Early Detection of Herding Behaviour during Emergency Evacuations

David Amores
Infrastructure Engineering, The University of Melbourne, Parkville, VIC 3010, Australia
damores@student.unimelb.edu.au

Maria Vasardani
Infrastructure Engineering, The University of Melbourne, Parkville, VIC 3010, Australia
mvasardani@unimelb.edu.au

Egemen Tanin
Computing and Information Systems, The University of Melbourne, Parkville, VIC 3010, Australia
etanin@unimelb.edu.au

Abstract

Social scientists have observed a number of irrational behaviours during emergency evacuations, caused by a range of possible cognitive biases. One such behaviour is *herding* — people following and trusting others to guide them, when they do not know where the nearest exit is. This behaviour may lead to safety under a knowledgeable leader, but can also lead to dead-ends. We present a method for the automatic early detection of herding behaviour to avoid suboptimal evacuations. The method comprises three steps: (i) people clusters identification during evacuation, (ii) collection of clusters' spatio-temporal information to extract features for describing cluster behaviour, and (iii) unsupervised learning classification of clusters' behaviour into 'benign' or 'harmful' herding. Results using a set of different detection scores show accuracies higher than baselines in identifying harmful behaviour; thus, laying the ground for timely irrational behaviour detection to increase the performance of emergency evacuation systems.

2012 ACM Subject Classification Information systems → Location based services, Computing methodologies → Spatial and physical reasoning

Keywords and phrases spatio-temporal data, emergency evacuations, herding behaviour

Digital Object Identifier 10.4230/LIPIcs.GIScience.2018.1

1 Introduction

Certain cognitive biases may govern the way people react and move during emergency evacuations and may result in irrational behaviours that can hinder operations and lead to slower evacuation times, perhaps even endangering lives. An example of a common and well-known behaviour is herding – "when under highly uncertain and stressful situations, an individual tends to follow others almost blindly" [19]. This behaviour sometimes helps people exit a building safely when the leader knows the way out (benign herding), but may otherwise lead people to dead ends (harmful herding). Early identification of such behaviour can aid in more timely, orderly, and ultimately more successful evacuations.

Considering these benefits, this work proposes an automatic method for the early detection of harmful herding behaviour, based on features extracted from the spatio-temporal characteristics of people's group (cluster) movements during emergency evacuations. Figure 1 depicts snapshots of a moving cluster of people during a building evacuation at different times, which displays harmful herding behaviour. Figure 1b shows the point in time when

10th International Conference on Geographic Information Science (GIScience 2018).
Editors: Stephan Winter, Amy Griffin, and Monika Sester; Article No. 1; pp. 1:1–1:15
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

(a) Group seemingly heading towards the exit. **(b)** Display of harmful herding behaviour.

Figure 1 Snapshots of group behaviour at two different time steps.

the group moves into a room instead of going for the exit. This is when a human observer with knowledge of the building layout would identify this behaviour as erratic and alert the people. The proposed method succeeds in analysing the group's movement trajectory and, more importantly, the group leader's trajectory, to make an *earlier* detection. The assumption is based on the herding behaviour's definition — people delegating wayfinding responsibility to the group's leader. If the leader's past trajectory displays erratic movement, chances that the group will head straight to the nearest exit decrease.

Our method comprises three steps. First, clusters of people traveling together are identified. Second, information about the identified clusters is collected, such as the cluster and cluster leader's moving trajectories, as well as the cluster's distance from the nearest exit. This information is compiled into a *feature vector*. Third, all feature vectors are classified as either *benign* or *harmful* behaviour, using an unsupervised learning classification method. The method is assessed against a ground truth, and also compared to human assessment. The ground truth knows at all times if the ultimate destination of each cluster is the exit or a dead end. The human assessment is performed by visually inspecting the cluster's trajectory and determining the point of wrong going (e.g., turning away from the exit). A set of scores is defined and used to assess the performance of the suggested method when detecting harmful behaviour.

Experiments based on simulated emergency evacuation scenarios show favorable results, as the method outperforms baseline cases and visual inspection in early detection of harmful behaviour. Using different cluster feature combinations, the results also allow for some interesting observations. For example, considering only the actual distance between the cluster and the nearest exit in fact hurts the classification, making it resemble a random one. Instead, the previous moving history of a cluster, rather than its mere distance from an exit, is a better indicator of harmful behaviour. Accordingly, the main contributions of this work are: (1) The identification of spatio-temporal cluster features that can be trusted to describe herding behaviour as either benign or harmful, and (2) a method that uses these features to early detect harmful herding behaviour during emergency evacuations, in an automated way.

The remainder of this paper is organised as follows. Section 2 summarises related research in behaviour detection including simulations, pattern recognition, and personalised evacuation systems. Section 3 discusses the concepts and previously defined behaviours on which our herding detection method is based. Section 4 presents the suggested methodology – clustering, feature extraction, and unsupervised behaviour classification – as a proof of concept for automatic herding behaviour detection. Section 5 discusses different experiments results using various spatio-temporal cluster feature combinations. In Section 6 we present the main findings and suggestions for future work.

2 Related work

Research on crowd behaviour and herding is extensive. A frequent outcome in such research is a simulation depicting more realistic behaviours. Movement patterns, such as hotspots, that arise because of people's biases are also analysed in both outdoor and indoor scenarios. State-of-the-art evacuation systems can use personalised warning messaging and routing directions. This section discusses literature in these areas.

