



# Quantifying Intrinsic Value of Information of Trajectories

Kien Nguyen  
kien.nguyen@usc.edu  
University of Southern California  
Los Angeles, California, USA

John Krumm  
jckrumm@microsoft.com  
Microsoft Research  
Redmond, Washington, USA

Cyrus Shahabi  
shahabi@usc.edu  
University of Southern California  
Los Angeles, California, USA

## ABSTRACT

A trajectory, defined as a sequence of location measurements, contains valuable information about movements of an individual. Its value of information (VOI) may change depending on the specific application. However, in a variety of applications, knowing the intrinsic VOI of a trajectory is important to guide other subsequent tasks or decisions. This work aims to find a principled framework to quantify the intrinsic VOI of trajectories from the owner's perspective. This is a challenging problem because an appropriate framework needs to take into account various characteristics of the trajectory, prior knowledge, and different types of trajectory degradation. We propose a framework based on information gain (IG) as a principled approach to solve this problem. Our IG framework transforms a trajectory with discrete-time measurements to a canonical representation, i.e., continuous in time with continuous mean and variance estimates, and then quantifies the reduction of uncertainty about the locations of the owner over a period of time as the VOI of the trajectory. Qualitative and extensive quantitative evaluation show that the IG framework is capable of effectively capturing important characteristics contributing to the VOI of trajectories.

## CCS CONCEPTS

- Information systems → Geographic information systems;
- Human-centered computing → Ubiquitous computing.

## KEYWORDS

location measurements, trajectory, value of information, information gain, measurement uncertainty, location inference

### ACM Reference Format:

Kien Nguyen, John Krumm, and Cyrus Shahabi. 2021. Quantifying Intrinsic Value of Information of Trajectories. In *29th International Conference on Advances in Geographic Information Systems (SIGSPATIAL '21)*, November 2–5, 2021, Beijing, China. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3474717.3483912>

## 1 INTRODUCTION

The availability of mobile devices with location-tracking capability has enabled individuals to generate a great amount of location data from various types of signals, e.g., GPS, Wi-Fi, or cell service. A

trajectory, which is a sequence of location measurements, is an important type of location data that contains valuable information about the owner's locations and movements. For example, a trajectory may indicate the owner's moving behavior [20], e.g., a trip to home, office, or favorite shops and time spent there; it may help identify irregularities such as vacation days or house moving; it is important for traffic, speed, and route inference [21].

Each trajectory gives information of some kind. The value of the information (VOI) varies depending on the use and user of the data. For example, the same trajectory may have one value for the owner and another for an enterprise, and that same trajectory may also have different values for different enterprises, such as a public or private, especially for an ad-targeting company.

However, a trajectory also contains an intrinsic, formulaic VOI, since it contains quantified locations of an individual over space and time. For example, it is reasonable to assume that a home-to-office trajectory with 1000 measurements from beginning to end of the trip tends to have more information than another home-to-office trajectory with only 2 measurements at home and office. Our goal is to *find a principled approach to quantify the intrinsic VOI of trajectories from the owner's perspective*. To the best of our knowledge, this is the first attempt at such a quantification.

Given a trajectory, there are natural questions about selling it [10], sharing it, using it to train machine learning, storing it, or examining it more closely. Quantifying its VOI helps measure how valuable, revealing, informative [19], distinct, and surprising [21] it is. A trajectory's VOI can be a critical piece of metadata that indicates its intrinsic value for a variety of tasks, both stand-alone and as part of a collection.

This problem presents several interesting challenges:

- First, a trajectory has many characteristics contributing to its VOI, and effectively capturing their complex relationships is non-trivial. Examples of these characteristics are the number of measurements, temporal duration, and how measurements distribute spatially and temporally.
- Second, a trajectory can be degraded for different purposes, e.g., measurements can be perturbed by adding random noise or completely removed to enhance the owner's privacy before releasing or selling. It is challenging to capture the effect of different types of degradation on the VOI of the trajectory.
- Third, the owner may assume different prior knowledge about a trajectory, which may change its VOI from the owner's perspective. For example, if the owner has never released any trajectory data before, meaning their trajectory would give more information about their locations than if they already released a perturbed version of the trajectory.

Some straightforward methods, e.g., using a trajectory's characteristics such as size or duration, fail to capture other characteristics or prior knowledge. Previous work based on Spatial Privacy

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*SIGSPATIAL '21*, November 2–5, 2021, Beijing, China

© 2021 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-8664-7/21/11...\$15.00

<https://doi.org/10.1145/3474717.3483912>

Pricing [12] can be adapted to sum the values of degraded measurements of the trajectory. However, it also fails to capture some characteristics (e.g, how measurements distribute over space) and prior knowledge, because it ignores the mutual information of the constituent points. The notion of correctness [17] also appears promising. For this method, based on a degraded version  $Z$  of a trajectory  $S$  and prior knowledge, a probabilistic prediction is made for each measurement. Then the correctness of the prediction, indicating how close it is to the actual measurement, is aggregated to derive the VOI of  $Z$ . However, this method is not applicable when  $Z = S$ , because correctness is not available when there is no actual measurement with which to compare the prediction.

The core idea behind our approach is that instead of computing the VOI of a trajectory from its discrete location measurements over time, we view VOI as how much the trajectory data helps reconstruct locations of the owner. More specifically, the VOI corresponds to how much the trajectory data helps reduce the uncertainty of estimating its owner’s locations in continuous time. We use information gain (IG) to quantify that reduction. Even though IG is a known measure [14], it is not obvious how it should be used to compare discrete-time trajectories with widely different characteristics: long and short, dense and sparse, clean and noisy. Our innovation here is that we transform each trajectory to a canonical representation, i.e., continuous in time with continuous estimates of mean and variance, which is necessary for comparing widely different trajectories and comparing new data against prior data. IG can only then be applied after all those steps are taken. We realize this transformation by employing a reconstruction method that can produce continuous-time probabilistic predictions along the trajectory.

Consequently, we propose an IG framework as a principled way to quantify the intrinsic VOI of a trajectory. The main idea is to quantify the reduction of uncertainty about the owner’s continuous-time locations, comparing a new trajectory to a previously released degraded version or to prior information. Thus the IG framework utilizes a reconstruction method to produce continuous probabilistic predictions over time, and then it calculates the reduction of uncertainty from those predictions compared to a reconstruction from prior knowledge. The uncertainty is measured by (differential) entropy. Our IG framework accepts any reasonable probabilistic reconstruction method. Gaussian process is used as the reconstruction method in this paper due to its popularity and flexibility, but is not necessarily the only choice.

Specifically, our contributions are:

- Propose the problem of quantifying intrinsic VOI of a trajectory from the owner’s perspective
- Define characteristics that should be captured by an appropriate quantification method
- Show how alternate methods fail, even when they appear reasonable initially
- Introduce a method to transform each trajectory to a canonical representation, allowing us to examine the VOI with a continuous location reconstruction and to compare trajectories with widely different characteristics
- Develop an IG framework over transformed trajectories as a principled approach capable of effectively capturing various trajectory characteristics, prior knowledge, and degradation

- Evaluate the proposed framework both qualitatively and quantitatively with extensive experiments on a large, real-world trajectory dataset

By providing a standard, comprehensive method for assessing the VOI of a trajectory, we enable a deeper understanding of trajectories and their utility for a wide variety of purposes.

