ARCHIVIO ISTITUZIONALE DELLA RICERCA

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

MarcoPolo: A Zero-Permission Attack for Location Type Inference from the Magnetic Field Using Mobile Devices

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:
Perez, B., Mehrotra, A., Musolesi, M. (2025). MarcoPolo: A Zero-Permission Attack for Location Type Inference from the Magnetic Field Using Mobile Devices. Springer Science and Business Media Deutschland GmbH [10.1007/978-981-97-8016-7_1].

Availability:
This version is available at: https://hdl.handle.net/11585/1034291 since: 2026-02-14

Published:

DOI: http://doi.org/10.1007/978-981-97-8016-7_1

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/).
When citing, please refer to the published version.

(Article begins on next page)

13 October 2026

MarcoPolo: A Zero-Permission Attack for Location Type Inference from the Magnetic Field using Mobile Devices

Beatrice Perez[1], Abhinav Mehrotra[2], and Mirco Musolesi[3,4]

[1] University of Massachusetts, Boston, MA, United States of America
beatrice.perez@umb.edu
[2] Samsung AI Center, Cambridge, United Kingdom
a.mehrotra1@samsung.com **
[3] University College London, London, United Kingdom
m.musolesi@ucl.ac.uk
[4] University of Bologna, Bologna, Italy

Abstract. Location information extracted from mobile devices has been largely exploited to reveal our routines, significant places, and interests, just to name a few. Given the sensitivity of the information it reveals, location access is protected by mobile operating systems and users have control over which applications can access it. We argue that applications can still infer the coarse-grain location information by using alternative sensors that are available in off-the-shelf mobile devices that do not require any permissions from the users.

In this paper we present a zero-permission attack based on the use of the in-built magnetometer, considering a variety of methods for identifying location-types from their magnetic signature. We implement the proposed approach by using four different techniques for time-series classification. In order to evaluate the approach, we conduct an *in-the-wild* study to collect a dataset of nearly 70 hours of magnetometer readings with six different phones at 66 locations, each accompanied by a label that classifies it as belonging to one of six selected categories. Finally, using this dataset, we quantify the performance of all models based on two evaluation criteria: (i) leave-a-place-out (using the test data collected from an unknown place), and (ii) leave-a-device-out (using the test data collected from an unknown device) showing that we are able to achieve 40.5% and 39.5% accuracy in classifying the location-type for each evaluation criteria respectively against a random baseline of approximately 16.7% for both of them.

Keywords: Location inference · Magnetic field · Mobile phones.

** This work was carried out by Beatrice Perez and Abhinav Mehrotra when they were affiliated with University College London.

1 Introduction

Mobile phones are equipped with a variety of sensor used by applications to obtain a user's contextual information (*e.g.*, location, humidity, acceleration, and network connectivity) to support a variety of applications and services [11]. However, the sensors embedded in smartphones have also unintentionally become the source of information leaks that might adversely impact the privacy of their owners. Indeed, information extracted through sensors can be used to identify users or devices [20, 42, 57, 31, 36, 54], and infer their behavioral patterns, interests, personal preferences [40], and even their health condition [59]. The fact that information can automatically be extracted by means of passive sensors is generally perceived negatively by users [10, 18, 45, 46].

In order to mitigate the privacy risks, mobile operating systems have adopted a permission-based paradigm where each application must request access to any of the protected resources on the phone, including camera, microphone, location, and contact list. Each permission is flagged as sensitive depending on the invasive nature of the resource and the importance of the data it might reveal; therefore, the permission to answer phone calls is more restricted as compared to the permission that allows for access to the camera of the device. Users can choose to disable system-wide access to a certain type of information, or control it at the application level where access to some information can be revoked.

Location is one of the information types that is protected by such permission systems. Location data has been shown to be sensitive for users as it can be used to track and profile individuals (*e.g.*, for advertising purposes). In particular, it can be used to infer a user's identity [19]; it can reveal a user's significant locations and points-of-interests (*e.g.*, home location, work location, morning coffee shop, events attended) along with their transportation routine and use them to predict trajectory and future locations [33]; and finally, it can be used to detect the general behavior of a single user and group users based on the similarity across their interests [32].

Restrictions placed on location data have led to the development of alternative methods that aim to derive location from sensors not protected under permission systems. Currently, Android and iOS application can access the gyroscope, accelerometer, and magnetometer sensors without requiring the user permission. These methods generally combine previously acquired environmental information (*e.g.*, produced by surveying a location) with real-time sensor readings to ascertain the user's current position [50]. In particular, Wang *et al.* combine the accelerometer and gyroscope in order to localize users [49] and Chen *et al.* combine WiFi signal strength with FM radio signals to determine a users' location [14].

In this paper, we show how one can leverage one of these sensors, which is not subjected to any permission, in order to infer the type of the location where the user is currently situated. We are not the first to propose such a solution. However, where previous studies focus on mapping locations to pinpoint position, we present a method that reduces the need for surveying the target locations (*i.e.*, locations where localization will take place). We also require that the identifica-

tion process relies exclusively on data that can be accessed without permissions, thus devising a *zero-permission localization attack*. Our approach exploits the readings captured by the magnetometer, present in most smartphones today, to infer a type of location from their magnetic signature. We present four different methods for time-series analysis of magnetic readings: (i) full signal matching, (ii) statistical descriptors, (iii) automated feature extraction, and (iv) shapelets analysis [55]. We compare the two traditional methods of statistical features and full signal pattern matching with two novel methods: convolutional neural networks through One Shot Learning and shapelets, that is a short time-series capturing repeated patterns observed in magnetic field readings. The use of different techniques for analyzing time-series data enables us to establish baselines and compare their performance when applied to this problem. An overview of the attack is presented in Figure 1. In order to evaluate our approach we collect a la-

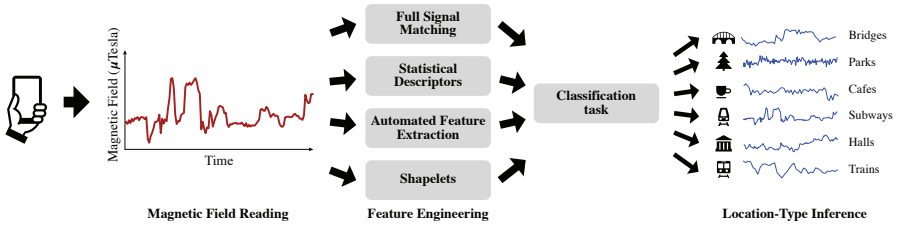


Fig. 1: Overview of the zero-permission attack leveraging the magnetometer of mobile devices to capture magnetic field readings.

beled dataset, used as ground truth, by conducting an in-the-wild measurement study and sampling the magnetic field at different locations over a metropolitan area. We collect readings from off-the-shelf smartphones at ten different types of locations: long-distance train stations, urban train stations (*i.e.*, subways), parks, bridges, coffee shops, halls, laundromats, bus stops, parking lots, and gyms. The places were chosen as alternatives for three groups: indoor environments, outdoor environments, and environments that contain clearly distinguishable events (*e.g.*, trains passing by). For each of the 10 location-types, we collect 10 minutes of magnetic field readings sampled at 1 Hz from five different phones at 11 locations. The entire dataset therefore consists of a total of 91 hours of readings. We evaluate the prediction performance of our four methods by first identifying a location-type from the magnetic readings of an unknown location and second, by identifying a location-type from the magnetic readings of an unknown device. We summarize our main contributions as follows:

– We collected over 91 hours of labeled magnetic field readings from various locations across a major metropolitan area. We show that location-types can be inferred from magnetic field readings available without any system permissions across a wide range of smartphones.

