Travel Mode Identification with GPS Trajectories using Wavelet Transform and Deep Learning

James J.Q. Yu, *Member, IEEE*

***Abstract*—Accurate identification in public travel modes is an essential task in intelligent transportation systems. In recent years, GPS-based identification is gradually replacing the conventional survey-based information-gathering process due to the more detailed and precise data on individual's travel patterns. Nonetheless, existing research suffers from deficient feature selection, high data dimensionality, and data underutilization issues. In this work, we propose a novel travel mode identification mechanism based on discrete wavelet transform and recent developments of deep learning techniques. The proposed mechanism aims to take GPS trajectories of arbitrary lengths to develop accurate travel mode results in both global and online identification scenarios. In this mechanism, raw GPS data is first pre-processed to compute preliminary motion and displacement attributes, which are input into a tailor-made deep neural network. Discrete wavelet transform is also adopted to further extract time-frequency domain characteristics of the trajectories to assist the neural network in the classification task. To evaluate the performance of the proposed mechanism, a series of comprehensive case studies are conducted. The results indicate that the mechanism can notably outperform existing travel mode identifications on a same data set with minuscule computation time. Furthermore, an architecture test is performed to determine the best-performing structure for the proposed mechanism. Lastly, we demonstrate the capability of the mechanism in handling online identifications, and the performance sensitivity of the selected attributes is evaluated.**

***Index Terms*—Travel mode identification, GPS trajectory, discrete wavelet transform, deep learning, feature selection.**

I. Introduction

TRAVEL mode identification is among the fundamental constituting components of intelligent transportation system (ITS) in future smart cities [1], [2]. Accurate public travel mode data is essential for governments, companies, and research institutes to better understand human behaviors and operate transportation systems [3]–[5]. Much research effort has been devoted in developing advanced transportation system control strategies based on the data, e.g., traffic signal control [6], regional transportation planning [7], and policy making [8], etc. Furthermore, travel mode identifications for individuals can also facilitate tailor-made user experience of mobile applications. For instance, travel mode information plays an important role in constructing activity-based customer models [3], which in turn can lead to personalized recommendations and advertisements.

This work is supported by the General Program of Guangdong Basic and Applied Basic Research Foundation under grant number 2019A1515011032.
The author is with the Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China (email: yujq3@sustech.edu.cn).

The knowledge of travel mode choices was conventionally obtained through in-person or online surveys. Nonetheless, such methods are both time-consuming and ineffective due to the low response rate and incomplete/inaccurate information provided [9]. In the past few decades, global positioning system (GPS) was widely adopted in civilian equipments and smart devices. Traditional travel mode surveys were gradually replaced by GPS data mining thanks to the more detailed information on individual's travel patterns over a prolonged period of time [10]. In addition, the enormous penetration of GPS-enabled devices in modern cities – most commonly smart phones – makes it possible for system operators to capture massive GPS trajectory data with minimal human-in-the-loop data gathering errors than surveys [9].

Typically, GPS devices record the positional characteristics of travels at a given interval, and multiple consecutive GPS records by a same device can be connected to create a GPS trajectory. In the meantime, GPS devices do not have any explicit knowledge on which travel mode is currently employed. This advocates research on data processing mechanisms which extract the hidden travel mode information from raw GPS trajectories. Common travel mode identification approaches adopt a two-step paradigm as in [11]. Firstly, each GPS trajectory is input into a data processing module, which subsequently outputs manually selected data attributes that are considered to be able to summarize its characteristics, e.g., mean and 95-percentile speed/acceleration values [3], etc. Then the extracted attributes are fed into a classifier, which is typically learning-based, to develop the inferred travel mode. In this process, both the selection of data attributes and implementations of the classifier are critical to the overall identification accuracy, and the research community is embracing a plethora of new attributes and classifiers, see [3], [9], [11]–[13] for some examples.

However, there remains a research gap in the current GPS-based travel mode identification approaches. On the one hand, careful investigation on how the GPS data attributes influence the identification accuracy is essential to achieve satisfactory system performance as introduced above. Nonetheless, the manually chosen attributes may not fully represent all critical characteristics of GPS trajectories in distinguishing travel modes [11]. A recent solution to this issue is to increase the number of data attributes extracted from one GPS trajectory aiming at a thorough summary of the trajectory features, see [3] for an example. Yet this approach adversely increases the data dimensionality, resulting in a more difficult classification task in the second step [9]. On the other hand, other research exploits deep learning approaches to automatically learn multiple levels of data representations from the raw trajectories without pre-defined attribute inputs. Such mechanisms are designed to only accept trajectories with a fixed and significant length, see [9] for an example. In this case, real-time travel mode identification is not possible without a notable waiting time. Furthermore, the identification accuracy may be undermined on trajectories that are longer than the designed input length, since early but (potentially) useful GPS records are discarded.

To close the research gap in existing travel mode identification approaches, we propose a novel intelligent travel mode identification mechanism based on discrete wavelet transform (DWT) and deep learning techniques. In the proposed mechanism, a long short-term memory (LSTM)-based deep neural network (DNN) is constructed to perform the data attribute extraction and travel mode classification tasks simultaneously. In addition, DWT is employed to provide auxiliary data features for DNN to better distinguish different modes thanks to its outstanding frequency-domain feature extraction capability. The main contributions of this paper are summarized as follows:

- We propose a new intelligent travel mode identification mechanism that can provide accurate and timely travel information. This paper is among the pioneer studies of using wavelet transform and recurrent neural network in travel mode identification.
- The proposed mechanism aims to take trajectories with arbitrary and sufficient lengths[1] to develop accurate results. It can be applied to global and online identifications.
- We conduct comprehensive case studies to assess the performance of the proposed mechanism. The results demonstrate satisfactory identification accuracy compared with the literature.
- Investigations are carried out to study the best structure of the proposed mechanism, and the sensitivity of data attributes is examined.

The remainder of this paper is organized as follows. In Section II, the background of travel mode identification research is presented. Section III elaborates on the proposed identification mechanism. We perform a series of case studies in Section IV to demonstrate the efficacy of the proposed mechanism. Finally, this paper is concluded in Section V with a summary of potential future research topics.

II. Background

Travel mode identification has attracted much research effort in recent years. The growing body of related literature has proposed a large number of mechanisms to detect commuters' transport modes online or offline based on various data sources, e.g., raw GPS trajectories, geographic information system, mobile phone sensors, and global system for mobile communications (GSM). Additionally, a wide range of methodologies have been applied to address the identification task, including but not limited to rule-based algorithms, fuzzy logic, decision tree and its variants, Bayesian belief network, support vector machine, and neural networks, etc. As this research aims to identify travel mode with only GPS trajectories using wavelet transform and deep learning approaches, we focus on introducing the research background of previous studies utilizing GPS data. Interested reader can refer to [14], [15] for complete surveys on related research.

