

An Experimental Evaluation of Grouping Definitions for Moving Entities

Lionov Wiratma

l.wiratma@uu.nl;lionov@unpar.ac.id

Dept. of Information and Computing Sciences, Utrecht
University, the Netherlands
Dept. of Informatics, Parahyangan Catholic University,
Indonesia

Marc van Kreveld

Maarten Löffler

Frank Staals

{m.j.vankreveld,m.loffler,f.staals}@uu.nl

Dept. of Information and Computing Sciences, Utrecht
University, the Netherlands

ABSTRACT

One important pattern analysis task for trajectory data is to find a *group*: a set of entities that travel together over a period of time. In this paper, we compare four definitions of groups by conducting extensive experiments using various data sets. The grouping definitions are different by one or more of three different characteristics: whether they use the measured sample points or the continuous movement, how distance is used to decide if entities are in the same group, and whether the duration of the group is measured cumulatively or as one contiguous time interval. We are interested in the differences between the definitions and comparisons to human annotated data, if available. We concentrate on pedestrian data and on different crowd densities. Furthermore, we analyze the robustness of the definitions and their dependence on different sampling rates. We use two different types of trajectory data sets: synthetic trajectories from a crowd simulation model, and real-life trajectories extracted from video surveillance. We present the results of the quantitative evaluations. For experiments with real-life trajectories, we augment them with a qualitative evaluation using videos that show groups in the trajectories with a color coding.

CCS CONCEPTS

• **Applied computing** → Law, social and behavioral sciences;
• **Theory of computation** → Computational geometry; • **Computing methodologies** → *Model development and analysis*.

KEYWORDS

Trajectories, collective motion, groups, experimental comparison

ACM Reference Format:

Lionov Wiratma, Marc van Kreveld, Maarten Löffler, and Frank Staals. 2019. An Experimental Evaluation of Grouping Definitions for Moving Entities. In *27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (SIGSPATIAL '19)*, November 5–8, 2019, Chicago, IL, USA. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3347146.3359346>

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.

SIGSPATIAL '19, November 5–8, 2019, Chicago, IL, USA

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6909-1/19/11...\$15.00

<https://doi.org/10.1145/3347146.3359346>

1 INTRODUCTION

The abundance of inexpensive devices equipped with tracking technologies, such as GPS-enabled mobile phones and RFID tags, allows easy recording of locations over a period of time. The quality of such data has increased considerably in recent years in sampling rate and geographic precision, and hence, vast amounts of movement data have become available. Consequently, this gives rise to an increasing interest to analyze them and furthermore, develop useful applications in many research fields: animal movement and behavior, traffic and transport, defense and surveillance, oceanographic observations, weather and natural phenomena, people behavior, health management, sports, and many others [14, 26].

Movement data of a single moving entity is typically described as a *trajectory*. Formally, a trajectory is a continuous mapping from a time interval $I = [t_{start}, t_{end}]$ to the space in which the entity is moving. Even though the movement is often continuous, a tracking device usually reports the entity's location only at specific moments with regular or irregular intervals in between. Therefore, trajectory data is often stored as an ordered sequence of discrete time-stamped locations. For example, trajectory $T = \{(p_1, t_1), (p_2, t_2), \dots, (p_\tau, t_\tau)\}$ represents the movement of an entity in two-dimensional space, where $p_i = (x_i, y_i)$ denotes the position of the entity at time t_i and τ is the total number of stored data points. Since the original movement is continuous, we must assume a position at any time between any two data points, and linear interpolation (constant velocity) is the simplest assumption. When the sampling rate was sufficiently high, i.e. the movement was sampled sufficiently often, we can assume that linear interpolation does not induce a significant error.

There are many different ways to analyze movement data. Previous research on algorithms for the analysis of trajectory data includes determining trajectory (and sub-trajectory) similarity [7, 33], segmenting a trajectory into a number of sub-trajectories based on certain criteria (e.g. speed or direction) [1, 9], detecting outliers [27, 51], finding popular places [4, 16], and clustering trajectories into a number of sets of trajectories [49]. Furthermore, we can analyze interactions between entities from their trajectories and try to determine particular movement patterns like leadership [2], commuting [6], chasing/avoidance behavior [11, 29], and more.

Collective movement patterns in which multiple entities travel together during a period of time are particularly relevant. For example, in ecology, researchers try to understand the behavior of groups of animals (e.g., [20, 35]). In veterinary science, researchers investigate whether the composition of animals in a group depends

on the health level of its members [10, 12], and in social psychology researchers analyze crowds to analyze human behavior [36]. In all these areas, identifying collective movement in trajectory data can provide critical new insights.

Many different definitions have been suggested to model the collective movement of a “sufficiently large” set of entities that travel “together” for a “sufficiently long” period of time: flocks [5, 15, 42], mobile groups [19], moving clusters [22], moving micro-clusters [31], herds [18], convoys [21], swarms [32], gatherings [50], traveling companions [39], platoons [30], groups [8], and refined groups [40]. It is beyond the scope of this paper to overview them all and explain their often subtle differences. We refer to the original papers, most of which introduce the new collective movement type, present one or more algorithms to compute it, and describe experiments where the new algorithms are run on some data sets, sometimes with a comparison to one earlier type.

Our Contribution. In this paper, we provide an extensive experimental study to find small groups in pedestrian data. This is an important case in analyzing the throughput in public spaces like shopping malls, parks and train stations, and in detecting suspect behavior in such spaces. Small groups can be as small as just two individuals. We compare four of the definitions, namely convoys (which in our setting are the same as traveling companions), swarms, groups, and refined groups. These four definitions differ in (i) how they model the input (as a continuous function or as discrete time stamps), (ii) how they model when entities are considered together, and (iii) how they measure if the entities are together long enough. Convoys (traveling companions) are a well-known type that considers groups whose composition does not change, where togetherness is consecutive, and assessment is done at the time stamps themselves. Groups and refined groups distinguish themselves from the other definitions on (i); treating time as a continuous phenomenon may make a difference for patterns that consist of relatively few timestamps, which is the case in our pedestrian settings. The refined groups definition is the only definition that measures togetherness within the group only, not having other non-group entities influence this. Swarms distinguish themselves by not requiring a contiguous grouping; interruptions are allowed. We discuss these four definitions in detail in the next section.

Early definitions that use a shape of the cluster (flocks) in the definition are disregarded because they exhibit the lossy flock problem [21]. Since we study small groups, allowing group composition change is not so suitable and hence we do not take definitions into account where the composition may change (like moving clusters and gatherings). Some other definitions are motivated mostly by vehicle data; we also do not consider these. Finally, we note that the four chosen definitions all use three main parameters: one for group size, one for group duration, and one for inter-distance. Therefore, comparing these definitions is more clean than including more complex definitions that need more parameters. There are no definitions that use fewer parameters.

