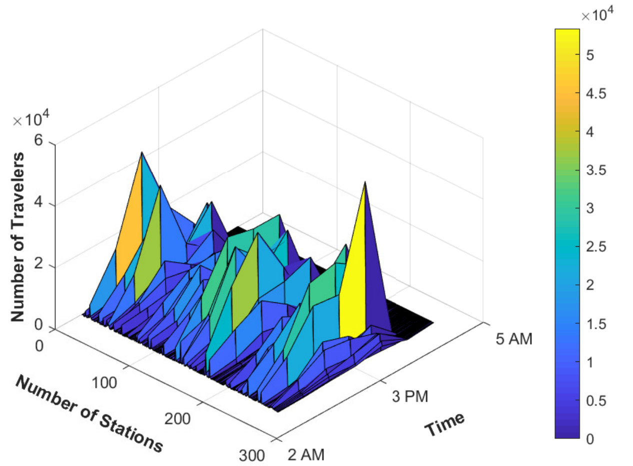
3) Bluetooth: Bluetooth is a widely-used wireless technology for data transmission between fixed and mobile devices in short distances. The Low-power Bluetooth Communication (LBC) with shorter wavelengths, UHF radio waves collect surrounding information. By regular scanning, Access Point Indicators (APIs) to locate nearby Bluetooth devices in real-time [12]. Phones with Bluetooth stores and pairs the list of Bluetooth devices. With the Bluetooth pairing, different age-groups can be determined and stored in the central railway database to advise certain actions to certain age-groups. In case vulnerable age-group (over 60s) traveler enters into a station, the list of Bluetooth devices it has encountered can be fetched, and a notification would be generated to the travelers to follow certain safe traveling routes, train carriages, station or platforms.

4) Ultra-Wideband (UWB): UWB is one of a radio technology that has short-range and high-bandwidth communication on a low energy level by using large antenna arrays and ultra-wide bandwidths, a decimetre level accuracy (minimum accuracy) in location systems becomes viable. Minimum accuracy with the precise range measurement to estimate the distance between target and reference base station can be achieved by using UWB technology, unlike Bluetooth or Wi-Fi where RF signal's time difference of arrival (TDOA) or Time of Flight provides more precision[2].

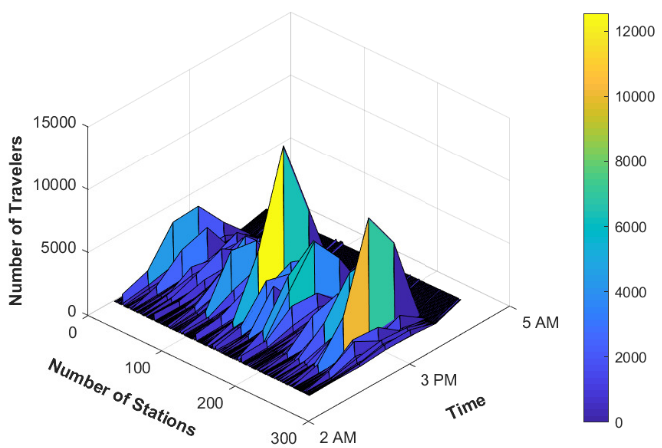
IV. Results & Recommendations

The strategy and overall life-cycle of travelers-tracing and profiling is to collect the information through travelers-tracing enablers provided on each station at the LUO network. Recorded data is sent to the railway database by using RRTN for daily monitoring. We have presented simulation results for daily train travelers in order to obtain ML-driven classification accuracy, precision, recall and F-measure scores by classifying mobility prediction of different age-groups. It can be seen that the SVM classifier outperformed than all other five classifiers with an overall classification accuracy of 86.43% and 81.96% in age-group 16-59 years and over 60 years. Below are the few recommendations that certain age-group travelers would

[2]Ultra-Wideband (UWB) Ofcom document can be found at: https://www.ofcom.org.uk/__data/assets/pdf_file/0015/25152/uwb.pdf



(a) Crowd density of age-group 16-59 years old.



(b) Crowd density of age-group over 60 years old.

Fig. 2. Simulation results for daily train travelers considering three dimensional (3D) approach for 1 week in the year 2017-18. Plotted age-groups are from 05:00 AM to 02:00 AM (21-hours) on multiple stations.

be designated in the context of RRTN realisation through safe mobility routing against COVID-19.