2.1 Simulations displaying social behaviours

Most of the computer-related works that study herding behaviour have the goal of producing simulations. A number of simulations that take into account the microscopic interactions during an evacuation is proposed in the literature (e.g., [8]). Agent-based models are a popular way of creating simulations that include social interactions between the agents. Interactions such as negotiation, following, or collision avoidance can be coded to reproduce common behaviours like herding [19], while cellular automata are frequently used in simulating evacuations [30]. Behaviours such as "freezing by heating", "faster is slower" and herding behaviour are identified in simulations using a social force model [11]. Although such models are successful in displaying social behaviours, including herding, their identification of such behaviours is done in a visual and manual manner. That is, there is a human checking for instances of behaviour, and papers usually include an image of the seen behaviour. Our method goes a step further by making the behaviour detection automatic.

2.2 Movement and behaviour detection

A number of methods are used for analysis and detection of movement and behaviour patterns. For example, trajectory prediction models using mobile data have been proposed in normal circumstances [17], and during disasters [23]. Such prediction is done with extensive prior knowledge about a person's movement habits. For example, they rely on social networking data to know a person's usual locations. Our model relies on real time and short trajectory knowledge for prediction, and focuses on specifically identifying irrational behaviours.

Hotspot detection is a useful mechanism for alerting stakeholders about people's concentrations. Many hotspot detection mechanisms have been developed for indoor evacuations [9] and crowd disasters [4]. While hotspot detection is useful, detection of other behaviours is rather scarce. The current work specifically targets the detection of harmful herding behaviour.

2.3 Personalised alert and evacuation assistants

Before the wide adoption of mobile technologies, alert systems targeted a large number of persons through mass media. An overview of past research regarding the warning stage of a disaster can be found in [7]. An overview of how warning response, adoption, and timing affects people's behaviours during disasters is given in [24]. With the rise of microblogging services, such as Twitter, further research was conducted in message personalisation. The proposed method aims at the wider use and integration of personalised alert messages produced by observing the real time behaviour of people during emergency evacuations.

Personalised warning messages and assistance is a possibility due to improved research on video tracking technologies and the use of mobile phones. A study in [3] underscores the research needed to send localised warning messages to people's cell phones during an imminent hazard. Furthermore, mobile phone sensors provide grounds for context-aware indoor navigation. A routing system is proposed in [28] that exploits cell phone sensors in order to have context knowledge in real-time, for example blocked exits. A robot-assisted evacuation method is proposed in [25] improving evacuation times and is tested in a simulated shopping mall environment. These approaches fail to take into account people's beliefs and biases, which may affect their successful adoption. The work in this paper takes a first step into examining people's behaviours, and extracts characteristics that can detect potentially harmful herding caused by cognitive biases. Some relevant work has looked into the role of leaders during emergencies [29]. The authors argue for the optimal number and position of evacuation assistants. However, they only take into account formally defined leaders, rather than leaders that naturally arise in groups of people during emergency evacuations. The latter type of leaders and their behaviour is examined in this work.

3 Background

This section discusses the concepts inherent to herding behaviour, and describes certain methods used in each of the three steps of our methodology: people cluster identification in evacuations, feature extraction to describe cluster behaviour, and a learning model for cluster classification.

3.1 The problem with herding behaviour

Herding behaviour is a cognitive bias examined in early psychology and sociology research [18, 5] comprising different contexts of everyday life. In the context of evacuations, herding behaviour is exhibited when people follow others, without knowing with certainty where the group is heading to. Although herding can successfully lead people towards a safe place, it can also lead them to prevent successful evacuations, as evidenced by past studies in bushfires [1] and indoor evacuations [9]. A study in [10] considers a balance between individualistic behaviour and herding behaviour to be optimal for indoor evacuations. This research focuses on identifying harmful herding behaviour. For language consistency, we distinguish **benign** herding behaviour – when people follow others successfully to safety – from **harmful** herding behaviour – when the group fails to find an exit.

3.2 Moving people clustering

The first step in our method is cluster identification during evacuation. We borrow ideas from previous works that have studied crowd clustering [21, 20] and groups of points moving together [14]. Clustering methods often use Euclidean distance for assigning members in a cluster. Nevertheless, several applications, including this work, require non-traditional distance measures, such as graph distance or similarity measures. The suggested method clusters people in a floor setting; therefore, people separated by a wall should not be assigned to the same cluster, even if their Euclidean distance is short. Spectral clustering takes a similarity matrix as input for identifying clusters [15]. Such a similarity matrix can be computed from any pair-wise distance metric of the instances – persons in our case. The way the similarity matrix is built in this work is explained in Section 4.1.

3.3 Feature extraction for conveying herding behaviour

The second step involves the collection of spatio-temporal information from clusters previously defined, to be encoded into a feature vector. The set of these features is used to describe the harming herding behaviour that a cluster might be displaying, and is one of this work's major contributions. Previous work in activity recognition and anomaly detection from trajectories provides inspiration for this model. As in many learning problems, feature engineering is a crucial step towards an effective model. Additionally, motion information representation is the basis in spatiotemporal analysis [13]. Consequently, several approaches encode trajectory information (e.g. distance between objects, acceleration) into their feature vector [32, 22].

The aim of this work is to produce a feature vector that describes herding behaviour. Relevant characteristics are:

- **Characteristic 1.** Forming groups.
- **Characteristic 2.** Moving towards or away from an exit.
- **Characteristic 3.** Delegating wayfinding to the leader and then moving collectively.