Our study is not just an analysis of the four definitions, but also of how the input, space, and time can be treated and how this affects the results. Therefore, it may give indications on the results for other definitions. We note that it is not the objective to find a “best” definition, since we typically do not have a ground truth. For

several of the data sets we do have human annotations of groups, so we can compare the four definitions to the human annotations, which are by their nature subjective.

We consider the following research questions:

- (1) How well do the above definitions correspond with what humans consider a “group”, and how do the characteristics mentioned (input, connectivity, and duration) influence this?
- (2) How does the number of groups, as reported by the various definitions, depend on the density of the entities?
- (3) How does the number of groups, as reported by the various definitions, depend on the sampling rate of the input trajectories?

To answer these questions we perform both a quantitative and a qualitative study. For the quantitative analysis we compute and compare the number of reported groups and the precision, recall, and F1 scores of the various definitions with respect to human annotation. Furthermore, we compute how well the various definitions correspond to *social formations* used in crowd simulation, a behavior scheme used to generate synthetic movement data that represents a group of friends moving in a crowd [24]. For the qualitative analysis we develop a novel visualization to show and compare groups in video footage showing the movement of the entities. In particular, our visualization allows easy comparison of the detected groups with human annotation or with any other group definition. In our evaluation we use six data sets, all with different characteristics. Four of these data sets are from real world trajectory data. The remaining two are completely synthetic trajectories from a crowd simulation model.

In a previous short paper we compared groups and refined groups with a less broad focus [46]. The research questions, results and analyses in this paper are new.

Organization. The remainder of the paper is organized as follows. In Section 2 we review the grouping definitions that we consider, and analyze how they differ in theory. We describe the methods for our experimental comparison and introduce our new visualization method in Section 3. The results of our experimental evaluation are presented in Section 4.

2 THE DEFINITIONS

The four definitions rely on three parameters to define a group: the size parameter (the number of entities in a group), the temporal parameter (the time interval in which those entities form a group), and the spatial parameter (the distance between entities in the group). We formalize these parameters to define a group G from a set of moving entities X during time interval I :

- G contains at least m entities.
- I has a duration at least δ .
- Every pair of entities $x, y \in G$ is *connected* during I .

For the size parameter, the required minimum of entities to form a group is the same for all four definitions.

For the temporal parameter, the swarms [32] definition handles it differently from the others since it measures the duration of a group cumulatively. Let T be a set of timestamps where at each timestamp, every pair of entities $\in G$ are connected. Then, swarm uses the size of T —the number of timestamps—to define the duration

Table 1: Differences between grouping definitions

	Input	Connectivity	Duration
Original Groups (OG)	continuous	free	consecutive
Refined Groups (RG)	continuous	within group	consecutive
Convoys (CO)	discrete	free	consecutive
Swarms (SW)	discrete	free	cumulative

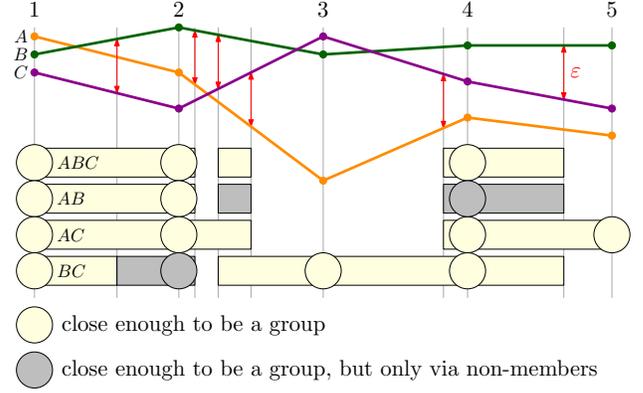
of G , rather than the duration of one contiguous time interval I . Note that with this property, swarm allows entities $\in G$ to leave G and join again later, as long as G is formed during at least k timestamps (which may be non-consecutive). For the other three definitions, the requirement for the temporal parameter is similar. We note, however, that the original and refined group definitions use interpolated positions between timestamps, the start and end time of the duration of a group will typically not be at any timestamp.

For the spatial parameter, we take a close look at each definition. The original group definition uses ϵ -connectivity between two entities as follows [8]: Two entities x and y ($x, y \in X$) are *directly ϵ -connected* if at any particular timestamp t , the Euclidean distance between x and y is at most ϵ (for some parameter $\epsilon > 0$). Furthermore, x and y are *ϵ -connected in X* at time t if there is a sequence $x = x_0, \dots, x_k = y$, with $x_0, \dots, x_k \in X$ and for all i , x_i and x_{i+1} are directly ϵ -connected at time t . This definition has the advantage that we need to consider the locations of all entities at time t only, to decide whether two of them are ϵ -connected.

One may claim that it is more natural if connectivity for x and y at time t can only be provided by entities who are in the same group, which is the approach taken by the refined group definition [40]. More specifically, to decide if x and y are ϵ -connected in a group G , we ignore all entities not in G and require a sequence $x = x_0, \dots, x_k = y$, with $x_0, \dots, x_k \in G$ where x_i and x_{i+1} are directly ϵ -connected at time t . Computing groups using this refined definition appears more complex since we cannot decide just from the locations at time t whether x and y are ϵ -connected. We need the location history and future as well.

The convoy [21, 39] and swarm [32] definitions use the concept of density connection [13], which is similar to the requirement of the spatial parameter for the original group, but with a slight difference. Let the ϵ -neighborhood of an entity $x \in X$ is defined by $N_\epsilon(x)$, the number of other entities in X that have the Euclidean distance at most ϵ ($\epsilon > 0$) from x (at any given timestamp t). Now, given a density threshold μ ($\mu > 0$), an entity $y \in X$ is *directly density-reachable* from x if $y \in N_\epsilon(x)$ and $|N_\epsilon(x)| \geq \mu$. Furthermore, y is *density-reachable* from x if a sequence of entities $\in X$ exists where each consecutive pair of entities in the sequence from x to y is directly density-reachable. Clearly, if $\mu = 1$ then the notion of (directly) density-reachable is exactly the same as the (directly) ϵ -connected in the original group definition. Henceforth, we only use $\mu = 1$ since $\mu > 1$ prevents the convoy and swarm definitions to identify groups that contain only two entities.

We summarize the differences between the definitions in Table 1. We note that no two of the definitions we consider in this paper are the same on all three aspects.

**Figure 1: Maximal groups according to: ($m = 2, \delta = 1$)**

CO: $ABC[1, 2], ABC[4], AC[4, 5], BC[1, 4]$

SW: $ABC(1, 2, 4), AC(1, 2, 4, 5), BC(1, 2, 3, 4)$

OG/RG: $ABC[1, 2.1], AC[1, 2.5], AC[3.8, 5], BC[2.3, 4.6]$

Maximal Groups. In the original and refined group definitions [8, 40], a group G is a *maximal group* during time interval I if there is no time interval $I' \supset I$ for which G is also a group and there is no $G' \supset G$ that is also a group during I . Moreover, swarms also has exactly the same concept as the maximal group, namely *closed swarm*. On the other hand, the definition of convoys only considers the ones that are maximal. Henceforth, we also use the term maximal group to describe the (maximal) convoy and the closed swarm. Figure 1 illustrates the concept for a small example. Note that the same set of entities can appear multiple times (at different moments in time) as a maximal group under all definitions except swarms.