Reference [11] is among the pioneer work that utilizes GPS trajectories for travel mode identification. In the proposed approach, long trajectories are first divided into multiple trip segments with distinct modes using a change-point-clustering segmentation scheme. Statistical features of the segments, e.g., mean and variance of speed, are used as inputs of traditional classification techniques, which develop the travel modes. This work laid the foundation stone of the popular two-step identification paradigm. In order to obtain better performance, researchers carefully tweaked the statistical features in follow up research, referring to [3], [16], [17] for some examples. To illustrate the effectiveness of the features, reference [18] presented a series of empirical studies and concluded that frequency-domain features of trajectories are critical in travel mode identification using GPS data.

On the other hand, using advanced learning techniques to automatically extract data features has received attention in recent years. Among the techniques, machine learning is arguably the most widely adopted due to its excellent distinctive feature extraction capability given sufficient training data [19]. Reference [3] proposed a random forest classifier combined with a rule-based classification method for travel mode identification. This work emphasizes the importance of socioeconomic attribute data on the accurate classification. However, the inclusion of socioeconomic data hinders its practical usage since such data may not be always available, and they did not lead to outstanding performance compared to other approaches. Reference [20] developed a deep neural network-based strategy to automatically extract supplementary data features from entity movement trajectories. The feature extraction process is enhanced by a tailor-made spatio-temporal information-preserving data transformation mechanism. Nonetheless, the detection accuracy is inferior: on the widely-adopted GeoLife [11] dataset, the accuracy barely reaches 67.9%. Reference [9] created a new convolutional neural network-based travel mode identification mechanism. By a series of data preprocessing schemes, the proposed mechanism is capable of interpreting the raw trajectory data with fundamental motion characteristics, which are employed in the neural net for classification. However, as convolution neural network (CNN) is adopted as the backbone of the proposed algorithm, the length of trajectories must be fixed. Additionally, the length employed in [9], i.e., 200, is too long to be considered universal as a significant portion of trajectories in real life and in GeoLife has only dozens of GPS records [11]. This limitation is also observed in the approach proposed by [21] which devised a hybrid deep learning model empowered by convolutional bi-directional LSTM for transportation mode identification, and [22] which presented a control gate-based recurrent neural network-based strategy for travel mode detection, which made use of additional mobile phone sensor information for better accuracy. Reference [23] trained a bi-directional LSTM classifier with

[1]As will be illustrated in Section IV-C, the minimal length requirement is reasonable even in real-time identification scenarios.

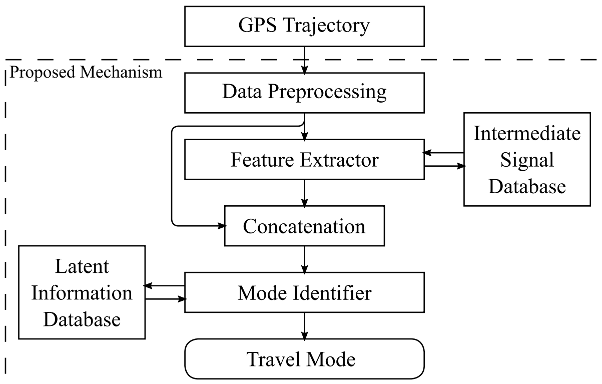


Fig. 1. Data flow of the proposed travel model identification mechanism.

both the raw trajectory data and the heterogeneous data sampling interval as an extra feature for better identification accuracy. Nonetheless, while trajectories are embedded into multiple groups of different sampling intervals, how they interact with the proposed model is not stated, rendering a restricted solution to trajectory identification. Most machine learning-based identification approaches do not notably excel those with manual feature selection in terms of identification accuracy [9]. To handle these drawbacks, we propose a new mechanism for travel mode identification. By utilizing LSTM as the backbone, the proposed mechanism can accept variable-length GPS trajectories, and the input features are carefully engineered so that no additional socioeconomic data are required. Additionally, DWT enables the proposed mechanism to extract frequency-domain data features, which notably boost the detection accuracy as will be shown in Section IV-B.

III. Travel Mode Identification Mechanism

As analyzed in Section II, existing research mostly relies on manually selected temporal characteristics in GPS trajectories for travel mode identification. Other work that tries to make use of intelligent systems to extract partial temporal characteristics cannot fully utilize the frequency-domain characteristics from the data. To fully exploit the hidden information exists in GPS trajectory data, we propose a DWT and LSTM-based travel mode identification mechanism in this section. We first elaborate the system structure and data flow of the proposed mechanism. Then we present the detailed formulation of the mechanism with brief introductions to employed techniques.

A. Structure of the Proposed Mechanism

The structure of the proposed mechanism is depicted in Fig. 1. This mechanism takes GPS trajectories, which are series of GPS records, to infer the travel mode. Since the raw GPS data only explicitly presents minimal travel information, it is firstly pre-processed to expose more relevant trajectory properties to subsequent calculations. Then the processed data is input into a feature extractor to obtain the frequency-domain features of the trajectory, and the intermediate results are cached in an intermediate signal database. Subsequently, the extracted features are input into a mode identifier together with the previously processed GPS data. Finally, the output of the identifier is considered an inference of the travel mode represented by the input GPS trajectory. This completes the whole identification task.

The proposed mechanism is designed to be capable of both online and global travel model identification. For online identification, the mechanism waits for at least ω GPS records in a trajectory before performing calculations. After $\omega - 1$ records, whenever a new one is available, the whole process is repeated to develop an updated result. To eliminate repetitive computations among consecutive identifications of a same trajectory, a latent information database is introduced to cache the intermediate values produced by the mode identifier. The information can later be retrieved if the mechanism has to infer the same trajectory with more GPS records. This scheme will be further elaborated in Section III-E.

In the proposed design, it is evident that all the three main blocks (data pre-processing, feature extractor, and mode identifier) play critical roles in accurately identifying travel modes. These blocks cooperate to derive the distinguishing temporal-correlated characteristics of the GPS trajectories, which are successively adopted to make classifications. In this paper, we follow the previous literature on travel mode identification to design the GPS data pre-processing algorithm. We employ DWT to extract the frequency-domain features thanks to its outstanding feature-summarizing capabilities demonstrated in other research, e.g., [24], [25]. Finally, we construct a DNN to further exploit the data features from trajectories automatically, and perform the mode identification task.

B. GPS Data Pre-processing

The motivation of the pre-processing step is intuitive as follows. Raw GPS data is typically collected over a period of time, and each record is presented as a 3-tuple comprised of the latitude, longitude, and absolute time of the GPS sampling device. However, it cannot be guaranteed that consecutive GPS records in a trajectory are sampled at a constant rate, rendering the isolated positions less useful in identification, see [11] for detailed explanations. A commonly adopted solution is to combine the position and time information to develop speed-related characteristics, which can be handled by other techniques more easily [9]. Additionally, pre-processing the raw data can incorporate human knowledge in travel mode identification tasks. This can alleviate difficulties in training the intelligent system in subsequent data processing blocks of the proposed mechanism. Data pre-processing is a widely-adopted pattern in other travel mode identification algorithms presented in the literature, e.g., [9], [11], [15].

We consider a series of GPS records $\{R_1, R_2, \cdots, R_\omega, \cdots R_n\}$ of length n, each of which is defined by $R_i = \langle \mathrm{lat}_i, \mathrm{lng}_i, t_i \rangle$ where $\mathrm{lat}_i, \mathrm{lng}_i, t_i$ are the latitude and longitude of the sampling device location, and the sampling time, respectively. When calculating the distance between two GPS locations R_i and R_j, we use the Vincenty's formulae [26] for its accuracy, denoted by $\mathrm{Vincenty}(\mathrm{lat}_i, \mathrm{lng}_i, \mathrm{lat}_j, \mathrm{lng}_j)$.