## 2 PROBLEM SETTING

This section introduces the problem of quantifying the intrinsic VOI of a trajectory from the owner’s perspective. We focus on the most basic form of a trajectory, which is a sequence of location measurements. Therefore, incorporating other information, such as census data or points of interest, is beyond the scope of this work.

The owner can be anyone having access and right to use the data, e.g., one whose phone recorded this trajectory, or a data collector who aggregates location data from individuals. Without loss of generality and to ease the discussion, from here on, we assume the owner is the individual whose device recorded the trajectory.

A trajectory  $S$  is a sequence of (potentially noisy) location measurements  $S = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{|S|}\}$ . Each measurement or data point  $\mathbf{x}$  (bold symbol) is a tuple  $\mathbf{x} = \langle lon, lat, t, \sigma \rangle$  where  $\mathbf{x}.lon$  and  $\mathbf{x}.lat$  are the longitude and latitude,  $\mathbf{x}.t$  is the timestamp, and  $\mathbf{x}.\sigma$  is the accuracy or uncertainty of the measurement. It is reasonable to assume Gaussian noise for location measurements [5] such as GPS points, so  $\mathbf{x}.\sigma$  can be considered as the standard deviation of independent Gaussian noise of longitude and latitude. Measurements in  $S$  are ordered by their timestamps, i.e.,  $\forall \mathbf{x}_i, \mathbf{x}_j \in S, 1 \leq i < j \leq |S|, \mathbf{x}_i.t \leq \mathbf{x}_j.t$ .

The owner can assume the potential recipient has some prior knowledge  $\Omega$  about the location  $\mathbf{x}_i$ , e.g., by gathering public data or from some data already released from the owner before. This prior knowledge is represented as a prior distribution  $P(\mathbf{x}_i|\Omega)$  and should be respected by the proposed methods. Intuitively, prior knowledge about a point would tend to reduce the VOI of the point.

In addition to releasing raw trajectories, owners can also offer their data at different quality levels, potentially enhancing their privacy while reducing the trajectory’s VOI, e.g, a trajectory can be degraded by adding noise or subsampling. More details about degradation are discussed in Section 2.3. The methods to quantify the VOI also need to reflect this quality degradation.

Our goal is to find a principled approach to quantify the VOI of a trajectory from the owner’s perspective. It should capture different characteristics of the raw trajectory, the prior knowledge, and the effect of various degradation processes. Next, we described the characteristics that an appropriate method should aim to capture. This is an extensive but not exhaustive list of desirable characteristics, thus we expect future work will study more characteristics. We will then illustrate how several baseline methods fail to capture these characteristics and how our new proposed framework succeeds.

### 2.1 Trajectory Characteristics

From the raw trajectory  $S$ , certain characteristics can be derived that should contribute to its VOI. These includes *size*, *duration*, *spatial distribution*, *temporal distribution*, and *measurement uncertainty*.

**Size.** The size  $|S|$  is the number of measurements of  $S$ . A larger size often means there is more data to learn about the owners’

locations and movements. Therefore, it is reasonable to assume that a trajectory with a larger size tends to have more information.

**Duration.** The duration of a trajectory  $S$  is the time difference between the first and last measurements of  $S$ , i.e.,  $D(S) = x_{|S|.t} - x_{1.t}$ . Similar to the size, a trajectory with a longer duration often means it may have more information about the owner's locations than one with a shorter duration.

**Spatial distribution.** The spatial distribution of  $S$  indicates how its measurements are distributed over space. For example, one trajectory may have all measurements at one place, while another trajectory may have measurements at multiple places (e.g., solely at home vs. at home, then at a coffee shop, then at the office). A trajectory visiting more places tends to give more information about locations of the owner. The spatial distribution of  $S$  can be measured using *spatial entropy*  $H_S(S)$  over a grid of cells, say  $10m \times 10m$ , covering the area of interest.  $H_S(S)$  is calculated by creating a histogram of number of measurements of  $S$  belonging to each cell, converting this histogram to probabilities, and then computing the Shannon entropy for the probabilities.

**Temporal distribution.** Similarly, the temporal distribution of a trajectory  $S$  indicates how its measurements are distributed over time. For example, one trajectory may have many measurements in the first few minutes and then a long gap before the next one. Another trajectory may be distributed more uniformly over time, e.g., one measurement every 30 seconds. Intuitively, when measurements distribute more evenly over a longer period of time, they tend to give more information. The temporal distribution of  $S$  can be measured using the *temporal entropy*  $H_T(S)$  calculated during  $[x_{1.t}, x_{|S|.t}]$  by computing a histogram of the number of measurements of  $S$  belonging to each temporal bin, say every 1 minute, from  $x_{1.t}$  to  $x_{|S|.t}$ , converting the histogram to probabilities, and then computing the Shannon entropy for the probabilities.

**Measurement uncertainty.** Measurements may have their own uncertainty depending on how they were taken, e.g., an uncertainty of several hundred meters with cell towers [2] or 1 to 5 meters for GPS-enabled smartphones [13]. Since Gaussian noise is a reasonable assumption for GPS [5], measurement uncertainty is defined as the standard deviation  $\sigma$  of independent Gaussian noise of longitude and latitude. In general, less accurate measurements (i.e., larger  $\sigma$ ) tend to give less information about locations of the owner.

## 2.2 Prior Knowledge

As mentioned, the owner can assume there exists some available prior knowledge about the location  $x_i$ , e.g., from public data or from previous releases from this owner. This knowledge is expressed as a prior distribution  $P(x_i|\Omega)$  for each timestamp of interest. In general, better prior knowledge often means the trajectory gives less information compared to the case with poorer prior knowledge.

## 2.3 Trajectory Degradation

The owner can offer their trajectories at lower quality. The process of lowering the quality of a trajectory is called *degradation*. The reason for the owner to produce degraded trajectories is that a recipient can still benefit from data at a certain quality depending on their specific applications [12]. For example, estimating a neighborhood-level origin-destination matrix may not need extremely accurate

measurements, thus the recipient with this application can potentially spend less to purchase lower quality trajectories. In general, a lower quality trajectory would give less information than a higher quality one. A degraded version of a trajectory  $S$  is denoted as  $Z$ .

In this work, we consider three types of degradation: *perturbation*, *truncation*, and *subsampling*. Other types of degradations and other variations of these degradations are beyond the scope of this paper and considered as part of future work.

**Perturbation.** The perturbation process degrades a trajectory  $S$  by adding independent random Gaussian noise with standard deviation  $\sigma_z$  to each  $x_i \in S$ . Formally, a perturbed trajectory  $Z$  of  $S$  consists of measurements  $z_i$  s.t.

$$z_i.lon = x_i.lon + \eta_1 \quad (1)$$

$$z_i.lat = x_i.lat + \eta_2 \quad (2)$$

$$z_i.t = x_i.t \quad (3)$$

$$z_i.\sigma = \sqrt{(x_i.\sigma)^2 + \sigma_z^2} \quad (4)$$

$$\eta_1, \eta_2 \sim \mathcal{N}(0, \sigma_z^2) \quad (5)$$

Noise magnitude  $z_i.\sigma$  is called the *total noise* of  $Z$ .

**Truncation.** The truncation process degrades a trajectory  $S$  by truncating  $S$  and only keeping a fraction  $\alpha_t \in [0, 1]$  of the first measurements of  $S$ . For example, if  $|S| = 100$ , an  $\alpha_t = 0.2$  means that the temporally first 20% of measurements are kept, which are  $\{x_1, \dots, x_{20}\}$ . The fraction  $\alpha_t$  is called the *truncation ratio*. The retained measurements are kept in their raw form. Also, there is at least one measurement retained.

