IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 26, 2019, accepted August 22, 2019, date of publication September 3, 2019, date of current version September 17, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2939203

A Survey of Human Mobility Models

GÜRKAN SOLMAZ[1,2], (Member, IEEE), AND DAMLA TURGUT[2], (Senior Member, IEEE)
[1]NEC Laboratories Europe, 69115 Heidelberg, Germany
[2]Department of Computer Science, University of Central Florida, Orlando, FL 32816-2362, USA

Corresponding author: Gürkan Solmaz (gurkan.solmaz@neclab.eu)

ABSTRACT Human mobility models are key components of various research fields including transportation, mobile networks, disaster management, urban planning, and epidemic modeling. Understanding human mobility has a major role in the realistic evaluation of new approaches to challenges in these fields. For the perspective of networked systems, simulations of the networks with human participants such as opportunistic social networks are highly dependent on human mobility. In this article, we summarize the state of the art for scientific research on human mobility and survey the currently used human mobility models. We discuss the commonly used metrics and data collection techniques. Furthermore, we include a taxonomy of the mobility models according to their main characteristics and classify them. We lastly discuss the general trends, applicability, further research directions and open problems of human mobility modeling.

INDEX TERMS Human mobility, mobility models, Internet of Things, smart cities.

I. INTRODUCTION

Realistic human mobility modeling has great potential benefits to societies. Traces generated by human mobility models can be used in the simulations of wireless ad hoc networks [1], epidemics, urban planning, transportation systems, and disaster response.

As smartphones have become increasingly common in recent years, they have taken on a central role in lives as the information and communication sources [2]. With their sensing capabilities, smartphones are widely used in various life activities such as health condition monitoring [3] and crowd monitoring in ordinary scenarios [4] and during disasters [5]. Moreover, researchers nowadays propose new and unconventional applications which are enabled by the use of smartphones such as applications of crowdsourcing [6] and opportunistic communication [7], [8]. Human mobility plays the utmost role in the performance of such applications. For instance, for an opportunistic social network which consists of smartphones, intercontact times of the nodes vary with the movement of the smartphone owners. Since the nodes in opportunistic social networks store and carry messages to each other, intercontact times of the nodes have a direct influence on the message delays such that longer intercontact times cause longer message delays. Another example effect of human mobility is related to crowdsourcing for environmental monitoring [9]. If people having smartphones with sensing capability frequently move and explore an area, crowdsourcing will produce a global estimation of the environment. On the other hand, if people stay in a smaller area for a long time without significant diffusion, they will have only the local information about the environment. Therefore, coverage of the crowdsourcing application highly depends on the movement of the people.

While realistically representing the deterministic aspects and non-deterministic aspects of human mobility stays a challenge, there is a certain need for increased research efforts in mobility modeling. Upcoming technologies such as the Internet of Things (IoT) and 5G can help more accurate understanding of human mobility and researchers can simulate many environments more realistically [10]. One of our goals is to grow attention to the field of human mobility modeling. A general overview of the processes that are followed by the mobility modeling researchers is given in Figure 1 and can be briefly summarized as follows. Mobility modeling involves the collection of real-life mobility data and filtering of the data as well as the modeling of the environment and the people's behavioral decisions. A mobility model can be verified with thorough analyses of both parts with various metrics and the model needs to be calibrated during this process. As the last main process, the implementation of a human mobility simulation generates synthetic mobility traces. The traces of the mobility simulation can be used for other simulations of networks, urban transportation planning, crowd management, disaster management, and so on.

The associate editor coordinating the review of this article and approving it for publication was Byung-Seo Kim.

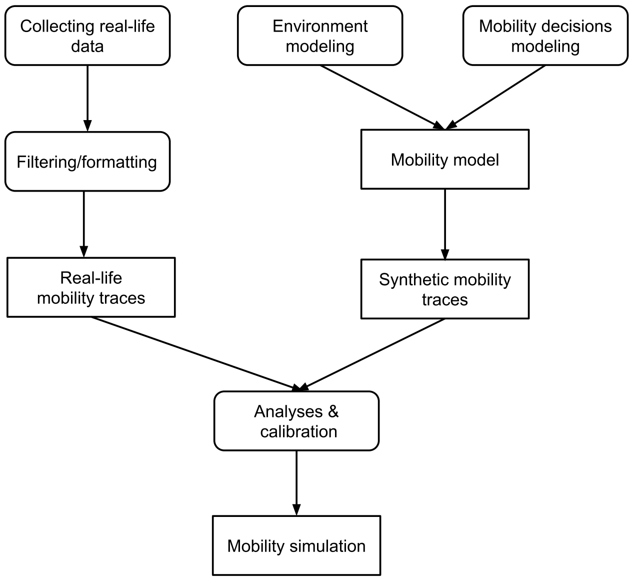

FIGURE 1. **Overview of the creation of a human mobility model. Rounded rectangle: A process, sharp rectangle: A product of the process(es) (data, model, or simulation).**

A. BACKGROUND AND OUTLINE

This article surveys human mobility models from the computer science perspective. There are two major targets:

- Classification of models
- Introduction to the field

The first target aims for the audiences who are the users of human mobility models. We classify the models to help them easily choose and use the most relevant human mobility model of their needs. As an example, an engineer who wants to simulate the human mobility inside an office environment can simply eliminate the mobility models which incorporate vehicle use as a main component. Section IV provides simple classifications of human mobility models for this purpose. The second target aims for researchers and developers to understand human mobility modeling in-depth to build their models based on the current state-of-art. For this purpose, the article provides an in-depth look at the main findings of the scientific research on human mobility analyses and technical features of the human mobility models.

There exist surveys related to mobility modeling in the literature. These studies particularly focus on the communications field. The surveys by Camp *et al.* [11] and Bai and Helmy [12] provide wider mobility scope in the sense that they focus on synthetic models (e.g., Random Waypoint Model). These surveys provide initial background on the mobility modeling in ad hoc networking research. In a more recent study, Treurniet [13] focuses on mobile networking and describes more recent and advanced mobility models in the same domain. Batabyal and Bhaumik [14] survey the mobility models and traces of vehicles, humans, and wild animals that are used in opportunistic routing algorithms in networks. Similarly, Dong *et al.* [15] survey handling mobility in networks and the mobility-aware MAC protocols. A recent survey by Hess *et al.* [16] is related to the engineering aspects of mobility modeling and data-driven processes of mobility model creation for mobile networks. They focus on processes such as data collection and validation of mobility models. Karamshuk *et al.* [17] provide a short overview of human mobility models for opportunistic networks. Our survey presents a comprehensive classification and technical background with a specific focus on human mobility models. It essentially provides a different perspective and classification based on emerging research areas of pedestrian mobility, vehicular networking, social networking, and disaster management.

The remaining of this article is organized as follows. We start with the findings led to certain advancements in the human mobility models in Section II. We describe commonly used metrics and data collection techniques in Section III. We classify the human mobility models based on emerging research fields and include a comprehensive taxonomy in Section IV. We summarize recently developed significant pedestrian walk models (without transportation) in Section V. We focus on the models based on social interactions and the social network theory in Section VI. We discuss the human mobility models which consider the use of the vehicles in Section VII and the mobility models considering the disaster scenarios in Section VIII. We present the general trends and applicability of the surveyed models in Section IX. Finally, we provide a discussion on the future research directions, open problems, expected advancements and challenges in human mobility in Section X.

II. UNDERSTANDING HUMAN MOBILITY

In this section, we describe the significant works and statistical analyses on human mobility which led to certain advancements in the field. We start our description with the studies related to human mobility in geographical scale. By geographical scale, we mean the human travels in geographical distances, which includes long-distance travels with airplanes, trains, cars, and other vehicles. Later, we describe the studies related to mobility in micro-scale. By micro-scale, we mean the people's mobility in a smaller area such as a building or a shared pedestrian way, such that the speed of a person changes due to crowd dynamics, traffic congestion, social interactions, or various other reasons.

As travels of the people are direct causes for the spread of epidemic diseases around the world, the statistical analyses on human mobility have fundamental importance to society. Brockmann *et al.* [18] focus on the scaling of human mobility in geographical scale. While large datasets of GPS traces of human mobility are not available, these researchers analyze the mobility by the circulation of the banknotes around the contiguous United States. Their observation dataset, which is obtained from a bill-tracking system, consists of 1,033,095 tracks (reports) of 464,670 dollar bills. They consider the geographical displacements between two consecutive reports of the same dollar bill for finding the travel distance, such that the second report's location x_2 and the first report's location x_1 are simply subtracted for finding

r, which is the geographical displacement and $r = |x_2 - x_1|$ and the time t elapsed between these two points. The banknotes are originated from three cities, Seattle, Jacksonville, and New York. They observe that most bank nodes are traveled shorter distances ($r \leq 10$km) in a period of 2 weeks for consecutive reports ($t \leq 14$ days). The percentages of banknotes which traveled short distances are listed as 52.7% for Seattle, 71.4% for Jacksonville, and 57.7% for New York. On the other hand, smaller percentages of banknotes traveled longer distances ($r > 800$km). These percentages are shown as 7.8% for Seattle, 2.9% for Jacksonville, and 7.4% for New York.

By analyzing large datasets of banknote circulation, Brockmann et al. show in [18] that flight distances (travel distance) of people has a power-law distribution, such that $P(\Delta r) \sim \Delta r^{-(1+\beta)}$ with the exponent value $\beta = 0.59 \pm 0.002$ (mean and standard deviation), where $P(r)$ is the probability of traveling a distance Δr in a Δt time interval. Having $\beta < 2$ corresponds to the Lévy walks [19], which is a random walk process for which step size Δr follows a power-law distribution. Lévy walk behavior shows that people mostly travel shorter distances as opposed to longer distances. The similar behavior is also observed in various animal species. For instance, Viswanathan *et al.* [20] observe that wandering albatrosses have the Lévy flights behavior.

González *et al.* [21] analyze two mobility datasets in their research. One of the datasets includes trajectories of 100,000 cell phone users for six months. From more than 6 million mobile phone users, 100,000 are randomly selected. Their trajectories include the locations of cell towers to which mobile phones are connected when they send or receive text messages or have phone calls. Overall, the dataset consists of more than 16 million displacements (Δr) entries. The second dataset involves 206 mobile phone users and their location every two hours are included for one week period. The second dataset is relatively smaller (compared to the first dataset) with 10,407 entries. González et al. observe that for both datasets, the displacement values follow a truncated power-law distribution,

$$P(\Delta r) = (\Delta r + \Delta r_0)^{-\beta} exp(-\Delta r/\chi) \tag{1}$$

with the exponent value $\beta = 1.75 \pm 0.15$, where cutoff values are $\chi_1 = 400$km for the first dataset and $\chi_2 = 80$km for the second dataset and $\Delta r_0 = 1.5$km. Hence, they find that the trajectories follow truncated Lévy walks. They also find inherent differences (heterogeneity) in trajectories of individuals which coexist with Lévy walks.

González et al. also analyze the gyration radiuses in [21]. The gyration radius of a user is the total travel distance of the user for a time interval of Δt. They determine the gyration radius distribution of all users in both datasets and found that the distribution fits the truncated-power law equation

$$P(g) = (g + g_0)^{-\beta} exp(-g/\chi) \tag{2}$$

with $g_0 = 5.8$km and cutoff value $\chi = 350$.

Their results show that Lévy flights observed in [18] is a result of population heterogeneity and the individual human mobility. However, individual human mobility has a significant regularity such that a person visits the same places such as home or workplace more frequently than other places. On the other hand, banknotes always diffuse such that it is given from one person to another while this does not apply to individuals. They find that the individual trajectories can be characterized by a two-dimensional probability distribution which is independent of gyration.

Song *et al.* [22] study the human mobility with the goal of finding if the mobility patterns are potentially predictable or not. They analyze a dataset of 50,000 cell phone users (with average call frequency $f \geq 0.5$ per hour) selected from approximately 10 million anonymous users for a period of 3 months. The dataset contains mobile phone tower trajectories as previously described.

Based on the observation of the entropy in individual human mobility trajectories, they find the maximum predictability Π_{max} for each individual which shows the future whereabouts of the person. Π_{max} represents the fundamental limit of predictability. They find that probability $P(\Pi_{max})$ is narrowly peaked approximately at $\Pi_{max} \approx 0.93$, which implies 93% predictability and no cell phone use appears to have less than 80% predictability. This analysis supports the previous study by González *et al.* [21] which suggests the regularity in the human mobility patterns. Hence, despite the spontaneity and changes in human mobility decisions, human mobility is found to be characterized by deep-rooted regularity.

So far, we mentioned about the studies related to geographical scale human mobility. Let us now discuss the research on micro-scale human mobility. Social force model (SFM) is proposed by Helbing and Molnár [23] for modeling the crowd dynamics with social forces. Social forces are caused by pedestrian densities or directions of the pedestrian streams. As the simulation of pedestrian dynamics during situations such as evacuation from buildings has critical importance, pedestrian behavior is separated into two different conditions: normal and panic situations. The pedestrian crowd dynamics in these conditions are modeled in [24]. Helbing and Johansson discuss the concepts such as social force concept, panic situation, freezing-by-heating, crowd turbulence, and emergence in [25]. They also describe the evolutionary approach for calibrating the parameters of SFM using video tracking data.