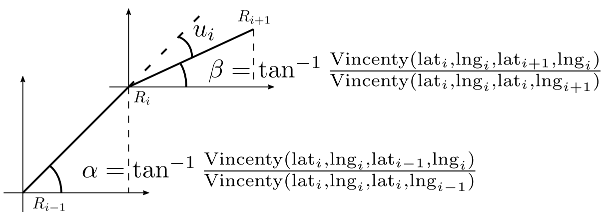


Fig. 2. An illustration of the turn u_i at GPS location R_i.

For two consecutive records R_i and R_{i+1}, the distance is calculated by

$$d_i = \text{Vincenty}(\text{lat}_i, \text{lng}_i, \text{lat}_{i+1}, \text{lng}_{i+1}) \tag{1}$$

Based on the distances in a trajectory, we can first construct three motion-related attributes as follows:

$$s_i = d_i/(t_{i+1} - t_i), \forall i \in \mathbb{N}^+, i < n;\ \ s_n = s_{n-1}; \tag{2a}$$
$$a_i = (s_{i+1} - s_i)/(t_{i+1} - t_i), \forall i \in \mathbb{N}^+, i < n;\ a_n = 0; \tag{2b}$$
$$k_i = (a_{i+1} - a_i)/(t_{i+1} - t_i), \forall i \in \mathbb{N}^+, i < n;\ k_n = 0; \tag{2c}$$

where s_i, a_i, k_i denote the speed, acceleration, and jerk at location R_i, respectively. Besides the commonly adopted speed and acceleration, we follow [9] and incorporate jerk in the attribute set, which is considered a significant characteristic during safety issues related to transportation [27]. In (2), both attributes are padded to make their lengths equal to n.

When interpreting raw GPS data into motion-related attributes, the displacement-related information is discarded, which can also be important in identifying travel modes. Intuitively, it is much easier for pedestrians to take sharp turns within few seconds than bicycles and buses. Therefore, we also calculate a new displacement attribute – "turn" – to account for the missing information as follows:

$$\begin{aligned} u_i = &\tan^{-1} \frac{\text{Vincenty}(\text{lat}_i, \text{lng}_i, \text{lat}_{i-1}, \text{lng}_i)}{\text{Vincenty}(\text{lat}_i, \text{lng}_i, \text{lat}_i, \text{lng}_{i-1})} \\ &- \tan^{-1} \frac{\text{Vincenty}(\text{lat}_i, \text{lng}_i, \text{lat}_{i+1}, \text{lng}_i)}{\text{Vincenty}(\text{lat}_i, \text{lng}_i, \text{lat}_i, \text{lng}_{i+1})}, \\ &\qquad\qquad \forall i \in \mathbb{N}^+, 1 < i < n; \\ u_n = &\ u_0 = 0, \end{aligned} \tag{3}$$

where u_i is the turn of trajectory at location R_i. The calculation of this equation is illustrated in Fig. 2, in which u_i corresponds to the angle difference between α and β.

Using (2) and (3), the raw GPS data can be pre-processed into an aligned time sequence vector of length n with four motion and displacement attributes. This data serves as the input of both the feature extractor and the mode identifier, which are introduced in the following sub-sections.

C. Trajectory Feature Extractor

During the data pre-processing, raw GPS data is transformed into aligned motion and displacement attribute vectors. While it is possible to solely rely on DNN to extract the hidden features among them, a better approach is to use established knowledge in signal processing when analyzing the data characteristics, which can later improve the identification performance of DNN. In this work, we adopt DWT to extract the hidden time-frequency domain features in the processed data. This technique convolves the input signal with discrete wavelets $\psi_{a,b}(t)$, which can be developed from pre-defined mother wavelets $\psi(t)$ at level a and location b as follows:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{2^a}} \psi(\frac{t}{2^a} - b),\ a, b \in \mathbb{Z}. \tag{4}$$

Given a time sequence signal $s(t)$, DWT transforms the input by wavelet $\psi_{a,b}(t)$ into the following signal:

$$d_{a,b}(s(t), \psi(t)) = \int_{-\infty}^{+\infty} s(t)\psi^*_{a,b}(t)\, dt = \langle s(t), \psi_{a,b}(t)\rangle, \tag{5}$$

where $\psi^*_{a,b}(t)$ is the complex conjugate of $\psi_{a,b}(t)$.

Furthermore, DWT can be interpreted in terms of a multi-resolution decomposition of the input signal [28], where a hierarchy of an approximation and M-level of detailed signals can be constructed:

$$\begin{aligned} s(t) = &\sum_b A_{M,b}\, 2^{-M/2} \varphi(\frac{t}{2^M} - b) + \\ &\sum_a^M \sum_b d_{a,b}(s(t), \psi(t))\, 2^{-a/2} \psi(\frac{t}{2^a} - b) \\ \triangleq &\ A_M(t) + \sum_a^M D_a(t), \end{aligned} \tag{6}$$

where $A_{M,b} = \langle s(t), \varphi_{M,b}(t)\rangle$ are the approximation coefficients at level M, and $\varphi(t)$ is a companion scaling function [28]. Using (6), input signal $s(t)$ is decomposed into an approximation signal $A_M(t)$ and M detail signals $D_a(t)$. This decomposition works as a sequence of high-pass and low-pass filters. Interested reader may refer to [28] for detailed theoretical analysis.

Different from previous research utilizing DWT for data analysis, e.g., [24], [25], we are more interested in the general trend of GPS trajectory attributes than the details. In addition, since the GPS records are typically sparse compared with electrical signals, the details of the ground truth signal may be in fact lost in GPS trajectories. Therefore, we discard all detail signals and only employ the approximation signal $A_M(t)$ of the four pre-processed attributes. We follow the analysis in [24] and adopt `db` (daubechies) and `sym` (symlets) mother wavelets to decompose the attributes. The selection of mother wavelets is based on the characteristics of the data set. When there are sufficient samples, both employed families are generally preferred contributed by their robustness and agnostic to heterogeneous data properties, e.g., signal length and sample size. Consequently, we select eight discrete wavelets from `db` and `sym` families to extract the time-frequency domain feature of the processed data as presented in Table I.