- Time - to ensure the time-specific mobility for age-specific travelers would be assigned for safeguarding vulnerable age-group from others. This would mean to allot off-peak day travel time-slots between 10 am to 12 noon and 2pm to 4pm with the guidelines available at stations and platforms. Staff can be asked to assist in this regard, alternatively guideline notifications can be sent through travelers-tracing enablers on user smartphones.
- Route - to provide safe routes at the stations and platforms by designating specific routes which can be color coded to prominence for travelers according to their age-groups.
- Train carriage - for optimum safety, one train would be divided into blocks of virtual train carriages with different age-group markings assist the vulnerable travelers.
- Station - to facilitate train travelers to board, alight or freight in specific stations of the train network for vulnerable age-group safety in the times of disease spread. Certain stations equipped with extra assistance and ancillary services as ticket sales, waiting rooms and baggage/freight to be used for different age-group travel in different times of the day.
- Platform - at least one track-side platform would be assigned to vulnerable age-group travelers with extra assistance and services such as luggage carts, user friendly electronic boards, emergency call facility, etc.
- Special services - to further ensure the safety of the travelers and train drivers, new limits to the number of travelers on board at any one time shall be introduced. In case of congestion at the stations, alternate bus services would be arranged to facilitate/help vulnerable age-group travelers mobility.

V. Conclusion

We have introduced a new strategy to address crowded train carriages, stations, and platforms that are highly susceptible to spreading the disease supported by mobility prediction accuracies using ML classifiers. Our strategy also exploits the technologies including WiFi, RFID, Bluetooth, and UWB to effectively track daily train travelers with different age-groups in the LUO network. Therefore, the different age-group traveler's mobility can efficiently be monitored in order to issue recommendations of designated safe track routes, certain train carriages, stations, and platforms.

References

[1] D. M. Dobkin. “The RF in RFID: Passive UHF RFID in Practice”. Newnes, Elsevier, 2013.

[2] S. Asad. et al., “Reinforcement Learning driven Energy Efficient Mobile Communication and Applications,” 2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Ajman, United Arab Emirates, 2019, pp. 1-7,

[3] S. Asad, J. Ahmad, S. Hussain, A. Zoha, Q. H. Abbasi, M. A. Imran, “Mobility Prediction-Based Optimisation and Encryption of Passenger Traffic-Flows Using Machine Learning”. Sensors 2020, 20, 2629.

[4] L. Zhang, G. Zhao, and MA Imran (2020). “Internet of Things and Sensors Networks in 5G Wireless Communications” (eds M.A. Imran).

[5] V. Gemmetto. et al., “Mitigation of infectious disease at school: Targeted class closure vs school closure,” BMC Infect. Dis., vol. 14, Aug. 2014.

[6] B. Wang, Y. Sun, T. Q. Duong, L. D. Nguyen and L. Hanzo, ”Risk-Aware Identification of Highly Suspected COVID-19 Cases in Social IoT: A Joint Graph Theory and Reinforcement Learning Approach,” in IEEE Access.

[7] Shah, S.A., Fan, D., Ren, A. et al. Seizure episodes detection via smart medical sensing system. J Ambient Intell Human Comput (2018). https://doi.org/10.1007/s12652-018-1142-3

[8] N. van Doremalen, T. Bushmaker, D. H. Morris, et al., “Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1,” N. Engl. J. Med., vol. 382, no. 16, pp. 1564–1567, Apr. 2020. http://www.nejm.org/doi/10.1056/NEJMc2004973

[9] Y. Bai, L. Yao, T. Wei, F. Tian, et al., “Presumed asymptomatic carrier transmission of COVID-19,” JAMA, vol. 323, no. 14, pp. 1406–1407, Apr. 2020. http://doi.org/10.1001/jama.2020.2565

[10] de Montjoye YA, Quoidbach J, Robic F, Pentland AS. “Predicting personality using novel mobile phone-based metrics In: Social Computing, Behavioral-Cultural Modeling and Prediction”. Springer; 2013. p. 48–55.

[11] W. M. Griggs, R. Verago, J. Naoum-Sawaya,et al. Localizing missing entities using parked vehicles: An RFID-based system. IEEE Internet of Things Journal, 2018, 5(5):4018–4030

[12] T. Altuwaiyan, M. Hadian, and X. Liang, “Epic: Efficient privacypreserving contact tracing for infection detection,” in 2018 IEEE International Conference on Communications (ICC). IEEE, 2018, pp.1–6.