The aforementioned studies have high impacts in the human mobility field since they inspire other research studies and could be used as a basis for the development of new human mobility models [26], [27]. While these studies all analyze human mobility, they belong to two different domain. First, we discussed the studies [18], [21], [22] belonging to the first domain. The first domain considers longer distance travels (e.g., from one city to another) as well as relatively shorter travels (e.g., movement from one base station to another). The second domain [25] deals with shorter travels and focus on the micro-scale pedestrian movement dynamics. The first domain analyzes the diffusion of people on the geographical scale, while the second considers temporal speed changes during movement of dense crowds. Therefore, these two domains are not conflicting, but rather they complement each other. Furthermore, the data collection methods also vary according to the domains where the second domain demands much higher granularity. For instance, the first domain can use the movement between base stations or even movement of money banknotes, while the second domain must be able to detect distances between pedestrians using cameras and accurate computer vision algorithms. In summary, the studies of the first domain target problems in more global scale such as epidemic modeling or country-wide transportation planning, while the studies of the second domain aim to solve problems such as crowd management and evacuation planning for disasters.

III. DATA COLLECTION AND VALIDATION

A. MOBILITY DATA COLLECTION

Various human mobility data collection methods exist in the literature. Among all, GPS traces are mostly preferred as they provide more spatial and temporal accuracy compared to the other methods. However, as in one of the aforementioned studies [18], even dollar bills can be tracked and used as mobility data and researchers can exploit different data types and find new facts about human mobility. For validation of the new human mobility models, types and sizes of the collected data have a direct influence on the metrics that are used for evaluation. For instance, using Bluetooth contact traces of several people may not help to find the number of waiting places per day, while GPS trajectory of the person can be easily used for revealing such results based on this metric.

Privacy is a major concern which needs to be properly addressed by the people who collect any type of mobility data. While some datasets claim to have anonymous users, these users are probably known by the people who collect the data in the first place. Therefore, data collection should involve all possible techniques for making the users anonymous at the same time of collecting their data. Encrypting personal information before storing in the database or adding noise to the dataset in a way that it makes it impossible to differentiate individuals' information from each other can be examples of such techniques. As the latter case of adding noise brings some distortion to the data, the accuracy should be well-justified before use for validation.

GPS traces provide accurate measurements of positions and velocities of people. However, they are restricted to outdoor places due to reduced signal strength, so that they are not useful for applications such as sensor networks inside buildings or large indoor environments such as airports or indoor amusement parks. Moreover, large datasets of GPS traces are not available to the scientific community. While some cloud-based mobile applications such as Google Maps can collect GPS traces from a vast number of people all over the world, these traces are not publicly available. Applications such as OpenStreetMap [28] that allow users to voluntarily upload individual GPS traces could be useful in the future with more participation. Some relatively small-sized human mobility GPS traces collected from specific environments such as university campuses can be found in the CRAWDAD [29] archive.

Call Detail Records (CDRs) are quite coarse because each record is spatially accurate to the granularity of space covered by a cell tower. Temporally, while the times of the phone calls are accurate, they do not cover discrete time intervals such that durations between consecutive phone calls vary. Furthermore, CDRs do not have semantic information such as which place is a person's home or workplace. However, large datasets having CDRs exist and their use led to a deeper understanding of human mobility. Wifi access point (AP) datasets are also used for understanding human mobility. Each entry of these datasets includes information such as the ID of the device, start time and duration of the access and location. These datasets are being used for evaluating network simulations. For instance, WiFi APs data collected from a building with people using smart-phones can be useful for evaluating the performance of opportunistic forwarding algorithms.

The last data collection technique we would like to discuss is the video tracking technique. A very large number of videos, from websites such as YouTube or Dailymotion, can be available for research purposes, especially for tasks such as understanding crowd dynamics and micro-mobility in places such as stadiums and city squares. Moreover, nowadays all major cities deployed security cams and constantly monitor crowded environments. Computer vision techniques can be applied for extracting certain behaviors out of these videos. One important concern of the video tracking technique is that making people anonymous requires too much effort. However, for public videos such as the videos on YouTube, the people who appear in the videos already does not have anonymity. Hence, with the advances in computer vision using public videos could be a way for extracting certain statistical characteristics of micro-mobility decisions.

B. VALIDATION METHODS AND METRICS

In this section, we discuss the validation methods and metrics which are used for the human mobility models. Most of the proposed mobility models include validation in a way that the mobility models' output trajectories are compared against other mobility models as well as traces from a real mobility dataset. The models providing better statistical matches are considered as valid models. This validation method provides valuable evaluation for the mobility models, especially for the ones which are scenario-specific models. For the generic mobility models, the validation requires using multiple mobility datasets from various environments. Another method of validation is the visual matching of the mobility patterns. By comparing the overall movement trajectories of different mobility models and real traces, the one with a

TABLE 1. Characteristics of the metrics.

Metric	Mov.-based	Link-based	Netw.-based	Spatial	Temporal	Social
Flight length	✓			✓		
Pause time	✓				✓	
Speed	✓			✓	✓	
Number of visits	✓			✓		
Visit frequency	✓				✓	
Mean-squared distance	✓			✓		
Intercontact time		✓			✓	✓
Recontact rate		✓			✓	✓
Contact duration		✓			✓	✓
Node density		✓		✓		
Node density variance		✓		✓	✓	
Pairwise distance		✓		✓		
Relative velocity		✓		✓	✓	
Message delay			✓		✓	
Data loss ratio			✓			
Transmission count			✓	✓		
Energy consumption			✓			

similar look in terms of diffusion and mobility patterns can be a better match.

We classify the human mobility metrics into three basic categories [30].

- Movement-based metrics
- Link-based metrics
- Network-based metrics

The first category of movement-based metrics is the result of individuals' mobility traces. Flight length, speed, pause time, number of places visited, visit frequency, mean-squared distance (MSD) can be considered in this context. The link-based metrics [30] focus on the effects of mobility to the spatiotemporal relations or similarities of the people. Intercontact time (ICT), recontact rate/count, contact duration, node density, the temporal variance of node density, pairwise distance, relative velocity are examples of the link-based metrics. The network performance-based metrics show the effects of human mobility on the performance of the networks. Various network-related metrics exist and used for comparison of mobility models. Message delay, data loss ratio, transmission count, energy consumption are examples of these metrics.

Another categorization of the metrics is the following.

- Spatial metrics
- Temporal metrics
- Social metrics

In this categorization, metrics such as flight length, speed, number of places visited and MSD can be considered as spatial metrics. Temporal metrics are the ones such as visit frequency, pause time. Social metrics are the metrics which are related to contacts of people. ICT, recontact rate, contact duration, node density can be considered as social metrics.

Let us briefly describe several metrics that are popularly used for validation of most human mobility models. Flight length is one of the most popular mobility metrics used in the literature. It is also referred to as jump size or jump length. It is found by averaging the distances between two consecutive waiting points (visiting places). Mean, variance, and standard deviations of flight lengths are important statistical characteristics such that researchers can observe diffusion characteristics and consistency of the models. MSD is another metric showing diffusion, which is found by computing the total travel distances of people from the beginning of the simulations. ICT is the time gap among two consecutive contacts of a pair. This is a social metric which has a direct impact on applications such as opportunistic communication. Problems such as transmission scheduling and routing highly depend on ICTs. Similarly, contact duration (i.e., contact time) between pairs affects the routing decisions. Pause time (i.e., waiting time or visiting time) is different than the contact duration in the sense that individuals' pause times are independent of other people.

Number of places visited (i.e., number of waiting points) metric is used for various types of human mobility models. For example, it can be used to evaluate if a mobility model represents the daily life of people living in the cities. If the number is too high (e.g., more than 20), the model loses its validity since a person mostly spends time at work, home, and maybe several more places in a day (mostly less than 5 and maximum up to 17 [31]). Frequency of visits shows the possibility and expected time of a person to visit the same place. As an example, for the same daily-life scenario, we can consider a person visits the workplace 2 times a day, one in the morning and one after lunch. For another mobility model which aims to imitate the person's movement in the office, we can consider him visiting his desk or cubicle multiple times in a day, while the same person visits a meeting room only once.

We list the characteristics of the significant metrics in terms of the two classifications in Table 1.

TABLE 2. Human mobility models by main characteristics.

Mobility model	Ref.	Class	Trace-based	Scn.-specific	Group mob.	Obstacle	Vehicle use	Social interact.	Pedestrian dynamics	Role-based	Real maps
SLAW	[26]	Pedestrian	✗	✗	✗	✗	✗	✗	✗	✗	✗
SMOOTH	[32]	Pedestrian	✗	✗	✓	✗	✗	✗	✗	✗	✗
MSLAW	[33]	Pedestrian	✗	✗	✗	✓	✗	✗	✗	✗	✓
TP	[1]	Pedestrian	✗	✓	✗	✗	✗	✓	✗	✗	✗
ParkSim	[34]	Pedestrian	✓	✓	✗	✓	✗	✓	✓	✗	✗
Street Mod.	[35]	Pedestrian	✗	✓	✗	✗	✗	✗	✗	✗	✓
SWIM	[36]	Macro	✗	✗	✗	✗	✓	✓	✗	✗	✗
STEPS	[37]	Macro	✗	✗	✓	✗	✗	✓	✗	✗	✓
WLAN	[38]	Macro	✓	✗	✗	✗	✗	✗	✗	✗	✗
SNT	[39]	Social	✗	✗	✓	✗	✗	✓	✗	✗	✓
HCMM	[40]	Social	✗	✗	✓	✗	✗	✓	✗	✗	✓
GeSoMo	[41]	Social	✗	✗	✗	✓	✗	✓	✗	✗	✗
SIMPS	[42]	Social	✗	✗	✗	✗	✗	✓	✓	✗	✗
HHW	[43]	Social	✗	✗	✗	✗	✗	✓	✗	✗	✗
N-Body	[44]	Social	✗	✗	✓	✗	✗	✓	✓	✗	✗
WHERE	[45]	Vehicle	✓	✗	✗	✗	✗	✗	✗	✗	✗
DP-WH.	[46]	Vehicle	✓	✗	✗	✗	✗	✗	✗	✗	✗
WDMM	[47]	Vehicle	✗	✗	✓	✓	✓	✓	✗	✗	✓
RLMM	[48]	Vehicle	✗	✗	✗	✗	✓	✗	✗	✗	✗
Orbit	[49]	Vehicle	✗	✗	✗	✗	✓	✗	✗	✓	✗
TVC	[50]	Vehicle	✗	✗	✗	✗	✓	✗	✗	✗	✗
ADMM	[51]	Vehicle	✗	✗	✗	✓	✓	✗	✗	✓	✓
MCM	[52]	Disaster	✗	✗	✓	✓	✗	✗	✗	✓	✓
TP-D	[27]	Disaster	✗	✓	✗	✓	✗	✓	✓	✗	✓
DA	[53]	Disaster	✗	✓	✗	✓	✓	✗	✗	✓	✗

IV. A TAXONOMY OF THE MOBILITY MODELS

In this section, we include a taxonomy of the significant human mobility models in the literature. The listed taxonomy is empirical in the sense that instead of creating classes and pushing existing mobility models into those classes, our method is based on studying and understanding the similarities between the model and what aspects make certain models similar or different. The existing models belong to certain fields for which the model is created. For instance, people from the transportation area create models which are fundamentally different than the ones who are in the wireless networking field. The former focuses on people's mobility considering their use of vehicles, whereas the latter considers the communication aspects (e.g., between smartphones). Similarly, some models mainly focus on individual movement decisions of pedestrians, whereas others aim to simulate the social groups such as families. Based on this empirical observation, we categorize the mobility models with the following four main classes.

- **Class 1:** Pedestrian walk models
- **Class 2:** Social network-based models
- **Class 3:** Vehicular models
- **Class 4:** Models for disaster scenarios

The above classification is based on the research areas of pedestrian mobility, social and opportunistic networking, vehicular networking/transportation systems, and disaster management. These areas have been studied in recent years by various groups in academia and industry.

While some models may contain certain aspects from multiple classes, we focus on the main characteristics of the models for the classification. For instance, a mobility model that simulates the bus routes in a metropolitan area is classified as a vehicular model, even the model may also contain movements of pedestrians from their houses to bus stops. As another example, some disaster models simulate pedestrian walk behavior; however, their main application is planning and management for disasters rather than understanding the nature of the walking behavior. This makes the minor aspects of such models simplistic (e.g., assuming pedestrian walk by RWP) and based on the use of off-the-shelf models, while most effort and complexity in these studies are found in their prime aspects. Therefore, we believe that this classification provides not only an empirical but also a logical representation of the human mobility models. While the initial classification is based on the targeted application scenarios, the complete feature sets can still be found in Table 2.

This article includes a section for each of these four classes, where we describe the mobility models that fall into these categories. Before going into the details of the models, we present an overview of possible classifications in Table 2.

TABLE 3. The listed characteristics and their contraries.

Characteristic	Contrary
Trace-based	Synthetic
Scenario-specific	Generic
Group mobility	Individual mobility
Obstacles	Clear area
Vehicle use	Walk models
Social interaction	Independent people
Pedestrian dynamics	Lacking micro mobility
Role-based	Homogeneous users
Real maps	Synthetic maps

The table is sorted according to the main classes that the models fall into.

