Group Anomaly Detection for Spatio-Temporal Collective Behaviour Scenarios in Smart Cities

Andreas Lohrer
Ludwig-Maximilians-Universität München
München, Bavaria, Germany
Christian-Albrechts-Universität zu Kiel
Kiel, Schleswig-Holstein, Germany
alo@informatik.uni-kiel.de

Johannes Josef Binder
Ludwig-Maximilians-Universität München
München, Bavaria, Germany
johannes.binder@campus.lmu.de

Peer Kröger
Christian-Albrechts-Universität zu Kiel
Kiel, Schleswig-Holstein, Germany
pkr@informatik.uni-kiel.de

ABSTRACT

Group anomaly detection in terms of detecting and predicting abnormal behaviour from entities as a group rather than as an individual, addresses a variety of challenges in spatio-temporal environments like e.g. traffic and transportation systems, smart cities, geoinformation systems, etc. They provide information about a commonly large number of individual entities. Examples for such entities would be airplanes and drones, vehicles, ships but also people, remote sensors and any other information source in interaction with the environment. However, as point anomaly detection is quite common for revealing the abnormal behaviour of individual entities, the collective behaviour of the individuals as a group remains completely uncovered. For example potential for traffic flow optimizations or increased local traffic guideline violations cannot be detected by one single drive but by considering the behavior of a group of vehicle drives in this area. With this work-in-progress we elaborate the potential of group anomaly detection algorithms for spatio-temporal collective behaviour scenarios in smart cities. We describe the group anomaly detection problem in the context of urban planning and demonstrate its effectiveness on a public real-world data set for urban rental bike rides and stations in and around Munich revealing abnormal groups of rides, which allows to optimize the rental bike accessibility to the population and with that to contribute to a sustainable environment.

CCS CONCEPTS

• **Computing methodologies** → **Anomaly detection**; **Search methodologies**; **Machine learning approaches**; • **Information systems** → **Spatial-temporal systems**.

KEYWORDS

group anomaly detection, collective anomaly detection, smart cities, urban planning, machine learning, artificial intelligence

IWCTS '22, November 1, 2022, Seattle, WA, USA

ACM ISBN 978-1-4503-9539-7/22/11…$15.00
https://doi.org/10.1145/3557991.3567801

ACM Reference Format:
Andreas Lohrer, Johannes Josef Binder, and Peer Kröger. 2022. Group Anomaly Detection for Spatio-Temporal Collective Behaviour Scenarios in Smart Cities. In *The 15th ACM SIGSPATIAL International Workshop on Computational Transportation Science (IWCTS '22), November 1, 2022, Seattle, WA, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3557991.3567801

1 SPATIO-TEMPORAL COLLECTIVE BEHAVIOR SCENARIOS

In times of an increasingly growing world population regions and living areas without information providing people, species, devices and sensors are becoming rare. Burdening the planet on the one hand-side spatially exhaustive and continuously available information enables for sustainable counter-steering measures like environmental monitoring, earth observation or other resources optimizing initiatives like smart cities, intelligent transportation systems and others. What all these application areas have in common, is the fact that they can support many of their goals effectively by considering information from individuals as a group instead of, or additionally to, as a single entity. The information of entities as members of a group represents together a collective behavior, which can be further differentiated as normal or abnormal group behavior. In spatio-temporal environments there are various scenarios in which abnormal group behaviors are desired to be uncovered. The following examples provide a short overview of groups with potential abnormal behavior. Images of human crowds like travelers, tourists, event visitors, etc., which are moving or behaving differently as expected can disclose infrastructure capacity shortages or pandemic developments after contact tracking. In the wildlife, remote sensors and images of herds, swarms or packs of animals or organisms can abnormally share or visit unexpected areas at unusual times, move together, etc. which allows to infer populations, living area circumstances or species interactions. Furthermore, GPS and operation sensors from fleets of airplanes, vehicles or ships showing abnormal fuel, energy or wearing part consumption can provide insights about route environments, changes according to driving guidelines or routes. Considering economical domains, the group behavior of salesmen, suppliers and customers (positioning, trajectories, location-based sales or purchases, etc.) at a shop, market or trade fair hall or at any other arbitrary urban region has the potential to reveal abnormal positive or negative environmental influences for purchases or sales. In logistic or grocery manufacturing locations allow RFID chips and other remote sensors to uncover

global issues related to maintenance, supplier quality or operation and hence to avoid negative impacts to the urban supply chain.

Trying to detect abnormal behavior as described in the aforementioned scenarios, traditional point anomaly detection algorithms would tend to fail since they are only classifying and scoring single instances as anomalies and neglect to consider behavioral dependencies within the group context.

For modeling and detecting group behavior in static and dynamic scenarios the framework of Toth et al.[8] refers to three sub-problems: 1) definition of group structures and relationships, 2) quantification of statistical properties and 3) change/deviation detection and scoring e.g. by hypothesis tests, discriminative or generative models. Group structures are either known upfront or obtained after clustering. In a subsequent step groups can be scored based on properties like density, orientation or shape, and also in combination with given contexts, e.g. time, location or other contextual subspaces.

Differently to statistical testing (e.g. [9]), established svm-based methods (e.g. [6], [7]) or deep learning approaches (e.g. [1], [5], etc.) learn group representations and score groups based on given distance measures or loss functions to distinguish between normal and abnormal groups.

In this work-in-progress we elaborate the potential of group anomaly detection (GAD) approaches for spatio-temporal collective behaviour scenarios in smart cities. We distinguish between instance- and distribution-based group anomaly detection, describe the group anomaly detection problem in the context of urban planning and demonstrate its effectiveness in scope of a case study on real-world mobility information.