Differences. Now, the differences shown in Table 1 affect how each definition specifies maximal groups from a set of trajectories. We demonstrate this using examples. First, we present an example in Figure 2, where a maximal group containing exactly the same entities may have different time durations, depending on which definition we use. Let two black entities x and y be the only entities that move; all red entities are stationary. Furthermore, trajectories of x and y consist of the shown positions at t_0, t_1, \dots, t_6 , and we set $\delta = 2$. Since x and y are not ϵ -connected (or density reachable) at t_5 , the pair $\{x, y\}$ is a convoy during the time interval $[t_0, t_4]$, while the swarm $\{x, y\}$ is formed during the timestamps of $\{t_0, t_1, t_2, t_3, t_6\}$. With the original group definition, x and y are a group starting at t_1 and ending at $t_{4.5}$. Finally, the refined group of $\{x, y\}$ can only start when the distance between x and y is $\leq \epsilon$ because they cannot be ϵ -connected through the red entities. Therefore, they can only start at t_2 and must end at t_4 .

Next, we show that the type of connectivity between entities in a group such as in the refined group definition can result in a completely different grouping. In Figure 3 [40], two entities a and h are moving in the same direction, opposite to the other entities. At any time during the time interval $I = [t_1, t_3]$, a and h are ϵ -connected through other entities. As a consequence, the convoy and the swarm definitions consider $\{a, h\}$ to be a group at timestamps t_1, t_2, t_3 , or during interval I for the original group definition. In

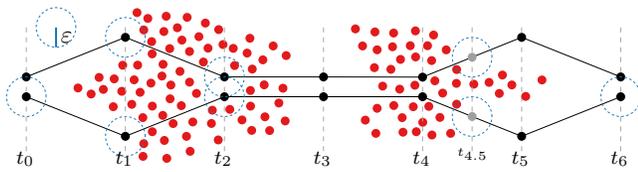


Figure 2: According to different definitions, the black entities are a group at different times

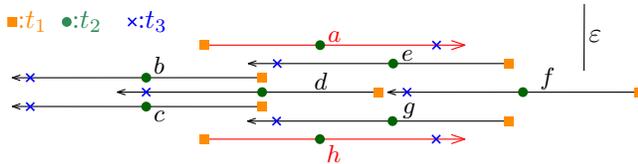


Figure 3: Entities a and h are not a refined group during $[t_1, t_3]$, but they are an original group, convoy, and swarm during $[t_1, t_3]$ or t_1, t_2, t_3 [40]

the refined group definition, $\{a, h\}$ is not a group during I because their connectivity is only through (changing) entities not in the group itself. There are several refined groups that include a and h that have a considerably shorter duration.

3 METHODS

To answer the research question described in Section 1, we conduct extensive experiments by computing all maximal groups from various trajectory data sets according to the different group definitions. We evaluate the results both quantitatively and qualitatively.

3.1 Experimental set-up

Data sets. To conduct our experiments, we need trajectory data. We use various data sets, which we divide into two *categories* based on the source of the trajectories.

- Real-life trajectories extracted from videos surveillance in a public area: the *NYC Grand Station* [47, 48], *Crowds by Examples* [25, 28, 38] and *Vittorio Emanuele II Gallery* data sets [3, 37, 38].
- Artificial trajectories generated by a computer simulation: the *Netlogo Flocking* data set [43, 44] and the *Utrecht University Crowd Simulation* data set [24].

We describe each data set in more detail along with the results of experiments using them in Section 4. The real-life data sets are captured from video surveillance; hence their raw coordinates are frame (pixel) coordinates from the videos. These coordinates are first converted to world coordinates using a homography matrix to be able to make fair distance comparisons. Most real-life trajectory data sets also come with a list of *human-annotated* groups; only the *NYC Grand Station* data set does not.

Implementations. In order to compute all maximal groups according to the different notions of groups, we need implementations. We used existing implementations where available, and implemented the remaining algorithms ourselves. In particular, we use

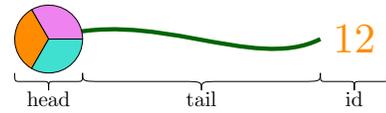


Figure 4: A moving entity is shown in a schematic manner.

- the Smart-and-Closed algorithm [39] to compute *convoy*s,
- the ObjectGrowth algorithm [32] to compute *swarm*s,
- our implementation from Buchin et al. [8] to compute *original groups*, and
- our implementation from Wiratma et al. [46] to compute *refined groups*.

Note that since the swarm algorithm has an exponential running time we were unable to compute all swarms for some of the parameter values in our experiments.

3.2 Evaluation set-up

Quantitative Evaluation. We analyze and evaluate the results from all experiments quantitatively. We compute and count all maximum groups in our data sets according to the four definitions while varying the parameters of the definitions: the distance ϵ , the minimum time duration δ , and the minimum group size m .

- For datasets with an alternative truth (either human-annotated or generated), we test which of the groups found by each definition match with the alternative truth¹ by providing precision, recall, and F1 scores.
- We vary the *density* of the environment by considering 200, 300, and 400 entities moving in the same bounded space. We compute the number of groups for each definition and study how it changes with the number of entities.
- We vary the sampling rate, or level of detail, of the trajectories by ignoring a fraction of the vertices in each trajectory. We count how many groups are identified by the different definitions, and analyze the consistency of these numbers.

Qualitative Evaluation. We also qualitatively evaluate the results of our experiments by visualizing the trajectories of pedestrians integrated in the videos from the data sets.

Conceptually, we represent each moving entity by a schematic figure that is overlaid on the video material; refer to Figure 4. Each entity consists of three parts. The *head* is a disk which shows the *current location* of the entity. The *tail* is a piece of curve which shows the *previous locations* of the entity during a set duration. The *id* is a unique identifier of the entity.

Grouping information is encoded by the *color* of the head and tail. We use the color of the *tail* to show a base grouping: depending on the data set this can be either the “ground truth”, a human-annotated grouping, or a grouping computed by one of the methods. Entities belonging to the same group have the same color, and every entity can only belong to at most one group, which cannot change over time. Entities that do not belong to any group in the “ground truth” have a white tail. The color of the *head* indicates the grouping as computed by the method currently under study. The computed

¹We avoid the terms “correctness” and “ground truth”: we can only test to what extent the groups found agree with an alternative. In particular, human-annotated data is likely to be influenced by personal interpretation and therefore not a ground truth.



Figure 5: Snapshot from a video.

groups can in principle overlap, and they do change over time. As a result, the head of an entity can have multiple colors, and the color of a head may change as time progresses.

The combination of colors of the tails and heads gives insight into the matching between annotated groups and groups based on a grouping definition. Note that colors are chosen at random; even when a method produces a group that exactly matches with one of the annotated groups, the color of the head may be different than the color of the tail.