Characteristic 1 is achieved by the clustering step. Characteristics 2 is encoded into the feature vector by calculating the distance change from the cluster towards or away from exits. For satisfying characteristic 3, the trajectory from the cluster's leader is analysed from previous time steps. How these features are formally obtained is explained in Section 4.

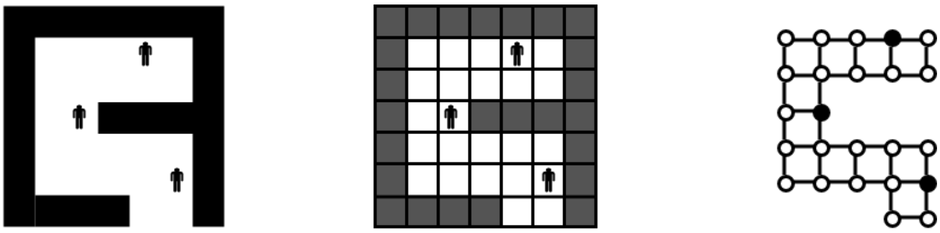
3.4 Learning model

The proposed approach uses an unsupervised learning method to identify the clusters heading towards a dead end. Machine learning methods are now a common practice for categorising a set of instances. Each instance comprises a set of features and may contain continuous or discrete values. As such, learning methods are used for the detection of differing behaviours or anomalies. Previous works for categorising trajectories and behaviours have used semi-supervised [22] and unsupervised [32] learning models by means of different clustering algorithms such as Gaussian mixture, or Latent Dirichlet Allocation (LDA).

The proposed method compares two different and widely used unsupervised learning algorithms: k-means clustering and hierarchical clustering. K-means clustering finds a centroid per cluster and uses a distance based metric to classify points based on the proximity to the centroid. Hierarchical clustering performs better on non-linear and high-dimensional data. Our method has high dimensionality as it uses up to 55 different features.

3.5 Data sources and simulation

The proposed method assumes known coordinate positions of each person for the duration of the emergency evacuation. As real data of this type are scarce, a simulation instead is used, while current complementary research efforts are developing technologies for real-time monitoring of evacuees [6]. Also, in order to focus on examining herding behaviour, the effects of indoor landmarks on way finding, or the limits of maximum evacuation times and multi-level building complexities are left for future consideration. The simulation is built based on the general guidelines provided in [19]. In that work, the authors construct a simulation that displays different “nonadaptive crowd behaviours”, including herding behaviour. They build an agent-based model in which agents display social interactions, such as negotiation or people-following. They define a set of possible actions and different types of profiles. The simulation used in this paper uses a subset of those actions and profiles for displaying the expected behaviour (herding). The following set of possible actions is used: (i) **Random walk** - heads towards a random direction in sight, (ii) **Seek** - if the exit is known, heads to the exit; otherwise, keep looking for the exit by going towards doors, and (iii) **Target following** - follow the nearest group of people.



(a) Floor layout. Black means walls, white empty space.

(b) Grid on top of floor layout.

(c) Graph G representing the floor layout.

Figure 2 A 7x7 floor layout discretization.

Table 1 Similarity matrix of sample pair distances.

	p1	p2	p3
p1	0	-4	-9
p2	-4	0	-5
p3	-9	-5	0

Accordingly, three profiles are used for agents. The exact probabilities are not provided in [19], so they are based on evacuation behaviour findings in [31] and [16]. Each profile contains the probabilities for the actions it can take (probabilities must sum up to 1).

- Adult: $random\_walk = 0.2$, $seek = 0.4$, $target\_following = 0.4$
- Child: $random\_walk = 0.3$, $seek = 0.2$, $target\_following = 0.5$
- Elderly: $random\_walk = 0.0$, $seek = 0.7$, $target\_following = 0.3$

4 Method

The methodology used to detect harmful herding behaviour comprises three steps. The purpose of detecting herding behaviour is to know if people may be headed towards a dead-end, or taking a much longer evacuation route. In this case, the behaviour belongs to a group of people rather than to individuals. As such, the method first identifies clusters at each time step. A feature vector is extracted from each of these clusters and an unsupervised learning method is used to predict the ones displaying herding behaviour. The following subsections describe each step in the methodology.

4.1 Clustering

The floor layout is discretised into a grid and represented by a graph G. Each grid cell that is not a wall is a node of G. Figure 2a shows a sample 7x7 floor layout, figure 2b shows a grid on top of it, and figure 2c shows the respective graph G. Black dots represent people, and each vertex in G is connected to the nodes up, down, left and right.

At each time step, each person is located at a node of G (as in Figure 2c) and the distances between each pair of persons is computed into a *similarity matrix*. The sample persons in Figure 2 are located at $(4, 5)$, $(2, 3)$, and $(5, 1)$, and we call them p_1, p_2, and p_3 respectively. The *similarity matrix* for the sample 3 persons is shown in Table 1.

The similarity matrix contains the distance of each pair of points multiplied by -1, to represent a *similarity* rather than *dissimilarity*. Computing shortest paths in a graph at each time step can be time consuming. Therefore, the paths are pre-computed and stored in a hash table $PATHS$ in memory, such that for pair p_1 and p_2 we can obtain its graph distance by calling $PATHS(p_1, p_2)$. An additional variable, $EXITS$ stores distances from each person to the nearest exit (e.g., $EXITS(p_1)$).

The *similarity matrix* is then input to the spectral clustering algorithm for cluster identification. The clustering step represents virtually the whole method's time complexity as $O(n^3)$, while next steps run in linear time or less. When clustering is performed at every time step, it might produce temporal errors. For example, people passing each other in opposite directions could temporarily be close together but shouldn't be considered part of the same cluster. To address this, we use a parameter τ that represents the number of time steps required for a group of people to be considered as 'traveling together'.