**Subsampling.** Similarly, the subsampling process degrades a trajectory  $S$  by uniformly subsampling  $S$  with probability  $\alpha_s \in [0, 1]$ , called the *subsampling ratio*. More specifically, a measurement  $x_i \in S$  is retained with probability  $\alpha_s$ ; otherwise,  $x_i$  is discarded. Hence, in expectation, a fraction  $\alpha_s$  of measurements of  $S$  are retained. The retained measurements are also kept in their raw form and at least one measurement is retained.

## 3 BASELINES

This section discusses baseline techniques to quantify the VOI of a trajectory. Each baseline is described along with which desirable characteristics from Section 2 they can capture. Table 1 summarizes which characteristics the baselines and our proposed method can faithfully represent. Our proposed method, based on information gain and shown in the last row of Table 1, is described in detail in Section 4. A cross (✗)/exclamation mark (⚠)/checkmark (✔) indicates that the method in that row cannot/can partially/can fully capture the characteristic in that column, respectively.

The following analysis shows how the baseline methods, while appearing initially reasonable, fail to represent some important characteristics of the VOI. For each characteristic, a method is first evaluated qualitatively. In some cases, it is obvious that a method can or cannot capture a characteristic, e.g., a size-based method can capture the size but not prior knowledge. In other cases, it may not be as clear. In those cases, the method is said to be capable of capturing the characteristic if there is a strong correlation between the output of the method and the characteristic. The correlation is examined using Spearman's rank correlation coefficient  $\rho$ , which quantifies strictly monotonic relationships between two variables

	Size	Duration	Spatial Distribution	Temporal Distribution	Measurement Uncertainty	Prior Knowledge	Degradation
Fixed Value	✗	✗	✗	✗	✗	✗	✗
Size-based	✓	⚠	✗	✗	✗	✗	⚠
Duration-based	⚠	✓	✗	✗	✗	✗	⚠
Travel Distance	⚠	⚠	⚠	⚠	✗	✗	⚠
Entropy-based	⚠	⚠	✓*	✓*	✗	✗	⚠
SPP-based [12]	✓	⚠	✗	✗	✓	✗	✓
Correctness-based [17]	⚠	⚠	⚠	⚠	⚠	✓*	✓*
Information Gain	✓	✓	⚠	✓	✓	✓	✓

**Table 1: Potential methods to quantify the intrinsic VOI of a trajectory and the characteristics they can or cannot capture. Our proposed method based on information gain can capture almost all of the desirable characteristics.**

and is relatively robust against outliers [16]. An absolute magnitude  $|\rho|$  in  $[0.4, 0.7)$  and  $[0.7, 1]$  indicates moderate and strong correlation, respectively. The cutoff points are based on previous work [16]. The values of  $\rho$  are calculated from a large real-world trajectory dataset described in Section 5.1.

### 3.1 Fixed Value Method

One potential method is to set the same value for all trajectories. While this method is straightforward and, in fact, was used in some surveys about values of location data from a seller’s perspective [4, 18], this method ignores the fact that each trajectory may contain a different VOI. For example, a 40-mile long commute from home to work can be very different from a short trip to a nearby store. Hence, this method does not capture any of the desirable characteristics.

### 3.2 Size-based Method

The size-based method uses the size  $|S|$  of a trajectory  $S$  to quantify the VOI of  $S$ . It is reasonable to say that a trajectory with more measurements tends to have more information. While this method may distinguish trajectories with many or few measurements, the size alone would fail to fully capture the VOI of a trajectory, because the size  $|S|$  depends heavily on the sampling rate and the duration. For example, the same trip from home to office if sampled every 1 second would have a size 5 times larger than if sampled every 5 seconds, while having roughly similar information about the locations of the person along the trip.

Consequently, the size-based method successfully captures the size characteristic of a trajectory, but not other aspects. For the duration characteristic, because it only depends on the first and last measurements of the trajectory, the size cannot fully capture it, e.g., a trajectory with two measurements can be arbitrarily short or long. However, the size can partially capture the duration if the sampling rate is relatively similar among trajectories. The correlation coefficient  $\rho$  between size and duration of trajectories in our dataset is 0.86, which indicates a strong correlation between them.

The size  $|S|$  also does not capture the spatial nor temporal distribution of  $S$ , because it does not indicate how the measurements are distributed. The same number of measurements can happen at nearly the same place/time if the sampling rate is high, or at

different places/times if the sampling rate is low. It is also clear that  $|S|$  does not indicate measurement uncertainty nor prior knowledge. In fact, any method using solely trajectory characteristics would fail to capture prior knowledge because prior knowledge is not considered in that method. The size can capture some degradation such as truncation, but not other degradation such as perturbation.

### 3.3 Duration-based Method

Another method is to use the duration  $D(S)$  of  $S$ , i.e.,  $D(S) = x_{|S|,t} - x_{1,t}$ , to quantify its VOI, as it is reasonable to assume that a temporally longer trajectory tends to have more information. However, using only duration would fail to fully capture the VOI since it ignores all information between the start and end points.

Similar to the size-based method explained before, this method can fully capture the duration but can only partially capture the size. The duration also does not represent how the measurements distribute over space and time, and is unable to take into account measurement uncertainty and prior knowledge. Duration can capture the effect of truncation but not that of perturbation nor subsampling, thus, only partially capturing the effects of degradation.

### 3.4 Travel Distance Method

This method computes the travel distance  $L(S)$  of  $S$  by summing the distances between each pair of consecutive measurements. It is reasonable to assume that a trajectory with longer distance tends to have more information than a shorter one.

With the same mode of transportation (e.g., with bikes or cars), travel distance likely reflects the duration. Thus, travel distance can partially capture duration and the size of  $S$ . With a longer distance, the trajectory tends to go through more places and time period. In our dataset, a strong correlation 0.89 and 0.78 between the travel distance and spatial and temporal entropy indicate that travel distance can partially capture the spatial and temporal distribution.

However, using travel distance would fail take into account measurement uncertainty or any prior knowledge. It can capture some degradation, e.g., truncation, but not others, thus, only partially capturing the effect of degradation.

### 3.5 Entropy-based Method

The spatial and temporal entropy  $H_S(S)$  or  $H_T(S)$  of  $S$ , computed from the Shannon entropy of the probabilities converted from the histogram of number of measurements of  $S$  belonging to each grid cell or temporal bin, can be good candidates to quantify the VOI given its extensive applications in information theory. A higher spatial/temporal entropy likely indicates that the trajectory gives more information about a person's locations over space/time.

Since  $H_S(S)$  or  $H_T(S)$  are used to measure the spatial and temporal distribution of  $S$ , respectively, a method combining both  $H_S(S)$  or  $H_T(S)$  can capture these two characteristics. However, it is unclear how they should be combined, which is why there are asterisks for the entropy-based row in Table 1.

While  $H_S(S)$  and  $H_T(S)$  do not fully capture the size and duration (e.g., a trajectory with only two points but far apart from each other can have a long duration but low entropy values), when  $S$  has a relatively high sampling rate,  $H_T(S)$  would be highly correlated with the size and duration. In our dataset, the correlation coefficients between  $H_T(S)$  and size and duration are 0.89 and 0.97, respectively, which indicate a very strong correlation.