Although researchers in various fields need to implement human mobility models or use the models' output traces in their simulations, they are not supposed to have expert knowledge in the area of mobility modeling. Table 2 serves as a starting point for these people to decide if a mobility model is a good match with the expectations before incorporating the mobility model in their simulations. For instance, if the simulation is for a specific scenario such as simulating the mobility of theme park visitors, readers should focus on the scenario-specific models. As another example, if the simulation needs the mobility of people in urban environments, the mobility models involving vehicle use and social interactions may be good matches.

While Table 2 contains nine main characteristics, these characteristics can be considered with their contraries, where the models with ✗ have the characteristic feature. The characteristics and their contraries are listed in Table 3.

Trace-based models are the models which heavily exploit the mobility traces collected from people for calibration, such that for modeling mobility of people in an environment actual traces should be fed into the model before simulating their movement. On the other hand, synthetic models can be run independent from the traces. Some of the synthetic mobility models are too simplistic having abstract parameters, However, for more realistic results, parameters of some synthetic models may still be calibrated based on the real-life mobility traces. Scenario-specific models have the aim for better representation of mobility considering the characteristics and constraints of different environments. Generic mobility models, on the other hand, aim to be useful in wider application scenarios. The models with group mobility, as opposed to individual mobility models, involve the behavior where people are assigned to groups and group members may travel from a place to another together. Some models consider the movement behavior of people in the presence of obstacles. On the other hand, some models assume that the environment is a clear area without obstacles or any type of restrictions for movement.

Models with vehicle use are mostly either simulate movement in urban environments or travel between distant places such as different cities. Walk models, on the other hand, aim to simulate the pedestrian movement in various environments such as urban environments or specific areas where vehicle use is limited (e.g., airports). Some human mobility models consider social interactions having a major role in mobility decisions of people. For example, family members or close friends would like to meet in common places such as a home or a restaurant and they move to these places at specific times of the day. Other models assume that individuals are independent as they decide to visit places. Pedestrian dynamics adds sophistication to the mobility models as some models try to model various dynamics such as the speed or direction changes due to crowds or pedestrian traffic. Other models lacking micro mobility neglect these changes as they mostly focus on the overall movement in a large area. User roles are introduced in role-based models where each person has different mobility behavior or responsibility. Mobility of a group with a leader who decides the places to visit as a group can be considered as an example of the role-based models. On the other hand, most models assume that everyone is equal (homogeneous). Lastly, some mobility models are based on the use of real maps. For instance, the map of a particular city can be used for realistic modeling in times of earthquakes. Since the use of real-maps brings more sophistication to the mobility modeling, most models are based on synthetic maps, which can be either a 2D clear area or a map consisting of synthetically generated obstacles (e.g., polygons).

Figure 2 shows an example scenario of using real maps in the mobility models. For a walking model of university students, we export the University of Central Florida (UCF) campus map using OpenStreetMap (left) and later the OSM data is processed for extracting the pedestrian ways. The processed model of the map is shown on the right figure, where pedestrian ways are shown by lines connecting the user-tag waypoints and the dots represent the waypoints.

In the following four sections, we will have an in-depth look into the mobility models that are listed in Table 2.

V. PEDESTRIAN WALK MODELS

In this section, we describe the human walk models which focus on the macro-mobility decisions of pedestrians such as deciding the next destination to visit, but not try to imitate behaviors causing dynamic variations in the speed or direction due to reasons such as pedestrian traffic. We call these type of models pedestrian walk models. These models are mostly used in network simulations as they provide statistical matches to the real movement traces of people in metrics such as flight length or ICT, which affect the coverage and performance of the networks with human participants.

Self-Similar Least-Action Walk (SLAW) [26], [54] is a generic human walk model which produces mobility traces having certain statistical features of human mobility. These five features are flight lengths and pause times with heavy-tail distribution, heterogeneously bounded movement regions, truncated power-law ICTs, self-similarity, and least-action trip planning. SLAW tries to emulate individual mobility and each individual's movement is independent of the other people, assuming that they do not have any social interactions. SLAW model uses fractal waypoints which are generated

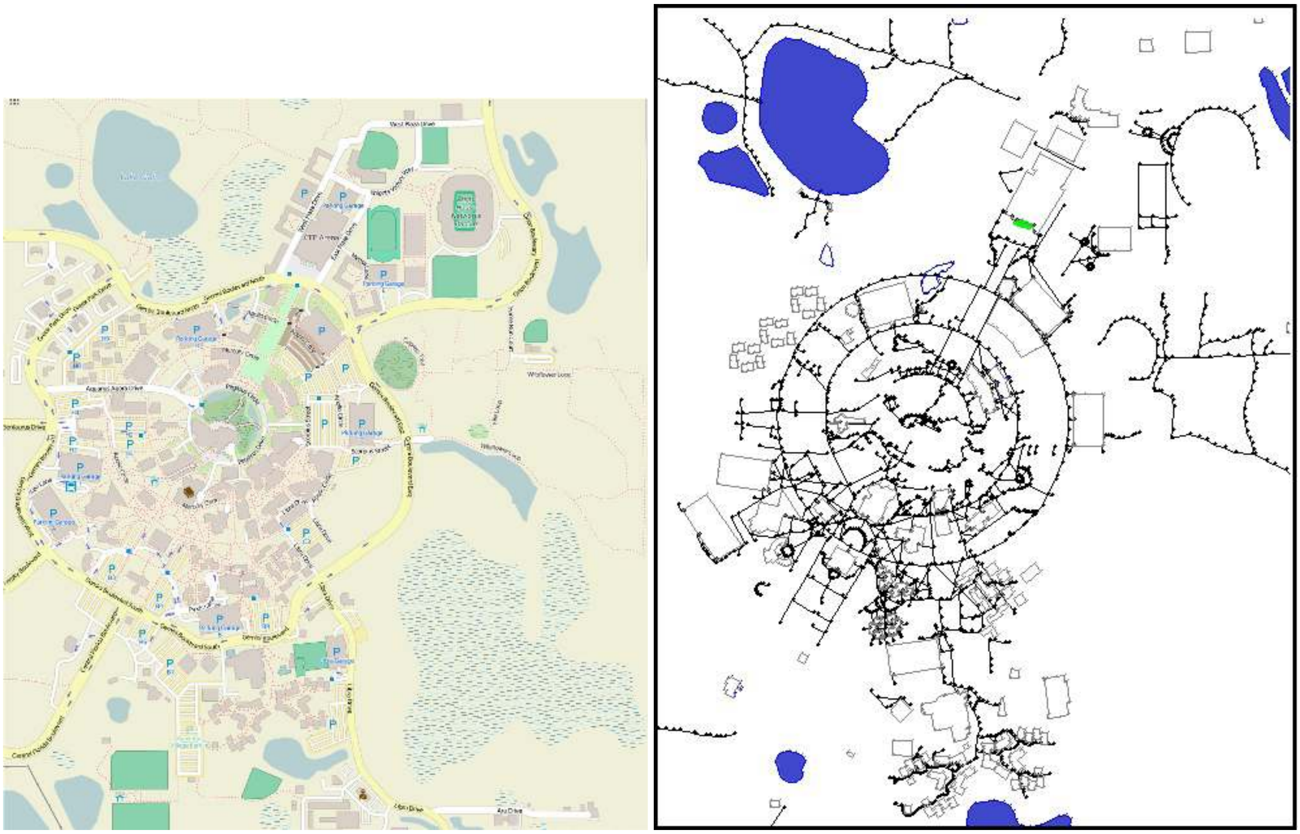

FIGURE 2. The map of the UCF campus. Left: The map extracted from OSM, right: The map of the processed model which involves the pedestrian ways.

over a 2D plane. As the gaps between the waypoints determine the flight lengths, Hurst parameter (H) of self-similar points controls the gap distribution characteristics. The waypoints are clustered and each individual randomly chooses a cluster as the movement region for a daily trip. Each individual starts their walk by randomly choosing a waypoint as the starting point and later makes movement decisions based on the least-action trip planning (LATP) algorithm. In LATP, a person decides the next destination by selecting a waypoint in the cluster and straightly moves to that destination. The equation in the LATP algorithm for assigning probability values for deciding the next destination waypoints are given as follows.

$$P(w) = \frac{d(v, w)^{-\alpha}}{\sum_{w_i \in W-W'} d(v, w_i)^{-\alpha}}, \tag{3}$$

where $P(w)$ is the probability of visiting the next destination waypoint w ($w \in W - W'$), v is the waypoint on top of which the person is currently waiting, and $d(v, w)$ represents the Euclidean distance among the two waypoints. W is set as the waypoints which are planned to be visited and W' is the set of already visited waypoints. The value of the parameter α is set according to the GPS traces collected from the specific outdoor environment for providing a statistical match. For instance, $\alpha = 3$ produces the difference less than only 2% between the synthetic mobility traces and the GPS traces for Disney World theme parks. LATP is shown to produce mobility traces that match the real GPS traces very well for the range $1 \leq \alpha \leq 3$.

Mostly, the person chooses the waypoint which is closer to the current waypoint where the person stands. However, there is some randomness such that each waypoint has a chance to be selected. Multiple individuals can share the same waypoints during their trips. Moreover, as they are independent individuals, their movement choices do not depend on each other's movement. SLAW model outputs have power-law ICTs and flight lengths with the heavy-tail distribution. Movement traces which are obtained from the SLAW simulation is compared with 226 daily human mobility traces with 12h average duration. The GPS traces are obtained from participants in 5 different environments: 2 university campuses, a state fair, and a theme park (Disney World). The model is also used for evaluation of various routing protocols such as PROPHET [55] and random forwarding [56] for delay-tolerant networks (DTNs).

The SLAW mobility model is used by various studies in the fields of networking and human mobility modeling. As we are interested in the human mobility models, we describe three human mobility models which used SLAW model as the baseline: SMOOTH, MSLAW, and TP.

SMOOTH [32] is proposed as a realistic and simple to implement human mobility model. SMOOTH aims to provide 7 features of human mobility: 1) truncated power-law distribution (TPL) of flight lengths, 2) TPL of ICTs, 3) TPL of pause times, 4) human behavior of choosing popular places to visit, 5) visiting the closer places first (least-action principle), 6) non-uniform distribution of people, 7) heterogeneous division of regions of mobility for different people (moving around communities). In the model, clusters of waypoints are formed in a way that each cluster represents a community (place of movement). Clusters have unequal sizes and the cluster sizes represent the popularity of the places. For instance, on a university campus, food courts are more popular in terms of the number of people visit every day, compared to places such as a specific department's building. People move in the region in groups and the region is defined as a cluster of waypoints.

For the individual's perspective, each person chooses a community according to the corresponding cluster sizes. Then, the person selects a subset of the waypoints in the cluster to visit. These two steps can be considered as the pre-planning phases. Later, the person visits the selected waypoints via LATP. The person's pause time on a waypoint is determined randomly by the power-law distribution. As opposed to SLAW having parameters such as the Hurst parameter, SMOOTH has simpler inputs so that people having no previous knowledge about the mobility models can still be able to use the model in their simulations. These inputs are listed as size of the area, number of people, number of waypoints, mobile nodes' transmission range, number of clusters, and minimum and maximum pause times. The model also expects inputs for alpha and beta parameters which are used by the LATP algorithm for choosing the next destination points and setting the pause times respectively. Furthermore, SMOOTH can imitate SLAW. Specified values for maximum sizes of a group and maximum distance of waypoints from each other correspond to ranges of Hurst parameter values which are listed in a table [32]. The SMOOTH model is validated using the GPS traces collected from the 5 aforementioned outdoor environments and in comparison to the SLAW model with metrics such as complementary cumulative distribution functions (CCDFs) of flight lengths and ICTs and average message delays. For network simulations, SMOOTH is a good alternative human mobility model as it is easy to implement and much more realistic compared to other commonly used mobility models such as RWP, which are already proven to be unrealistic [57].

The SLAW mobility model assumes no obstacle for the human movement, such that after deciding the next destination, a person can straightly move to the next waypoint without any disturbance, so the speed and the direction do not change. Map-based SLAW (MSLAW) [33] mobility model introduces geographical restrictions to the SLAW model. The algorithm of MSLAW overall follows similar steps (e.g., fractal points generation) of the SLAW's algorithm, but these steps are modified to include the map-based geographic restrictions. For instance, waypoints are not created on top of the rivers, forests, or other inaccessible areas. MSLAW has a modified version of the LATP algorithm for deciding the next destinations. In SLAW, the distance $d(v, w)$ between two waypoints v and m was given as the Euclidean distance, while in the modified version, the distance is given according to the optimal route lengths. Considering the optimal route $(v_i^*)_{1\leq i\leq j}, j \leq 2$ from v to w,, such that $v = v_1^*$ and $w = v_j^*$, the total distance of the optimal route is given as

$$d(v, w) = \sum_{i=1}^{j-1} \left\| v_i^* - v_{i+1}^* \right\| . \quad (4)$$

In this equation, v and w are the actual waypoints which were generated in the initial fractal point generation phase. However, the points in between these two points $(v_2^*, v_3^*, \ldots v_{j-1}^*)$ may or may not be the waypoints, as they are only chosen as part of the optimal route for avoiding obstacles.