- We compare the performance of four methods for time-series classification in identifying location-types. We perform an in-depth analysis of the choice of the values of the parameters for each methods, providing a methodology for their selection.
- Through our proof-of-concept implementation and in-the-wild study across different locations, we demonstrate the feasibility of the location attack.

2 Overview

In this section we introduce the key concepts and present the main motivations for our work.

2.1 Key Concepts

In writing this work, we take ownership of some words and create new concepts. We define them as follows:

Location-Type. As defined in [47], localization is the problem of ascertaining the position of an agent relative to a map. In this work, we address the problem of localization from a novel perspective, where our primary interest is to understand the environmental similarities between places. Location-type is the name we assign to the top level of a two-tier hierarchy (the lower level being the distinct places). Locations are grouped based on common environmental characteristics including, building structure and materials, human movement patterns, and events (defined later in this section).

Magnetic Field. The magnetic field is the combination of geomagnetic and electromagnetic phenomena. Geomagnetic fields describe the naturally occurring (magnetic) field emitted by the planet's core and the local variations caused by ferromagnetic materials such as iron, cobalt, or nickel [23, 27]. As an example, the steel structure of a building and the movement of metal objects (*e.g.*, a car or a train) will distort the planet's field.

Time-Series. A time-series is an ordered set of values (*i.e.*, the magnetic field) sampled at a fixed interval that, together, form a single observation or reading [21]. The length of the time series is equal to the number of values available in that observation. Subsets or samples of a time-series with more than one element (which will be associated to shapelets later) are time-series observations in their own right, with the distinction that the length of the subset must be no greater than the length of the originating observation(s).

Events. We define events as the dynamic extension of landmarks for time-series data. While a landmark represents a set of structural characteristics of an environment (*e.g.*, the corner of a building, the presences of a fountain, the electrical wiring of a room), an event represents a transient occurrence (*e.g.*, the movement of a train or car, riding on an escalator, or walking by a person). Unlike landmarks, which are unique to a place, events can be identified in places described by a location-type. In our analysis, we make the assumption that locations that

belong to the same type are characterized by similar events, which will be associated to specific *shapes* in the time-series of the magnetic field.

2.2 Motivation

Information about a user's physical location provides private insights about the users themselves. Indeed, applications installed on the users' smartphones can access location data provided by location services using the GPS chip and the cellular network through permission-based controls offered by the operating system. It is left to the user's discretion to grant location access to any application that requests it [4, 3, 35].

On the other hand, side-channel, zero-permission attacks have been exploited to infer users' location using alternate approaches without requesting access to the corresponding GPS and cellular permissions. They rely on constructing, for each target location, a map of the environmental characteristics measured by the sensors embedded in the smartphone. These sensors, whose access does not require any specific permissions, include primarily the magnetometer, the accelerometer and the gyroscope [26, 34, 38, 39, 58]. In practice, given a target location (*e.g.*, an area or a building), a map can be built by recording sets of sensor measurements (*e.g.*, acceleration from the accelerometer and/or magnetic field from the magnetometer) that are associated to precise coordinates within the location being surveyed.

As opposed to the GPS module, the magnetometer is a low-power sensor and accessing readings from a magnetometer does not require any notification to users (or permissions) in either Android or iOS. Moreover, once the information is processed it can be used to protect the privacy of users: while the GPS provides detailed location information (accurate to centimeters of the true position), a generic localization method based on magnetic field could potentially be used to verify that a user is within the premises of a certain type of place or in a specific environment without disclosing exact information. Both marketing and verification applications could obtain the information they need without being invasive with regards to user location.

3 Related Work

Our work is related to the general problem of localization. According to the taxonomy presented by Thrun *et al.*, we address a global localization problem in a dynamic environment where we passively monitor users [47]. In this context, localization of an autonomous entity, referred in the literature as "agent", can only be carried out through sensor readings, either by matching current readings against an up-to-date map or through geometric calculations that determine the unknown position of the agent against a known marker. In the following, we discuss different approaches and contextualize this work with respect to the literature.

3.1 Model-Based Localization

Model-based techniques include methods where the movement of the user is calculated from a sensor, including GPS, as well as cellular and WiFi signals. They are by far the most common and wide spread techniques used for localization.

Mobile operating systems give developers the option to localize users based exclusively on cellular radio signal strength. The position of the user is inferred by extracting features from a signal and comparing them to an established ground truth. As an example, network signal strength is one of the most commonly used features and the actual position of the user is computed from the triangulation between multiple cellular base stations and the signal received from each station by the device [48].

These methods can be applied to indoor environments as well. Localization from WiFi access points using received signal strength has been found to be accurate enough to track users inside buildings with the capability of distinguishing between adjacent rooms [12, 15, 53]. The method only requires WiFi signal strength and the layouts of the building so as to localize a user. The primary limitation of these methods is that the signal strength calculation is typically carried out on the device and continuous tracking drains the resources of a phone [15].

Model-based techniques are used as a means of error correction for dead-reckoning schemes. Urban dead-reckoning is the process by which a user's location is tracked by measuring the side-effects of motion, usually through the use of the compass and the gyroscope. It has been shown that urban dead reckoning accumulates errors of up to 100 meters in 6 minutes of collected data [49]. Indoor localization methods often combine dead reckoning with model-based localization to maintain acceptable tracking accuracy [51, 52].

The error correction mechanism is obtained from a variety of sources. For example, Haverinen and Kemppainen use GPS readings to mark the entrance of the building and again collect active GPS readings from open areas or windows to validate and correct their location [27]. Wang *et al.* collect magnetic field readings along with gyration and acceleration information from different users to find landmarks and correct the error for individual users [49].

3.2 Fingerprint-Based Localization

Fingerprint-based techniques include all methods where a site (or any other geographical location) is surveyed in order to build a map which is then used to pinpoint the location of a user. Fingerprint-based techniques are costly with respect to collection and maintenance of the maps. Active areas of research in this field are mainly centered on reducing the overall financial cost necessary for building maps primarily through leveraging crowd-sourcing techniques where data is passively collected by a large population [44]. In the following, we detail the approaches specific to the type of environment, whether it is indoor or outdoor. This type of techniques have been used successfully to fingerprint devices themselves (see, for example, [42, 57]).