However, since entropy reflects the uncertainty, when the measurements have higher uncertainty (i.e., larger  $x.\sigma$ ), both  $H_S(S)$  and  $H_T(S)$  tend to increase. This is opposite of what one might expect, because a more uncertain measurement means less information. Thus, entropy does not capture measurement uncertainty nor degradation such as perturbation. Computing  $H_S(S)$  and  $H_T(S)$  also ignores all prior knowledge. Another issue is that the entropy can change significantly when arbitrary parameters for computing entropy (i.e., size of grid cells or length of time bins) change.

### 3.6 Spatial Privacy Pricing-based Method

This method is based on the previous work on Spatial Privacy Pricing [12] (SPP) where each measurement has the same value defined by the owner, and the value can be reduced when the measurement is perturbed by noise. The VOI of the trajectory is then calculated by summing up the values of each individual measurement.

With the ability to change the value based on the noise in the measurements, this method can capture the measurement uncertainty and represent degradation. Summing the individual values means this method is similar to the size-based method, but is also sensitive to the perturbation parameters. Therefore, this method has similar characteristics as the size-based method, which means fully capturing size, partially capturing duration, and unable to capture spatial and temporal distributions and prior knowledge.

### 3.7 Correctness-based Method

This method is based on the *correctness* of reconstructing the raw measurements from available data [17]. Roughly speaking, when calculating the VOI of a degraded version  $Z$  of  $S$ , the owner can assume the recipient is attempting to reconstruct each measurement  $x_i \in S$  based on  $Z$ . The correctness is the expected error between the actual points  $x_i$  and the probabilistically reconstructed points  $\hat{x}_i$ . If the expected error is lower, then it would be reasonable to assume a higher VOI of  $Z$ .

While this concept was proposed for a different problem setting, some techniques can be used to adapt it to our problem. In fact,

our proposed framework, discussed later, also has a reconstruction step that can capture prior knowledge and degradation. Thus a correctness-based method can capture the prior knowledge and degradation. The asterisks in the correctness-based row in Table 1 indicate that a correctness-based method needs some modifications for the problem setting and techniques to be fit for our problem.

The main drawback of using correctness is that it requires some ground-truth measurements of  $S$  being available to evaluate the correctness of the reconstructed trajectory made from the degraded version  $Z$  and prior knowledge. So, it cannot measure the VOI of the full, raw trajectory  $S$ , because there is no ground-truth measurement exists to evaluate correctness in that case. Consequently, since the full, raw trajectory cannot be fully captured, which means there are often some characteristics not fully captured (e.g., subsampling changes the size and truncation changes the duration), this method can only partially capture characteristics of a trajectory.

Another issue is that this method relies on the correctness of discrete-time predictions, and it is unclear how the correctness of each prediction should be aggregated to obtain the correctness of the whole trajectory. For example, the correctness derived from an 80% subsampled version is evaluated on each measurement of the remaining 20% data, while the correctness derived from a 60% subsampled version is evaluated on the remaining 40%. It is unclear how the correctness should be modified to reasonably quantify both, and how the correctness of each prediction should be aggregated.

### 3.8 Other Potential Quantities

There are other potential quantities contributing to the VOI of a trajectory. However, these are not considered as baselines because they are either orthogonal (i.e., they can be used in conjunction with other methods) or complicated (i.e., finding these quantities requires techniques beyond the scope of this work). Examples are *time period*, *subjective sensitivity*, and *visits*.

**Time period.** The time period that a trajectory  $S$  was taken, e.g., weekdays or weekend, can be a factor contributing to its VOI. For example, a person may often have commute-related trajectories during weekdays but more leisure-related trajectories on weekends. While time period does not represent actual locations, it can be used in conjunction with other methods, e.g., an owner can have different values for trajectories during weekdays compared to weekends because of privacy concerns.

**Subjective sensitivity.** Each person may have their own sensitivity for different types of locations, which may lead to different sensitivity for different trajectories. For example, one may feel their workplace is more sensitive than their favorite coffee shop, thus having a higher sensitivity for the trajectory from home to work than the one to the coffee shop. While this information can contribute to the VOI of a trajectory, incorporating it requires additional information about the location measurements, which is not the focus of this work. When this information is available, it can be used in conjunction with the proposed methods in this work to better quantify the VOI tailored to the owner's subjective reasoning.

**Visits.** This method bases on the number of visits of a trajectory to quantify the VOI. Its main drawback is how to define a visit. It is often not feasible for the owner to manually define all visits for all of their trajectories. On the other hand, complicated techniques

to automatically find visits [20] are often used to segment a long sequence of measurements into trajectories, which is beyond the scope of this work. These techniques also often require additional information and/or a complex set of parameters where a small change of some parameters may result in a significantly different number of visits, which is not desirable.

## 4 THE INFORMATION GAIN FRAMEWORK

This section describes the proposed framework to quantify the VOI of a trajectory. The framework is based on the notion of *information gain (IG)* which quantifies the reduction of uncertainty when new information is available. Adopting IG to trajectories is not straightforward, and we enable it by transforming each trajectory to a canonical representation. This transformation is done by employing a reconstruction method that can produce continuous-time probabilistic predictions along the trajectory. The Gaussian process (GP) is used in this work as the reconstruction method and discussed in the next section, but we emphasize that our IG framework accepts any reasonable method of producing probabilistic location inferences.

### 4.1 Trajectory Information Gain

We define information gain  $IG_T(Z, \Omega)$  of a degraded version  $Z$  of a trajectory  $S$  as the total reduction of uncertainty about locations of the owner over time period  $T$  compared to the uncertainty from the prior knowledge  $\Omega$ . For a typical human trajectory, there are potentially several reasonable choices for  $T$ . In this work,  $T$  is defined as the entire day covering the trajectory, because a typical trajectory would not extend beyond a day. We propose  $IG_T(Z, \Omega)$  to quantify the intrinsic VOI of  $S$  and/or its degraded version  $Z$ . This section describes how  $IG_T(Z, \Omega)$  is derived. In Section 5, we will explain how  $IG_T(Z, \Omega)$  satisfies almost all of the criteria in Table 1, making it a better choice than the previous methods we described.

Recall that the owner can assume that the recipient obtained some prior knowledge  $\Omega$ , e.g., from public data or from previous noisier release of the same trajectory  $S$ . From  $\Omega$ , the owner can assume that the recipient can derive a prior probability distribution  $P(\mathbf{x}|\Omega)$  for the location of the owner at a specific time  $\mathbf{x}.t$ .

The owner can then assume that after receiving  $Z$ , the recipient can use a model to reconstruct (or predict/interpolate) locations of the owner in continuous time. This assumption is made, because without knowledge of the recipient's intent, the owner should act conservatively and assume the recipient will exploit the new data fully, such as with a maximally accurate reconstruction. Such an inference can be represented as a posterior distribution  $P(\mathbf{x}|Z, \Omega)$ .

Subsequently, the information gain  $IG_t(Z, \Omega)$  at timestamp  $t = \mathbf{x}.t$  indicates the reduction of uncertainty from  $P(\mathbf{x}|\Omega)$  to  $P(\mathbf{x}|Z, \Omega)$ . The uncertainty is measured by *differential entropy* [3] (or continuous entropy), because both the prior and posterior distributions are likely continuous distributions in space.