Set of parameters used in the MSLAW mobility model include the Hurst parameter as well as the LATP distance weight parameter α, along with common mobility parameters such as the number of people, simulation time, and the size of the simulation area. Implementing MSLAW requires similar steps of the SLAW implementations, but it is relatively harder to implement the simulation of this model compared to the SMOOTH model. MSLAW is not validated by the simulations in comparison to real-life GPS traces. However, the model is evaluated with several metrics such as ICTs, contact durations, and recontact rates against RWP, SLAW, and a map-based random mobility model RaST [58]. The mobility traces generated by the MSLAW model are more convenient than the traces of the SLAW model for scenarios such as human walks in urban environments.

Theme park (TP) [1], [59] mobility model is a scenario-specific human walk model to model the large and crowded theme park environments. Human mobility in theme parks has different characteristics compared to daily mobility. First of all, in daily life, people use vehicles for transportation, while spending a day in a theme park they mostly walk. Moreover, spatial regularity in human movements does not exist in the theme parks, because people mostly prefer to go to unseen locations (attractions) instead of going to an already visited attraction.

Modeling the theme park environment consists of five phases. The first phase is the generation of fractal points. In the second phase, the fractal points are clustered for finding the densest areas with a modified version density-based algorithm (DBScan) [60], which allows inputs such as the number of clusters and the percentage of non-clustered points. In the third phase, attractions are modeled by using the clusters and the non-clustered points are marked as noise points. Inclusion of the visitor model forms the landmark and considered as the fourth phase. Lastly, multiple landmarks are connected by a graph with non-directed edges that represent the roads. Landmarks considered separate to isolate the areas where people only walk, while on top of the roads they are assumed to travel using their vehicles or transportation services provided by the theme park operators such as trains or buses. A queueing-theoretical approach is used for modeling the popular places (attractions) according to their types. For instance, M/D/n queues are used for modeling the service and the waiting lines of roller-coasters and M/M/1 queues are used for modeling the service of restaurants. While each visitor makes decisions independent from each other, their mobility is affected by the environment, such that their waiting times in attractions depend on the service rates and the people already waiting in the lines. For modeling the mobility decisions, a modified version of LATP algorithm is used in a way that attractions have more probability to be selected then the non-clustered noise points. In the algorithm, the probability of a noise point n to be selected as the next destination is given by

$$P(n) = \frac{d(v,n)^{-\alpha}}{\sum_{n_i \in N-N'} d(v,n_i)^{-\alpha} + \sum_{a_j \in A-A'} w_j \cdot d(v,a_j)^{-\alpha}}, \quad (5)$$

where N is the set of all noise points and N' is the set of already visited noise points, A is the set of all attractions and A' is the set of already visited attractions. w is the weight value of an attraction and it is equal to the number of fractal points clustered in the corresponding cluster, which was initially used for modeling the attraction. Similarly, the probability of selecting an attraction a is given by the following equation.

$$P(a) = \frac{w_a \cdot d(v,a)^{-\alpha}}{\sum_{n_i \in N-N'} d(v,n_i)^{-\alpha} + \sum_{a_j \in A-A'} w_j \cdot d(v,a_j)^{-\alpha}}, \quad (6)$$

The only difference of the attraction a with a noise point n is that it has a weight value w_a which multiplies the probability of being selected as the next destination. As a result, attractions in the model are popular places where people mostly prefer to visit. While TP mobility model is proposed for modeling the theme park environments, it is possible to extend it with minor modifications for crowded areas with limited vehicle use such as airports, shopping malls, and so on.

An activity-driven scenario-specific human mobility model is developed as the ParkSim simulator. This mobility model [34] is also used for simulating the movement of theme park visitors. They represent the area by a theme park layout extracted using OpenStreetMap [28]. OSM maps are parsed by their simulator such that the map is divided into walking and activity areas. The visitor is assumed to move in the walking areas which are walkways or plazas. On the other hand, people spend time in activity areas such as indoor or outdoor attractions, restaurants, or live outdoor performances. The mobility model considers both macro and micro-mobility of the pedestrians. Walking among activity areas, spending time in these areas are waiting in the lines (queueing) are all considered as part of the macro mobility. Considering a person as an agent in the simulation, walking, spending time, and queueing are considered as three different states of the agent. Activity areas have different types such as events or restaurants with different priorities and they cause different mobility behaviors of the pedestrians. For instance, visiting an event means selecting a random spot in the event area and waiting there. Priorities are used for defining the popularity of the places.

In the mobility model implemented on the ParkSim simulator, micro-mobility is defined as the mobility between two activity areas. Micro mobility consists of routing decisions, collision avoidance, queueing, and inter-area mobility. For routing from one area to another, Dijkstra's shortest path algorithm is used. Queues are defined by width and length such that width and length are used to calculate the number of cells where each cell is assumed to have one person. Inter-area mobility is defined as the mobility of people inside an activity area. RWP and random sit-point models are used for simulating the inter-area mobility. 670 GPS traces collected from Disney theme park visitors are used for calibrating the mobility model's parameters.

Street Model [35] is an analytical model, based on queueing theory, for representing the pedestrian mobility in city environments. It considers two-way street segments where nodes arrive from both endpoints of the segments by a Poisson distribution. The walking speed of people is decided randomly by a probability density function. The main purpose of the Street Model is to simulate content distribution in urban areas based on network connectivity, contact rate, and contact duration for network performance and connectivity. Stockholm's downtown area is simulated using the real map of the area to create artificially the street segments. This mobility model is a purely mathematical and synthetic such that real GPS traces are not used for either calibration or validation of the model.

A simple-to-implement and easy-to-tune mobility model, small world in motion (SWIM) [36], is based on the intuitive movement choices of people according to the popularity and the proximity of the places. In this model, each person is assigned to a place called the person's home considered as the base and the proximities of places are considered by the base. For instance, if there exist two popular restaurants and one of them is closer to home, the closer one is mostly chosen. The locations of the homes are created by uniform random distribution on an empty two-dimensional area. The places such as restaurants, schools, and offices have weight values representing their popularity. As opposed to most human mobility models, individuals are not assumed to have global knowledge of the map. However, they are assumed to know the popularity of the places for the last time they visited. Thus, weights of the places are relative and it is given for a person i as

$$w(p) = \delta \cdot d(H_i, p) + (1-\delta) \cdot n(i,p), \quad (7)$$

where p is the place with a chance of being visited by its weight w_p, H_i is the location of the home of the visitor i, and $n(i,p)$ is the number of people in the place p when the person i visited last time. δ is a constant in the range [0, 1] which adjusts the importance of the proximity $d(H_i, p)$ or the popularity $n(i,p)$. Another difference in the SWIM model compared to models such as SLAW, MSLAW, and SMOOTH is that the speed of a person is not fixed. It is assumed that the person's speed is inversely proportional to the distance $d(c, p)$, where c is the current location of the person and p is the destination. This is a simplistic assumption, while it has some truth such that the vehicle use is possible for longer distances. However, for instance, the vehicle's speed should not fluctuate by the distance.

Spatio-TEmporal Parametric Stepping (STEPS) [37] is a human mobility model aims to satisfy the statistical properties such as TPL of flight lengths and pause times. The model is based on two principles, which are preferential attachment and attraction. Preferential attachment is having less probability of visiting distant locations (the inverse proportion between distance and probability) and attraction is the principle that having a higher probability to visit closer places to the preferred location instead of moving farther. A Markovian model is used having states to represent locations such as home, office, shop, and others. The model is represented with a graph model where nodes represent the states and the edges represent the transition mobility pattern. The pause times in the locations and the transitions between them are drawn by the power-law distribution. The simulation area is divided into squared grids with the assumption that each grid (zone) represent a place such as a building or a school. Each node is assigned to a preferential zone.

The last pedestrian walk model we discuss is based on the wireless user traces from Dartmouth College with 13 months duration. This model is called the WLAN [38] mobility model. The mobility of people from an access point (AP) to another is used for finding the paths people follow. The GPS traces are collected between APs and all paths are estimated by a Kalman filter for approximating the GPS data. In the approximated user tracks, pause times fit a log-normal distribution such that most people wait a short period whereas fewer people wait for longer times. Moreover, speeds of the people also fit a log-normal distribution, having an average of 4.5km/h.

In the WLAN model, the ratio between the mobile and stationary users are used for the movement decision of each person. Each person is selected as a mobile or a stationary user according to this ratio. For a mobile user, start time and initial region are selected by the start time and start region distributions. Then, a regional transition matrix consisting of transition probabilities is used for deciding the next region (destination). The waypoints are uniformly distributed and a Gaussian distribution is used for choosing the number of waypoints for a particular movement. The speed of a mobile user between waypoints is also selected according to the overall speed distribution. After a mobile user reaches the next destination, the person waits for a pause time for that particular region and chooses the new next destination. While the model uses real user traces from a campus environment for extracting the mobility model, mobility characteristics may differ from one place to another. Hence, the model should be tested in other scenarios.

Although all pedestrian walk models propose representations of pedestrian routes on a geographical scale, there are clear variances between these mobility models in terms of their complexity, granularity or accuracy, validation, and applicability. Models such as SLAW, SMOOTH, MSLAW, SWIM, and Street Model are easy to implement but their granularity is not sufficient for specific scenarios such as theme parks or city squares. On the other hand, TP and ParkSim provide more accurate representations of people's movement in the theme parks while they require certain efforts for implementation. The granularity difference is significant. For instance, compared to WLAN model, which is based on people's mobility from one AP to another, models such as ParkSim provide much better precision. Moreover, pedestrian dynamics (e.g., the effect of the crowd) is not incorporated in SLAW, MSLAW, and WLAN. For instance, in ParkSim model, movements are affected by street segments (e.g., width), people in the streets, and intersections, while in SLAW, crowd movement density and pedestrian flows do not affect the individuals. Another clear difference is the validation and calibration based on GPS traces. Street Model is simply lacking any validation, while most other models (e.g., SLAW, TP, ParkSim) are calibrated for their scenarios using real traces. SLAW and the models that are inspired by SLAW (MSLAW, SMOOTH, TP) apply to areas where neither vehicles nor obstacles exist. On the other hand, ParkSim and Street Model apply to areas where clear street segments or pedestrian roads can co-exist. Furthermore, SWIM, TP, and SLAW are also applicable for environments where the popularity of places directly affects mobility (e.g., hot-spots).

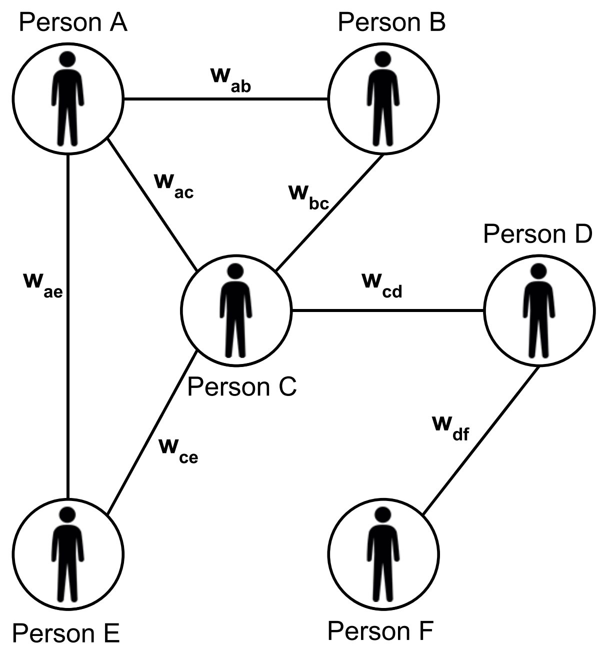

FIGURE 3. An illustration of the graph models representing social networks.

VI. SOCIAL NETWORK-BASED MODELS

Social networks are used as bases for various human mobility models. In this article, we call such mobility models as the social network-based models. Some of the models are built on top of the findings of social sciences. The other models are simply based on the intuitive expectations from social interactions. According to the social network-based models, mobility decisions of people highly depend on their ego-centric friendship networks. For instance, a person mostly chooses to meet with his or her close friends or family over other people. Most models define a social network which is modeled with a graph theoretical model. An example illustration of the graph models is shown in Figure 3. In this figure, 6 people ($Person\ A$, $Person\ B$, $\cdots$, $Person\ F$) form a social network. An edge between two people represents friendship and each friendship between pairs has a weight value (e.g.,w_{ab}, w_{ce}). In the simulation of these models, movement decision probabilities are based on these weight values. Assuming that the friendship between two people has the same value for each other's perspective, mostly weighted unidirectional graphs are preferred to represent the social networks. Moreover, mostly dynamic graph models, where the weight values and edges are updated in discrete time intervals, are preferred because friendship values change by time.

Let us start describing significant mobility models which use social relationships for determining human mobility. In [39], social network theory is used as a baseline for the design of a human mobility model, which we call the Social Network Theoretical (SNT) mobility model. Design of the model consists of four phases: 1) modeling social relationships, 2) detecting community (group) structures, 3) placing the communities on the simulation area, 4) modeling the mobility dynamics.