We applied this scheme to all our data sets and generated videos for various parameter settings; see Figure 5 for an example. Our implementation of the visualization is based on the work by Maurice Marx [34]. In the remainder of this paper, we supply some snapshots of interesting configurations, as well as links to specific video fragments. The complete collection of videos from this paper can be found on our website [45].

4 EXPERIMENTAL EVALUATION

In this section we evaluate the results of our experiments. We focus our evaluation on the differences of the four definitions, and thus on the maximal groups that are reported, rather than the differences between the algorithms and their implementation. All implementations are non-optimized prototypes and therefore, comparing statistics like running times is meaningless.

4.1 Comparisons with Annotated Groups and Social Formations

We aim to establish how well the definitions capture the human intuition of grouping. In our first experiment, we compute the groups, as reported by the various definitions, and compare them to human annotation. We then report the *precision* (the percentage of the groups reported by the algorithm that also occur in the annotated data), the *recall* (the percentage of the human-annotated groups also found by the definition), and the corresponding *F1*-score. In our second experiment, we use a crowd simulation framework to generate trajectories including a set of entities traveling in a “social formation”. Intuitively, this is a group. We test if the various definitions identify these entities as a group.

Table 2: Information on the trajectories in the data sets and parameters used in the experiments. Here, g denotes the number of annotated groups, and G their size range. The video length is specified in minutes and seconds; the values for ε and δ are in meters and frames, respectively.

	VEIIG	CBE
video length	05:00	03:36
FPS	8	25
#entities	630	434
avg τ	189.82	400.12
g	207	115
G	2-7	2-4
ε	0.963	{1.22,1.52}
δ	{17,38,58}	{36,57,78}

4.1.1 Annotated Groups. We use two data sets consisting of real-life trajectories from video surveillance: Vittorio Emanuele II Gallery (VEIIG) and Crowds by Example (CBE). See Table 2 for details of each data set. Besides trajectories of pedestrians, these data sets are supplemented with homography matrices and lists of groups that are annotated manually by the authors. The annotations specify only *which* entities appear in a group, not *when*, or *how long* the entities form a group. Moreover, unlike in the four definitions, an entity occurs in at most one group in the human annotation.

For each definition, we count how many maximal groups match exactly with the annotated groups and evaluate the correctness using the precision, recall, and F1-score. For each data set we set the minimum required number of entities m to 2. The values for the inter-entity distance ε are chosen based on a study by Solera et al. [38], who analyze the average distance between people in the same group in a human crowd. Finally, we determine three different values for required minimum time δ that a group is together, based on the group annotations. In particular, we assume that a set of people cannot form a group when not all members are present in the video. Hence, we compute the time interval during which all members of an annotated group are present, and define the duration of the group to be the length of this interval. The minimum such duration over all groups gives us one choice of δ . The other two are chosen based on the average such duration $\bar{\delta}$ and the standard deviation σ . In particular, we pick $\bar{\delta} - \sigma$ and $\bar{\delta} - \frac{1}{2}\sigma$.

Vittorio Emanuele II Gallery. The data set is taken from the video surveillance in a hallway inside the Vittorio Emanuele II Gallery in Milan, Italy. The flow of entities in the video is mostly bidirectional. The results of our experiments are in Table 3.

First, we note that for all definitions, the precision values are relatively small. This can be explained by the fact that all definitions (except for swarm) consider a set of entities that is together during two disjoint but sufficiently long time intervals as two different (maximal) groups. Also, a group of 3 (or more) entities is often also found as one or two subgroups of 2 entities with slightly longer duration. Therefore, we focus on relative precision values. This also holds for the other data sets in our experiments.

We observe that in this data set, the refined group corresponds better to human annotation than the others based on their F1-score.

Table 3: Comparative results on the Vittorio Emanuele II data set with human annotation.

		Precision	Recall	F1 Score
$\delta = 17$	OG	0.199	0.952	0.329
	RG	0.223	0.947	0.361
	CO	0.202	0.952	0.333
	SW	0.125	0.957	0.221
$\delta = 38$	OG	0.357	0.884	0.509
	RG	0.414	0.884	0.564
	CO	0.362	0.884	0.514
	SW	0.238	0.932	0.379
$\delta = 58$	OG	0.444	0.778	0.565
	RG	0.503	0.778	0.611
	CO	0.451	0.778	0.571
	SW	0.315	0.870	0.463

OG = Original Groups, RG = Refined Groups, CO = Convoys, SW = Swarms



Figure 6: (Left) The pedestrians in a group are not within distance ϵ long enough. **(Center)** The pedestrian in the middle is close to the left pedestrian and not close to the right one during this snapshot, but the reverse was the case earlier in the video. The annotation has all three in a group. **(Right)** The pedestrian on the left is always close to a group but the annotation does not include it.

This is mostly since the refined group definition has the highest precision out of the definitions considered. The swarm definition has the best recall value, while the maximal groups by the original group and convoy definitions find the same number of annotated groups; refined group misses one in total. The higher recall of swarm is related to the lower precision: swarm outputs many more groups, some of which correspond to human annotation. These may be groups with interrupted duration.

Our qualitative review shows several reasons why the grouping definitions cannot match all annotated groups, see Figure 6. One main reason is that the members of an annotated group are not within ϵ distance for a duration δ . This results in (i) annotated groups not recognized at all, or (ii) grouping definitions only found subgroups of annotated groups. There are also situations where entities are always within distance ϵ , but they were not annotated as a group. It possible to increase the recall by increasing ϵ , but the precision is likely to go down.

Figures 7 and 8 show several scenarios that help to explain the low precision of all four definitions.

Crowds by Example. The Crowds by Example (CBE) data set records pedestrian outside a university building. The flow of pedestrians is different than in the GVEII data sets: pedestrians move in

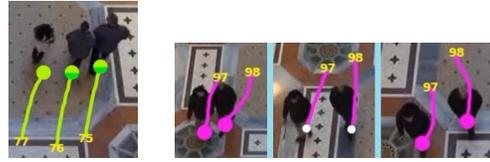


Figure 7: (Left) In this group of 3, the two pedestrians on the right appear earlier and disappear later in the videos, making a maximal group of 2 that is a subgroup of the 3. **(Right)** Frames in sequence from left to right show a group that separated for a while, resulting in two different maximal groups by the grouping definitions, except for swarm.

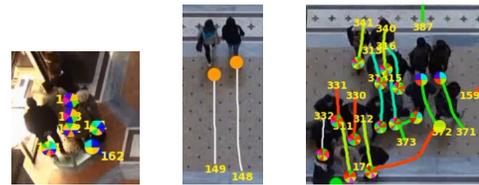


Figure 8: (Left) Two groups of pedestrians standing close together, making different maximal groups when they walk again. **(Center)** A group found by all four definitions that was not annotated. **(Right)** In a dense environment, many more groups are produced by all grouping definitions.