The differential entropy  $h(X)$  of a random variable  $X$  with probability density function  $P$  whose support is a set  $\mathcal{X}$  is defined as

$$h(X) = - \int_{\mathcal{X}} P(x) \log P(x) dx \quad (6)$$

Several popular probability distributions have a closed-form expression for their differential entropy, e.g. if a  $X$  is a Gaussian random

variable with distribution  $\mathcal{N}(\mu, \sigma^2)$ , its differential entropy is

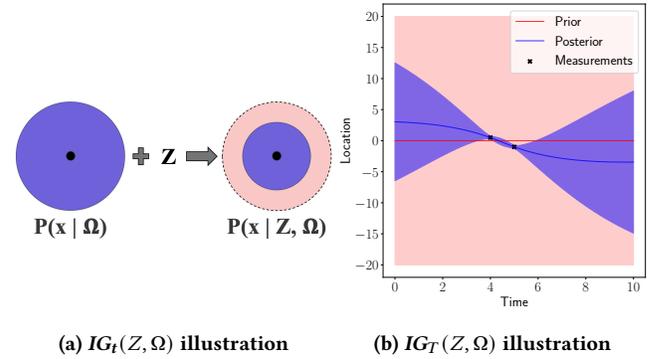
$$h(X) = \frac{1}{2} \log 2\pi e \sigma^2 \quad (7)$$

The entropy values of latitude and longitude of  $\mathbf{x}$  are calculated separately using Equation 6, and summed to get  $h(\mathbf{x}|\Omega)$  and  $h(\mathbf{x}|Z, \Omega)$ .

Thus the IG at time  $t$ ,  $IG_t(Z, \Omega)$ , can be calculated as

$$IG_t(Z, \Omega) = h(\mathbf{x}|\Omega) - h(\mathbf{x}|Z, \Omega) \quad (8)$$

Figure 1a illustrates  $IG_t(Z, \Omega)$  where the black dot shows the prediction mean and the blue area shows the uncertainty as the circle with a radius which is twice the standard deviation of the predicted distribution  $P(\mathbf{x}|\Omega)$ . If  $Z$  is available, the uncertainty is reduced to the smaller blue circle on the right representing  $P(\mathbf{x}|Z, \Omega)$ .  $IG_t(Z, \Omega)$  quantifies the reduction and is illustrated as the red ring.



**Figure 1: Illustrations of information gain (a) at a single timestamp and (b) over a time period. The reduced uncertainty is shown in the red area minus the blue area.**

Note that  $\mathbf{x}$  does not need to be in any trajectory, because the probabilistic reconstruction is continuous in time. For example,  $\mathbf{x}$  can be in between two measurements  $\mathbf{x}_i, \mathbf{x}_{i+1} \in S$ , thus, knowing  $\mathbf{x}_i$  and  $\mathbf{x}_{i+1}$  would help reduce the uncertainty of where  $\mathbf{x}$  can be.

Subsequently, the information gain  $IG_T(Z, \Omega)$  over a time period  $T$  can be calculated by integrating  $IG_t(Z, \Omega)$  for all time  $t \in T$ , i.e.,

$$IG_T(Z, \Omega) = \int_{t \in T} IG_t(Z, \Omega) dt \quad (9)$$

Figure 1b illustrates  $IG_T(Z, \Omega)$  computed over a time period  $T = [0, 10]$  with prior  $P(\mathbf{x}|\Omega) = \mathcal{N}(0, 100)$  for each timestamp. Over this period, the red line shows the prediction mean of the prior distributions, and the red area shows the prior uncertainty. When  $Z$ , shown as two black crosses, is available, the uncertainty is reduced to the blue area. The amount of such reduction is shown as the red area minus the blue area and is quantified using  $IG_T(Z, \Omega)$ .

### 4.2 Reconstruction Method

To compute  $IG_T(Z, \Omega)$ , a probabilistic reconstruction method is needed to reconstruct locations of the owner given  $\Omega$  and/or  $Z$ . A Gaussian process (GP) is used in this work as the reconstruction method because of its flexibility to incorporate different types of information. However, we emphasize that reconstruction is not the

focus of this paper. The IG framework accepts any reasonable probabilistic reconstruction method. Thus, we briefly discuss important aspects of the GP. More details about GPs can be seen in [15].

For a scalar function  $f(t)$ , a GP implies that any subset of points sampled from  $f(t)$  is distributed according to a multidimensional Gaussian. Two independent GPs are created for longitude and latitude prediction. The input for a GP are pairs  $\langle t, f(t) \rangle$  where  $f(t)$  is longitude or latitude of the measurement at time  $t$ . The output for a set of timestamps  $\mathcal{T}^*$  are predictions  $f^*(t)$  for each  $t \in \mathcal{T}^*$ .

A GP depends on a scalar covariance kernel/function  $k(t, t')$  defining how much a measurement at time  $t$  correlates with a measurement at time  $t'$ . In general, the correlation decreases to zero as  $|t - t'|$  gets larger. A good kernel can help incorporate different types of information into the model, especially when kernels can be combined together. In our implementation, two common kernels are summed together: the main kernel is a Matérn kernel that captures the relationship between measurements as well as the prior knowledge, and a white kernel to capture measurement uncertainty.

The formula of the Matérn kernel used in this work is

$$k_{\text{Matérn}}(t, t') = \sigma_f^2 \left( 1 + \frac{\sqrt{3}}{l} |t - t'| \right) \exp \left( - \frac{\sqrt{3}}{l} |t - t'| \right) \quad (10)$$

Setting  $\sigma_f$  to the standard deviation of the prior distribution can help capture the prior knowledge, so that when the model becomes more uncertain, the standard deviation of the prediction will gradually reach this value. The length scale  $l$  is trained from data.

The formula of a white kernel is

$$k_{\text{white}}(t, t') = \begin{cases} \sigma_m^2 & \text{if } t = t' \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

Setting  $\sigma_m$  to  $x.\sigma$  or  $z.\sigma$  helps capture the measurement uncertainty.

The final kernel is

$$k(t, t') = k_{\text{Matérn}}(t, t') + k_{\text{white}}(t, t') \quad (12)$$

Finally, a GP also depends on a mean function  $m(t)$  which defines the expected mean values of the measurements. In our adaption, the mean function  $m(t)$  of a GP is set to be the regression line obtained by running a linear regression on the data in  $\Omega$ .

## 5 EVALUATIONS OF THE IG FRAMEWORK

This section provides an evaluation of how the IG framework can capture each characteristic from Table 1, both qualitatively and quantitatively. The quantitative evaluation is performed on the Geolife dataset, which is a large, real-world trajectory dataset.

### 5.1 Dataset

The Geolife dataset [22] is used for experiments to quantitatively justify claims in the paper. This is a trajectory dataset collected by 182 people carrying GPS loggers and GPS-phones in the Beijing area from April 2007 to August 2012. A trajectory is represented by a sequence of measurements containing the latitude, longitude, and timestamp, with a variety of sampling rates. With this large, real-world dataset, covering a large span in both space and time, different types of devices, and a variety of sampling rates, we expect it is representative of the observations in other real-world datasets.

Measurements in the Geolife dataset are filtered to retain only ones within the Beijing area with longitude from 116.20 to 116.55

degrees and latitude from 39.80 to 40.06 degrees. This area covers almost all measurements and, in total, more than 16 million measurements were retained from all 182 people. The latitude/longitude coordinates in each measurement are converted to local Euclidean coordinates  $(x, y)$  in meters with the reference origin  $(lon_0, lat_0)$  arbitrarily chosen as the center of the aforementioned area. In local Euclidean coordinates, the area has a lower left coordinate of  $(-15000, -15000)$  and an upper right coordinate of  $(15000, 15000)$ .