In the first phase, the social relationships are modeled using a weighted nondirected graph, such that each node represents a person and each edge represents the social interaction between two people. The graph is represented by a two dimensional interaction matrix M, where each element $m_{ij} \in [0, 1]$ represents a social interaction between person i and j. M is symmetric since the edges are considered as nondirected. The matrix M can be created using real relationship data (e.g., phone calls, text messages collected from smartphones) or synthetically created for the simulation purpose. A boolean connectivity matrix Q with the same size is created in way that if $m_{ij} > \Theta$, then $q_{ij} = 1$, otherwise $q_{ij} = 0$, where Θ is a constant threshold value. In other words, the connectivity matrix marks the strong social interactions between pairs. Q is later used for detecting the community structures in the social network with an algorithm based on the centrality of nodes.

In the second phase, simulation area is considered to be divided into grids and each community C is randomly placed on top of a grid G_{xy}, where x and y represents the index of the grid based on rows and columns. Each grid is considered as a different place. In the third phase, people's mobility choices are determined. Considering the community C_{xy} placed on the grid G_{xy}, initially each person in the community randomly chooses a location in the associated grid as the next destination. After a person i moves to the first destination, he or she can go to another place independently from other members of the community. The selection of the new place is based on the social attraction A of the person i to the different places, which is given by

$$A(i, G_{xy}) = \frac{\sum_{j \in P_{xy}} m_{ij}}{\|P_{xy}\|}, \tag{8}$$

where P_{xy} represents the set of people which are located in the grid G_{xy} at the time of the decision and $\|P_{xy}\|$ is the number of people in the grid at that time. The next destination is set as a random point in the selected G_{xy}. The decision of the next destination is made according to the probability that is given as follows

$$P(G_i) = \frac{A(p, G_i) + \eta}{\sum_{j=1}^{x*y}(G_j + \eta)}, \tag{9}$$

where η is a random value and $\eta > 1$ which is used to ensure each grid has a probability of being selected as the next destination, such that $P(G_i) > 0, \forall i$. The social network is reconfigured after a period, which is given as input to the mobility model so that new communities can be formed. While the new communities are not distributed to the simulation area again as in the initial phase, people make their mobility decisions for the next destinations according to the new social relationships after the reconfiguration.

Human Cell Mobility Model (HCMM) [40] is an extension of the previously discussed SNT model. The HCMM model includes spatial (location) attractions to the already existing social (community) attractions of the SNT model. While people are still organized into the social communities and the area is structured with cells as in the SNT model, in HCMM each person also has one cell which is pre-defined as the home cell. As a person belongs to a social community, the person's home cell not only affects the movement choices of the person but also affects the movement choices of the other members in the same community. To model this behavior, the social attraction $A(i, G_{xy})$, which defines the probability of the person i for a particular cell (grid) G_{xy} is given as

$$A(i, G_{xy}) = \frac{\sum_{j \in H_{xy}} m_{ij}}{\|H_{xy}\|}, \tag{10}$$

where H_{xy} is the set of people whose home cell is G_{xy} and $\|H_{xy}\|$ is the number of people satisfying this condition. Additionally, for each movement decision each person has a probability $P(G_{external})$ of staying in an external cell and $1 - P(G_{external})$ probability for choosing home as the next destination. After the next destination cell is decided, the person chooses a point in the particular cell by uniform random distribution and moves to that new destination point.

SNT and HCMM are not the only models that are based on social network models for generating human mobility traces. The general social mobility model (GeSoMo) [41] uses social network models (SNMs) as input for generating the social mobility model. The GeSoMo model takes spatial and temporal regularities of human mobility and the group movements into account. In the model, people are represented as the nodes and places (e.g., restaurant) are represented as the anchors. The anchors are the places where the nodes are attracted and meet. After meeting and spending time in an anchor, the nodes move to their next anchor. This movement may be done as a group in the case of the nodes having strong social ties. This makes GeSoMo different than most human mobility models since the other models focus on individual movement decisions.

Let $G = (V, E)$ be the graph that represents the social network, where V is the set of the nodes and E is the set of social relationships between pairs. The graph is dynamic in the sense that each edge e_{ab} between the nodes $a, b \in V$ is active for some period of time $\tau_{ab} \subseteq T$, where T is the total simulation time. The GeSoMo model has the conformance criterion between the input social network G and the meetings M of the resulting mobility model.

With the help of a parameter, the model adjusts itself to inconsistencies with the social network. The movement decisions of the people depend on the attraction and repulsion, such that nodes tend to visit the places with more attractive and less repulsive forces. The model also includes the visit frequency as a factor for providing temporal regularity. Moreover, the overall attraction of the group changes the probability of a person to choose a specific anchor, which creates the group movement behavior.

Attractive and repulsive social forces are used by various social network-based human mobility models as well as micro-mobility models such as SFM. Most of the time, a comfort range which is defined as a circular area for each person, such that the people in the comfort range cause attractive and/or repulsive forces. Figure 4 illustrates the basics of this popularly used attraction&repulsion concept. In this figure, the person is located in the center of its comfort zone and 3 people are inside the comfort range. 2 strangers (shown by grey human figures) cause repulsive forces $\vec{F}_R$ and the only friend (shown by green, hands-on human figure) causes an attractive force $\vec{F}_A$. The other people outside the comfort zone do not have any impact on the person's mobility. In such a case, the person is expected to move along the direction of $\vec{F}_A$. Another example, for the micro-mobility modeling perspective, the forces causes a change in the acceleration so that if the person is currently moving toward the strangers, the velocity of the person decreases and finally the person stops or starts going toward the opposite direction.

Sociological Interaction Mobility for Population Simulation (SIMPS) [42] focuses on the causes which govern human mobility at its roots by the leveraging sociological findings of people's social interactions. The two main sociological findings are listed as follows.

1) A person has a social interaction need which is quantifiable with a level depending on the personal status such as age or social class.
2) People make acquaintances for fulfilling their social interaction needs.

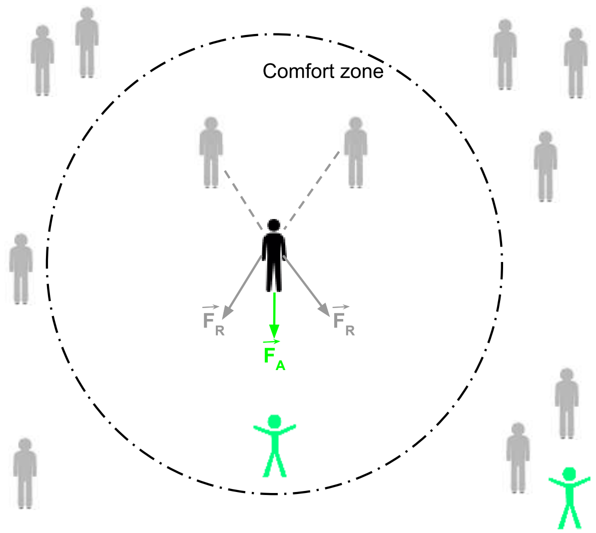

FIGURE 4. Comfort zone of a person, an attractive force by a friend and repulsive forces by two strangers.

The SIMPS model classifies social interactions basically into two as the interactions with acquaintances and non-acquaintances and translates them into a behavioral model for people's movement decisions. Each person has a specific sociability level and they have two simple complementary behaviors which are socializing and isolating. Socializing is the behavioral movement towards the acquaintances and isolating is the behavior of escaping non-acquaintances. The social ego network of acquaintances is represented by a graph with directed edges. Each social encounter is used as feedback for comparing the person's individual needs and the current socialization volume. Moreover, each person i is assumed to have a social comfort range c_i given as

$$c_i = [s_i(1 - t_i)), s_i(1 + t_i))], \tag{11}$$

where $t_i \in [0, 1]$ is the tolerance level and s_i is the sociability level of the person. The individual i tends to socialize at this comfort range. Moreover, based on Hall's communication theory of proxemics [61], each person has a social distance (a circle with approximately 3.5m radius (R)) and the person's surround in this area changes the perception. As the person progressively notices the others' presence in the circle, the perceived surround ps_i of the person i is updated by the below formula.

$$ps_i = \frac{\|P_i(t)\| + ps_i}{2}, \tag{12}$$

where $P_i(t)$ is the set of people present in the circle having the condition $d(i, j) \leq R$, $\forall j \in P_i(t)$ ($d(i, j)$ is the distance between two individuals). ps_i serves as part of the feedback for the movement decision process. Social motion influence depends on the sum of the attractive forces by the acquaintances and the repulsive forces by non-acquaintances (strangers). Lastly, motion execution is achieved by changes in velocity and acceleration. The model can represent micro-dynamics in the movement such as the velocity change.

However, for simplicity purpose, the model does not try to imitate behaviors such as collision avoidance.

Another model which is based on the social network theory is called the Heterogeneous Human Walk (HHW) [43] model. In its initial phase, the HHW model constructs a k-clique graph which represents overlapping communities. This graph model can be used for satisfying common statistical features of real social networks such as having dense connections between people in a particular community, having sparser connections between people in distinct communities, and power-law degree distribution of the nodes (people). Overlapping communities is considered such that one person may belong to one or multiple social communities such as a classroom, hobby group or a research team. Moreover, the social communities that the person belongs to change during the day. For instance, during night time at home, the person belongs to the family, while in the morning at the laboratory the same person belongs to the laboratory group. Moreover, each person has a local degree for each community, shown as L_i^c for the person i in the local community c. The local degree L_i^c shows the popularity of the person in the community. For a k-clique community, $L_i^c \geq (k-1), \forall i$. The local degrees are heterogeneous such that the values $L_i^c \geq (k-1)$ follow a power-law distribution.

After the construction of the k-clique graph, the HHW model assumes a 2D plane divided into grids. Each community has its zone having associated cells for the community. Each individual is associated with a set of cells in the community's zone and initially placed randomly on top of these associated zones. Later, the person moves toward a random point in the same associated zone. The pause time in the random point follows a power-law distribution. This process repeats itself while community structure and the associated zones of the people dynamically change. The mobility decisions of individuals also adapt to these changes. Individuals with higher local degrees (popular people) have more cells associated with them. Therefore they have more chance to meet people from other communities. This heterogeneity between the individuals serves as the baseline of the HHW model.

N-Body [44] is another social network-based model to represent wider application scenarios as a generic human mobility model. In the model, the social relations of N people are calculated according to their pairwise closeness and randomness over time. The friendship and the steadiness are defined and used by the nodes to decide to stay away from each other or form groups. Initially, people choose random destination points as in RWP, but their movement depends on the forces

$$\vec{G}_i = \vec{A}_i + \sum_{j=0}^{N} \vec{F}_{ij}, \tag{13}$$

where $\vec{F}_{ij}$ is the force from node j to node i. $\vec{A}_i$ represent the force for the attraction by the destination point. The person i adapts its velocity according to the sum of these attraction and inter-nodal forces. When two or more nodes have friendship and come close to each other, it causes a group forming tendency and the inter-nodal forces make them pull each other and form a clog. This brings the group mobility behavior to the N-Body model. After starting to move as a group, their destinations compete against each other and only one of them wins, so that the group moves mostly to the closest destination first, spends time together, and then moves to the next destination. During the movement of the nodes, a small movement that causes a change in the sum of the forces may cause the group to be separated. However, in the case that the forces among the group members are stronger, the group can even wait together in a destination. As can be understood, the speeds of the people dynamically change between 0 and the maximum speed, as people adjust their velocity due to the forces.

VII. HUMAN MOBILITY MODELS WITH VEHICLE USE

Work and Home Extracted REgions (WHERE) [45] is a human mobility model which is developed considering metropolitan areas such as the mobility of people in Los Angeles or New York. The baseline idea is that people spend most of their time at home or workplace. Their initial model is a two-place model called WHERE2, in which a person only moves from home location to work which has a particular distance from home. Home locations, workplaces, distances between them, and the time spent in these places are inferred from a user database of CDR traces and public sources, which can be considered as spatiotemporal samples. Extension to the initial WHERE2 model is achieved in a way that a place such as a restaurant can be added to make it a three-place model WHERE3, and the model can be increasingly more complex as new places added as extensions to WHERE4, WHERE5, and so on. Compared to most human mobility models, this model is simplistic. On the other hand, it aims to solve the population movement problem in a large metropolitan scale which includes many factors such as using transportation services. The model's accuracy is based on drawing information of work and home places and calibration by the help of the CDR traces. The model is used for creating synthetic CDR traces based on the spatial and temporal distribution of the people. DP-WHERE [46], where DP stands for differential privacy, adds controlled noises to the set of empirical probability distributions which are used by the previous WHERE model. The aim of DP-WHERE is adding privacy to the model without loss of accuracy and make it impossible to infer sensitive information of people so that the data with noisy distributions can be released to the public without creating privacy concerns.