From (6), it can be observed that the decomposition level M influences the approximation signal. There is a maximum level of decomposition for each discrete wavelet determined by the signal and wavelet decomposition filter lengths:

$$\overline{M} = \lfloor \log_2 \frac{n}{F - 1} \rfloor, \tag{7}$$

TABLE I
SELECTED WAVELETS AND THEIR FILTER LENGTHS

Wavelet	Filter Length	Wavelet	Filter Length
`db1` and `sym1`	2	`db2` and `sym2`	4
`db3` and `sym3`	6	`db4` and `sym4`	8

where F is the filter length. In this work, we set the decomposition level to its maximum value in order to obtain fine-grained decomposed signals, which is also recommended by the literature, e.g., [24], [29]. The details of selected discrete wavelets are also presented in Table I. Combining this table and (7), it is clear that ω must be no less than 14 in order to at least have one decomposition level in all of the selected wavelets, i.e., $\omega \geq 14$. Consequently, totally eight approximation signals can be calculated for each GPS trajectory comprising 14 GPS records or more. Nonetheless, the decomposed signals are in general too long as inputs of the subsequent DNN [24]. To handle this issue, we manually select a series of statistical features of these approximation signals to represent the characteristics of the original GPS trajectory attributes. Specifically, we calculate the maximum, minimum, mean, and standard deviation values of all the eight $A_{M,b}$ signals to provide input information for the subsequent DNN. These statistics have demonstrated their superiority in the literature, see [24], [25], [29], [30] for examples. Finally, given a GPS trajectory, 4 attributes $\times\, 8$ wavelets $\times\, 4$ statistics $= 128$ features can be calculated. The feature values are used to construct a feature vector, which is input into a DNN-based travel mode identifier to develop the result.

D. Travel Mode Identifier

DNN is a type of machine learning techniques which employs multiple hidden neuron layers to emulate the highly non-linear mathematical relationships between the input and output provided in a training data set [19]. It has been widely adopted in many areas of research, and is recognized as one of the most effective data classifiers [31], [32]. In this work, we construct a DNN to handle travel mode identification tasks.

The schema of the proposed DNN structure is presented in Fig. 3. In this network, there are three major components, namely, LSTM, residual links, and fully connected neuron layers (FCL). LSTM maps the t-th record and all previous ones in the input, i.e., $\{\langle s_i, a_i, k_i, u_i\rangle | \forall i \leq t\}$, to an output vector h_t using the following calculations [33]:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \tag{8a}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \tag{8b}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \tag{8c}$$

$$c_t = f_t * c_{t-1} + i_t * \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \tag{8d}$$

$$h_t = o_t * \tanh(c_t), \tag{8e}$$

where $x_t = \langle s_i, a_i, k_i, u_i\rangle$ is the input vector, f_t, i_t, o_t are the intermediate gate states, c_t is the intermediate cell state, $\sigma(\cdot)$ is the sigmoid function, and we use symbol $*$ to denote Hadamard product. All W and b entries in the equations denote the network weights and biases to be trained, which are called network parameters together. From the equations, it can be observed that to calculate h_t, both the previous output h_{t-1} and cell state c_{t-1} are employed, and variables h_t and c_t are also involved in the subsequent h_{t+1} calculation. In this process, the temporal feature characteristics can be extracted from the input data, which is essential in identifying travel modes from trajectories.

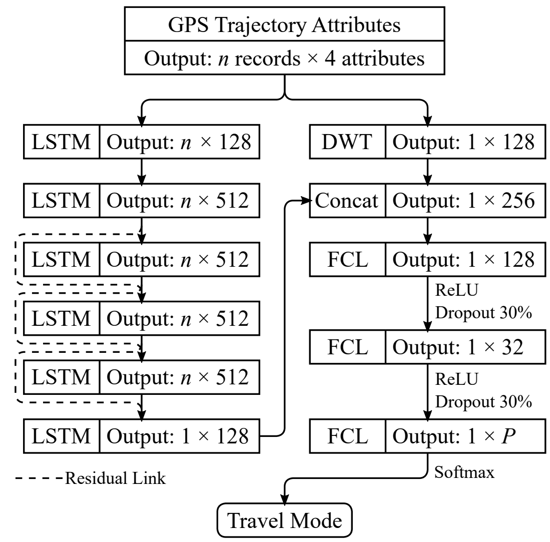


Fig. 3. Schema of the proposed travel mode identifier.

Furthermore, we introduce residual links inspired by residual network, which aims to resolve the accuracy saturation problem exists in DNN [34]. The design principle is quite simple: adding an additional identity function to the existing network structure such that the new structure with a residual link is easier to optimize than the original one [34]. As presented in Fig. 3, each residual link in the proposed DNN bypasses a layer of LSTM. The new mapping function between x_t and h_t can be re-written as follows:

$$h_t = \mathrm{LSTM}(x_t, h_{t-1}, c_{t-1}) + W_r x_t + b_r, \tag{9}$$

where $\mathrm{LSTM}(\cdots)$ refers to the LSTM calculations defined in (8), and W_r and b_r are trainable network parameters. The introduction of residual links can effectively improve the performance of the proposed identifier, which will be demonstrated by empirical studies in Section IV-B.

After being extracted by the LSTM layers, the temporal features of the input GPS trajectory attributes are concatenated with the time-frequency domain features previously obtained by DWT. The new feature vector is input into three consecutive FCL to perform the identification, each of which can be mathematically expressed as follows:

$$y = \mathrm{actv}(W_d x + b_d), \tag{10}$$

where $\mathrm{actv}(\cdot)$ is the layer-wise activation function, and W_d and b_d are trainable network parameters. We adopt Rectified Linear Units (ReLU), i.e., $f(x) = \max(0, x)$, as the activation function of the first two layers due to its simplicity and non-linearity [19]. Since we are interested in identifying which travel mode can be inferred from the GPS trajectory out of a total P modes, the last layer is activated by the $\mathrm{softmax}$ function [19].

In the proposed DNN, the values of all network parameters need to be fine-tuned in order to emulate the non-linear relationship between input trajectory attributes and the output travel mode. Nonetheless, the enormous number of such parameters makes it impractical to be simultaneously optimized without significant overfitting problems [19], [35]. In this work, we adopt an effective scheme called "dropout" to handle the problem. Dropout randomly removes the output of a number of neurons at a controlled probability during training process, such that these neurons are not involved in the computing graph of the network. Hence, the classification result does not have strong correlations with arbitrary individual neuron parameters. This scheme has demonstrated superior performance in alleviating the overfitting issue in deep learning thanks to its capability of reducing the co-adaptation relationship among the network parameters [19], [35]. Dropout is applied to the first two FCLs in the networks, each of which has a $30\,\%$ random dropout probability.

E. Caches for Online Travel Mode Identification

Using the data pre-processing scheme, trajectory feature extractor, and travel mode identifier elaborated in previous subsections, it is evident that the system can determine a travel mode given a raw GPS trajectory with a sufficient length. This meets the requirement by the global travel mode identification, in which batches of such trajectories are input into the model and they are not updated or modified subsequently.

Meanwhile, online travel mode identification takes a step further. The system needs to provide pre-mature identifications, and later update the result given more GPS records in a same trajectory which are originally not available. While it is possible to process the new trajectories without using prior knowledge, such information can notably reduce the repetitive computations, rendering a faster identification speed. In the proposed mechanism, we use two caches to store the results of computation steps in DWT and DNN for later use. In particular, the intermediate approximation signals calculated during the multi-resolution decomposition in DWT can be re-used if a new input signal is to be decomposed with the same wavelet [28]. These intermediate results are stored in an intermediate signal database. Similarly, the output and cell state of all LSTM layers in the proposed DNN can be re-used according to (8), which are stored in a latent information database. For instance, to evaluate the result of an n-step input without prior information, all five sub-equations have to be calculated for n times. If a new $(n+1)$-step input with the same first n steps as the previous one is to be evaluated, each sub-equation is calculated only once given c_n and h_n, which are previously computed. In theory, $n+1$ times speedup can be achieved with the database.