Measurements of each individual are separated into trajectories by iterating through the ordered measurements and creating a new trajectory whenever the time gap between the current measurement and the next measurement is more than  $\tau$  seconds. The maximum time gap  $\tau$  is arbitrarily chosen at 300 seconds or 5 minutes. The actual value of  $\tau$  does not significantly change the experimental results. We also emphasize that finding trajectories is not the focus of this work. Thus, more sophisticated methods to find trajectories from measurements can also be used as an alternative to this segmentation approach. The final dataset has 45,831 trajectories.

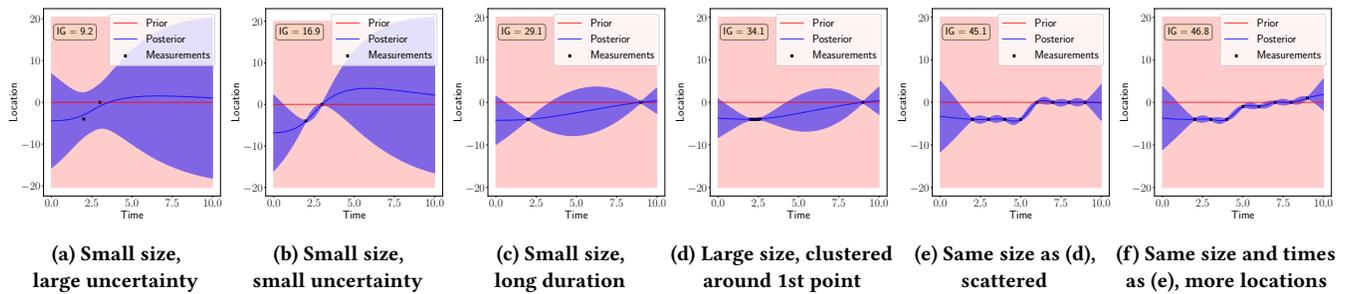
### 5.2 Experiment Setup

The experiments were conducted on the aforementioned dataset with various sets of parameters. The results, such as the correlation coefficient or regression line, are computed from all trajectories. Several outliers are removed for visual presentation purpose but still included in all computations. The default size of grid cells/temporal bins to calculate spatial/temporal entropy is 10 meters/one minute.

The raw measurements do not include uncertainty. However, because they were recorded by GPS-equipped devices and the noise for GPS-equipped smartphones is about 3 meters [13], each measurement is assumed to have  $x_i.\sigma = 3$ . For perturbation degradation, the noise is added such that the total noise (i.e.,  $z_i.\sigma$ ) is in  $\{3, 10, 100, 200, 300, 400\}$  meters. The ratios for truncation and subsampling degradation (i.e.,  $\alpha_t$  and  $\alpha_s$ ) are  $\{0.8, 0.6, 0.4, 0.2, 0.05\}$ , which mean  $\{80\%, 60\%, 40\%, 20\%, 5\%\}$  of the raw trajectory. The subsampling is implemented so that the subsampled data with a higher ratio is a superset of the subsampled data with a smaller ratio.

A measurement is assumed to be inside the Beijing area. This choice is reasonable for this dataset. However, it is not crucial to the framework where any reasonable prior knowledge can be used (e.g., measurements can be anywhere on Earth) and would only change the scale of IG. This prior knowledge is represented by a Gaussian distribution with extremely high variance where the mean is at the center of the area (i.e., coordinates  $(0, 0)$ ) and the standard deviation is  $\sigma_0 = 7500$  meters (i.e., the distance from the center to an edge of the area is twice this standard deviation). This is an *uninformative* prior and called the *Gaussian prior*, i.e., for each timestamp, the location is assumed to be  $\mathcal{N}(0, \sigma_0^2)$  for each dimension.

It is also natural to consider any previously released data as prior knowledge when computing the VOI of the data the owner still retains. So we also consider previous releases as priors. This also demonstrates an advantage of the IG framework, where any previous release can be naturally considered when computing the VOI of the owner's remaining data. To illustrate these cases, depending on the nature of a degradation, different *informative* priors based on previous releases are considered. For perturbation, two additional priors are *400m noise* and *300m noise* priors, illustrating the cases



**Figure 2: An illustration of how IG can capture trajectory characteristics. The figures from left to right showing the effect of measurement uncertainty, duration, size, temporal, and spatial distribution of a trajectory.**

where the owner released their trajectories at those noise levels before. For example, with the 400m noise prior, when computing the IG for a degraded version  $Z_{100}$  with 100m noise of a trajectory  $S$ , the owner may consider the prior knowledge is the degraded version  $Z_{400}$  with 400m noise of  $S$ . Thus, there are three priors for perturbation in total: the uninformative *Gaussian prior* and two informative priors 400m noise and 300m noise. Similarly, there are three priors for truncation/subsampling: the uninformative *Gaussian prior* and two informative 5% trajectory and 20% trajectory priors, illustrating the cases when the owner released trajectories with 0.05 or 0.2 truncation/subsampling ratio before. Other prior knowledge, e.g., previously purchased trajectories, can also be considered as priors, provided that they can be expressed as prior location distributions.

The GP is trained on the prior  $\Omega$  (e.g., 300m prior), except when the prior is the uninformative Gaussian prior, when it is then trained on the new degraded version  $Z$ . The degraded version  $Z$  and prior  $\Omega$  are combined and then provided to the GP. For perturbation, measurements  $Z$  and  $\Omega$  of each timestamp are combined using inverse-variance weighting [7]. For truncation and subsampling, because the higher-ratio degraded version is always a superset of a lower-ratio degraded version (e.g.,  $Z_{5\%} \subset Z_{20\%} \subset Z_{40\%}$ ), the combined data is  $Z$ . This is a reasonable approach for GP. There can be other methods to better combine a prior and new data, especially if another prediction model is used instead of GP; however, this is not the focus of the paper. While we expect similar trends, especially from models similar to GP such as a Kalman filter [9] or a particle filter [6], absolute values from other models can be slightly different.

For GP kernels,  $\sigma_0 = 7500$  is provided to the Matérn kernel. Total noise  $z.\sigma$  is provided to the white kernel. Length scale  $l$  can take values in  $[0.01, 10]$  and is trained separately for each degraded version. IG over the time period is computed using numerical integration with the trapezoid rule. Finally, logarithms are base 2.

### 5.3 IG Capturing Trajectory Characteristics

For qualitative evaluation, Figure 2 illustrates the IG for different cases in one spatial dimension. Starting from left to right with a trajectory having two noisy measurements close to each other in Figure 2a, the IG increases as expected when the measurements are less noisy as in Figure 2b. When the duration increases, illustrated by two measurements being farther apart in Figure 2c, the IG also increases because these two measurements can also help *reduce uncertainty for the longer time in between these two measurements*.

When more measurements are available as in Figure 2d, these points help reduce more uncertainty around them, thus increasing the IG further. The IG keeps increasing when the measurements are temporally distributed more evenly as shown in Figure 2e because, for most of the duration along the trajectory, there are temporally nearby measurements, thus helping reduce uncertainty. Finally, in Figure 2f, the IG also slightly increases when measurements appear in between the start and end locations (shown on the vertical axis), thus making the movement less abrupt. However, if the locations jump unpredictably, the uncertainty tends to increase compared to, e.g., when the person stays at the same place, thus decreasing the IG for that case. Therefore, the IG can only partially capture the spatial distribution and can capture the size, duration, temporal distribution, and measurement uncertainty, as reflected in Table 1.