Indoor Environments. Localization in an indoor environment is linked to a map with high level of detail. Even for small spaces this presents a challenge in terms of the volume of data that must be available for the task. In fact, it is important to consider that sensor values might vary with altitude as well as with horizontal displacement. Furthermore, selecting a sensor that provides consistent differences in enclosed environments increases the complexity of the task. Potential solutions to these problems are manual collection of ambient readings or strategic placement of location beacons (and sensors) throughout the area of study under consideration [16, 23, 49, 52]. Chung *et al.* use the magnetic field to determine the position of a user in a corridor. They manually collect magnetic field readings every 60 cm and find that localization is accurate to 1.64 m for 90% of the test observations [16]. They also test their system in elevators and in the atrium of a building and find that there are measurable differences across these locations. Following the work by Chung *et al.*, Haverinen and Kemppainen test whether the difference in magnetic field readings can be used to locate an agent across a larger area (in their work, they test four buildings). They find that these readings have low variability over time while being spatially distinct and build a localization system based on them [27]. Finally, Wang *et al.* show that magnetic field readings can be used to discriminate between activities such as standing or walking. In particular, they use changes in the magnetic field to identify when a user is moving in an escalator [49].

Outdoor Environments. Outdoor localization approaches are classified based on the technique used to establish the map. One common fingerprinting method from outdoor environments is matching the visual cues obtained from the camera of the phone to a map of geo-tagged images [48]. Narain *et al.* use the combination of the motion and position sensors available in smartphones (primarily the accelerometer and gyroscope) to infer the trajectory of a moving car [38]. In particular, they propose a method to reconstruct the route taken by the car based on the physical characteristics of the roads (*i.e.*, speed, bumps, stop signs, and curvature), the junctions between roads, and the angle of each turn was recorded. With this information, they are able to match the estimated candidate trajectory to the map of the (known) city using OpenStreetMap and estimate the actual route taken by the user. Similarly to our work, the authors propose a zero-permission attack to determine location information about users from motion sensors. However, we focus on a different localization task that aims to identify the places that users have visited solely from the magnetometer sensor instead of the trajectories they have taken. Moreover, we employ an event-based approach that identifies events and matches events to locations thereby reducing the need for up-to-date maps. Finally, [7], the authors present a system where a location ID is transmitted via low power magnetic coil and received by permissionless apps. This system is different from ours, since MarcoPolo does not rely on any external device for identification.

3.3 Contribution to the State of the Art

The main challenge faced by outdoor localization techniques comes from the amount of data required for accurate comparison as well the ability to search this data efficiently in order to find the user's current location. On the other hand, the main challenge of indoor localization techniques is the design of fingerprinting methods that are costly in both manpower and time because they require fine spatio-temporal granularity [52]. Moreover, the surveyed signal used to construct the map may change over long periods of time. Consequently, maps have to be repeatedly updated. Similarly, the calibration process on the device that is necessary to accurately measure the signal and the calculation of distances between devices and (signal) sources requires complex computations that might adversely impact the performance of smartphones [49].

To the best of our knowledge, our work is the first that explicitly maps location-types in terms of their characteristics, such as the structural similarity between buildings, the isolation of repeated events, the size and distribution of people and crowds, and any similarity across the tasks undertaken at each location. We combine these characteristics to generate a *magnetic signature* for each location-type which we exploit to identify new locations in the city. The signatures for each location-type are composed of a set of shapelets as defined in [24] [28] [37] [55] [56]. The classification of a new observation into a location-type is a function of the likelihood of a shapelet being present in that observation. To summarize, we propose a novel, low-cost, and passive methodology for location inference. In particular, we present supervised classification algorithms that take as input the time-series (observations) for all classes and predict the probability and label (*i.e.*, location-type) of a new observation.

4 Methodology

In this section we detail the different approaches that we use in our methodology to predict location-types from magnetic field readings. In particular, our methodology addresses two problems: (1) classification from time-series; and (2) deriving location-types from magnetic field readings.

4.1 Techniques for Feature Extraction

Measurements over time provide a more complete view of reality, they allow us to uncover relationships between consecutive measurements. We propose that collecting and processing longitudinal magnetic field readings will prove to be useful in the task of localization. However, time-series classification, clustering, and prediction are identified in the literature as hard problems. Challenges in time-series analysis include: depending on the length of the time series, resource consumption (*i.e.*, computational complexity) of the method; definition of a similarity metric in data that is noisy and prone to outlying values and shifts; and finally, high-dimensionality of observations. An in-depth discussion of these challenges can be found in [56, 21, 2, 5]. In this work we compare four methods for

time-series analysis: (i) full signal matching; (ii) statistical descriptors; (iii) automated feature extraction through the use of neural networks; and (iv) classification through the extraction of shapelets, in both time and frequency domains. Details of these methods are discussed in the remainder of this section.

Full Signal Matching. Pattern matching is a method for time-series classification where each test sample is matched against all labelled training samples. The test sample is then assigned to the label of the closest training sample (*i.e.*, with the shortest distance to it). Consequently, this approach requires a pairwise distance calculation from every signal in the testing set to every signal in the training set, which might become intractable as the number of signals grows. In this analysis, we used four different distance measures for performing classification task: Dynamic Time Warping (DTW), Euclidean Distance, Cosine Distance, and Bhattacharyya Distance [29]. We selected these distance measures as they extract different information from the corresponding input signals: DTW involves both signals and tries to find the best fit between them in a cross-correlation computation, the Euclidean Distance measures the magnitude of the separation between the sampled signals, the Cosine Distance measures the angular separation between the two vectors being compared, and finally, the Bhattacharyya Distance captures the divergence between the probability distributions of each of the signals. One of the key limitations of this approach is that it is sensitive to small variations between the two signals, including differences in the mean for matching signals, which could negatively impact the classification performance. In contrast, we aim at identifying similarities between observations collected in noisy environments. Therefore, in the context of our application, we are expecting this method to have poor prediction accuracy. Nevertheless, we use it as the baseline metric for our performance evaluation.

Statistical Descriptors. The second approach we consider consists in the extraction of representative statistical descriptors (i.e., features) from the time-series. The extracted features are used to represent key characteristics of the data. It is worth noting that the classification performance using these features depends on the degree to which the sampled data represents the population of interest. Good statistical models require a sufficient sample size. We selected eight commonly used statistical features: the median, the amplitude, the energy of the signal, the magnitude and frequency of the natural frequency, the spectral centroid, and the magnitude and frequency of maximum power. Note that we sub-sample the input data and compute these features in both the time and frequency domain. It is worth noting that other alternative features might be extracted. Also for ensuring the replicability of the experiments, we used the open source versions of three state-of-the-art classification algorithms provided by the scikit-learn library [41] to construct our prediction models: k-Nearest Neighbors (kNN) [17], Random Forests (RF) [8], Extreme Gradient Boost (XGB) [13]. kNN is one of the most popular and effective unsupervised learning algorithms,

whereas RF and XGB are widely used for their interpretability and performance, respectively.

Automated Feature Extraction. Algorithms based on neural networks remove the burden of manual extraction of features in order to train prediction models [22]. In particular, in this study, we use Siamese Networks [30] — a specific type of neural network architecture that aims to learn to differentiate between two inputs rather than classifying the inputs into given classes. This network architecture is comprised of two identical neural networks each taking one of the two inputs and the last layers of the two networks are then fed to a contrastive loss function to compute the similarity between the inputs [25]. Two neural networks are identical if they have the same configuration in terms of parameters and weights. Since the weights across the two networks are shared, there are fewer parameters to train, which in turn means that less amount of training data is required. This also reduces the chance of overfitting. Since our input is in the form of a time series, we use 1-D convolutional layers to extract patterns from the data, which are passed from the feed-forward layers to obtain the results for the output layer that is used for computing similarity. More specifically, our network consists of C convolutional layers (CNN layers) and a feed-forward layer that maps the features (extracted by the CNN layers) to the output layer with 100 nodes (*i.e.*, the number of final features to be extracted). We considered values of $C \in [2, 3, 4]$. We do not consider higher values of C given the size of the training set under consideration.