Working Day Movement Model [47], which we shortly call WDMM in this article, is a map-based model and it represents the human mobility in the daily life of people. According to the model, people wake up in the morning, spend time at work during the day and come back to their home in the evenings. The model is considered for use in the simulations of delay tolerant networks (DTNs). The overall goal of the model is to capture major characteristics of people's movements while neglecting some micro movement decisions such as going to someplace from work for a short time or movement between rooms in the home. The model is respected as a combination of 4 activities:

- Home activity
- Office activity
- Evening activity
- Transportation

The home activity model is considered for the times, where a person is assumed to have a home location on the map and he or she statically wait there before waking up for the next day. On the other hand, the office activity model is relatively sophisticated and the movement in the office is modeled as follows. The person enters the office from the door, which is a specific point on the map. After entering, the person moves towards the desk with a walking speed (defined as a parameter value). After spending time near the desk, the person chooses a random point in the office and moves there. After spending time on the random point, the person moves back to the desk and the movement continues as described during the day between the desk and random points. The pause times near the desk and the random points follow a Pareto distribution. This model ensures that people having desks closer to each other meet more frequently since they spend approximately half of the day working in their desks.

The evening activity model involves the group mobility behavior, such that each person has a favorite meeting spot such as a restaurant or a shopping mall and the person is assigned to a group based on the favorite meeting spot. The assignment of the meeting spots, groups, and creation of new meeting locations are uniformly random. The person moves to the meeting spot by the transportation model and waits there until all the group members meet. When all members are present, they start walking as a group by a map-based walk model, defined as a random walk on top of the streets. After traveling a distance as a group, they spend a long time and each individual walks back home. The distance they travel and the time they spend is defined by parameters.

WDMM has a heterogeneous transportation model, which consists of three submodels: walking, car, and bus submodels. Walking submodel uses Dijkstra's shortest path algorithm for route and people have a constant speed. Some of the people have cars and they can travel faster using the same shortest path. A person without a car is assumed to know the bus route. The person can use one of the busses having the same route. Buses can carry multiple people. The distances in between home to the closest bus stop or bus stop to the office are traveled by walk. Map of the urban environments such as Helsinki is used by the WDMM model.

The real-life mobility model (RLMM) [48] considers the WDMM model as a baseline while it has certain characteristic differences. Distinguishing the days when people work and the weekend days is an example of such differences. While this seems to create more sophistication to an already sophisticated mobility model, the RLMM model is more simplistic, having assumptions such as people stay static at work. RLMM defines five states of a person: at home, at work, at a popular place, regular travel, and alternative travel [48]. At home, at work and a popular place, each person is static. Regular travel is executed among these locations and pause times and commute distances are defined as parameters. The alternative travel state brings randomness such that the person breaks the daily routine and moves to random and mostly longer distances during the weekends. The travel speed is by default defined as a static value according to the cases of pedestrian movement (1.25m/s), cycling (5m/s), slow vehicle movement (8.3m/s), and fast vehicle movement (11.1m/s). Moreover, the speed can be changed dynamically if needed, while it certainly adds more randomness to the model. Other major differences of the RLMM model are, compared to the WDMM model, not involving social interactions and not using real-maps.

ORBIT [49] is a simple macro-scale mobility framework with the assumption that people's movement follows orbital mobility patterns in various scales. For instance, a student's mobility in the campus is considered as an orbital movement, where the student travels between hubs such as classroom and restaurant in a cycle. Moreover, the orbits have different levels, for instance, considering the movement in the campus as the Level 1 Orbit, the overall daily movement of the student is considered as the Level 2 Orbit. In Level 2, the orbital movement consists of travels between home, school, and gym in an ordinary daily scenario. Level 3 orbit is considered as visiting three cities in order, where one of the visited cities is the hometown. While this orbital structure provides a practical abstraction to human mobile modeling, it must be validated for efficient use with real-traces.

The ORBIT model separates the movement of a person between the hubs and the movement inside the hubs (intra-hub). The intra-hub movement follows the RWP model with a specified intra-hub speed range and intra-hub pause time. The movement between the hubs (orbit movement) follows a point-to-point (P2P) linear model. In the P2P model, a person who wants to move to the destination hub chooses a random point inside the destination hub and linearly moves from the current point to the randomly selected point. The speed is defined by the inter-hub range parameter. The question of how to choose the hubs as the next destinations differs according to the scenario and different scenarios are modeled with different orbit models such as random orbit, uniform orbit, restricted orbit, and overlaid orbit models [49]. In the random orbit model, hubs are created randomly and the people randomly choose the hubs to move. In the uniform orbit model, the orbit is considered to be uniformly divided into grids where each grid represents a hub and people still choose the hubs randomly. In the restricted orbit model, each person is restricted to move inside only one particular hub for some time. Lastly, the overlaid orbit model is similar to the restricted orbit. The only exception is having a user role such that each person is assigned to a set of hubs instead of one particular hub. While the model does not bring the user roles to model, this feature still has role-based behavior. For instance, modeling a workplace as an orbit, each employee stays near their particular room or the area where the employee's teamwork (hubs), while the director is supposed to visit different areas (hubs) of the workplace to track the progress in different teams. While ORBIT aims to model specific scenarios such as a conference or a campus environment, we do not consider ORBIT as a scenario-specific human mobility model. While the scenarios are specific, the mobility model itself is generic, claiming to represent almost all possible movement scenarios of people. Conversely, while the ORBIT model does not aim to represent the movement of vehicles, we consider the model as a model with vehicle use as the Level 3 orbit requires long-distance travels among cities.

Time variant community (TVC) [50] model is proposed based on the observation of two features: skewed location visiting preferences and time-dependent periodical behavior of people. The observation is based on WLAN traces collected from two university campuses and corporate buildings. By location visiting preference we can simply understand the pause times in different locations. By time-dependency, we can understand that people visit different locations depending on the time of the day. Another observation is the periodical re-appearance of people, such that they connect to the same APs in the dataset. The TVC model is one of the mobility models which are mathematically tractable. This property does not exist in some models since they are sophisticated. On the other hand, some mobility models, which are simple and can be easily theoretically treated, are missing mathematical analyses.

Let us describe two terms community and epoch in the TVC model. As opposed to the other mobility models, a community in TVC does not represent a group of people sharing similar interests. Instead, a community defines a squared region in the simulation field. Community region is decided by an exponential distribution with an average length parameter. Community regions may intersect with each other and a community may even cover the entire area. The individual's mobility consists of a sequence of epochs. During each epoch, the person decides a new community, moves to the community region, and spends time in that region. The movement inside the community is a random direction movement within the community region. The direction of the movement and the speed are uniformly random. After completing the epoch, the person waits stationary for a uniformly random pause time, which is the initial phase of the new epoch (the period of decision for the next community). When the next community is selected, the person chooses a random point in the next community and moves on a straight shortest path with a randomly chosen speed. The community selection is based on the probabilities of the communities for the individual. The selection process is a time-variant Markov chain to capture the spatial and temporal dependencies in the individual's mobility. The selection of new communities is structured in a way that it captures the temporal preferences and the periodicity of human mobility. Various scenarios could be modeled using the TVC model by setting the number of communities, community regions, and time periods parameters. Three example scenarios are considered in the study, which aims to fit WLAN traces, vehicle mobility traces, and human encounter traces. The model does not explicitly involve vehicle models such as cars and buses or imitate their behavior. However, as the model's parameters can be adjusted for fitting it to the vehicular movement, we consider the model as another human mobility model with vehicle use. Another main feature of the model is being a generic model as the model represents application scenarios varying from conference traces to the traces collected from corporate buildings.

As the last human mobility model with vehicle use, we discuss the agenda-driven mobility model [51]. We shortly refer to the model in this article as ADMM. The ADMM model is considered for its use in various wireless networks including opportunistic networks and vehicular networks. ADMM is based on the person's social activities, which are considered as the driving force behind the mobility decisions. The model introduces the concept agenda to human mobility modeling. According to this concept, each person has an agenda that guides the individual movement. Other than the agenda, the model is also based on the geographical locations of the places (maps).

Each person has a personal agenda based on the social role. For instance, a student goes to school by a vehicle at a travel distance of 5 miles, while a repairman visits multiple places during the day. To extract the social roles of people, data from the National Household Travel Survey (NHTS) is used and the model drives statistics based on a large number of records. Moreover, real maps from the GIS database can be used for modeling the movement environment. Based on the statistics from NHTS and the real maps, agendas and maps are synthetically created. First, the map which consists of streets and avenues are created. Second, the addresses on top of the map are defined as the buildings' locations. People move along the roads and spend time in the buildings which are marked with the addresses. The speeds of the people are based on the road traffic and the speed limits of the roads that are taken by the people. The agenda consists of items and each item indicates the time and the location of the person for a particular future activity. Each activity has an expected start and end times. From the NHTS dataset, 35 different types of activities are extracted (e.g., going to work, shopping, or hanging out). Social roles are defined according to 14 occupations. Each occupation corresponds to a unique agenda type. Finally, the routes between any two consecutive activity locations on the agenda are found by Dijkstra's shortest path algorithm. While maps are created synthetically, we still classify the ADMM model as a role-based model with the use of real maps as it is possible to incorporate real maps. Since the model is based on databases for activities, social roles, and also real maps, it is expected to realistically represent human mobility. However, the model's outcomes are compared only with the RWP model and comparison with more realistic human mobility models is required for validation.

VIII. HUMAN MOBILITY IN DISASTERS

The last set of human mobility models we discuss are the disaster mobility models. The main goal of these models is to simulate the movement behavior of people during the times of natural or man-made disasters [27].

The mission-critical mobility model (MCM) [52] aims to model human mobility in the presence of obstacles for mission-critical applications such as networks in times of disasters. Considering an example opportunistic network application, the nodes are carried by human participants such as firemen or policemen. MCM is a generic model, such that it is considered for all types of disaster scenarios. In the model, the destination points are chosen uniformly random by each person. In the case of no obstacle between the current point and the destination point, the person directly moves to the destination, creating a trajectory which is a straight line. In the case of having obstacles in-between, the person passes each obstacle by choosing a directly visible vertex that is closest to the final destination. The vertex location is set as the intermediate destination to pass by the obstacle. The person similarly passes multiple obstacles one by one. The movement algorithm can be considered as greedy since the person always chooses the visible vertex which is closer to the final destination, trying to minimize the total time it takes to walk. Note that the algorithm does not give the optimal shortest-path, as there are many ways of going to the same final destination, which may have shorter overall distances. However, considering the people's movement choices, people also do not always manage to choose the shortest path, instead one can claim that they mostly make greedy movement decisions for having less movement effort. The movement speed is a random value between the minimum and maximum boundaries and the speed is set every time after deciding a new next destination.

MCM has the group mobility behavior, such that each group has a leader who decides the next destination for all group members. Moreover, people have different roles in the disaster area. They are considered in two roles: emergency workers and medical staff. In the MCM model, destination points are set based on the randomly generated event locations. For each event, there is a particular pause time, which represents the time it takes to handle the event. The events are categorized as normal, serious, or complex events and the group leaders do their choices based on the types of the events. For instance, groups of different types need to be involved in handling complex events. The events are handled according to First in First out (FIFO) order. The MCM model is the newer and more extensive version of the previous Human Obstacle Model (HUMO) [62], such that the model is analyzed and validated with real-life human mobility traces.

A disaster area mobility model which is simulated for theme park environments (TP-D) [27], [30] focuses on the human walks during the times of natural or man-made disasters. The TP-D model is based on the evacuation scenario of a large and crowded area with limited vehicle use such as theme parks or airports. As the escape behavior of people has different mobility characteristics than their ordinary movement, the model aims to imitate this behavior. The model assumes that there is no external help from other people such as security operators. Furthermore, the model assumes that communication systems are not available during a disaster, which is an expected result of damages on the infrastructure due to hazards. The model uses real maps extracted from OpenStreetMap [28]. Uniformly random disaster zones are generated on top of the maps. The disaster area is considered as a combination of pedestrian ways, lands, and obstacles. The macro mobility of the pedestrians depends on the waypoints which are located on top of the pedestrian ways. The pedestrians are initially randomly distributed on the waypoints and every pedestrian has a target point, which is one of the exit doors so that by reaching the target they can have access to transportation services such as buses or ambulances. People are considered pedestrians sharing the same roads. This causes them to create pedestrian traffic. The crowd dynamics and the flow of the people during an evacuation are modeled with SFM [25]. Hence, the model takes social interactions into account for micro-mobility decisions such as slowing down. The micro-mobility is considered as the movement of a pedestrian between two consecutive waypoints.

While most disaster mobility models aim to solve the mobility problem for urban scenarios, these scenarios are complicated as they involve the use of vehicles, evacuation from buildings, and people with various roles such as firemen or police. The TP-D model, however, focuses on the human walk problem and aim to generate realistic movement traces for specific scenarios, which can be useful for simulating networks resilient to disasters. Furthermore, solving some aspects of the overall disaster mobility problem could be useful in the future for other more sophisticated scenarios such as the evacuation of metropolitan regions.