Table 4: Comparative results on the CBE data set; $\epsilon = 1.22$

		Precision	Recall	F1 Score
$\delta = 36$	OG	0.172	0.565	0.264
	RG	0.181	0.565	0.274
	CO	0.167	0.600	0.261
	SW	0.176	0.652	0.277
$\delta = 57$	OG	0.243	0.461	0.318
	RG	0.254	0.452	0.325
	CO	0.245	0.470	0.332
	SW	0.269	0.574	0.366
$\delta = 78$	OG	0.307	0.339	0.322
	RG	0.315	0.339	0.327
	CO	0.291	0.357	0.321
	SW	0.326	0.522	0.401

OG = Original Groups, RG = Refined Groups, CO = Convoys, SW = Swarms

various directions with varying speed. In this data set, vertices are sampled once every 6 frames. For experiments using this data set, we set ϵ based on Proxemics Theory [17], rather than the theory by Solera et al. [38] which seems to suggest an unrealistically small value for ϵ (namely 0.41m). Instead, the maximum far phase for a personal distance between pair of individuals from Proxemics Theory gives $\epsilon = 1.22m$.

Although swarms have the same discrete handling of the input, it performs better on recall because it appears there are groups with interrupted duration that are not found by the other definitions.

In dynamic crowds, we expect that entities from the same group will not be close to each other all the time, which is one reason why

Table 5: Comparative results on the CBE data set; $\varepsilon = 1.52$

		Precision	Recall	F1 Score
$\delta = 36$	OG	0.131	0.817	0.226
	RG	0.138	0.817	0.236
	CO	0.130	0.835	0.225
	SW	0.097	0.861	0.174
$\delta = 57$	OG	0.180	0.722	0.288
	RG	0.203	0.722	0.317
	CO	0.182	0.713	0.290
	SW	0.135	0.800	0.231
$\delta = 78$	OG	0.224	0.609	0.328
	RG	0.260	0.591	0.361
	CO	0.228	0.617	0.333
	SW	0.164	0.757	0.270

OG = Original Groups, RG = Refined Groups, CO = Convoys, SW = Swarms

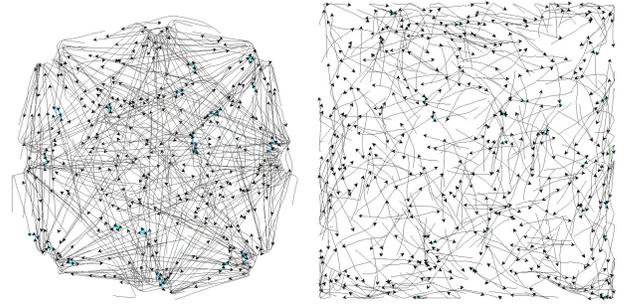
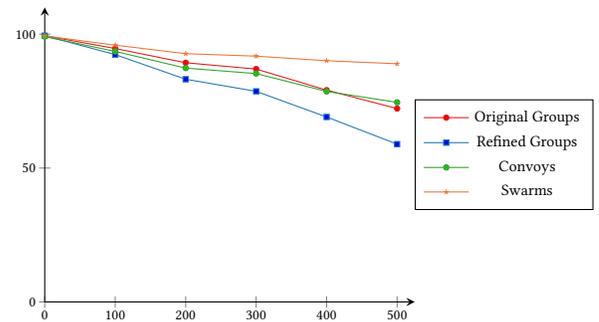
swarms can have a high recall score. Hence, we do the same experiment with different ε , to see how the other definitions that find maximal groups by a longest consecutive timestamp will perform. We set $\varepsilon = 1.52$ and present the results in Table 5. As expected, we see and increase in recall and decrease in precision for all definitions. However, the F1 results from the swarm get worse and other definitions now perform better for all δ .

Our qualitative evaluation of the Crowds by Example data shows similar situations as for the VEIIG data set.

4.1.2 Comparison to Generated Social Formations. We now compare the grouping definitions to other notions of grouping. In particular, we consider “socially-friendly formations” in crowd simulation [24].

We use a crowd simulation framework developed at Utrecht University [41], aimed at generating realistic crowd behavior. The framework allows agents (entities) to traverse a virtual environment, using global “indicative” routes on an underlying navigation mesh [23]. The framework supports generating routes for a set of entities that attempt to stay in a socially-friendly formation throughout the motion (and to re-establish such a formation when it is lost) [24]. We use the framework to generate trajectories in which there is exactly one such a socially-friendly formation G . We then test how well the definition of maximal groups capture G .

The input trajectories are generated as follows: we construct a virtual environment in which we place a set P of nine “points of interest”. We choose the eight corners of an octagon and its center point as P . Each entity has a global route visiting these points of interest. These global routes are picked randomly, i.e., when an entity gets close to its point of interest we randomly select a new target point of interest. We fix a set of four entities G that behave like a social group, make sure that they start at the same point in P , and follow the same global route. All remaining $n' = n - 4$ entities do not have any social group behavior and pick their global route individually. Figure 9 (left) shows an example of trajectories in this data set. For each choice of $n' \in \{100, 200, 300, 400, 500\}$, we run the simulation for 1000 time steps, thus producing trajectories with 1000 vertices each. Furthermore, for each n' , we repeat the simulation ten times. Note that even though the entities in G use

**Figure 9: The trajectories from the Crowd Simulation data set (left) and the Netlogo Flocking data set (right).****Figure 10: The percentage of time during which G is a maximal group as a function of the number of entities n' (averaged over ten simulations).**

the same global route, the individual trajectories will differ as the entities in G have to avoid colliding with other entities.

Finding Social Formations. We compute the maximal groups on the resulting data sets, selecting $\varepsilon = 1.6$ and $m = 4$, and we compare when G is a maximal group according to all definitions. Figure 10 shows the percentage of time during which G is a maximal group as a function of the number of dummy entities n' , and thus of the density in the environment. First note that, even though the entities in G use the same global route, this is not a guarantee that the entities remain close together (and thus form a group) throughout the entire time interval considered. Indeed, we see that as we increase the number of dummy entities, the entities in G may be forced to spread out in order to avoid collisions. We see that for all densities considered, G occurs as a maximal group longest in the swarm definition and shortest in the refined group. An explanation for this is that in all definitions except for the refined group, the entities in G can stay in a group longer by using the non-group entities to remain connected. Note, however, that this does not guarantee that G is a maximal group for a longer period of time, since these dummy entities could also create a larger maximal group $H \supset G$ that prevents G from being maximal. As expected, swarm has a larger percentage where G is a group than the other definitions, especially when the environment becomes more dense.