Over time, clusters may add members, lose members, split, or even completely dissolve. Therefore, identifying a cluster over time requires some flexibility about its members. Thus, we define the equivalence between cluster C_1 from time step t and cluster C_2 from time step $t+1$, if $|C_1 \cap C_2| \geq 2 \implies C_1 \equiv C_2$, and we define the age of cluster C as $T_C = |C_t, ..., C_{t+n}|$ where $C_i \equiv C_{i+1} \forall i \in \{i | t \leq i \leq t+n\}$. With that, the age constraint for cluster C to be considered herding is $T_C \geq \tau$. For the experiments described in Section 5, a visual inspection of the moving clusters showed a τ value of 5 ensures a group of people are moving together. This paper doesn't cover the effect that varying values of τ can have on the discovery of moving clusters. For more thorough techniques on this area the reader is referred to [12].

4.2 Behaviour Definition and Feature Vector

Once the clusters of people traveling together are identified, a cluster feature vector is extracted from each. The method relies on the group's leader past trajectory as one of the most important features for behaviour description. So before listing the feature candidates, a formal spatio-temporal definition for 'leader' is provided.

Leader identification

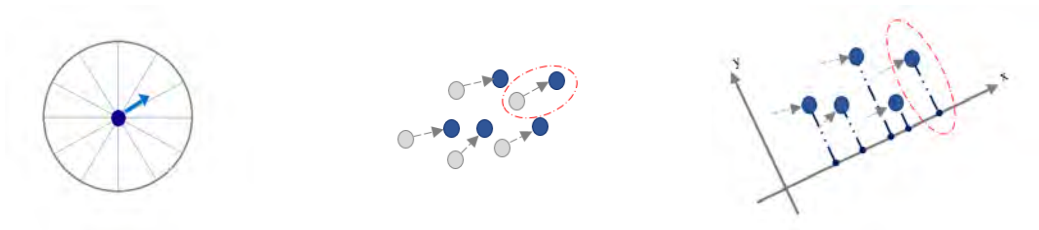
In plain terms, the leader is the person guiding the group. However, that definition is not enough for identifying the leader in spatio-temporal data. A simple definition of cluster leader is used where the leader is considered the most salient point in the cluster's moving orientation, as depicted in Figure 3b. More elaborate methods for leader identification fall out of the scope of this study, but the interested reader is directed to [2].

To obtain the cluster's orientation, a 12-direction discretised space is used (Figure 3a). The discretised angle of point m, is $\angle m$. Then, given a cluster C with n members $m_0, ..., m_n$, the orientation of C is defined as the mode of the discretised angles of its members: $\angle C = Mo(\angle m_0, ..., \angle m_n)$. Once $\angle C$ is computed, a plane rotation of $\angle C$ is performed, as shown in Figure 3c, and every member m_i is projected into the x-axis. From there, leader l_p is the member with the p-largest projection in the x-axis.

Feature candidates

In order to comply with the characteristics of herding behaviour listed in Section 3.3, three kinds of feature candidates (FC) are extracted from cluster C with members $m_0, ..., m_n$:

- $\boldsymbol{FC_1}$ **- Cluster's distance to exit** (dist_to_exit) – The average shortest distance to the closest exit for each member m of C. Distances from m to the nearest exit are stored in the $EXITS$ hash table. So $dist\_to\_exit(C)$ determines this feature's value.



(a) 12-direction discretised space.

(b) Cluster movement. Darker dots are points at time t, lighter ones are the same points at time $t-1$.

(c) Projection of cluster members into a rotated x-axis for finding the leader.

Figure 3 Graphical definitions of orientation and cluster leader.

- $\boldsymbol{FC_2}$ **- Cluster's distance change towards exit** (dist_change_i) – The change in the average shortest distance from the i previous time steps to the current one. If the change is negative it means the cluster is getting closer to the exit. This is computed by checking the positions of the members in the previous time step and using the stored distances in the $EXITS$ hash table.
- $\boldsymbol{FC_3}$ **- Leader's trajectory** (leader_l_away_steps_i) – This field refers to the number of steps the group leader l has taken away from the exit in the last i time steps. For instance, leader_1_away_steps_5 (i.e., $l = 1$ and $i = 5$) counts how many of the previous 5 steps leader l took away from the exit. The value would range from 0 to 5 in this example, and from 0 to i in general.

At every time step, clusters are identified and features extracted. The set of features to use can be the full set described, or a subset of it. The experiments in Section 5 use different subsets of the features explained here. Every feature set is stored and used in the unsupervised learning method explained in the next subsection.

4.3 Unsupervised Learning

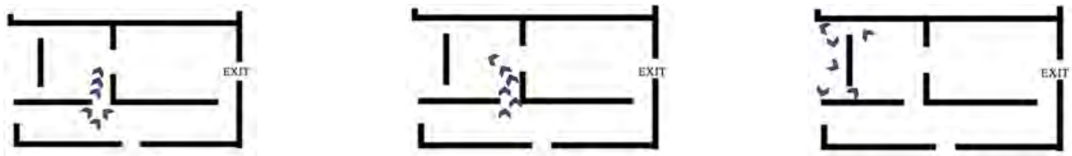
The final step of the method is applying a learning algorithm for classifying clusters displaying benign or harmful herding behaviour. Thus, two classes are defined: *benign* and *harmful*. As mentioned in Section 3.4, two unsupervised learning algorithms are used: k-means (KM) and hierarchical clustering (HC). Additionally, three baselines are used for thorough comparison:

- Zero rule: Classifies every instance as the most popular one. In this case, it will classify everything as harmful.
- Random: Classifies each instance randomly as either harmful or benign.
- Random with distribution: Classifies similar to the Random baseline but uses prior knowledge about the distribution of harmful and benign instances.