For quantitative evaluation using the Geolife data, the correlation coefficients of the IG with size, duration, and temporal entropy are 0.85, 0.97 and 0.96, respectively, which show very strong correlation. The correlation coefficients of the IG and spatial entropy is 0.68 which is also close to a strong correlation. Figure 3 shows the histograms of the IG and these characteristics on a log scale of the number of trajectories, along with a linear regression line indicating a clear trend between the IG and these characteristics. The regression uses Huber loss [8] to mitigate the effect of outliers. Quantitative evaluation shown in Figure 4 (explained in detail later) also shows that the IG can capture measurement uncertainty, because when a trajectory has large measurement uncertainty (i.e., less informative), the IG tends to decrease. Computing IG took around 2-4 seconds per trajectory on a single machine.

### 5.4 IG Capturing Prior Knowledge and Degradation

The impact of prior knowledge and degradation can be effectively captured by the IG framework. In general, when a trajectory is degraded more (e.g., larger total noise), the IG tends to decrease, matching the intuition that a lower quality trajectory gives lower information. Similar trends show in both the absolute value  $IG(Z, \Omega)$  and the ratio  $\frac{IG(Z, \Omega)}{IG(S, \Omega)}$  of the IG compared to the raw trajectory  $S$ .

We first discuss the results for perturbation. Figure 4 shows the IG value  $IG(Z, \Omega)$  and the IG ratio  $\frac{IG(Z, \Omega)}{IG(S, \Omega)}$  in percentage with different perturbations and prior knowledge. As discussed in Section 5.2, the total noise ranges from 3 to 400 meters, and the prior knowledge

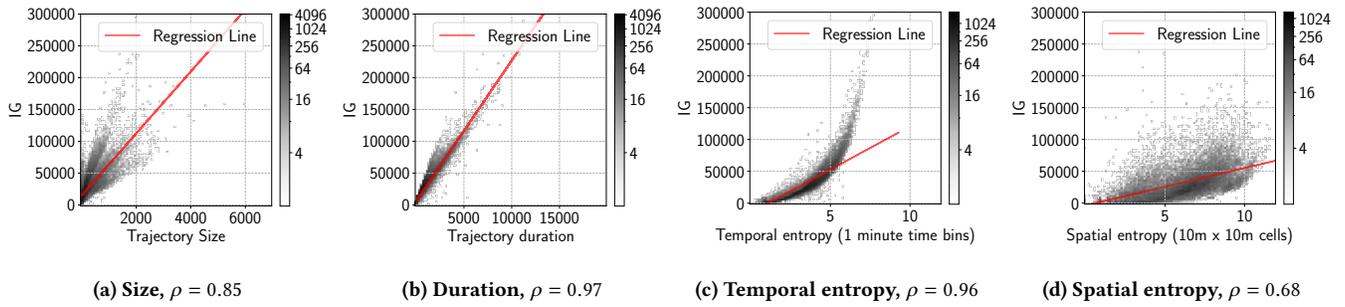


Figure 3: Histogram of IG and (a) size, (b) duration, (d) spatial, (c) temporal entropy, and their correlation coefficients.

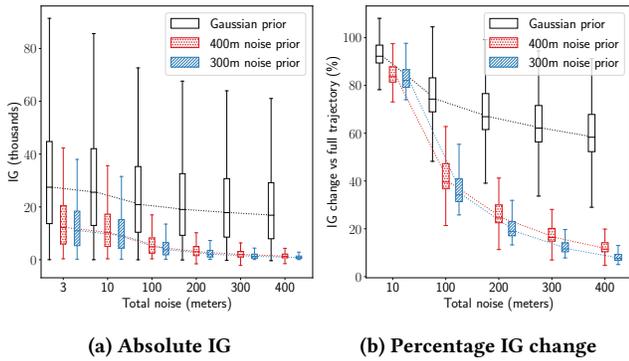


Figure 4: (a) Absolute IG and (b) percentage IG change compared to original trajectories for different values of total noise and different types of prior knowledge.

scenarios are *Gaussian*, *400m noise*, and *300m noise* priors. Each box plot in Figure 4a (or 4b) shows the IG value (or the IG ratio) of  $Z$  with the total noise shown on the x axis given particular prior knowledge. For example, the red, dotted box plot at total noise 100 in Figure 4a shows the IG values of a degraded version  $Z_{100}$  of  $S$  with total noise 100 meters, given that the prior knowledge is another previously-released degraded version  $Z_{400}$  of  $S$  with total noise 400 meters. It is clear that when total noise becomes larger, i.e., more degradation, the IG value and the IG ratio tend to decrease.

With a more informative prior, the absolute IG value tends to be smaller. For example, the 400m noise prior is a more informative prior than the Gaussian prior, thus the absolute IG value of the same  $Z$  with the 400-noise prior is smaller than with the Gaussian prior. This also supports intuition: if one already has good information about the location of a person, getting more data does not increase such information as much as when one only has little information. This is an important property that does not exist in other baselines.

Truncation and subsampling degradation also show similar observations as those from perturbation. Figures 5 and 6 show the IG value and the percentage of IG change for truncation and subsampling with different prior knowledge, respectively. The truncation/subsampling ratios range from 0.8 to 0.05, which means 80% down to 5% of the original trajectory is retained. Because of the 0.05 (or  $\frac{1}{20}$ ) ratio, these figures show the results from almost

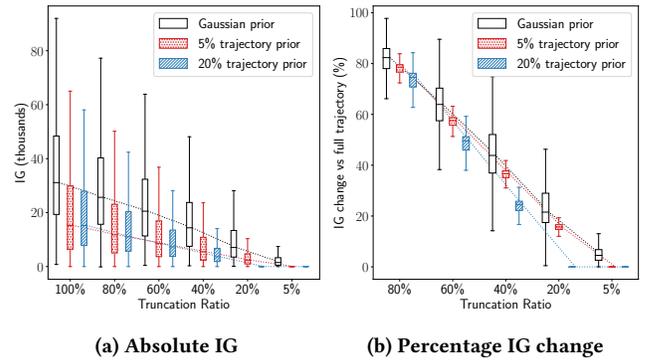


Figure 5: (a) Absolute IG and (b) percentage IG change compared to original trajectories for different truncation ratios and types of prior knowledge.

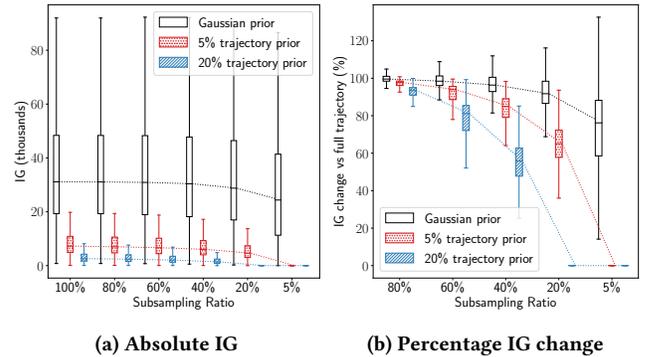


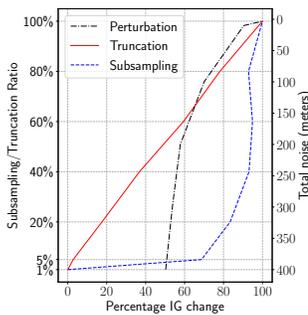
Figure 6: (a) Absolute IG and (b) percentage IG change compared to original trajectories for different subsampling ratios and types of prior knowledge.