Table 6: The average number of maximal groups for $m = 10$ and $\delta = 10$, and the standard deviation in the Netlogo data set for 10 generated sets.

		average (10 sets)				std. dev.			
$n = 200$	ϵ	OG	RG	CO	SW	OG	RG	CO	SW
	4		0	0	0	0	0	0	0
5		2.9	2.2	3.0	10.8	3.99	2.86	3.63	16.27
6		33.9	25.9	36.6	229.3	23.48	17.95	21.65	252.86
$n = 300$	ϵ	OG	RG	CO	SW	OG	RG	CO	SW
	4	3.0	1.7	3.2	9.1	3.58	2.00	4.60	11.00
5		56.3	38.4	61.7	229.0	12.17	6.50	12.79	77.89
6		396.1	304.5	418.6	5017.3	64.15	47.61	53.61	2363.13
$n = 400$	ϵ	OG	RG	CO	SW	OG	RG	CO	SW
	4	23.1	12.3	31.1	112.2	9.63	7.98	9.83	45.00
5		411.6	259.0	410.8	5299.6	108.00	64.43	106.38	2466.9
6		1905.7	1357.0	1830.4	-	250.68	204.28	226.49	-

OG = Original Groups, RG = Refined Groups, CO = Convoys, SW = Swarms

Research question (1): correspondence to human annotation and other models. Over all experiments, the refined groups have a slightly higher F1 score in the correspondence to human annotation than the original groups and convoys definitions, but they are usually close. The higher F1 score is caused by a better precision. The swarms definition sometimes corresponds better and sometimes worse to human annotation. It appears to depend on the precise parameter settings. We also observe that human annotation is likely somewhat subjective. When analyzing the retrieval of a socially-friendly formation, we observe that all definitions notice interruptions of the groups with increasing density. The swarms definition suffers least from this and the refined groups definition most.

4.2 Dependence on Density

We generated several data sets using an adapted version of the Net-Logo Flocking model [44]. This data is convenient because we can easily produce data sets of varying densities. In the adapted model the entities start to turn when they approach the border (instead of wrapping around), and there is a small random component in the new direction of the entities. This same model was used by Buchin et al. [8] to test the definition of original groups.

In all our experiments, the size of the environment in which the entities move is fixed and set to 256×256 units. See Figure 9 (right) for a general impression of the moving entities in these data sets. We consider different densities by varying the number of entities n to be 200, 300 or 400, and generate data sets with 500 time stamps each. For each generated data set, we compute all maximal groups for all four definitions, with a fixed $\delta = 10$ and $m = 10$, but using three values of ϵ , namely 4, 5, and 6. We chose to vary ϵ because this distance value is related to density, the characteristic under investigation. Each experiment is performed 10 times and the average and standard deviation are computed. The results of these experiments are shown in Table 6. There are no results for swarm when $\epsilon = 6$ and $n = 400$ due to the computation time needed: the swarm algorithm has exponential running time.



Figure 11: The Grand Station and trajectories of pedestrians.

Generally, for all definitions the number of maximal groups increases as the density increases or when ϵ increases, which is not a surprise. We see that in most settings the refined group produces fewer maximal groups than the other definitions, and swarm produces more. All definitions show roughly a 20-fold increase from $n = 200$ to $n = 300$ when $\epsilon = 5$. From $n = 300$ to $n = 400$, the swarm definition has a larger than 20-fold increase while the other three definitions have a less than 10-fold increase. For $\epsilon = 4$, the values are too small for such observations. For $\epsilon = 6$, we notice that the increase for swarm from $n = 200$ to $n = 300$ is much larger than for the other three definitions. Hence, it seems that swarm has a larger increase in the number of maximal groups than the other three definitions when the density increases. We notice a similar effect when ϵ increases rather than the density.

Research question (2): dependence on density. As expected, all grouping definitions find more groups when the density of entities increases or when connectedness is satisfied at larger distances. The swarm definition has a larger increase in the number of maximal groups than the other definitions. Since we do not have a ground truth or alternative truth, we cannot draw further conclusions from these observations.

4.3 Dependence on Sampling Rate

The purpose of our last experiment is to examine how different sampling rates of trajectories affect the maximal groups produced by each definition. We conduct experiments by gradually removing vertices from trajectories, thus decreasing their sampling rate. For each new data set consisting of trajectories with a lower sampling rate, we count how many maximal groups result.

The data set consists of trajectories from pedestrians inside the Grand Central Terminal in New York City, USA (see Figure 11²). The data set contains 6000 video frames at which data points are generated manually. This is once every 0.8 seconds. There are 12,684 pedestrians, with an average of 105.52 pedestrians in each frame. For our experiment, we choose two sets of 800 consecutive frames that have a high density. The first set contains 2591 trajectories while the second contains 3313 trajectories. The average number of vertices in a trajectory are 46.57 and 46.85, respectively.

²The background image and movement data are from [47].

Table 7: The number of maximal groups from 2592 trajectories in the Grand Station data set with different sampling rate

		25%	50%	100%
$\delta = 8s$	OG	174	178	170
	RG	173	177	169
	CO	140	158	177
	SW	153	180	249
$\delta = 12s$	OG	127	118	116
	RG	127	118	117
	CO	111	122	121
	SW	115	138	199
$\delta = 16s$	OG	97	96	96
	RG	97	97	96
	CO	92	94	97
	SW	98	111	162

OG = Original Groups, RG = Refined Groups, CO = Convoys, SW = Swarms

Table 8: The number of maximal groups from 3313 trajectories in the Grand Station data set with different sampling rate

		25%	50%	100%
$\delta = 8s$	OG	276	268	257
	RG	269	262	255
	CO	222	256	264
	SW	229	259	379
$\delta = 12s$	OG	211	206	203
	RG	206	200	200
	CO	190	203	204
	SW	196	215	304
$\delta = 16s$	OG	172	168	159
	RG	167	163	157
	CO	160	156	161
	SW	166	183	252

OG = Original Groups, RG = Refined Groups, CO = Convoys, SW = Swarms

First, we create a homography matrix to map frame coordinates from the data set into ground coordinates. We choose $\epsilon = 0.76m$ for a personal distance between pairs, based on the maximum *close phase* from Proxemics Theory [17]. We vary the required minimum duration for a maximal group $\delta \in \{8, 12, 16\}$ in seconds. Finally, we consider different sampling rates for the two sets of trajectories by taking 25%, 50%, 100% of the vertices of the trajectories. Some trajectories may be removed because less than 2 vertices remain. The results of our experiments are in Tables 7 and 8.

We notice that the number of original groups and refined groups is stable or increases slightly when reducing the sampling rate. In contrast, the number of convoys decreases slightly when reducing the sampling rate, and the number of swarms decreases substantially. This trend is related to the cumulative version of the time duration of swarms. Imagine a swarm with several disconnected time intervals on its time duration (with a sampling rate of 100%). By reducing the sampling rate, some of these time intervals will likely to disappear, or their length is reduced. Therefore, the swarm may not meet the required δ anymore. On the other hand, the other

definitions which use consecutive timestamps are not affected much by this situation.

Research question (3): dependence on sampling rate. In general it is preferable when a definition of grouping is not influenced too much by the sampling rate, so in this respect the original and refined group definitions perform a bit better than convoys and much better than swarms.