Comparing to a "dumb rule" classifier, such as Zero rule, ensures the proposed method meets minimum requirements, while comparing to the random baselines ensures it does not perform randomly. Comparisons with baselines ascertain credibility and robustness.

4.4 Evaluation Method

To evaluate the suggested method, every instance is associated with a label – *benign* or *harmful* – describing its behaviour. An instance refers to a cluster from its identification until its dissolution. Figure 4 shows a cluster in different stages of its lifespan.



(a) Identification. **(b)** Harmful herding. **(c)** Dissolution.

Figure 4 Three stages in the lifespan of a cluster of 6 persons: (a) the group is identified as such, (b) the human annotator identifies the group is taking the wrong turn, (c) the cluster dissolves after an exit is not found.

The ground truth holds every instance's label based on the *ultimate* cluster's destination. That is, if the cluster ends up in a dead-end or clearly goes in the wrong direction, it is labeled as *harmful*, whereas if it ends up exiting the building or closer to the exit, it is labeled as *benign*. The suggested method is evaluated on its ability to detect harmful herding behaviour but also on detection timeliness, as it is expected to make detections early on. Therefore, three checkpoints along the lifespan of a cluster are defined (Figure 4):

- **Checkpoint 1 (CP1)** – At cluster identification. This is when the cluster is identified by the clustering method defined in Section 4.1 (Figure 4a).
- **Checkpoint 2 (CP2)** – At human detection point. That is, when the human tester first realises that the cluster is headed towards the wrong direction (Figure 4b).
- **Checkpoint 3 (CP3)** – At cluster dissolution. This is when the clustering method defined in Section 4.1 stops identifying the former cluster members as one (Figure 4c).

Then, to assess detection timeliness, five scores – called **detection scores** – are defined using the checkpoints:

- **Early detection** (ED) – Number of harmful instances detected before CP1.
- **Detection** (D) – Number of harmful instances detected between CP1 and CP2.
- **Late detection** (LD) – Number of harmful instances detected between CP2 and CP3.
- **No detection** (ND) – Number of harmful instances not detected at all.
- **False warnings** (FW) – Number of benign instances detected as harmful at any time.

Additionally, **unified scores**, allowing a comparison between the method's detection times and the visual inspection (VI), are defined:

- **Before VI** (BVI) – The number of harmful instances detected before CP3 (faster than VI), plus the number of benign instances not identified as harmful. Formally, $BVI = ED + D + (TB - FW)$, where TB is the total number of benign instances in the ground truth.
- **After VI** (AVI) – The number of harmful instances detected after CP3 (slower than VI), plus the non-detected instances, plus the false warnings. Formally, $AVI = LD+ND+FW$

It is worth noting that false warnings tend to be sensitive, as a single harmful detection in a whole benign trajectory would yield a false warning. For that reason, a tolerance variable is introduced. Each of the detection scores checks for at least *one* harmful prediction. Using the tolerance variable t, the detection scores have to check for at least t harmful predictions, before classifying it as harmful.

The manual labeling is performed visually by a human observer. Although not optimal, this labeling allows for performing a proof-of-concept evaluation method against human judgment. In the future, a more thorough labeling mechanism such as domain expert labeling, or labeling from multiple annotators can be used.

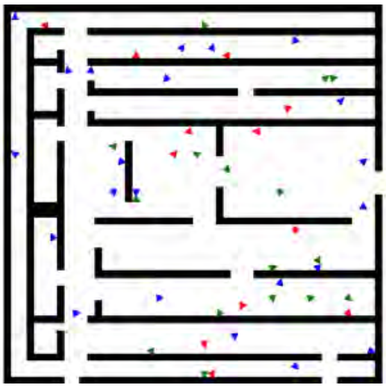


Figure 5 Initial setup of the simulation.

5 Experiments

Data for the experiments are generated by running the simulation described in Section 3.5. Location data for each agent at each time step are recorded in a text file. The text file is used as the input to the suggested method. The simulation is realised using the GAMA[1] simulation software. In the simulation, 50 agents are placed on a 50x50 grid. The layout of the grid resembles a building layout with walls and exit doors. Figure 5 shows the initial setting for the simulation to run.

Having obtained the simulation data, the main objective of these experiments is to test which features in the cluster feature vector describe best the herding behaviour. Feature sets are built using feature candidates (FC_{1-3}) described in Section 4.2, as follows:

- Feature set 1 (FS_1) – The information this feature set contains is the cluster's distance to the exit (FC_1), the cluster's previous movements (FC_2 with $i = 5$), and the trajectories of 3 leaders (FC_3 with $l = 3$ and $i = 20$).
- Feature set 2 (FS_2) – In this feature set, distance to the exit (FC_1) is not used, for checking its relevance. Considered are: cluster movement (FC_2 with $i = 5$) and leader trajectory (FC_3 with $l = 1$ and $i = 20$).
- Feature set 3 (FS_3) – Leader information (FC_3) is not considered, to check its relevance. Considered are only distance to exit (FC_1) and cluster movement (FC_2 with $i = 5$).