Disaster area (DA) [53] mobility model is based on the two disasters that happened in Germany, which are the Wuppertal Railway Crash in 1999 and the Bruehl Roller-Coaster Fire in 2001. For the civil protection in disasters, the movements are driven by tactical reasons. For instance, a group leader for the rescue operation directs workers to move to some places. In the model, the disaster area is considered as consisting of separated regions into disaster incident site, casualty treatment site, transport site, and hospital site. The affected and injured people are considered to be found and rescued in the incident site, taken to the casualty treatment site, and moved to hospitals from the transport zone. The main characteristics of the DA model are having different roles for people, being heterogeneous area-based model, and avoiding obstacles. The group mobility behavior is considered as a future extension to the model. The model also involves the vehicular behavior for transportation services such as ambulances. The disaster simulation area is a synthetic map considered as a static region with obstacles which are modeled with polygons. Visibility graphs are used for optimal path planning and Dijkstra's shortest path algorithm is used on top of this graph, where edges are the Euclidean distances. Each separated area (tactical area) has entry and exit locations. Each person is assigned to an area. Some people are considered as stationary nodes, meaning that they only move inside their assigned area. The others are considered as transport nodes having the ability to carry patients to the next area. In the incident zone, all people are considered as transport nodes, while in casualty treatment area there are only stationary nodes. The movements inside the areas are modeled with the RWP model [63].

IX. GENERAL TRENDS AND APPLICABILITY

Let us briefly discuss the general trends of human mobility modeling and the applicability of the proposed models. First of all, there has been an ongoing practice of using more realistic human mobility models as opposed to simple to implement random movement models. In particular, due to recent advancements in the mobile ad hoc networks, opportunistic networks, wireless sensor networks (WSNs) with mobile elements (sensors, actors, sinks), the networking research field has contributed to the design and development of realistic human mobility models [64]–[67]. We expect this trend to grow also in the field of transportation research [68].

With the increased data collection by the popular use of smartphones and cloud-based mobile applications, there have been studies related to social inferences [69] and social networking [70]. For instance, personal information such as call logs, contact lists, and connected base stations collected from smartphones (e.g., MIT Reality Mining dataset [71], Dartmouth StudentLife dataset [72]) can be analyzed to find the friendship network of individuals based on the smartphone data [73] and mobility information.

Another recent development in the human or crowd mobility and more general urban mobility (including people and their vehicles) is in the visualization of the human mobility patterns [74]–[76]. This aspect of human mobility is expected to be enabled considering the advancements in smart cities, IoT, and 5G networks which bring massive data collection from crowds in urban environments. Real-time crowd monitoring and incident management are currently among the emerging fields in the IoT area [77].

The applicability of the mobility models is the major issue that shapes the general direction in human mobility modeling. Considering the various environments with specific needs, environment or scenario-specific modeling is necessary for applicability. Therefore, the general trends are towards scenario-specific and data-driven models [16] (based on traces collected from the specific environment), as opposed to generic and pure synthetic models. Applicability is mostly subject to the requirements of specific scenarios (e.g., granularity requirement). Especially for newly developed human mobility models, the analytical and experimental results and comparisons with real-life human mobility traces for the specific scenarios can be used as evidence for the applicability of new models in these scenarios.

X. CONCLUDING REMARKS AND FUTURE RESEARCH DIRECTIONS

In the previous sections, we analyzed human mobility and organized human mobility models in four main categories. With the increased everyday use of mobile devices and the new technologies as well as globalization that enables cheaper distant travels, ongoing research on understanding human mobility has been gaining more attention. However, understanding all the features of human mobility in different scales involves various open problems. For instance, from the evacuation of buildings with various sizes to ordinary mobility in the office environment or from intercontinental travels to walks in urban environments, different behaviors need to be taken into consideration. Thorough understanding of human mobility will have certain positive impacts on the societies in the world for handling various problems such as networking, transportation, disaster management, epidemic prevention, and so on. Therefore, human mobility modeling is a topic that requires more resources such as big mobility datasets. Mobility modeling is currently an interdisciplinary area studied by research centers in computer science, civil engineering, and physics departments.

Currently, real-life human mobility traces (e.g., GPS traces) are collected by cloud-based applications such as Google Maps and mobile network operators from a vast majority of people in the USA and other developed countries. While these traces are stored in large databases, privacy is becoming a major issue still in need of a solution. However, using methods such as encryption and being transparent on how data is stored, privacy concern can be potentially addressed. Moreover, encrypted mobility datasets should be available to researchers to achieve significant advances in human mobility modeling.

One obvious necessity of the field is the lack of human mobility models in different scales and specific scenarios. Although we discuss various generic and scenario-specific mobility models in this article, the existing models are not able to cover all ranges of scales or specific urban environments such as major cities. Therefore, there is a certain need for scenario-specific and realistic human mobility models and their incorporation into the simulations. While different mobility models have approached the problem from different angles, such as modeling movement based on social interactions or based on the least action principle, mobility modeling is still open for new perspectives. Some of these environments might be restricted to vehicle use such as airport terminals and theme parks and focus only on human mobility. Additionally, people's roles in their societies and findings of recent social studies could be leveraged for the development of novel mobility models.

Reliable validation of the newly proposed human mobility models depends on several aspects. First, the traces of the mobility models should be tested against real-life traces.

For accuracy purposes, GPS traces are preferred over other types of data. Second, the assumptions of the models (e.g., assuming a constant speed) should be well-justified for the models and the scenarios. Third, the validation should include different perspectives such as statistical matches, the comparisons with significant datasets, trajectory pattern matchings, and consistency tests.

The advancements in the fields such as IoT and 5G networks are going to enable vast data collections from people and their "things". People's use of new services based on IoT and 5G can make a big difference in the data collection, while it brings new challenges such as big data analytics and context awareness. For instance, each device people use can be registered through network service providers to IoT platforms. For instance, in a smart ski resort in a mountain area, movement data can be collected from skiers through the devices attached to their skis and can be forwarded to a nearby smart city to be used for transportation and accommodation planning based on crowd densities. However, the data must be well understood (context awareness) and it must be analyzed in a way that it leads to intelligent decision making (big data analytics). Moreover, as an emerging target of IoT, global interoperability can provide data mashups from various IoT platforms all over the world. For instance, all the smart cities in the world can share their human mobility data for many uses in research, government, and various industries. However, while different domains provide human mobility data with different granularity, data coming from these different sources and platforms must be usable by all parties involved. Moreover, analyses of vast data coming from various environments and conditions will be more complex than the current human mobility analyses that are based on simple data collection techniques. Hence, with the advancements of new services, we expect human mobility to become even more prominent while it will have new major challenges shortly due to the advancements in IoT and 5G networks.

Smart cities have become a major investment area for governments, industry, and academia. Mobility has a direct impact on smart cities. Various domains such as public safety, transportation, surveillance, disaster evacuation planning, and connectivity are affected by human mobility. To enable services efficiently in these different domains, human mobility must be better understood. For this purpose, new human mobility models and simulations can be developed based on the needs of these cities. Moreover, the realization of smart cities and 5G networks enable vast data collections from pedestrians as well as vehicles. The vast amount of data can be analyzed offline by data mining techniques. Some services require real-time processing of human mobility data. For instance, in the case of an emergency such as a fire in a stadium, the fire stations, and the police department need to be aware of the conditions by visualizing real-time data analytics results for the pedestrian flows, possible traffic congestions, and any other necessary information for the safe evacuation. Due to such scenarios, real-time data stream processing will be another major challenge for human mobility data analytics.

An important issue for future mobility models is that even if they are realistic and sophisticated, they should be easy to develop and use. For instance, having mobility simulations with user-friendly interfaces would be very helpful. Furthermore, as the parameter values may change from one scenario to another, the parameters should be easy to understand and change. For instance, for a typical user of the mobility model, setting values of parameters named as α and β is much less meaningful than setting the value of expected walking speed of a person.

Finally, although it is not the main concern for most of the mobility model simulations that are conducted offline, complexity is still an important issue. With the increased sophistication of the mobility models, time and space complexity tend to increase. Another issue is the optimal implementation of the mobility models to prevent errors and memory leaks. We observe that some mobility simulations support simulations of only limited numbers of people (e.g., up to 100 people). Therefore, in the future, complexity and optimization issues need to be taken into consideration when developing new human mobility models.

REFERENCES

[1] G. Solmaz, M. I. Akbas, and D. Turgut, "A mobility model of theme park visitors," *IEEE Trans. Mobile Comput.*, vol. 14, no. 12, pp. 2406–2418, Dec. 2015.

[2] A. Smith, "US smartphone use in 2015," PEW Res. Center, Washington, DC, USA, Tech. Rep., 2015. [Online]. Available: https://www.pewinternet.org/wp-content/uploads/sites/9/2015/03/PI_Smartphones_0401151.pdf

[3] J. Torous, R. Friedman, and M. Keshavan, "Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions," *JMIR mHealth uHealth*, vol. 2, no. 1, p. e2, 2014.

[4] F.-J. Wu and G. Solmaz, "CrowdEstimator: Approximating crowd sizes with multi-modal data for Internet-of-Things services," in *Proc. ACM MOBISYS*, Jun. 2018, pp. 337–349.

[5] F. I. Sadiq, A. Selamat, and R. Ibrahim, "Human activity recognition prediction for crowd disaster mitigation," in *Intelligent Information and Database Systems* (Lecture Notes in Computer Science), vol. 9011, N. Nguyen, B. Trawiński, and R. Kosala, Eds. Cham, Switzerland: Springer, 2015.

[6] G. Chatzimilioudis, A. Konstantinidis, C. Laoudias, and D. Zeinalipour-Yazti, "Crowdsourcing with Smartphones," *IEEE Internet Comput.*, vol. 16, no. 5, pp. 36–44, Sep. 2012.

[7] G. Solmaz and D. Turgut, "Tracking evacuation of pedestrians during disasters," in *Proc. IEEE GLOBECOM*, Dec. 2015, pp. 1–6.

[8] S. S. Bacanli, G. Solmaz, and D. Turgut, "Opportunistic message broadcasting in campus environments," in *Proc. IEEE GLOBECOM*, Dec. 2015, pp. 1–6.

[9] M. Stevens and E. D'Hondt, "Crowdsourcing of pollution data using smartphones," in *Proc. Workshop Ubiquitous Crowdsourcing*, 2010, pp. 1–4.

[10] F.-J. Wu and G. Solmaz, "We hear your activities through Wi-Fi signals," in *Proc. IEEE WF-IoT*, Dec. 2016, pp. 251–256.

[11] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Commun. Mobile Comput.*, vol. 2, no. 5, pp. 483–502, Sep. 2002.

[12] F. Bai and A. Helmy, "A survey of mobility models," in *Wireless Adhoc Networks*, vol. 206. Los Angeles, CA, USA: Univ. Southern California, 2004, p. 147.

[13] J. Treurniet, "A taxonomy and survey of microscopic mobility models from the mobile networking domain," *ACM Comput. Surv.*, vol. 47, no. 1, pp. 14:1–14:32, 2014.

[14] S. Batabyal and P. Bhaumik, "Mobility models, traces and impact of mobility on opportunistic routing algorithms: A survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1679–1707, 3rd Quart., 2015.