Shapelets (in Time Domain). Shapelets are small local patterns in a time-series that are highly predictive of a class [37]. In the context of our work, each shapelet corresponds to an *event* that is found in at least 90% of the observations belonging to a certain *location-type* (with 11 distinct locations and 5 devices we require an event to be present in 50 observations before a candidate is considered a shapelet). Our assumption is that for each class there exists at least one shapelet contained in all observations of that class, and we consider shapelets of different sizes where longer shapelets are more valuable in terms of class separation. One benefit of this technique is that the analysis of candidate shapelets is independent for each class. This means that the length of a shapelet is also a feature we are considering in the input and that the maximum length is determined on a per class basis (*i.e.*, one way to distinguish between classes that may have similar events is the duration of that event). As an example, the movement of a long-distance train and a subway are similar events but long-distance trains are longer which might correspond to longer shapelets. The methodological contribution of this work incorporates shapelets into probabilistic classification algorithms. Our method takes labeled time series observations and returns the location-type of an unlabeled observation. The algorithm is divided in three steps: first, it extracts shapelets from the training observations (to build a shapelet dictionary); then, uses the training data to describe each class

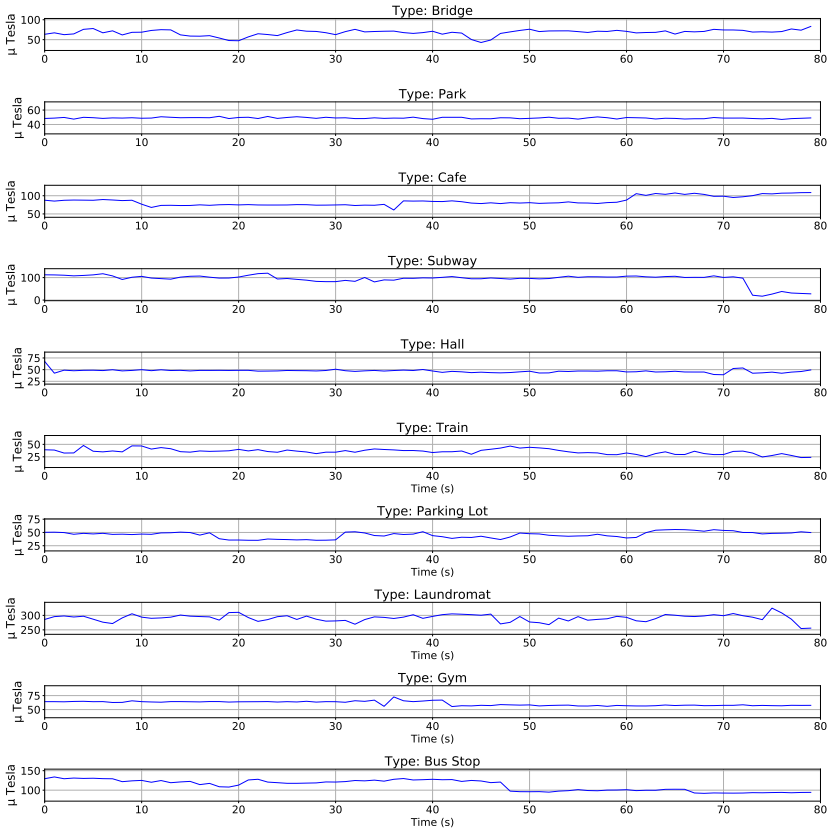


Fig. 2: Visual representation of a *shapelet*. In the figure, each sequence is the longest common time-series present in at least 90% of the observations for each location-type. The length of the shapelet can vary per class as it is possible to observe in the plot. Some might be shorter than others.

in terms of the discovered shapelets; and finally, it classifies test observations into the known classes.

Shapelets (in Frequency Domain). It is common practice in signal processing to study signals in the frequency domain [9]. Following the same procedure described in Section 4.1, we generate a bag of shapelets (with all the possible sub-signals from the training set) and transform each one to the frequency spectra before computing the correlation between all signals. Doing the piecewise transformation before extracting the shapelets might result in better classification if the sources of the signal are monotonic. We integrate the frequency analysis method, known as the Generalized Correlation Coefficient (GCC) [43] with the shapelet-based classifier and compare the accuracy of both methods.

We follow the same clustering procedure and apply hierarchical clustering to the distance matrix generated using GCC. The shapelet dictionary is extracted from the inspection of each cluster. During classification, we use GCC to compute the correspondence between each of the elements in the shapelet dictionary and the signal used as input. This process is repeated (separately) for the observations in the test set and evaluated using the same algorithms.

4.2 Criteria for Assessing Prediction Models

All Places, All Devices. This evaluation approach aims at answering the question as to whether the models are capable to correctly classify a known location from a known device. In the data, we take the 10-minute readings collected from each device and divide it into ten 1-minute segments. We then proceed to a 70/30 cross evaluation by randomly selecting 3 segments for the testing set and the remaining 7 for the training set, i.e., there is no overlapping between training and test sets. We repeat this process and average over 30 iterations using the analysis methods detailed in Section 4.1. Under this evaluation criterion, we assume that we have (previous) data from the device being tested at the location from which the new observation is tested.

Leave-a-Place-Out. This evaluation approach aims at examining the location-type prediction performance from magnetic field readings belonging to locations that have not been seen by the model but belonging to the existing set of location-types. For instance, in this approach, we aim to classify a magnetic field reading that belongs to a train station but in a station different to the ones present in the training set. With this approach, we want to investigate whether the need to map each location can be eliminated and, thus, the location-type can be determined based on its characteristics. In order to carry out this type of evaluation, we remove from each location-type a single distinct location at a time. We use the removed location as the test set and all the readings from the remaining locations in the training set. This process is repeated until each location is assigned once to the testing set.

Leave-a-Device-Out. In this final evaluation approach, we are interested in determining whether the models can predict the location-type from the reading of a new device (*i.e.*, an unknown user). To carry out this evaluation, we select all readings from one device and use them in the testing set. The training set is then composed of all readings of the remaining devices. We use cross validation to evaluate the performance of the method.

5 Dataset

Our data collection was carried out in three stages: (i) application development, (ii) data collection, and (iii) analysis and classification.

5.1 Data Collection

Our hypothesis in this project is that similar locations will be characterized by similar magnetic fields. We are attempting to determine whether a fingerprint exists for a particular location-type and use it to predict the class of unknown locations. This task requires on-site data collection at a number of locations across the city. The measurement devices must therefore be mobile and, if possible, able to be crowdsourced. For their sensors and ubiquity, we use cell phones to collect magnetic field readings at each location.

We developed two applications, for Android and iPhone, and used 14 devices to collect measurements (6 Android phones running Android 7.1.2, 8 iPhones running iOS 12). Both applications collect the three dimensional magnetic field readings along with location, linear acceleration, user activity (for Android phones), and phone identifiers. In the evaluation (*i.e.*, testing) of each model we only considered the five devices for which we had data at all locations.