IV. Case Studies

In this work, we propose a novel travel mode identifier based on DWT and DNN. In this section, we conduct a series of comprehensive case studies to assess its performance. In particular, we first compare the proposed mechanism with other identification approaches in recent literature and other machine learning techniques. Subsequently, we conduct a preliminary hyperparameter test to illustrate how to design the DNN architecture to yield the best system performance and the necessity of DWT. Then we investigate the online identification accuracy of GPS data streams to test the system robustness in online travel mode identification cases. Finally, we study the impact of the computed data attributes by the pre-processing step on the system performance.

In the case studies, we adopt $32\,444$ real-world crowd-sourced GPS trajectories by 69 users in GeoLife project [11], [16] in a period of over five years. Each trajectory is comprised of a sequence of GPS records with heterogeneous and infrequent sampling intervals from $1\,\mathrm{s}$ to $5\,\mathrm{s}$. We generally follow the previous work [9] to segment the raw data and assign travel modes. Specifically, we consider the five ground-based travel mode in the data set, i.e., walk, bike, bus, car[2], and train. Meanwhile, unlike [9], we do not employ a further data cleansing approach that makes use of transportation network or social information, e.g., road speed limit, human's available cycling power, etc. The un-cleansed data can in fact test the robustness of the proposed travel mode identification mechanism, and the previously employed data cleansing may not be available for many transportation systems or in online identification process.

To train the network parameters of the proposed DNN with supervised learning, we first divide all $32\,444$ trajectories into two data sets by the ratio of 3:1, which accords with the common practice, see [36], [37] for examples. The set with $24\,333$ trajectories, called training set, is used to train the neural network with Adam optimizer [38] and multinomial cross-entropy loss [19]. The other set with 8111 trajectories, called testing set, is used to evaluate the system performance after the DNN is well-trained. This design is mainly for cross-validation, and over-fitting problem can be easily identified [19].

All case studies are conducted on computing servers with two Intel Xeon E5 CPU and 128 GB RAM. The simulation is developed in Python, and the proposed DNN is modeled with PyTorch [39]. Eight nVidia GTX 1080 Ti GPUs are employed on each server for algebraic and DNN computing acceleration. Lastly, ω is set to 16, whose sensitivity can be illustrated by the system performance of online identifications in Section IV-C.

A. Identification Accuracy

We first investigate the accuracy of identification by the proposed mechanism, which is among the most important performance metrics in determining travel modes based on GPS trajectories. We use the previously defined training set to tune the network parameters in the constructed DNN, and adopt the testing set to evaluate the accuracy. The simulation results are presented in Table II. In this table, the multi-class identification results are demonstrates in the form of confusion matrices with respect to the testing and training sets. In addition, the recall and precision values for each individual travel mode are calculated, which imply the accuracy of the

[2]In the original data set, car and taxi are counted interchangeably. Hence, we combine the two classes as one.

TABLE II
CONFUSION MATRICES OF THE PROPOSED TRAVEL MODE IDENTIFICATION MECHANISM

		Testing Set						Training Set					
		Predicted Mode					Recall (%)	Predicted Mode					Recall (%)
		Walk	Bike	Bus	Car	Train		Walk	Bike	Bus	Car	Train	
Real Mode	Walk	2528	44	19	12	4	96.9	7557	118	51	29	10	97.3
	Bike	62	1279	20	14	5	92.7	174	3892	65	38	19	92.9
	Bus	11	37	1658	86	28	91.1	36	131	4983	247	75	91.1
	Car	9	26	65	1019	22	89.3	25	77	202	2977	68	88.9
	Train	10	23	41	52	1037	89.2	26	68	130	147	3188	89.6
Precision (%)		96.5	90.8	92.0	86.1	94.6	**92.7**	96.7	90.8	91.6	86.6	94.9	**92.9**

TABLE III
COMPARISON OF TRAVEL MODE IDENTIFICATION APPROACHES

Approach	Accuracy (%)
Proposed	**92.7**
DT-heuristic [11]	76.2
CNN [9]	84.8
Image-based DNN [20]	67.9
k-NN (in [9])	63.5
SVM (in [9])	65.4
RF (in [9])	78.1
MLP (in [9])	59.4
RF-based [3]	82.0

TABLE IV
TRAVEL MODE IDENTIFICATION APPROACHES WITH DIFFERENT DATA SET AND/OR ADDITIONAL INFORMATION

Approach	Accuracy (%)	Other Information Required
Proposed	92.7	-
SVM-based [10]	97.1	Smart phone sensors ($\geq$5 Hz)
RF-based [3]	89.3	-
RF-based [3]	93.1	Socioeconomic attributes
Statistical [40]	79.3 to 94.7	Instant speed and heading

identifier on specific modes and the precision of the identifier if a specific mode is predicted, respectively. The overall identification accuracy is also presented in bold for both sets.

From the simulation results it is quite clear that the proposed mechanism yields satisfactory travel mode identification accuracy. Additionally, "walk" mode can be more precisely identified compared with other modes. This is contributed by the more training instances of this mode [9]. Nonetheless, identifications on the other modes are still accurate despite the imbalanced training data thanks to the well designed neural network structure. Lastly, it can also be observed that modes "walk" and "bike" are more commonly mis-identified as the other, and so are "bus" and "car". This is because these modes may share similar motion and displacement characteristics.

It is also of interest to investigate the performance improvement of the proposed mechanism compared with previous results in the literature and other machine learning techniques. To illustrate this, we include the identification results of a series of methodologies on the same data set, namely, decision tree (DT)-based heuristic [11], CNN [9], image-based DNN [20], k-nearest neighborhood (k-NN), support vector machine (SVM), random forest (RF), and multilayer perceptron (MLP). In addition, reference [3] proposed two RF-based travel mode identifiers, one of which only uses GPS trajectories. We implement this approach and adopt the GeoLife data set to test its performance. The accuracy of all other compared approaches are adopted from the respective publication.

The comparison are demonstrated in Table III. We can easily draw a conclusion from this table that the proposed mechanism can significantly outperform all compared approaches on the same GeoLife data set. In particular, a 7.9% identification accuracy improvement from 84.8% to 92.7% can be witnessed beyond the state-of-the-art in the literature, i.e., [9]. This is due to the different design principles of the neural networks. In [9] the CNN-based approach presents time-sequence data as different features in a convolution layer. While this design can guide the network to consider the whole time horizon simultaneously, the temporal correlation in the input time sequence is discarded. On the contrary, this correlation is considered by both the wavelet transform process and the LSTM layers in our proposed design, resulting in improved performance. This is also a main reason for the superior identification accuracy compared with other approaches.