[15] Q. Dong and W. Dargie, "A survey on mobility and mobility-aware MAC protocols in wireless sensor networks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 88–100, 1st Quart., 2013.
[16] A. Hess, K. A. Hummel, W. N. Gansterer, and G. Haring, "Data-driven human mobility modeling: A survey and engineering guidance for mobile networking," *ACM Comput. Surv.*, vol. 48, no. 3, 2016, Art. no. 38.
[17] D. Karamshuk, C. Boldrini, M. Conti, and A. Passarella, "Human mobility models for opportunistic networks," *IEEE Commun. Mag.*, vol. 49, no. 12, pp. 157–165, Dec. 2011.
[18] D. Brockmann, L. Hufnagel, and T. Geisel, "The scaling laws of human travel," *Nature*, vol. 439, no. 7075, pp. 462–465, 2006.
[19] M. F. Shlesinger, G. M. Zaslavsky, and U. Frisch, "Lévy flights and related topics in physics," in *Proc. Int. Workshop Held Nice*, vol. 450, 1995, pp. 3–540.
[20] G. M. Viswanathan, V. Afanasyev, S. V. Buldyrev, E. J. Murphy, P. A. Prince, and H. E. Stanley, "Lévy flight search patterns of wandering albatrosses," *Nature*, vol. 381, no. 6581, pp. 413–415, 1996.
[21] M. C. González, C. A. Hidalgo, and A.-L. Barabási, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.
[22] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.
[23] D. Helbing and P. Molnár, "Social force model for pedestrian dynamics," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 51, no. 5, p. 4282, 1995.
[24] D. Helbing, I. J. Farkas, P. Molnar, and T. Vicsek, "Simulation of pedestrian crowds in normal and evacuation situations," *Pedestrian Evacuation Dyn.*, vol. 21, no. 2, pp. 21–58, 2002.
[25] D. Helbing and A. Johansson, "Pedestrian, crowd and evacuation dynamics," in *Encyclopedia of Complexity and Systems Science*, R. Meyers, Ed. New York, NY, USA: Springer, 2009.
[26] K. Lee, S. Hong, S. J. Kim, I. Rhee, and S. Chong, "SLAW: Self-similar least-action human walk," *IEEE/ACM Trans. Netw.*, vol. 20, no. 2, pp. 515–529, Apr. 2012.
[27] G. Solmaz and D. Turgut, "Theme park mobility in disaster scenarios," in *Proc. IEEE GLOBECOM*, Dec. 2013, pp. 377–382.
[28] M. Haklay and P. Weber, "OpenStreetMap: User-generated street maps," *IEEE Pervasive Comput.*, vol. 7, no. 4, pp. 12–18, Oct. 2008.
[29] D. Kotz, T. Henderson, I. Abyzov, and J. Yeo. (Sep. 2009). *CRAWDAD: The Dartmouth/Campus Dataset (v. 2009-09-09)*. [Online]. Available: http://crawdad.org/dartmouth/campus/20090909
[30] G. Solmaz and D. Turgut, "Pedestrian mobility in theme park disasters," *IEEE Commun. Mag.*, vol. 53, no. 7, pp. 172–177, Jul. 2015.
[31] C. M. Schneider, V. Belik, T. Couronné, Z. Smoreda, and M. C. González, "Unravelling daily human mobility motifs," *J. Roy. Soc. Interface*, vol. 10, no. 84, 2013, Art. no. 20130246.
[32] A. Munjal, T. Camp, and W. C. Navidi, "SMOOTH: A simple way to model human mobility," in *Proc. 14th ACM Int. Conf. Modeling, Anal. Simulation Wireless Mobile Syst.*, 2011, pp. 351–360.
[33] M. Schwamborn and N. Aschenbruck, "Introducing geographic restrictions to the SLAW human mobility model," in *Proc. IEEE 21st Int. Symp. Modelling, Anal. Simulation Comput. Telecommun. Syst.*, Aug. 2013, pp. 264–272.
[34] V. Vukadinovic, F. Dreier, and S. Mangold, "A simple framework to simulate the mobility and activity of theme park visitors," in *Proc. WSC*, Dec. 2011, pp. 3248–3260.
[35] V. Vukadinović, Ó. R. Helgason, and G. Karlsson, "An analytical model for pedestrian content distribution in a grid of streets," *Math. Comput. Model.*, vol. 57, nos. 11–12, pp. 2933–2944, 2013.
[36] S. Kosta, A. Mei, and J. Stefa, "Small world in motion (SWIM): Modeling communities in ad-hoc mobile networking," in *Proc. 7th Annu. IEEE Commun. Soc. Conf. Sensor, Mesh Ad Hoc Commun. Netw. (SECON)*, Jun. 2010, pp. 1–9.
[37] A. D. Nguyen, P. Sénac, V. Ramiro, and M. Diaz, "STEPS—An approach for human mobility modeling," in *NETWORKING 2011*. Berlin, Germany: Springer, 2011, pp. 254–265.
[38] M. Kim, D. Kotz, and S. Kim, "Extracting a mobility model from real user traces," in *Proc. INFOCOM*, vol. 6, Apr. 2006, pp. 1–13.
[39] M. Musolesi and C. Mascolo, "Designing mobility models based on social network theory," *ACM Mobile Comput. Commun. Rev.*, vol. 11, no. 3, pp. 59–70, Jul. 2007.
[40] C. Boldrini and A. Passarella, "HCMM: Modelling spatial and temporal properties of human mobility driven by users' social relationships," *Comput. Commun.*, vol. 33, no. 9, pp. 1056–1074, 2010.
[41] D. Fischer, K. Herrmann, and K. Rothermel, "GeSoMo—A general social mobility model for delay tolerant networks," in *Proc. 7th IEEE Int. Conf. Mobile Ad-Hoc Sensor Syst. (IEEE MASS)*, Nov. 2010, pp. 99–108.
[42] V. Borrel, F. Legendre, M. D. D. Amorim, and S. Fdida, "SIMPS: Using sociology for personal mobility," *IEEE/ACM Trans. Netw.*, vol. 17, no. 3, pp. 831–842, Jun. 2009.
[43] S. Yang, X. Yang, C. Zhang, and E. Spyrou, "Using social network theory for modeling human mobility," *IEEE Netw.*, vol. 24, no. 5, pp. 6–13, Mar./Oct. 2010.
[44] C. Zhao and M. L. Sichitiu, "N-body: Social based mobility model for wireless ad hoc network research," in *Proc. 7th Annu. IEEE Commun. Soc. Conf. Sensor, Mesh Ad Hoc Commun. Netw. (SECON)*, Jun. 2010, pp. 1–9.
[45] S. Isaacman, R. Becker, R. Cáceres, M. Martonosi, J. Rowland, A. Varshavsky, and W. Willinger, "Human mobility modeling at metropolitan scales," in *Proc. 10th Int. Conf. Mobile Syst., Appl., Services*, 2012, pp. 239–252.
[46] D. J. Mir, S. Isaacman, R. Cáceres, M. Martonosi, and R. N. Wright, "DP-WHERE: Differentially private modeling of human mobility," in *Proc. IEEE Int. Conf. Big Data*, Oct. 2013, pp. 580–588.
[47] F. Ekman, A. Keränen, J. Karvo, and J. Ott, "Working day movement model," in *Proc. 1st ACM SIGMOBILE Workshop Mobility Models*, 2008, pp. 33–40.
[48] M. Gorawski and K. Grochla, "The real-life mobility model: RLMM," in *Proc. 2nd Int. Conf. Future Gener. Commun. Technol. (FGCT)*, Nov. 2013, pp. 201–206.
[49] J. Ghosh, S. J. Philip, and C. Qiao, "Sociological orbit aware location approximation and routing (SOLAR) in MANET," *Ad Hoc Netw.*, vol. 5, no. 2, pp. 189–209, 2007.
[50] W. J. Hsu, T. Spyropoulos, K. Psounis, and A. Helmy, "Modeling spatial and temporal dependencies of user mobility in wireless mobile networks," *IEEE/ACM Trans. Netw.*, vol. 17, no. 5, pp. 1564–1577, Oct. 2009.
[51] Q. Zheng, X. Hong, J. Liu, D. Cordes, and W. Huang, "Agenda driven mobility modelling," *Int. J. Ad Hoc Ubiquitous Comput.*, vol. 5, no. 1, pp. 22–36, 2009.
[52] C. Papageorgiou, K. Birkos, T. Dagiuklas, and S. Kotsopoulos, "Modeling human mobility in obstacle-constrained ad hoc networks," *Ad Hoc Netw.*, vol. 10, no. 3, pp. 421–434, 2012.
[53] N. Aschenbruck, E. Gerhards-Padilla, and P. Martini, "Modeling mobility in disaster area scenarios," *Perform. Eval.*, vol. 66, no. 12, pp. 773–790, 2009.
[54] K. Lee, S. Hong, S. J. Kim, I. Rhee, and S. Chong, "SLAW: A new mobility model for human walks," in *Proc. INFOCOM*, Apr. 2009, pp. 855–863.
[55] A. Lindgren, A. Doria, and O. Schelén, "Probabilistic routing in intermittently connected networks," *ACM SIGMOBILE Mobile Comput. Commun. Rev.*, vol. 7, no. 3, pp. 19–20, 2003.
[56] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Efficient routing in intermittently connected mobile networks: The multiple-copy case," *IEEE/ACM Trans. Netw.*, vol. 16, no. 1, pp. 77–90, Feb. 2008.
[57] J. Yoon, M. Liu, and B. Noble, "Random waypoint considered harmful," in *Proc. 22nd Annu. Joint Conf. IEEE Comput. Commun. (INFOCOM)*, vol. 2, Mar./Apr. 2003, pp. 1312–1321.
[58] N. Aschenbruck and M. Schwamborn, "Synthetic map-based mobility traces for the performance evaluation in opportunistic networks," in *Proc. 2nd Int. Workshop Mobile Opportunistic Netw.*, 2010, pp. 143–146.
[59] G. Solmaz, M. I. Akba , and D. Turgut, "Modeling visitor movement in theme parks," in *Proc. IEEE LCN*, Oct. 2012, pp. 36–43.
[60] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. KDD*, 1996, vol. 96, no. 34, pp. 226–231.
[61] E. T. Hall, R. L. Birdwhistell, B. Bock, P. Bohannan, A. R. Diebold, M. Durbin, M. S. Edmonson, J. L. Fischer, D. Hymes, S. T. Kimball, W. La Barre, J. E. McClellan, D. S. Marshall, G. B. Milner, H. B. Sarles, G. L. Trager, and A. P. Vayda, "Proxemics [and comments and replies]," *Current Anthropol.*, vol. 9, nos. 2–3, pp. 83–108, Apr./Jun. 1968.
[62] C. Papageorgiou, K. Birkos, T. Dagiuklas, and S. Kotsopoulos, "An obstacle-aware human mobility model for ad hoc networks," in *Proc. IEEE Int. Symp. Modeling, Anal. Simulation Comput. Telecommun. Syst. (MASCOTS)*, Sep. 2009, pp. 1–9.
[63] N. Aschenbruck, M. Frank, P. Martini, and J. Tölle, "Human mobility in MANET disaster area simulation—A realistic approach," in *Proc. 29th Annu. IEEE Int. Conf. Local Comput. Netw.*, Nov. 2004, pp. 668–675.

[64] S. Hahn, D. M. Rose, J. Sulak, and T. Kurner, "Impact of realistic pedestrian mobility modelling in the context of mobile network simulation scenarios," in *Proc. IEEE 81st Veh. Technol. Conf. (VTC Spring)*, May 2015, pp. 1–5.

[65] A. J. Mashhadi, S. B. Mokhtar, and L. Capra, "Habit: Leveraging human mobility and social network for efficient content dissemination in delay tolerant networks," in *Proc. IEEE Int. Symp. World Wireless, Mobile Multimedia Netw. Workshops (WoWMoM)*, Jun. 2009, pp. 1–6.

[66] D. Zhang and J. P. G. Sterbenz, "Robustness analysis of mobile ad hoc networks using human mobility traces," in *Proc. 11th Int. Conf. Design Reliable Commun. Netw. (DRCN)*, Mar. 2015, pp. 125–132.

[67] K. Maeda, A. Uchiyama, T. Umedu, H. Yamaguchi, K. Yasumoto, and T. Higashino, "Urban pedestrian mobility for mobile wireless network simulation," *Ad Hoc Netw.*, vol. 7, no. 1, pp. 153–170, 2009.

[68] B. Pan, Y. Zheng, D. Wilkie, and C. Shahabi, "Crowd sensing of traffic anomalies based on human mobility and social media," in *Proc. 21st ACM SIGSPATIAL Int. Conf. Adv. Geograph. Inf. Syst.*, 2013, pp. 344–353.

[69] D. Wang, D. Pedreschi, C. Song, F. Giannotti, and A.-L. Barabasi, "Human mobility, social ties, and link prediction," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2011, pp. 1100–1108.

[70] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2011, pp. 1082–1090.

[71] N. Eagle and A. Pentland, "Reality mining: Sensing complex social systems," *Pers. Ubiquitous Comput.*, vol. 10, no. 4, pp. 255–268, 2005.

[72] R. Wang, F. Chen, Z. Chen, T. Li, G. Harari, S. Tignor, X. Zhou, D. Ben-Zeev, and A. T. Campbell, "StudentLife: Assessing mental health, academic performance and behavioral trends of college students using smartphones," in *Proc. 2014 ACM Int. Joint Conf. Pervas. Ubiquitous Comput.*, 2014, pp. 3–14.

[73] M. I. Akba , R. N. Avula, M. A. Bassiouni, and D. Turgut, "Social network generation and friend ranking based on mobile phone data," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2013, pp. 1444–1448.

[74] F. Wu, M. Zhu, X. Zhao, Q. Wang, W. Chen, and R. Maciejewski, "Visualizing the time-varying crowd mobility," in *Proc. SIGGRAPH Asia Vis. High Perform. Comput.*, 2015, Art. no. 15.

[75] L. Wu, Y. Zhi, Z. Sui, and Y. Liu, "Intra-urban human mobility and activity transition: Evidence from social media check-in data," *PLoS ONE*, vol. 9, no. 5, 2014, Art. no. e97010.

[76] W. Zeng, C.-W. Fu, S. M. Arisona, and H. Qu, "Visualizing interchange patterns in massive movement data," *Comput. Graph. Forum*, vol. 32, pp. 271–280, Jun. 2013.

[77] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.

GÜRKAN SOLMAZ received the B.S. degree in computer engineering from Middle East Technical University (METU), Turkey, and the M.S. and Ph.D. degrees in computer science from the University of Central Florida (UCF), USA. He is currently a Senior Researcher with the IoT Research Group, NEC Laboratories Europe, Heidelberg, Germany. His research interests include mobility, mobile computing/networking, and cloud-edge systems aspects of the IoT with a particular focus on crowd mobility in smart cities as well as autonomous vehicles and drones in new and the unconventional IoT applications. He is a member of the ACM.

DAMLA TURGUT received the B.S., M.S., and Ph.D. degrees from the Computer Science and Engineering Department, The University of Texas at Arlington. She is currently a Professor with the Department of Electrical Engineering and Computer Science, University of Central Florida. Her research interests include wireless ad hoc, sensor, underwater, and vehicular networks, as well as considerations of privacy in the Internet of Things. She is also interested in applying big data techniques for improving STEM education for women and minorities. She is a member of the ACM and the Upsilon Pi Epsilon Honorary Society. Her recent honors and awards include being selected as an iSTEM Faculty Fellow, from 2014 to 2015, and being featured in the UCF Women Making History series, in March 2015. She was a co-recipient of the Best Paper Award at the IEEE ICC 2013. She serves as a member of the Editorial Board and of the Technical Program Committee of ACM and the IEEE journals and international conferences.