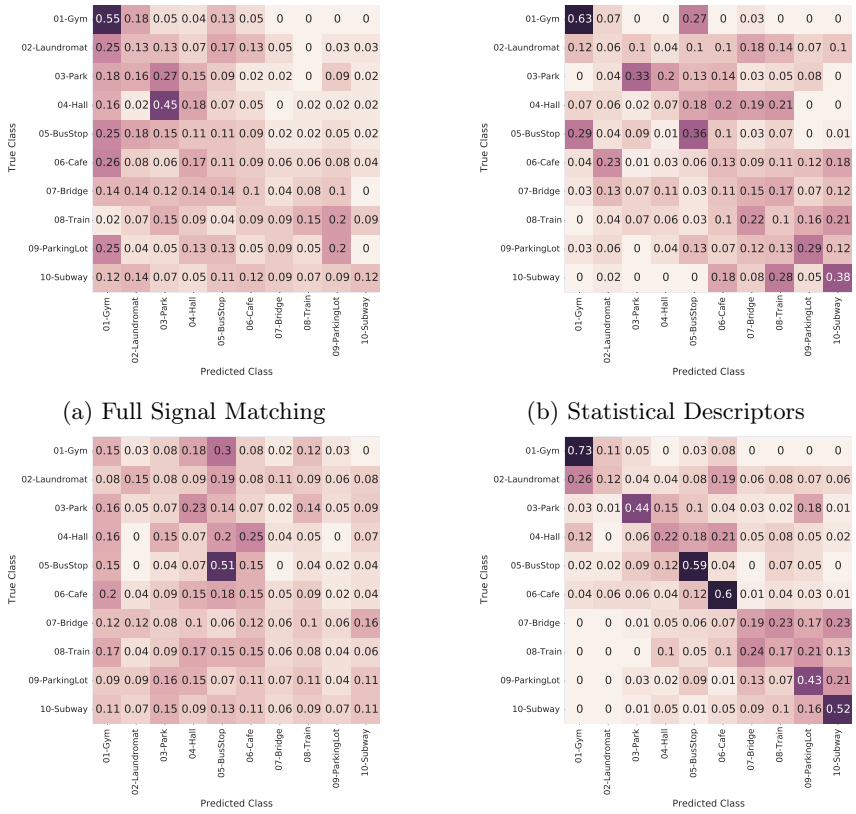
5.2 Description of the Dataset

The final dataset consists of 550 magnetic-field measurements: we collected readings from 5 off-the-shelf smartphones at 110 different locations (10 location-types and 11 locations per type allowing us to train on at least 10 locations for each location-type). Each magnetic reading contains at least 10 minutes of data sampled at 1 Hz. In our case, we are not interested in reconstructing the function and sampling at this rate allows us to record changes without overtaxing the resources of the phone. We designed a collection methodology such that the data that was collected mimic realistic life situations. Volunteers were instructed to hold one phone while walking around the location as they would normally do for a period of 10 minutes. The samples were taken at different times of the day over a period of three months by different volunteers in an attempt to capture the inherent characteristics of a location rather than any bias introduced by the volunteers.

The *types* of places were chosen to fit three groups: indoor environments (*i.e.*, halls, gyms, laundromats, and coffee shops); outdoor environments (*i.e.*, parks and bridges); and transportation hubs (*i.e.*, subway stations, long-distance train stations, parking lots, bus stops). Locations, excluding the parks, are contained in an area of 9.7 mi^2. In any case, parks are contained within the greater metropolitan area of the city taken into consideration in this study.

6 Results

In this section we present the evaluation of the four time-series classification methods, which we discussed in Section 4.1, constructed to infer the location-types using magnetic field readings. As discussed in Section 4.2, we evaluate each method using three criteria: (*i*) all places, all devices, (*ii*) leave-a-place-out, and (*iii*) leave-a-device-out.



(a) Full Signal Matching

(b) Statistical Descriptors

(c) Automated Features

(d) Shapelets

Fig. 3: Confusion matrices for the methods evaluated for the leave-a-place-out scenario for the best classifier/configuration.

6.1 All Places, All Devices Evaluation

This evaluation corresponds to a scenario where the classifier has been trained with data originating from the device being tested at the location in question (*i.e.*, we have labeled data for all devices at every location). The separation between training and testing occurs over the temporal component of each observation. We converted each observation into ten 1 minute segments and selected 3 random segments for the testing set. The remaining 7 segments become the training set. We repeated this process 30 times to reduce the impact of the random selection in the results. This evaluation criterion resulted in 100% accuracy for location-type identification from all devices.

In this scenario, the attack requires having labeled data from all devices at each target location (*i.e.*, building a map per device with all locations). In a deployed system this would require having real-time labeled data to use as the

basis of the map. Because of the pervasiveness of mobile devices this could be possible however, this would be the equivalent of the brute-force approach. In the following sections we present two scenarios where we explore how to generalize the method for new devices and places.

6.2 Leave-a-Place-Out Evaluation

In this evaluation scenario, we used the data of a single location (from all devices) as the test set and the data for the remaining 109 places as train set. We repeated this until we use the data for each of the 110 places for testing, then aggregated the results of all iterations to determine the overall performance.

In order to better understand the performance of all methods, we present the confusion matrices for the four methods under consideration in Figures 3(a)-(d). In each matrix, the average of the diagonal values corresponds to the overall accuracy of the indicated classifier. In particular, we present the average accuracy for each classifier constructed for the feature extraction methods and evaluated using leave-a-place-out evaluation criterion. Overall, our results show that the classifier constructed with shapelets outperforms all others.

The classifiers based on statistical descriptors and shapelets are optimized using three algorithms (*i.e.*, kNN, RF, and XGB). Our results show that both RF and XGB achieve the highest accuracy (with a negligible difference between them), whereas kNN has the worst performance. On the other hand, full signal matching classification was optimized through four different distance metrics (*i.e.*, DTW, Euclidean, Cosine, and Bhattacharya distances). We observe that DTW is the measure that performs best. Finally, we optimized automated feature classification for different number of convolution layers (*i.e.*, 2, 3, and 4 layers with 4, 8, and 16 filters, respectively). The results for this optimization show that the model with 3 and 4 convolution layers achieve the best accuracy. We report these in Table 1.

We further investigated whether the two best methods (*i.e.*, statistical descriptors and shapelets) extract and exploit different information from the data. To this end, we constructed a new method that combines both of these feature sets as input. If both methods extract the same information, then the accuracy of the combined model should not improve. We compared the results of the new combined feature classifier against statistical descriptors and shapelets used as baselines. The results in Figure 4 shows that there it is possible to observe a performance improvement. We constructed the new combined feature set method by using the kNN, RF, and XGB classifiers. We report these results in Table 2.

6.3 Leave-a-Device-Out Evaluation

The readings collected by mobile devices have some error due to the quality of the sensors, their age, and usage [6]. Each device has the software necessary to compensate for the error and provide measurements close to the actual value. As such, this error-correction enabled us to crowd-source the data collection in order to process and extract more robust descriptors for each *location-type* as we

Table 1: The performance of all classifiers for the leave-a-place-out evaluation.