40,000 trajectories that have at least 20 measurements. As discussed in Section 5.2, the prior knowledge scenarios are *Gaussian prior*, *5% trajectory prior* and *20% trajectory prior*. The IG also tends to decrease when the trajectory is degraded more (i.e., lower ratio) and when the prior gets better (e.g. 5% vs. 20% priors).

The decrease in IG with truncation is much stronger than with subsampling. The reason is that knowing a measurement not only

reduces the uncertainty at the time of the measurement but also for the time around that measurement. Thus when measurements are more evenly distributed, they help reduce more uncertainty, which means getting higher IG, as illustrated in Section 5.2. With the same ratio, which means the same number of measurements are retained, a truncation is similar to the case when measurements are close to each other (Figure 2d) while (uniform) subsampling is similar to the case when measurements are more evenly distributed (Figure 2e). This property also highlights the difference and advantage of IG for quantifying the VOI of a trajectory compared to other baselines.

Another observation is that by effectively capturing the impact of degradation, one can find the equivalence classes of different types of degradation for the same trajectory. For instance, given  $S$ , which truncation ratio produces the same VOI compared to perturbing  $S$  at total 300 meter noise? Figure 7 shows an example of such VOI equivalence for the first trajectory of the first owner in our dataset. The intersection between the (interpolated) truncation and perturbation lines indicates that truncating this trajectory at 65% gives roughly similar VOI to perturbing it with total 130m noise.



**Figure 7: An example of VOI equivalence classes of different types of degradation for a trajectory.**

## 6 RELATED WORK

Quantifying the value of location data has been an active line of research. There have been several surveys of how individuals value their location data [4, 18]. Aly et. al [1] showed how the value of a location data point can be quantified from the buyer’s perspective in a geo-marketplace. Also in a geo-marketplace context, Nguyen et. al [12] proposed a framework allowing sellers to offer a location data point at different qualities for different prices. However, previous work focused on the monetary value of location data. To the best of our knowledge, this is the first work that attempts to quantify the intrinsic VOI of a trajectory in their most basic form, which is a sequence of location measurements.

There is also extensive work on quantifying location privacy, which also attempted to reconstruct the locations of individuals. Krumm [11] surveyed a variety of computational location privacy schemes and emphasized the importance of finding a single quantifier for location privacy. Shokri et. al [17] proposed correctness as the metric for quantifying location privacy. Location privacy and the VOI of a trajectory can be related, e.g., a trajectory with high VOI may reveal more about a person than one with lower VOI,

thus leaking privacy. However, location privacy does not exactly correspond to the intrinsic VOI of a trajectory: the former concerns more about semantic locations and is often more subjective, while the latter concerns more about the location coordinates over time. Correctness and subjective sensitivity, which are related to location privacy, were also discussed in detail in Section 3.7.

## 7 CONCLUSION AND FUTURE WORK

The intrinsic VOI of trajectories plays an important role in many applications. We proposed the IG framework and qualitatively and quantitatively demonstrated its capability to effectively capture important characteristics of raw trajectories, prior knowledge, and various types of degradation. This shows that the IG is an appropriate framework to quantify the intrinsic VOI of trajectories. There are several potential directions for future work such as incorporating other features, or supporting other degradation types.

**Acknowledgements:** This research has been funded in part by NSF grants IIS-1910950 and CNS-2027794, the USC Integrated Media Systems Center, and unrestricted cash gifts from Microsoft. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of any of the sponsors such as the NSF.

## REFERENCES

- [1] Heba Aly, John Krumm, Gireeja Ranade, and Eric Horvitz. 2018. On the value of spatiotemporal information: Principles and scenarios. In *SIGSPATIAL '18*.
- [2] Mike Y Chen, Timothy Sohn, Dmitri Chmelev, Dirk Haehnel, Jeffrey Hightower, Jeff Hughes, Anthony LaMarca, Fred Potter, Ian Smith, and Alex Varshavsky. 2006. Practical metropolitan-scale positioning for gsm phones. In *UbiComp'06*.
- [3] Thomas M Cover and Joy A Thomas. 2012. *Elements of Information Theory*.
- [4] Dan Cvrcek, Marek Kumpost, Vashek Matyas, and George Danezis. 2006. A study on the value of location privacy. In *WPES'06*. 109–118.
- [5] Frank Van Diggelen. 2007. GNSS Accuracy: Lies, Damn Lies, and Statistics. *GPS world* 18, 1 (2007), 26–33.
- [6] Arnaud Doucet, Nando De Freitas, and Neil Gordon. 2001. An introduction to sequential Monte Carlo methods. In *Sequential Monte Carlo methods in practice*.
- [7] Joachim Hartung, Guido Knapp, and Bimal K Sinha. 2011. *Statistical meta-analysis with applications*. Vol. 738. John Wiley & Sons.
- [8] Peter J Huber. 1992. Robust estimation of a location parameter. In *Breakthroughs in statistics*. Springer, 492–518.
- [9] Rudolph Emil Kalman et al. 1960. A new approach to linear filtering and prediction problems. *Journal of basic Engineering* 82, 1 (1960), 35–45.
- [10] Yaron Kanza and Hanan Samet. 2015. An online marketplace for geosocial data. In *SIGSPATIAL '15*. 1–4.
- [11] John Krumm. 2009. A survey of computational location privacy. *Personal and Ubiquitous Computing* 13, 6 (2009), 391–399.
- [12] Kien Nguyen, John Krumm, and Cyrus Shahabi. 2020. Spatial Privacy Pricing: The Interplay between Privacy, Utility and Price in Geo-Marketplaces. In *SIGSPATIAL '20*. 263–272.
- [13] NCO PNT. 2021. *GPS Accuracy*. Retrieved May 08, 2021 from <https://www.gps.gov/systems/gps/performance/accuracy/>
- [14] J. Ross Quinlan. 1986. Induction of decision trees. *Machine learning* 1, 1 (1986).
- [15] Carl Edward Rasmussen and Christopher K. I. Williams. 2005. *Gaussian Processes for Machine Learning*. The MIT Press.
- [16] Patrick Schober, Christa Boer, and Lothar A Schwarte. 2018. Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia* (2018).
- [17] Reza Shokri, George Theodorakopoulos, Jean-Yves Le Boudec, and Jean-Pierre Hubaux. 2011. Quantifying location privacy. In *IEEE S&P 2011*. IEEE, 247–262.
- [18] Jacopo Staiano, Nuria Oliver, Bruno Lepri, Rodrigo de Oliveira, Michele Caravella, and Nicu Sebe. 2014. Money walks: a human-centric study on the economics of personal mobile data. In *UbiComp'14*. 583–594.
- [19] Daksh Varshneya and G Srinivasaraghavan. 2017. Human trajectory prediction using spatially aware deep attention models. *arXiv preprint 1705.09436* (2017).
- [20] Mingxuan Yue, Yaguang Li, Haoze Yang, Ritesh Ahuja, Yao-Yi Chiang, and Cyrus Shahabi. 2019. Detect: Deep trajectory clustering for mobility-behavior analysis. In *2019 IEEE International Conference on Big Data (Big Data)*. IEEE, 988–997.
- [21] Yu Zheng. 2015. Trajectory data mining: an overview. *TIST* 6, 3 (2015), 1–41.
- [22] Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. 2009. Mining interesting locations and travel sequences from GPS trajectories. In *WWW'09*. 791–800.