Furthermore, we also compare the proposed mechanism with other recent identification approaches using other data sets and optionally more information in [3], [10], [40] for readers' reference. Their overall identification performance is presented in Table IV. From the results, it is clear that although more information is required to develop travel mode results, the approaches in [3], [40] can only achieve a similar identification accuracy (within 2%). Considering that the testing data set in these references are relatively small, i.e., 914 for [3] and 324 for [40], the robustness of these approaches remains unknown. Additionally, the RF-based approach without additional information in [3] may serve as a baseline to compare Tables III and IV. As the approach can only score 82.0% accuracy, it seems that the data set used in this work is more difficult to be classified than data in the literature. Meanwhile, the SVM-based approach proposed in [10] can achieve an identification accuracy of 97.1%, providing relatively high-frequency sensor data of smart phones. The performance improvement is established based on a higher data transmission throughput requirement, which may not be always possible. In addition, our proposed GPS trajectory-based approach is not limited to smart phone devices. Therefore, both this work and [10] devise valid travel mode identification algorithms, which may be suitable for different real-world applications.

Lastly, we record the required computation time of the proposed mechanism. For all 8111 trajectories, the average identification time is less than $10\,\mathrm{ms}$, and the system can be trained within $56\,\mathrm{min}$. According to the record in [9], the CNN based approach listed in Table III can be trained

TABLE V
A SELECTION OF WELL-PERFORMING DNN ARCHITECTURES AND THEIR TRAVEL MODE IDENTIFICATION ACCURACY

Model	LSTM×3	Res.	LSTM	Res.	LSTM	Res.	LSTM	Res.	LSTM	FCL	FCL	FCL	FCL	Accuracy (%)
Proposed	128/512/512	Yes	512	Yes	512	Yes	-	-	128	128	-	32	P	**92.7**
A	128/512/512	-	-	-	-	-	-	-	128	128	-	32	P	79.5
B	128/512/512	Yes	-	-	-	-	-	-	128	128	-	32	P	79.6
C	128/512/512	-	512	-	-	-	-	-	128	128	-	32	P	87.3
D	128/512/512	Yes	512	Yes	-	-	-	-	128	128	-	32	P	88.0
E	128/512/512	-	512	-	512	-	-	-	128	128	-	32	P	92.1
F	128/512/512	Yes	512	Yes	512	Yes	512	Yes	128	128	-	32	P	92.4
G	128/512/512	Yes	512	Yes	512	Yes	-	-	128	-	-	32	P	90.9
H	128/512/512	Yes	512	Yes	512	Yes	-	-	128	128	128	32	P	92.3

TABLE VI
IDENTIFICATION ACCURACY WITHOUT DWT

Model	Testing Accuracy (%)	Training Accuracy (%)
Proposed	**92.7**	92.9
Proposed w/o DWT	88.0	89.6
Model E w/o DWT	87.2	89.1
Model F w/o DWT	87.6	89.3
Model H w/o DWT	87.7	89.6

in approximately $25\,\mathrm{min}$. Nonetheless, the model training process is typically conducted offline, and well-trained model can be used for travel mode identification without the need of further training. Therefore, both $25\,\mathrm{min}$ and $56\,\mathrm{min}$ offline training time are acceptable. Considering that the identification error rate is reduced from 15.2% (CNN) to 7.3%, the proposed mechanism still outperforms the other approaches.

B. Identifier Architecture

In this work, we propose a DWT and DNN-based travel mode identification mechanism. In the proposed DNN, we adopt several layers of LSTM and FCL to perform the data classification task. When designing a neural network, a common problem is to determine the number of layers to develop satisfactory results. In this sub-section, we conduct a preliminary DNN hyperparameter test to determine a well-performing DNN structure. In addition, we also investigate the necessity of DWT in the proposed mechanism.

We first test the architecture of the proposed DNN. Table V presents a selection of well-performing structures. In this table, "Res." denotes a residual link for the previous LSTM layer, and the values in the table represents the number of output features in the corresponding layer. The output of DWT is always concatenated with the output of the last LSTM layer as the input of the first FCL. All architectures are trained with the same training set, and the identification accuracy is assessed by the same testing set. From the table we can come to a series of conclusions. First, including more LSTM layers in the architecture generally improves the system performance until the seventh one, which decreases the accuracy from 92.7% to 92.4% (model F). This accords with the widely recognized concept of deep learning that while more neural network layers can improve the capability of extracting hidden data features, deeper architectures are more difficult to train and prone to over-fitting [19]. Additionally, adding residual links can slightly improve the performance. This is due to the fact that residual links can flatten the network parameter searching space, rendering a faster training speed [34]. Hence, better accuracy can be expected given finite training time and epochs. Finally, three layers of FCL after LSTM and DWT output concatenation is sufficient to perform the travel mode classification. More or less layers undermine the identification accuracy, though not substantially.

We also conduct a second analysis on the impact of DWT. Specifically, the best performing architectures from Table V, i.e., the proposed one, as well as models E, F and H, are selected to construct travel model identification mechanisms but without DWT. The output of LSTM layers is directly input into the first FCL without the original concatenation step. The new mechanisms are trained and tested with the same data set, and the results are presented in Table VI, in which "w/o" stands for "without". The simulation result indicates that DWT can notably improve the system performance on identifying travel modes. This can be observed by comparing the testing accuracy values with the corresponding ones presented in Table V. Furthermore, there is another interesting observation on the training accuracy without DWT that the gap between training and testing accuracy is larger when DWT is not involved in the computation. The observations lead to a conclusion that the hidden time-frequency domain features extracted by DWT are among the distinguishing ones for travel modes. Indeed, DWT is necessary in the proposed mechanism to further improve the system performance.

C. Online Travel Mode Identification

In previous case studies, the GPS trajectories in the testing set are considered as completely known when performing the classification, which corresponds to the global travel mode identification scenario. In the meantime, other real-world applications heavily rely on real-time accurate identification of travel modes given preliminary or partial GPS trajectories. While the proposed mechanism can handle streams of incoming GPS records thanks to DWT and LSTM, it is of interest to investigate how it performs in such cases, especially when there are few existing records in GPS trajectories.

In this sub-section, we conduct an online travel mode identification study using the same testing data set as in previous simulations. Instead of inputting the whole sequence of GPS records into the mechanism, we only use the first ω^* records to develop a pre-mature identification result. The result is then compared with the real travel mode to determine the system performance. All 8111 trajectories in the testing set are employed in this process, and the identification results among

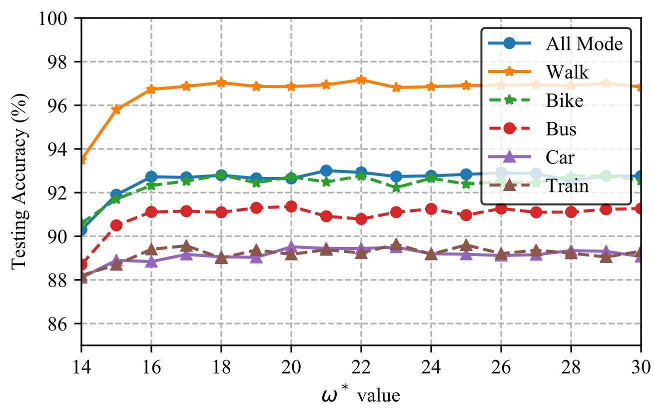


Fig. 4. Testing accuracy of online travel mode identification with different lengths of GPS trajectory.