(a)Full Signal Matching		**(b)Statistical Descriptors**		**(c)Automated Features**		**(d)Shapelet**	
Distance measure	*Accuracy*	*Classifiers*	*Accuracy*	*CNN Layers*	*Accuracy*	*Classifiers*	*Accuracy*
DTW	**0.1872**	kNN	0.1170	2	0.0901	kNN	0.2690
Euclidean	0.1750	RF	0.2040	3	**0.0907**	RF	**0.4045**
Cosine	0.1713	XGB	**0.2100**	4	0.0902	XGB	0.3920
Bhattacharyya	0.1005						

Table 2: Leave-a-Place-Out Classification

Combined Feature Set	
Classifiers	*Accuracy*
kNN	0.1860
RF	**0.3907**
XGB	0.3442

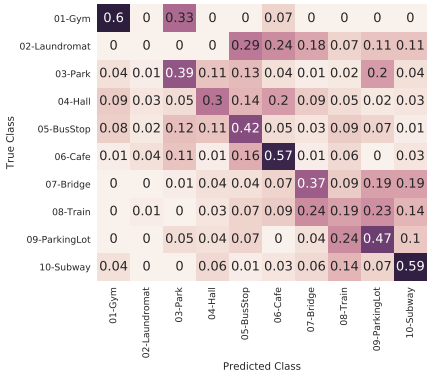
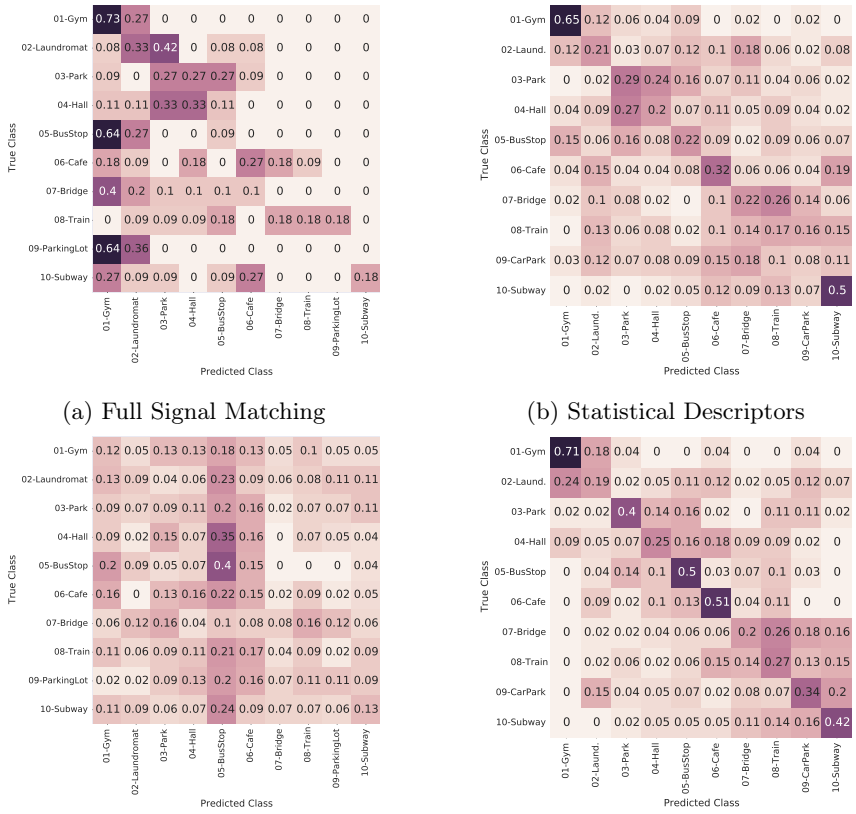


Fig. 4: Confusion Matrix for the best configuration: Leave-a-Place-Out.

would be able to observe the same places under different environmental settings. We evaluated our methods using a leave-one-device-out scenario as discussed in Section 4.2.

In Figures 5(a)-(d) we present the confusion matrices for the best performing models for each four methods presented in Section 4.1. Similar to the leave-a-place-out scenario, we optimized the methods based on statistical descriptors and shapelets using the same three classifiers (*i.e.*, kNN, RF, and XGB). Our results remained consistent: both RF and XGB achieve the highest accuracy (with a negligible difference between them). These results are presented in Table 3. Full signal matching classification was optimized through four different distance metrics (*i.e.*, DTW, Euclidean, Cosine, and Bhattacharya distances); we found that the method that works best uses Euclidean distance as its metric. Finally, we optimized the automated features classifier for different number of convolution layers (*i.e.*, 2, 3, and 4 layers with 4, 8, and 16 filters respectively). The results are again consistent with the previous analysis: the model with 3 and 4 convolution layers achieve the best accuracy. These results are reported in Table 4.

We investigated whether the two best methods (*i.e.*, statistical descriptors and shapelets) extract and exploit different information from the magnetic field readings. We constructed a new classifier that combines both feature sets as input and compares the results against statistical descriptors and shapelets based



(a) Full Signal Matching

(b) Statistical Descriptors

(c) Automated Features

(d) Shapelets

Fig. 5: Confusion matrices for the methods evaluated for the leave-a-device-out scenario for the best classifier/configuration.

methods (used as baselines). The evaluation results show that the classifier based on the combined features achieves an accuracy improvement of 7% as compared to the best of the previous methods. The increase in performance by the combined method can be attributed to the fact that the statistical descriptors exclude the outlying values from a dataset, whereas they are included by the shapelets in the temporal patterns [21]. This shows the consistency of the two methods (*i.e.*, shapelets and statistical descriptors) for extracting complementary information from the time-series data. We present these results in Figure 6.

7 Threats to Validity

The proposed method depends on the possibility of accessing the magnetometer through zero-permission or through involuntary access to it through an app.

Table 3: Leave-a-Device-Out Classification

Combined Feature Set	
Classifiers	*Accuracy*
kNN	0.1776
RF	**0.4256**
XGB	0.4200

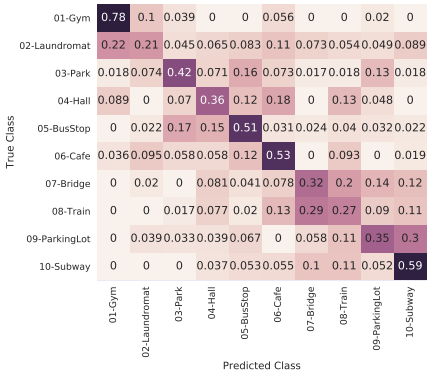


Fig. 6: Classification for the combined feature set for the two most accurate techniques: shapelets and statistical descriptors

Table 4: Performance of all classifiers for the leave-a-device-out evaluation.