The values for the number l of leaders and number i of steps to check from past trajectory were chosen based on the behaviour definition and by performing several preliminary tests of the method with a number of combinations. Three experiments are performed, summarised in Table 2. Every experiment runs the classification step using both k-means (KM) and hierarchical clustering (HC):

- **Experiment 1** sets the tolerance value to 1, the number of instances to 31 and all three feature sets are compared.
- **Experiment 2** is similar to Experiment 1, but using a tolerance value of 2.
- **Experiment 3** is used to check whether the algorithm would benefit from having more instances to cluster by increasing the number N of instances and using the best performing feature set – FS_2 as seen later – with tolerance $t = 1$.

[1] http://gama-platform.org/

Table 2 Parameters used in every experiment.

Experiment 1			Experiment 2			Experiment 3		
$tolerance = 1$			$tolerance = 2$			$tolerance = 1$		
FS_1	FS_2	FS_3	FS_1	FS_2	FS_3	FS_2		
$N = 31$			$N = 31$			$N = 21$	$N = 31$	$N = 52$

Table 3 Results of Experiment 1, using tolerance $t = 1$. Showing detection scores (ED, D, LD, ND), false warnings (FW), and unified scores (BVI, AVI) of k-means (KM) and hierarchical clustering (HC) using different feature sets (FS_i). Baselines are shown beside them for comparison

	FS_1		FS_2		FS_3		Baselines		
	KM	HC	**KM**	**HC**	KM	HC	ZR	R	RD
ED	48%	52%	57%	76%	100%	100%	100%	86%	81%
D	24%	33%	43%	24%	0%	0%	0%	14%	19%
LD	19%	10%	0%	0%	0%	0%	0%	0%	0%
ND	10%	5%	0%	0%	0%	0%	0%	0%	0%
FW	36%	55%	36%	55%	100%	100%	100%	100%	100%
BVI	69%	72%	**88%**	**81%**	66%	66%	66%	66%	66%
AVI	31%	28%	12%	19%	34%	34%	34%	34%	34%

6 Results Analysis

Tables 3 and 4 show the complete results of Experiment 1 and 2, respectively. The tables contain detection and unified scores (Section 4.4) for a thorough comparison. The tables present the results of k-means and hierarchical clustering for all three feature sets (FS_1, FS_2, FS_3). The three baselines defined in Section 4.3 – Zero Rule (ZR), Random (R), Random with distribution (Rd) – are placed next to the results for comparison.

Ideally, a method would detect every harmful herding behaviour early on ($ED = 100\%$). Even though the baselines have a perfect or near-perfect ED score — since ZR classifies everything as harming ($ED = 100\%$) — they also have a 100% false warning rate (FW), which renders these baselines unreliable. Hence, the consolidated BVI score is a better indicator of overall performance, as it penalises either low detection, or high false warning rates. Figures 6 and 7 show the BVI score in Experiments 1 and 3, respectively, while the main findings of the analysis are as follows:

Leader trajectory is the best herding predictor. Overall, the best performing feature set is FS_2 with either k-means, or hierarchical clustering with a $BVI = 88\%$ and $BVI = 81\%$ with $t = 1$ (Figure 6) respectively, and $BVI = 81\%$ and $BVI = 84\%$ with $t = 2$. These algorithms all perform well above the baselines. These positive results suggest the features chosen, namely the leader trajectory and the recent cluster movement, were appropriate. When comparing Experiments 1 and 2, as tolerance increases, the FW score decreases as expected, but the overall BVI is not improved.

Distance to exit is not meaningful. Low results of FS_3 suggest the distance to the exit (the feature not present in FS_2) is not a trusting feature, as it makes the classifier act randomly. This is probably the reason for the lower performance of FS_1 compared to FS_2, as it contains the *dist_to_exit* feature. This observation is reasonable, given that long distance from the exit does not necessarily mean the group is lost or not heading towards the exit.

Table 4 Results of Experiment 2, using tolerance $t = 2$. Showing detection scores (ED, D, LD, ND), false warnings (FW), and unified scores (BVI, AVI) of k-means (KM) and hierarchical clustering (HC) using different feature sets (FS_i). Baselines are shown beside them for comparison.

	FS_1		FS_2		FS_3		Baselines		
	KM	HC	**KM**	**HC**	KM	HC	ZR	R	RD
ED	33%	38%	29%	43%	62%	62%	62%	52%	57%
D	24%	29%	57%	52%	33%	33%	33%	38%	33%
LD	29%	24%	14%	5%	5%	5%	5%	10%	2%
ND	14%	10%	0%	0%	0%	0%	0%	0%	0%
FW	36%	36%	27%	36%	100%	100%	100%	100%	100%
BVI	59%	66%	**81%**	**84%**	62%	62%	62%	59%	59%
AVI	41%	34%	19%	16%	38%	38%	38%	41%	41%

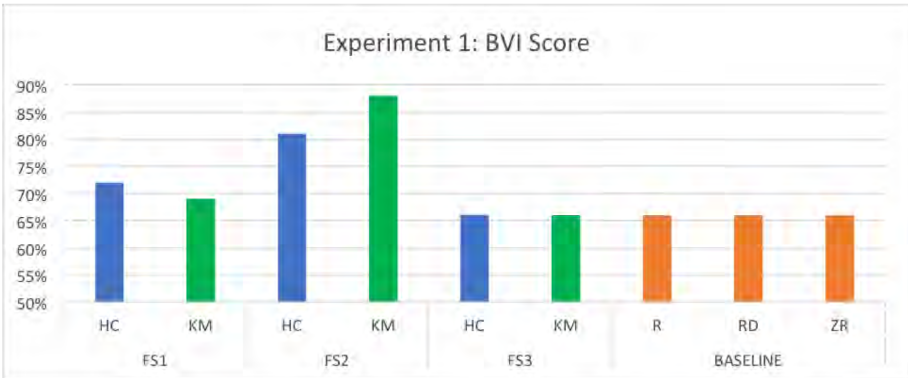


Figure 6 Experiment 1 results summary. BVI score is displayed comparing the suggested method to the baselines.