$\{\omega^* \in \mathbb{N}^+ | 14 \leq \omega^* \leq 30\}$[3] for each and all travel modes are summarized in Fig. 4.

From the results it is clear that the proposed mechanism can develop satisfactory identification results when $\omega^* \geq 16$. This is because the approximation signals developed by wavelets with longer filter length, e.g., `db4` and `sym4`, is too short to represent significant time-frequency domain features of the original signal when $\omega^* \in \{14, 15\}$. In such cases, the mechanism mostly relies on DNN to perform the identification, rendering a relatively inferior performance. Despite this observation, the accuracy of $\omega^* \in \{14, 15\}$ still exceeds 90%, which notably outperforms other travel mode identification approaches in the literature as indicated in Table III. On the other hand, the accuracy is not significantly improved or undermined when ω^* is greater than 16. This indicates that the proposed mechanism requires at least 16 GPS records in a trajectory to achieve its best identification accuracy, which corresponds to up to $57\,\mathrm{s}$ in the employed GeoLife data set.

Additionally, among the five travel modes in the data set, low-speed ones (walk, bike, and bus) are relatively more prone to short GPS trajectories. While the identification accuracy for car and train is also subpar with $\omega^* \in \{14, 15\}$, the performance decrease is not as notable as the other three modes. In spite of this, all modes can achieve stable and satisfactory accuracy with $\omega^* \geq 16$. The case study indicates that the proposed mechanism is suitable for online travel mode identification.

D. Selection of Data Attributes

In Section III-B, four motion and displacement attributes are defined based on the raw GPS data. According to Fig. 1, the data serve as the only input of the subsequent computations. Hence, we are interested in investigating if all the attributes can contribute to the identification performance of the proposed mechanism. In this sub-section, we conduct a series of simulations to test the sensitivity of these attributes. Specifically, we first create several combinations of these attributes as presented in Table VII. Each combination is independently tested using the proposed mechanism and GeoLife data set, and the testing set identification accuracy is summarized. Note that the number of neurons in the first layers of LSTM and FCL are modified to accommodate the changing number of attributes.

[3]The minimal value, i.e., 14, is selected in accordance with the analysis in Section III-C.

TABLE VII
ATTRIBUTE COMBINATIONS AND ACCURACY

Combination	s_i	a_i	k_i	u_i	Accuracy (%)
Proposed	Yes	Yes	Yes	Yes	**92.7**
A	Yes	Yes	-	-	90.6
B	Yes	Yes	Yes	-	91.5
C	-	Yes	Yes	-	88.4
D	Yes	-	-	Yes	91.0
E	Yes	Yes	-	Yes	91.7
F	-	Yes	Yes	Yes	90.1

The simulation results lead to some insights into the selection of data features. On the one hand, the speed and turn attributes play critical roles in identifying travel modes. Combination B and the proposed one both witnesses significant performance improvements over combinations C and F, respectively, thanks to the inclusion of speed attribute. Similar trends can also be observed on combination E and the proposed one comparing with combinations A and B due to introducing the turn attribute u_i, respectively. On the other hand, the above observation does not rule out the necessity of acceleration a_i and jerk k_i. In general, the identification accuracy increases with more attributes included in the computation according to Table VII. Hence, all the four motion and displacement attributes developed in the pre-processing step of the proposed identification mechanism contribute to accurately determining travel modes based on GPS trajectories.

Furthermore, we also investigate some other statistical features of the DWT approximation signals, namely, skewness, kurtosis, energy, and entropy. Specifically, we construct two new sets of attributes. The first one includes the original four as described in Section III-C together with skewness and kurtosis, and the second one includes all eight attributes. The former achieves a 92.5% overall accuracy with a $61\,\mathrm{min}$ training, while the latter develops a 92.8% overall accuracy with a $77\,\mathrm{min}$ training. Compared with the benchmark performance of 92.7% accuracy and $56\,\mathrm{min}$training, neither of the new attribute sets can develop better accuracy with comparable training time.

V. CONCLUSIONS

In this paper, we propose a new travel mode identification mechanism with only GPS trajectories based on discrete wavelet transform (DWT) and deep learning techniques. This mechanism takes a series of raw GPS records as input to infer the corresponding travel mode. Specifically, the input data is first pre-processed to summarize four motion and displacement attributes, namely, speed, acceleration, jerk, and turn. These attribute vectors then go through DWT in order to obtain the hidden time-frequency domain features. At the same time, the vectors are also input into a tailor-made deep neural network (DNN), which jointly considers the previously extracted features by DWT and develops the final travel mode inference. Different from previous work, the proposed mechanism fully utilizes temporal correlations in GPS trajectories to train an intelligent system for identification, and the length of trajectories does not need to be fixed. In addition, the whole identification process can be conducted in real-time thanks to the computationally efficient nature of DWT and DNN employed in the mechanism.

To evaluate the performance of the proposed mechanism, a series of comprehensive case studies are carried out. We first test the identification accuracy on a real-world crowd-sourced data set with $32\,444$ GPS trajectories, and compare the results with state-of-the-art travel mode identification approaches in the literature. The comparison indicates that the proposed mechanism can provide notably more accurate identifications on the same data set, i.e., 92.7% versus 84.8%. In addition, we conduct a structure test to reveal the best-performing architecture for the mechanism. Subsequently, we investigate the performance of the proposed mechanism on classifying online GPS trajectories. The results show that satisfactory accuracy can be achieved with reasonably short trajectories. Finally, a feature sensitivity test is conducted, which indicates that all four data attributes contribute to the identification process.

The future work can be divided into two parts. On the one hand, it is possible to extend the mechanism and employ more advanced artificial intelligence and machine learning techniques. On the other hand, the current mechanism relies on labeled GPS trajectories in the training set. It is an interesting research direction to investigate schemes to use unlabeled GPS trajectories in training the mechanism.

REFERENCES

[1] J. L. Adler and V. J. Blue, “Toward the design of intelligent traveler information systems,” *Transportation Research Part C: Emerging Technologies*, vol. 6, no. 3, pp. 157–172, Jun. 1998.

[2] J. Scheiner and C. Holz-Rau, “Travel mode choice: affected by objective or subjective determinants?” *Transportation*, vol. 34, no. 4, pp. 487–511, Jul. 2007.

[3] B. Wang, L. Gao, and Z. Juan, “Travel mode detection using GPS data and socioeconomic attributes based on a random forest classifier,” *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 5, pp. 1547–1558, May 2018.

[4] J. J. Q. Yu and A. Y. S. Lam, “Autonomous vehicle logistic system: joint routing and charging strategy,” *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 7, pp. 2175–2187, Jul. 2018.

[5] J. J. Q. Yu, “Two-stage request scheduling for autonomous vehicle logistic system,” *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 5, pp. 1917–1929, May 2019.

[6] N. Ding, Q. He, C. Wu, and J. Fetzer, “Modeling traffic control agency decision behavior for multimodal manual signal control under event occurrences,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2467–2478, Oct. 2015.

[7] N. Eluru, V. Chakour, and A. M. El-Geneidy, “Travel mode choice and transit route choice behavior in Montreal: insights from McGill University members commute patterns,” *Public Transport*, vol. 4, no. 2, pp. 129–149, Nov. 2012.