(a)Full Signal Matching		**(b)Statistical Descriptors**		**(c)Automated Features**		**(d)Shapelets**	
Distance measure	*Accuracy*	*Classifiers*	*Accuracy*	*CNN Layers*	*Accuracy*	*Classifiers*	*Accuracy*
DTW	0.1852	kNN	0.1776	2	0.1025	kNN	0.2495
Euclidean	**0.2407**	RF	**0.2998**	3	**0.1291**	RF	0.3552
Cosine	0.1481	XGB	0.2926	4	0.1274	XGB	**0.3593**
Bhattacharyya	0.1204						

We also assume that the magnetometer's refresh rate is high enough to allow shapelet extraction. This might change in the future and/or might be different in the future. It is worth noting that, for example, rate limitation for a variety of position sensors has been introduced in Android 12 [1]. In addition, different places might be characterized by similar geo-magnetic profiles. Our tests have been limited in terms of variety of places and we cannot exclude that this might happen in certain circumstances. More in general, it might happen that the environment is highly noisy. This might also affect the applicability of the proposed method in general.

Finally, in terms of resources, an application characterized by a disproportionate use of memory (used to store readings) and battery (for network transmission continuous observations) might look suspicious. In contrast, a system where shapelets are computed offline and transferred to the phone for classification will be practically unnoticeable to the user. Indeed, after training, classification in itself is a resource-efficient task. In our implementation, once computed, all shapelets (*i.e.*, the features) required approximately 50 kB of storage.

Table 5: Leave-a-Device-Out: Frequency domain.

Frequency Domain			
	Shapelets	*Statistical Descriptors*	*Combined*
kNN	0.1750	0.1701	0.1640
RF	**0.2030**	0.3095	0.2634
XGB	0.1675	**0.3147**	**0.2790**

8 Summary of the Contributions

We show that magnetic field measurements can be leveraged to identify types of places with approximately 40% accuracy. We infer location-types both from unseen places and unseen devices. To the best of our knowledge, we are the first to present such an approach. Most of the literature in localization bases location on maps: either by matching the exact values at one location or calculating the distance from some known position. We instead focus on the physical features of the environment: their magnetic signatures. In order to perform the experimental evaluation of the proposed attack, we performed an in-the-wild measurement study. We collected over 91 hours of data from 10 location-types across a major metropolitan area and present our results from a set of 110 distinct locations. The collection times at each location varied in terms of time of day, day of the week, and, for some locations, the readings between phones differ by up to a month. We expect that by allowing such variation, we mitigated the impact of external influences on the generalizability of our results.

Overall, this paper shows that environmental characteristics can be leveraged to infer the *location-type* achieving coarse-grained localization. We have presented two scenarios, one where the test location has never been visited and another where the phone that is collecting the measurements has never been seen. We believe that this constitutes a substantial contribution as the ability to determine coarse-grained location information without the need to map all the locations of interest. Moreover, we find that, as a passive, zero-permission attack it can be used to track users, inferring their routines.

Acknowledgments. Mirco Musolesi was supported by the Italian Ministry of University and Research (MUR) through the project PRIN 2022 "Machine-learning based control of complex multi-agent systems for search and rescue operations in natural disasters (MENTOR)" funded by the European Union - NextGenerationEU.

References

1. Sensors overview. `https://developer.android.com/about/dashboards/index.html`, accessed: 2024-07-01
2. Aghabozorgi, S., Shirkhorshidi, A.S., Wah, T.Y.: Time-series Clustering – A Decade Review. Information Systems **53**, 16 – 38 (2015)

3. Android: Location Permissions. https://developer.android.com/reference/android/Manifest.permission (2022), accessed: 2024-04-20
4. Apple: Apple Developer: Determining the Availability of Location Services. https://developer.apple.com/documentation/corelocation/determining_the_availability_of_location_services (2018), accessed: 2024-04-20
5. Bagnall, A., Lines, J., Bostrom, A., Large, J., Keogh, E.: The Great Time Series Classification Bake Off: A Review and Experimental Evaluation of Recent Algorithmic Advances. Data Mining and Knowledge Discovery **31**(3), 606–660 (2017)
6. Baldini, G., Dimc, F., Kamnik, R., Steri, G., Giuliani, R., Gentile, C.: Identification of Mobile Phones using the Built-in Magnetometers Stimulated by Motion Patterns. Sensors **17**(4), 783 (2017)
7. Block, K., Noubir, G.: My Magnetometer Is Telling You Where I've Been? A Mobile Device Permissionless Location Attack. In: WiSec. p. 260–270. Association for Computing Machinery, New York, NY, USA (2018)
8. Breiman, L.: Random Forests. Machine Learning **45**(1), 5–32 (Oct 2001)
9. Brook, D., Wynne, R.: Signal processing: principles and applications. Edward Arnold, London, England, UK (1988)
10. Brush, A., Krumm, J., Scott, J.: Exploring End-User Preferences for Location Obfuscation, Location-Based Services, and the Value of Location. In: ACM UbiComp. pp. 95–104. Copenhagen, Denmark (Sep 2010)
11. Campbell, A.T., Eisenman, S.B., Lane, N.D., Miluzzo, E., Peterson, R., Lu, H., Zheng, X., Musolesi, M., Fodor, K., Ahn, G.S.: The Rise of People-Centric Sensing. IEEE Internet Computing Special Issue on Mesh Networks (June/July 2008)
12. Chang, N., Rashidzadeh, R., Ahmadi, M.: Robust Indoor Positioning using Differential WiFi Access Points. IEEE Transactions on Consumer Electronics **56**(3), 1860–1867 (Aug 2010)
13. Chen, T., Guestrin, C.: XGBoost: A Scalable Tree Boosting System. In: ACM KDD. San Francisco, CA, USA (Aug 2016)
14. Chen, Y., Lymberopoulos, D., Liu, J., Priyantha, B.: FM-based Indoor Localization. In: ACM MobiSys. Low Wood Bay, UK (Jun 2012)
15. Chintalapudi, K., Padmanabha Iyer, A., Padmanabhan, V.N.: Indoor Localization Without the Pain. In: ACM MobiCom. Chicago, IL, USA (Sep 2010)
16. Chung, J., Donahoe, M., Schmandt, C., Kim, I.J., Razavai, P., Wiseman, M.: Indoor Location Sensing Using Geo-magnetism. In: ACM MobiSys. Bethesda, MD, USA (Jun 2011)
17. Cover, T., Hart, P.: Nearest Neighbor Pattern Classification. IEEE Transactions on Information Theory **13**(1), 21–27 (Jan 1967)
18. Cvrcek, D., Kumpost, M., Matyas, V., Danezis, G.: A Study on the Value of Location Privacy. In: ACM WPES. pp. 109–118. Alexandria, VA, USA (Oct 2006)
19. De Montjoye, Y.A., Hidalgo, C.A., Verleysen, M., Blondel, V.D.: Unique in the Crowd: The Privacy Bounds of Human Mobility. Scientific reports **3**, 1376 (2013)
20. Dey, S., Roy, N., Xu, W., Choudhury, R.R., Nelakuditi, S.: AccelPrint: Imperfections of Accelerometers Make Smartphones Trackable. In: NDSS '14. San Diego, CA, USA (2014)
21. Fu, T.C.: A Review on Time Series Data Mining. Engineering Applications of Artificial Intelligence **24**(1), 164–181 (Feb 2011)
22. Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y.: Deep Learning, vol. 1. MIT press Cambridge (2016)