5 CONCLUSION

We experimentally evaluated four definitions for grouping in trajectory data: (original) groups, refined groups, convoys, and swarms. We tried to establish how well these definitions correspond to the human intuition of a group, how the number of groups depends on the density of the entities in their environment, and how the number of groups depends on the sampling rate of the trajectories. In our experiments, the groups, refined groups, and convoys perform similar in terms of recognizing all sets of entities that were a group according to the human annotations, with the refined groups typically having the highest F1 score. On occasion the swarm definition outperforms the other definitions. To be more conclusive in these experiments, we first of all need better human annotation, and second of all, test more data sets and settings.

We observe that the definitions that consider the trajectories to be continuous mappings from time to space (original groups and refined groups) are more stable than the definitions considering the trajectories as discrete input (convoys and swarms) when we consider the number of reported groups under reductions of the sampling rate.

In general, it appears that swarm is most different among the definitions, suggesting that taking group duration cumulatively has a larger effect on grouping than the discrete or continuous handling of the data, or the type of connectedness (see Table 1).

We expect that, by counting duration cumulatively rather than consecutively, the swarm definition is more robust to noise than the other methods, but at the same time finds more doubtful groups that arise from several short, by-chance encounters. Other, more robust grouping definitions can be developed and compared, which would depend on a fourth parameter that describes how noise is handled. Examples are platoons [30] and robust groups [8]. The extra parameter makes proper experimentation harder, however.

In terms of qualitative assessment, we developed a style of video annotation that allows us to compare two different grouping definitions. It is best suited for comparisons to groups from human annotation. Videos using this visualization can be found on our website [45].

ACKNOWLEDGMENTS

M.v.K. and M.L. are partially supported by The Netherlands Organisation for Scientific Research on the Commit2Data project “Geometric Algorithms for the Analysis and Visualization of Heterogeneous Spatio-temporal Data” (no. 628.011.005). M.L. is partially supported by The Netherlands Organisation for Scientific Research on grant no. 614.001.504. F.S. is partially supported by The Netherlands Organisation for Scientific Research on grant no. 612.001.651. L.W. is supported by The Ministry of Research, Technology and Higher Education of Indonesia (no. 138.41/E4.4/2015).