[8] S. Shewmake and L. Jarvis, “Hybrid cars and HOV lanes,” *Transportation Research Part A: Policy and Practice*, vol. 67, pp. 304–319, 2014.

[9] S. Dabiri and K. Heaslip, “Inferring transportation modes from GPS trajectories using a convolutional neural network,” *Transportation Research Part C: Emerging Technologies*, vol. 86, pp. 360–371, Jan. 2018.

[10] X. Su, H. Caceres, H. Tong, and Q. He, “Online travel mode identification using smartphones with battery saving considerations,” *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 10, pp. 2921–2934, Oct 2016.

[11] Y. Zheng, Q. Li, Y. Chen, X. Xie, and W.-Y. Ma, “Understanding mobility based on GPS data,” in *Proc. International Conference on Ubiquitous Computing*, Seoul, Korea, 2008, pp. 312–321.

[12] P. A. Gonzalez, J. S. Weinstein, S. J. Barbeau, M. A. Labrador, P. L. Winters, N. L. Georggi, and R. Perez, “Automating mode detection for travel behaviour analysis by using global positioning systems-enabled mobile phones and neural networks,” *IET Intel. Transport Syst.*, vol. 4, no. 1, pp. 37–49, Mar. 2010.

[13] A. Bolbol, T. Cheng, and I. Tsapakis, “A spatio-temporal approach for identifying the sample size for transport mode detection from GPS-based travel surveys: a case study of London’s road network,” *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 176—187, 2014.

[14] L. Wu, B. Yang, and P. Jing, “Travel mode detection based on GPS raw data collected by smartphones: a systematic review of the existing methodologies,” *Information*, vol. 7, no. 4, pp. 1–19, Nov. 2016.

[15] A. C. Prelipcean, G. Gidófalvi, and Y. O. Susilo, “Transportation mode detection – an in-depth review of applicability and reliability,” *Transport Reviews*, vol. 37, no. 4, pp. 442–464, Oct. 2017.

[16] Y. Zheng, L. Liu, L. Wang, and X. Xie, “Learning transportation mode from raw GPS data for geographic applications on the web,” in *Proc. International Conference on World Wide Web*, Beijing, China, 2008, pp. 247–256.

[17] Z. Sun and X. J. Ban, “Vehicle classification using GPS data,” *Transportation Research Part C: Emerging Technologies*, vol. 37, pp. 102–117, 2013.

[18] H. Mäenpää, A. Lobov, and J. L. M. Lastra, “Travel mode estimation for multi-modal journey planner,” *Transportation Research Part C: Emerging Technologies*, vol. 82, pp. 273 – 289, 2017.

[19] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*, F. Bach, Ed. Cambridge, MA: MIT Press, Nov. 2016.

[20] Y. Endo, H. Toda, K. Nishida, and A. Kawanobe, “Deep feature extraction from trajectories for transportation mode estimation,” in *Proc. Advances in Knowledge Discovery and Data Mining*, Cham, Switzerland, Apr. 2016, pp. 54–66.

[21] J. V. Jeyakumar, E. S. Lee, Z. Xia, S. S. Sandha, N. Tausik, and M. Srivastava, “Deep convolutional bidirectional LSTM based transportation mode recognition,” in *Proc. ACM International Joint Conference and International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, Singapore, Oct. 2018, pp. 1606–1615.

[22] T. H. Vu, L. Dung, and J.-C. Wang, “Transportation mode detection on mobile devices using recurrent nets,” in *Proc. 24th ACM International Conference on Multimedia*, Amsterdam, The Netherlands, Oct. 2016, pp. 392–396.

[23] H. Liu and I. Lee, “End-to-end trajectory transportation mode classification using Bi-LSTM recurrent neural network,” in *Proc. 12th International Conference on Intelligent Systems and Knowledge Engineering*, Nanjing, China, Nov. 2017, pp. 1–5.

[24] J. J. Q. Yu, Y. Hou, A. Y. S. Lam, and V. O. K. Li, “Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks,” *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1694–1703, Mar. 2019.

[25] J. J. Q. Yu, Y. Hou, and V. O. K. Li, “Online false data injection attack detection with wavelet transform and deep neural networks,” *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3271–3280, Jul. 2018.

[26] T. Vincenty, “Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations,” *Survey Review*, vol. 23, no. 176, pp. 88–93, 1975.

[27] O. Bagdadi and A. Várhelyi, “Development of a method for detecting jerks in safety critical events,” *Accident Analysis & Prevention*, vol. 50, pp. 83–91, 2013.

[28] S. G. Mallat, “A theory for multiresolution signal decomposition: the wavelet representation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.

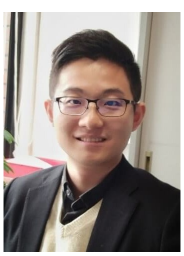
[29] D. Chen, S. Wan, and F. S. Bao, “Epileptic focus localization using discrete wavelet transform based on interictal intracranial EEG,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 5, pp. 413–425, May 2017.

[30] D. P. Mishra, S. R. Samantaray, and G. Joos, “A combined wavelet and data-mining based intelligent protection scheme for microgrid,” *IEEE Trans. Smart Grid*, vol. 7, no. 5, pp. 2295–2304, Sep. 2016.

[31] G. E. Hinton and R. R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006.

[32] J. J. Q. Yu, D. J. Hill, V. O. K. Li, and Y. Hou, “Synchrophasor recovery and prediction: A graph-based deep learning approach,” *IEEE Internet Things J.*, in press.

[33] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.
[34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2016, pp. 770–778.
[35] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *arXiv preprint arXiv:1207.0580*, Jul. 2012.
[36] J. J. Q. Yu, A. Y. S. Lam, D. J. Hill, Y. Hou, and V. O. K. Li, "Delay aware power system synchrophasor recovery and prediction framework," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3732–3742, Jul. 2019.
[37] J. J. Q. Yu, D. J. Hill, A. Y. S. Lam, J. Gu, and V. O. K. Li, "Intelligent time-adaptive transient stability assessment system," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 1049–1058, 2018.
[38] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," in *Proc. International Conference on Learning Representations*, San Diego, CA, Dec. 2015.
[39] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in PyTorch," in *Proc. Advances in Neural Information Processing Systems*, Long Beach, CA, Dec. 2017.
[40] B. Assemi, H. Safi, M. Mesbah, and L. Ferreira, "Developing and validating a statistical model for travel mode identification on smartphones," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 7, pp. 1920–1931, July 2016.



James J.Q. Yu (S'11–M'15) received the B.Eng. and Ph.D. degree in Electrical and Electronic Engineering from the University of Hong Kong, Pokfulam, Hong Kong, in 2011 and 2015, respectively. He was a post-doctoral fellow at the University of Hong Kong from 2015 to 2018. He is currently an assistant professor at Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China, and an honorary assistant professor at Department of Electrical and Electronic Engineering, the University of Hong Kong. He is also the chief research consultant of GWGrid Inc., Zhuhai and Fano Labs, Hong Kong. His research interests include smart city and urban computing, deep learning, intelligent transportation systems and smart energy systems. He is an Associate Editor of the IET Smart Cities journal.