Increasing number of instances improves performance. Figure 7 shows the results of Experiment 3, depicting how the scores change given an increasing number N of cluster instances. BVI score increases as N increases (except for k-means in $N = 31$), implying that the suggested method benefits from a higher number of instances. FS_2 is used in this experiment as it was the best performing feature set in the previous experiments.

7 Conclusions and future work

This paper presents a method for automatic, early detection of harmful herding behaviour using spatio-temporal information from clusters of people. The method comprises three steps. First, groups of people moving together are identified using clustering algorithms with added constraints. Second, relevant spatio-temporal information from the identified clusters is collected. Second, the extracted features are combined to spatially and temporally describe a herding behaviour. To achieve this, the position changes of the cluster and the cluster leader's movement trajectory are examined. The method assumes the leader's trajectory to be a most relevant feature for identifying the behaviour. Third, the observed clusters are classified as displaying either benign or harmful herding behaviour, using an unsupervised learning method.

The experimental results show promise towards advancing the understanding of herding behaviour effects. Seven different scores are defined to assess the method's ability to detect

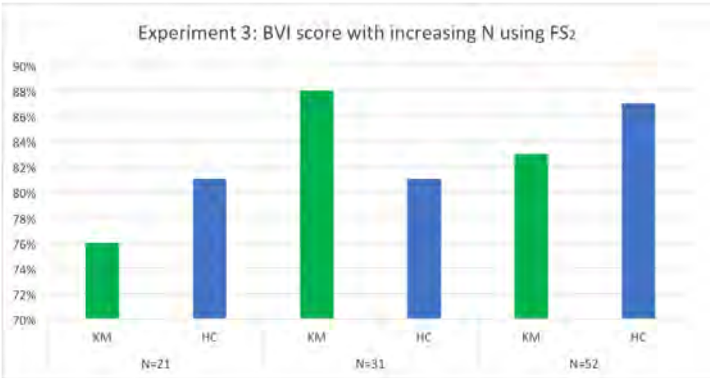


Figure 7 Experiment 3 results. FS_2 is used with an increasing number N of instances.

harmful behaviours and compare it to a human observer. In every experiment run, both algorithms (k-means and hierarchical clustering) are superior to the three baselines used. Different combinations of features were tested. The major findings are:

1. Features regarding leader trajectory and recent distance changes from the cluster to the exit best predict harmful herding behaviour, yielding above 80% of the BVI unified score in the experimens.
2. Distance to the exit (without considering movement) harms the prediction when added into the feature set, making it classify randomly.
3. Even though increasing the method's tolerance does not produce better results overall, it does decrease the amount of false warnings. This is useful in systems where issuing warnings is expensive, so additional confidence is needed.
4. The method benefits from large cluster instances in the data, which means that it scales well in environments with big crowds that need to evacuate in an emergency situation. Higher values of N, however, mean more time-consuming manual labeling for evaluation.

The harmful herding behaviour identification method can be further improved. Different graphs can be used, such as the visibility graph [26] or a bigraph [27] in the clustering step. The features extracted from the clusters can be improved by looking into more in-depth analysis of who the leader of a group is, rather than identifying the topmost one as such. A supervised learning method for behaviour classification can be compared to its unsupervised counterpart. An approach that would replace both learning approaches is a rules-based one where, given thorough domain knowledge, strict rules can be placed for the prediction of the harmful herding behaviour. Pertaining to the evaluation method, perhaps the most immediate step forward is the use of a real evacuation scenario datasets. Finally, herding is only one of several behaviours elicited by cognitive biases during disasters. Other biases such as the normalcy bias, confirmation bias, planning fallacy [1], may lead to equally harming behaviours during emergency evacuations. Consequently, future work may focus on identifying other behaviours, or even providing a bigger unified framework for irrational behaviour detection.

References

1 Carole Adam and Benoit Gaudou. Modelling human behaviours in disasters from interviews: Application to melbourne bushfires. *Journal of Artificial Societies and Social Simulation*, 20(3):12, 2017. URL: `http://jasss.soc.surrey.ac.uk/20/3/12.html`, `doi:10.18564/jasss.3395`.

2 Mattias Andersson, Joachim Gudmundsson, Patrick Laube, and Thomas Wolle. Reporting leaders and followers among trajectories of moving point objects. *GeoInformatica*, 12(4):497–528, 2008.

3 Hamilton Bean, Jeannette Sutton, Brooke F Liu, Stephanie Madden, Michele M Wood, and Dennis S Mileti. The study of mobile public warning messages: A research review and agenda. *Review of Communication*, 15(1):60–80, 2015.

4 Arianna Bottinelli, David TJ Sumpter, and Jesse L Silverberg. Emergent structural mechanisms for high-density collective motion inspired by human crowds. *Physical review letters*, 117(22):228301, 2016.

5 James Coleman. Foundations of social theory. *Cambridge, MA: Belknap*, 1990.

6 Davide Dardari, Pau Closas, and Petar M Djurić. Indoor tracking: Theory, methods, and technologies. *IEEE Transactions on Vehicular Technology*, 64(4):1263–1278, 2015.