REFERENCES

- [1] Aris Anagnostopoulos, Michail Vlachos, Marios Hadjieleftheriou, Eamonn Keogh, and Philip Yu. 2006. Global distance-based segmentation of trajectories. In *Proc. of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 34–43.
- [2] Mattias Andersson, Joachim Gudmundsson, Patrick Laube, and Thomas Wolle. 2008. Reporting Leaders and Followers among Trajectories of Moving Point Objects. *Geoinformatica* 12, 4 (2008), 497–528.
- [3] Stefania Bandini, Andrea Gorrini, and Giuseppe Vizzari. 2014. Towards an integrated approach to crowd analysis and crowd synthesis: A case study and first results. *Pattern Recognition Letters* 44 (2014), 16–29.
- [4] Marc Benkert, Bojan Djordjevic, Joachim Gudmundsson, and Thomas Wolle. 2010. Finding Popular Paths. *International Journal of Computational Geometry & Applications* 20, 1 (2010), 19–42.
- [5] Marc Benkert, Joachim Gudmundsson, Florian Hübner, and Thomas Wolle. 2008. Reporting flock patterns. *Computational Geometry* 41, 3 (2008), 111–125.
- [6] Kevin Buchin, Maike Buchin, Joachim Gudmundsson, Maarten Löffler, and Jun Luo. 2011. Detecting Commuting Patterns by Clustering Subtrajectories. *International Journal of Computational Geometry & Applications* 21, 3 (2011), 253–282.
- [7] Kevin Buchin, Maike Buchin, Marc van Kreveld, and Jun Luo. 2011. Finding long and similar parts of trajectories. *Computational Geometry* 44, 9 (2011), 465–476.
- [8] Kevin Buchin, Maike Buchin, Marc van Kreveld, Bettina Speckmann, and Frank Staals. 2015. Trajectory grouping structure. *Journal of Computational Geometry* 6, 1 (2015), 75–98.
- [9] Maike Buchin, Anne Driemel, Marc van Kreveld, and Vera Sacristán. 2011. Segmenting trajectories: A framework and algorithms using spatiotemporal criteria. *Journal of Spatial Information Science* 3, 1 (2011), 33–63.
- [10] Frances Colles, Russell Cain, Thomas Nickson, Adrian Smith, Stephen Roberts, Martin Maiden, Daniel Lunn, and Marian Stamp Dawkins. 2016. Monitoring chicken flock behaviour provides early warning of infection by human pathogen *Campylobacter*. *Biological Sciences* 283, 1822 (2016), 20152323.
- [11] Fernando de Lucca Siqueira and Vania Bogorny. 2011. Discovering Chasing Behavior in Moving Object Trajectories. *Transactions in GIS* 15, 5 (2011), 667–688.
- [12] Marcus Doherr, Tim Carpenter, David Wilson, and Ian Gardner. 1999. Evaluation of temporal and spatial clustering of horses with *Corynebacterium pseudotuberculosis* infection. *American Journal of Veterinary Research* 60, 3 (1999), 284–291.
- [13] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. 1996. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proc. of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*. 226–231.
- [14] Joachim Gudmundsson, Patrick Laube, and Thomas Wolle. 2012. Computational Movement Analysis. In *Handbook of Geographic Information*, Wolfgang Kresse and David Danko (Eds.). Springer, 725–741.
- [15] Joachim Gudmundsson and Marc van Kreveld. 2006. Computing Longest Duration Flocks in Trajectory Data. In *Proc. of the 14th Annual ACM International Symposium on Advances in Geographic Information Systems*. 35–42.
- [16] Joachim Gudmundsson, Marc van Kreveld, and Frank Staals. 2013. Algorithms for hotspot computation on trajectory data. In *Proc. of the 21st SIGSPATIAL International Conference on Advances in Geographic Information Systems, SIGSPATIAL 2013*. 134–143.
- [17] Edward Hall. 1992. *The Hidden Dimension*. Anchor Books.
- [18] Yan Huang, Cai Chen, and Pinliang Dong. 2008. Modeling Herds and Their Evolutions from Trajectory Data. In *Proc. of the 5th International Conference of Geographic Information Science, GIScience*. 90–105.
- [19] San-Yih Hwang, Ying-Han Liu, Jeng-Kuen Chiu, and Ee-Peng Lim. 2005. Mining Mobile Group Patterns: A Trajectory-based Approach. In *Proc. of the 9th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining, PAKDD'05*. 713–718.
- [20] Armand Jacobs, Cédric Sueur, Jean Louis Deneubourg, and Odile Petit. 2011. Social Network Influences Decision Making During Collective Movements in Brown Lemurs (*Eulemur fulvus fulvus*). *International Journal of Primatology* 32, 3 (2011), 721–736.
- [21] Hoyoung Jeung, Man Lung Yiu, Xiaofang Zhou, Christian Jensen, and Heng Tao Shen. 2008. Discovery of convoys in trajectory databases. *Proceedings of the VLDB Endowment* 1, 1 (2008), 1068–1080.
- [22] Panos Kalnis, Nikos Mamoulis, and Spiridon Bakiras. 2005. On Discovering Moving Clusters in Spatio-temporal Data. In *Proc. of the Advances in Spatial and Temporal Databases, 9th International Symposium, SSTD 2005*. 364–381.
- [23] Ioannis Karamouz, Roland Geraerts, and Mark Overmars. 2009. Indicative routes for path planning and crowd simulation. In *Proc. of the 4th International Conference on Foundations of Digital Games, FDG 2009*. 113–120.
- [24] Angelos Kremyzas, Norman Jaklin, and Roland Geraerts. 2016. Towards social behavior in virtual-agent navigation. *SCIENCE CHINA Information Sciences* 59, 11 (2016), 1–17.
- [25] UCY Computer Graphics Lab. [n.d.]. Crowd Data. <https://graphics.cs.ucy.ac.cy/research/downloads/crowd-data/>. Last accessed: August 1, 2019.
- [26] Patrick Laube. 2014. *Computational Movement Analysis*. Springer.
- [27] Jae-Gil Lee, Jiawei Han, and Xiaolei Li. 2008. Trajectory Outlier Detection: A Partition-and-Detect Framework. In *Proc. of the 24th International Conference on Data Engineering, ICDE 2008*. 140–149.
- [28] Alon Lerner, Yiorgos Chrysanthou, and Dani Lischinski. 2007. Crowds by Example. *Computer Graphics Forum* 26, 3 (2007), 655–664.
- [29] Francesco Lettich, Luis Otávio Alvares, Vania Bogorny, Salvatore Orlando, Alessandra Raffaetà, and Claudio Silvestri. 2016. Detecting avoidance behaviors between moving object trajectories. *Data & Knowledge Engineering* 102 (2016), 22–41.
- [30] Yuxuan Li, James Bailey, and Lars Kulik. 2015. Efficient mining of platoon patterns in trajectory databases. *Data & Knowledge Engineering* 100 (2015), 167–187.
- [31] Yifan Li, Jiawei Han, and Jiong Yang. 2004. Clustering moving objects. In *Proc. of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 617–622.
- [32] Zhenhui Li, Bolin Ding, Jiawei Han, and Roland Kays. 2010. Swarm: Mining Relaxed Temporal Moving Object Clusters. *Proceedings of the VLDB Endowment* 3, 1 (2010), 723–734.
- [33] Hechen Liu and Markus Schneider. 2012. Similarity measurement of moving object trajectories. In *Proc. of the 3rd ACM SIGSPATIAL International Workshop on GeoStreaming, IWGS@SIGSPATIAL 2012*. 19–22.
- [34] Maurice Marx. [n.d.]. Python 2d/3d Trajectory Visualization Library. <https://github.com/marximus/trackviz/>. Last accessed: August 1, 2019.
- [35] Harvey Miller, Somayeh Dodge, Jennifer Miller, and Gil Bohrer. 2019. Towards an integrated science of movement: converging research on animal movement ecology and human mobility science. *International Journal of Geographical Information Science* 33, 5 (2019), 855–876.
- [36] Kiran Rachuri, Mirco Musolesi, Cecilia Mascolo, Peter Rentfrow, Chris Longworth, and Andrius Aucinas. 2010. EmotionSense: A Mobile Phones Based Adaptive Platform for Experimental Social Psychology Research. In *Proc. of the 12th ACM International Conference on Ubiquitous Computing*. 281–290.
- [37] Francesco Solera. [n.d.]. group-detection. <http://imagelab.unimore.it/group-detection/>. Last accessed: August 1, 2019.
- [38] Francesco Solera, Simone Calderara, and Rita Cucchiara. 2016. Socially Constrained Structural Learning for Groups Detection in Crowd. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38, 5 (2016), 995–1008.
- [39] Lu An Tang, Yu Zheng, Jing Yuan, Jiawei Han, Alice Leung, Wen-Chih Peng, and Thomas La Porta. 2013. A framework of traveling companion discovery on trajectory data streams. *ACM Transactions on Intelligent Systems and Technology* 5, 1 (2013), 3:1–3:34.
- [40] Marc van Kreveld, Maarten Löffler, Frank Staals, and Lionov Wiratma. 2018. A Refined Definition for Groups of Moving Entities and Its Computation. *International Journal of Computational Geometry & Applications* 28, 2 (2018), 181–196.
- [41] Wouter van Toll, Norman Jaklin, and Roland Geraerts. 2015. Towards Believable Crowds: A Generic Multi-level Framework for Agent Navigation. In *ICT.OPEN 2015*. 10.
- [42] Marcos Vieira, Petko Bakalov, and Vassilis Tsotras. 2009. On-line discovery of flock patterns in spatio-temporal data. In *Proc. of the 17th ACM SIGSPATIAL International Symposium on Advances in Geographic Information Systems, ACM-GIS 2009*. 286–295.
- [43] Uri Wilensky. [n.d.]. NetLogo. <http://ccl.northwestern.edu/netlogo>. Last accessed: August 1, 2019.
- [44] Uri Wilensky and William Rand. 2015. *An introduction to agent-based modeling: modeling natural, social, and engineered complex systems with NetLogo*. MIT Press.
- [45] Lionov Wiratma. [n.d.]. Visualization of Grouping Definitions. <https://tiny.cc/groupingvideos/>. Last accessed: August 1, 2019.
- [46] Lionov Wiratma, Maarten Löffler, and Frank Staals. 2018. An Experimental Comparison of Two Definitions for Groups of Moving Entities (Short Paper). In *Proc. of the 10th International Conference on Geographic Information Science, GIScience 2018*. 64:1–64:6.
- [47] Shuai Yi. [n.d.]. Pedestrian Walking Path Dataset. https://www.dropbox.com/s/7y90xsql0lyv8d/cvpr2015_pedestrianWalkingPathdataset.rar. Last accessed: August 1, 2019.
- [48] Shuai Yi, Hongsheng Li, and Xiaogang Wang. 2015. Understanding pedestrian behaviors from stationary crowd groups. In *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015*. 3488–3496.
- [49] Guan Yuan, Penghui Sun, Jie Zhao, Daxing Li, and Canwei Wang. 2017. A review of moving object trajectory clustering algorithms. *Artificial Intelligence Review* 47, 1 (2017), 123–144.
- [50] Kai Zheng, Yu Zheng, Nicholas Jing Yuan, Shuo Shang, and Xiaofang Zhou. 2014. Online Discovery of Gathering Patterns over Trajectories. *IEEE Transactions on Knowledge and Data Engineering* 26, 8 (2014), 1974–1988.
- [51] Jie Zhu, Wei Jiang, An Liu, Guanfeng Liu, and Lei Zhao. 2017. Effective and efficient trajectory outlier detection based on time-dependent popular route. *World Wide Web* 20, 1 (2017), 111–134.