23. Gozick, B., Subbu, K.P., Dantu, R., Maeshiro, T.: Magnetic Maps for Indoor Navigation. IEEE Transactions on Instrumentation and Measurement **60**(12), 3883–3891 (Dec 2011)
24. Grabocka, J., Schilling, N., Wistuba, M., Schmidt-Thieme, L.: Learning Time Series Shapelets. In: ACM KDD. New York, NY, USA (Aug 2014)
25. Hadsell, R., Chopra, S., LeCun, Y.: Dimensionality Reduction by Learning an Invariant Mapping. In: IEEE CVPR. New York, NY, USA (Jun 2006)
26. Han, J., Owusu, E., Nguyen, L.T., Perrig, A., Zhang, J.: ACcomplice: Location Inference using Accelerometers on Smartphones. In: IEEE COMSNETS (Jan 2012)
27. Haverinen, J., Kemppainen, A.: Global Indoor Self-localization Based on the Ambient Magnetic Field. Robotics and Autonomous Systems **57**(10), 1028–1035 (Oct 2009)
28. Hills, J., Lines, J., Baranauskas, E., Mapp, J., Bagnall, A.: Classification of Time Series by Shapelet Transformation. Data Mining and Knowledge Discovery **28**(4), 851–881 (Jul 2014)
29. Kailath, T.: The Divergence and Bhattacharyya Distance Measures in Signal Selection. IEEE Transactions on Communication Technology **15**(1), 52–60 (1967)
30. Koch, G., Zemel, R., Salakhutdinov, R.: Siamese Neural Networks for One-Shot Image Recognition. In: ICML Deep Learning Workshop. vol. 2. Lille, France (Jul 2015)
31. Kumar, V., Paul, K.: Device fingerprinting for cyber-physical systems: a survey. ACM Computing Surveys **55**(14s), 1–41 (2023)
32. Li, Q., Zheng, Y., Xie, X., Chen, Y., Liu, W., Ma, W.Y.: Mining User Similarity based on Location History. In: ACM SIGSPATIAL. p. 34. ACM, Irvine, CA, USA (Nov 2008)
33. Liao, L., Patterson, D.J., Fox, D., Kautz, H.: Learning and Inferring Transportation Routines. In: AAAI. San Jose, CA, USA (Jul 2004)
34. Michalevsky, Y., Schulman, A., Veerapandian, G.A., Boneh, D., Nakibly, G.: PowerSpy: Location Tracking Using Mobile Device Power Analysis. In: USENIX Conference on Security Symposium. pp. 785–800. Washington, D.C., USA (Aug 2015)
35. Microsoft: Windows Devices: Geolocation. `https://docs.microsoft.com/en-us/uwp/api/Windows.Devices.Geolocation` (2018), accessed: 2024-04-20
36. Mohamed, R., Farrukh, H., Lu, Y., Wang, H., Celik, Z.B.: Istelan: Disclosing sensitive user information by mobile magnetometer from finger touches. Proceedings on Privacy Enhancing Technologies (2023)
37. Mueen, A., Keogh, E., Young, N.: Logical-Shapelets: An Expressive Primitive for Time Series Classification. In: ACM KDD. San Diego, CA, USA (Aug 2011)
38. Narain, S., Vo-Huu, T.D., Block, K., Noubir, G.: Inferring User Routes and Locations Using Zero-Permission Mobile Sensors. In: IEEE S&P. San Jose, CA, USA (May 2016)
39. Nawaz, S., Mascolo, C.: Mining Users' Significant Driving Routes with Low-Power Sensors. In: ACM SenSys. Memphis, TN, USA (Nov 2014)
40. Noulas, A., Musolesi, M., Pontil, M., Mascolo, C.: Inferring interests from mobility and social interactions. In: NIPS Workshop on Analyzing Networks and Learning with Graphs. pp. 2–88 (2009)
41. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V.: Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research **12**(Oct.), 2825–2830 (2011)
42. Perez, B., Musolesi, M., Stringhini, G.: Fatal Attraction: Identifying Mobile Devices Through Electromagnetic Emissions. In: WiSec (2019)

43. Rabiner, L.R.: Digital Processing of Speech Signal. Digital Processing of Speech Signal (1978)
44. Rai, A., Chintalapudi, K.K., Padmanabhan, V.N., Sen, R.: Zee: Zero-effort Crowdsourcing for Indoor Localization. In: ACM MobiCom. Istanbul, Turkey (Aug 2012)
45. Shih, F., Liccardi, I., Weitzner, D.: Privacy Tipping Points in Smartphones Privacy Preferences. In: ACM CHI. pp. 807–816. Seoul, Republic of Korea (Apr 2015)
46. Song, C., Qu, Z., Blumm, N., Barabási, A.L.: Limits of Predictability in Human Mobility. Science **327**(5968), 1018–1021 (2010)
47. Thrun, S., Burgard, W., Fox, D.: Probabilistic Robotics. MIT Press (2005)
48. Vo, Q.D., De, P.: A survey of fingerprint-based outdoor localization. IEEE Communications Surveys Tutorials **18**(1), 491–506 (2016)
49. Wang, H., Sen, S., Elgohary, A., Farid, M., Youssef, M., Choudhury, R.R.: No Need to War-drive: Unsupervised Indoor Localization. In: ACM MobiSys. Low Wood Bay, UK (Jun 2012)
50. Wang, J., Ghosh, R. K.and Das, S.K.: A Survey on Sensor Localization. Journal of Control Theory and Applications **8**(1), 2–11 (Feb 2010)
51. Woodman, O., Harle, R.: Pedestrian Localisation for Indoor Environments. In: ACM UbiComp. Seoul, Korea (Sep 2008)
52. Wu, C., Yang, Z., Liu, Y., Xi, W.: WILL: Wireless Indoor Localization without Site Survey. IEEE Transactions on Parallel and Distributed Systems **24**(4), 839–848 (Apr 2013)
53. Yang, J., Chen, Y.: Indoor Localization Using Improved RSS-Based Lateration Methods. In: IEEE GLOBECOM. Honolulu, HI, USA (Nov 2009)
54. Yapar, Ç., Levie, R., Kutyniok, G., Caire, G.: Real-time outdoor localization using radio maps: A deep learning approach. IEEE Transactions on Wireless Communications **22** (2023)
55. Ye, L., Keogh, E.: Time Series Shapelets: A New Primitive for Data Mining. In: ACM KDD. Paris, France (Jun 2009)
56. Zakaria, J., Mueen, A., Keogh, E.: Clustering Time Series using Unsupervised Shapelets. In: IEEE ICDM. Brussels, Belgium (Dec 2012)
57. Zhang, J., Beresford, A.R., Sheret, I.: Sensorid: Sensor calibration fingerprinting for smartphones. In: IEEE S&P. pp. 638–655 (2019)
58. Zhang, L., Liu, J., Jiang, H., Guan, Y.: SensTrack: Energy-Efficient Location Tracking with Smartphone Sensors. IEEE Sensors Journal **13**(10), 3775–3784 (2013)
59. Zhou, X., Demetriou, S., He, D., Naveed, M., Pan, X., Wang, X., Gunter, C.A., Nahrstedt, K.: Identity, Location, Disease and More: Inferring Your Secrets from Android Public Resources. In: ACM CCS. Berlin, Germany (Nov 2013)