7 Nicole Dash and Hugh Gladwin. Evacuation decision making and behavioral responses: Individual and household. *Natural Hazards Review*, 8(3):69–77, 2007.

8 Marcus Goetz and Alexander Zipf. Using crowdsourced geodata for agent-based indoor evacuation simulations. *ISPRS International Journal of Geo-Information*, 1(2):186–208, 2012.

9 Dirk Helbing, Lubos Buzna, Anders Johansson, and Torsten Werner. Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions. *Transportation science*, 39(1):1–24, 2005.

10 Dirk Helbing, Illés Farkas, and Tamas Vicsek. Simulating dynamical features of escape panic. *Nature*, 407(6803):487, 2000.

11 Dirk Helbing, Illes J Farkas, Peter Molnar, and Tamás Vicsek. Simulation of pedestrian crowds in normal and evacuation situations. *Pedestrian and evacuation dynamics*, 21(2):21–58, 2002.

12 Panos Kalnis, Nikos Mamoulis, and Spiridon Bakiras. On discovering moving clusters in spatio-temporal data. In *International Symposium on Spatial and Temporal Databases*, pages 364–381. Springer, 2005.

13 Teng Li, Huan Chang, Meng Wang, Bingbing Ni, Richang Hong, and Shuicheng Yan. Crowded scene analysis: A survey. *IEEE transactions on circuits and systems for video technology*, 25(3):367–386, 2015.

14 Xiaohui Li, Vaida Ceikute, Christian S Jensen, and Kian-Lee Tan. Effective online group discovery in trajectory databases. *IEEE Transactions on Knowledge and Data Engineering*, 25(12):2752–2766, 2013.

15 Jialu Liu and Jiawei Han. Spectral clustering. In Charu Aggarwal and Chandan Reddy, editors, *Data Clustering: Algorithms and Applications*. CRC Press, 2013.

16 Min Liu and Siu Ming Lo. The quantitative investigation on people's pre-evacuation behavior under fire. *Automation in construction*, 20(5):620–628, 2011.

17 Xin Lu, Linus Bengtsson, and Petter Holme. Predictability of population displacement after the 2010 haiti earthquake. *Proceedings of the National Academy of Sciences*, 109(29):11576–11581, 2012.

18 Alexander Mintz. Non-adaptive group behavior. *The Journal of abnormal and social psychology*, 46(2):150, 1951.

19 Xiaoshan Pan, Charles S Han, Ken Dauber, and Kincho H Law. Human and social behavior in computational modeling and analysis of egress. *Automation in construction*, 15(4):448–461, 2006.

20 Daniel Roggen, Martin Wirz, Gerhard Tröster, and Dirk Helbing. Recognition of crowd behavior from mobile sensors with pattern analysis and graph clustering methods. *arXiv preprint arXiv:1109.1664*, 2011.

21 Shobhit Saxena, François Brémond, Monnique Thonnat, and Ruihua Ma. Crowd behavior recognition for video surveillance. In *International Conference on Advanced Concepts for Intelligent Vision Systems*, pages 970–981. Springer, 2008.

22 Rowland R Sillito and Robert B Fisher. Semi-supervised learning for anomalous trajectory detection. In *BMVC*, volume 1, pages 035–1, 2008.

23 Xuan Song, Quanshi Zhang, Yoshihide Sekimoto, and Ryosuke Shibasaki. Prediction of human emergency behavior and their mobility following large-scale disaster. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 5–14. ACM, 2014.

24 John H Sorensen and Barbara Vogt Sorensen. Community processes: Warning and evacuation. In *Handbook of disaster research*, pages 183–199. Springer, 2007.

25 Bo Tang, Chao Jiang, Haibo He, and Yi Guo. Human mobility modeling for robot-assisted evacuation in complex indoor environments. *IEEE Transactions on Human-Machine Systems*, 46(5):694–707, 2016.

26 Alasdair Turner, Maria Doxa, David O'sullivan, and Alan Penn. From isovists to visibility graphs: a methodology for the analysis of architectural space. *Environment and Planning B: Planning and design*, 28(1):103–121, 2001.

27 Lisa A Walton and Michael Worboys. A qualitative bigraph model for indoor space. In *International Conference on Geographic Information Science*, pages 226–240. Springer, 2012.

28 Jing Wang, Haifeng Zhao, and Stephan Winter. Integrating sensing, routing and timing for indoor evacuation. *Fire Safety Journal*, 78:111–121, 2015.

29 Xiaolu Wang, Xiaoping Zheng, and Yuan Cheng. Evacuation assistants: An extended model for determining effective locations and optimal numbers. *Physica A: Statistical Mechanics and its Applications*, 391(6):2245–2260, 2012.

30 Song Wei-Guo, Yu Yan-Fei, Wang Bing-Hong, and Fan Wei-Cheng. Evacuation behaviors at exit in ca model with force essentials: A comparison with social force model. *Physica A: Statistical Mechanics and its Applications*, 371(2):658–666, 2006.

31 CM Zhao, Siu Ming Lo, SP Zhang, and M Liu. A post-fire survey on the pre-evacuation human behavior. *Fire Technology*, 45(1):71, 2009.

32 Bolei Zhou, Xiaogang Wang, and Xiaoou Tang. Understanding collective crowd behaviors: Learning a mixture model of dynamic pedestrian-agents. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 2871–2878. IEEE, 2012.