The contributions of this work-in-progress can be summarized as follows:

- introduction of GAD as machine learning approach for spatio-temporal collective behavior detection in smart cities
- case study: distribution-based GAD as optimization approach for urban planning
- public real-world dataset for GAD on spatio-temporal mobility information of urban bike rentals

The remaining chapters of this paper are Section 2 introducing the GAD preliminaries, followed by Section 3 demonstrating the potential of distribution-based GAD for urban planning in smart cities by a case study before Section 4 describes experiments and results. The paper summarizes with a conclusion in section 5.

2 PRELIMINARIES

In this section we describe the differences between point-based and distribution-based group anomalies (cf. Foorthuis [3] types VII-i and VII-j) and provide a common understanding for the definition of the group anomaly detection problem.

2.1 Point- and Distribution-based Group Anomaly Detection

Chalapathy et al.[1] describe a "point-based anomalous group [as] a collection of individual pointwise anomalies that deviate from the expected pattern". In contrast, a point-wise anomaly is considered as a single abnormal data point or outlier. A further definition of Xiong et al. [10] describes a "point-based group anomaly [as] a group of individually anomalous points". These are also known as Micro-Clusters. Differently to point-based group anomalies the distribution-based anomalies, whose detection is of main interest in scope of this work to optimize urban planning, allow also normal points as part of abnormal groups. In the work [1] Chalapathy et al. mention that "distribution-based group anomalies [...] are seemingly regular however their collective behavior is anomalous". In addition to that the work of Xiong et al. [10] describes "a distribution-based anomaly [as] a group where the points are relatively normal, but as a whole they are unusual". As an example for type VII-j), Figure 2 shows in red the distributions for our case study.

2.2 Group Anomaly Detection Problem

For a better understanding of the common Group Anomaly Detection Problem a formal definition is provided in the following according to the notion of Chalapathy et al. [1] and Kuppa et al.[5]:

Group Anomaly Detection requires groups $\mathcal{G} = \{G_m\}_{m=1}^{M}$ where the mth group contains N_m instances with

$$G_m = (X_{nv}) \in \mathbb{R}^{N_m \times V} \tag{1}$$

where X_{nv} is the vth feature (v = 1,2,...,V) of instances n (n = 1,2,..., N_m) in the group G_m, $\mathbb{R}$ is a continuous value domain. The total number of individual instances is $N = \sum_{m=1}^{M} N_m$.
In GAD the behavior or properties of the mth group is captured by a characterization function denoted by $f : \mathbb{R}^{N_m \times V} \rightarrow \mathbb{R}^D$ where D is the dimensionality on the transformed feature space.

After a characterization function is applied to a training dataset, group information is combined using an aggregation function $g : \mathbb{R}^{M \times D} \rightarrow \mathbb{R}^D$

A group reference is composition of characterization and aggregation functions on the input groups with

$$\mathcal{G}^{(ref)} = g[\{f(G_m)\}_{m=1}^{M}] \tag{2}$$

Then a distance metric $d(\cdot,\cdot) \geq 0$ is applied to measure the deviation of a particular group from the group reference function. The distance score $d(G^{(ref)}, G_m)$ quantifies the deviance of the mth group from the expected group pattern where larger values are associated with more anomalous groups.

3 CASE STUDY: DISTRIBUTION-BASED GAD AS OPTIMIZATION APPROACH FOR URBAN PLANNING

3.1 Domain and Motivation

Smart cities[2] are commonly defined by the six aspects people, living, economy, mobility, environment and governance. This case study has its focus especially on the mobility aspect with the goal of demonstrating the effectiveness of distribution-based group anomaly detection for revealing optimization potential for urban planning.

In this scope the study investigates in the optimization of rental bike rides and returns in the area of Munich. The Munich Transport Association (MVG)[1] offers with "MVG Rad"[2] a bike rental service

[1] https://www.mvg.de
[2] https://www.mvg.de/services/mvg-rad.html

to the population allowing its users to rent a bike from a pool of 4500 bikes by booking it via the related service app. According to the MVG rental guidelines[3] there is a differentiation between free and non-free return regions[4] (cf. Figure 1). In free return regions users are allowed to return the bike at one of the 300 bike rental stations but also at well visible self-selected places. In non-free return regions bike returns are only allowed at a bike rental station. Otherwise the user gets charged with an extra fee for the return of the bike by MVG. Although these guidelines reasonably care for service accessibility the question arises if there is further potential for improvement considering the actual spatial rental bike ride and return behavior of different bike ride groups.

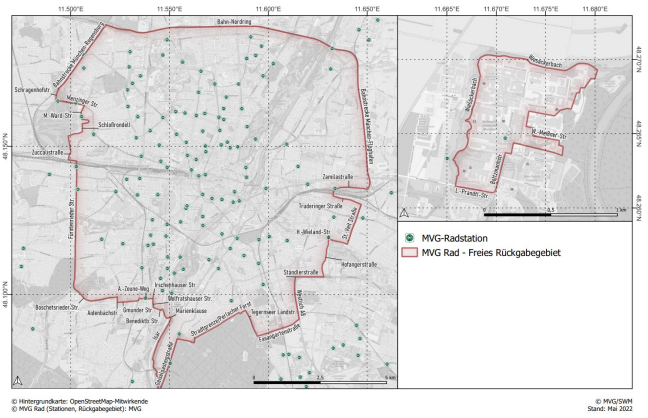

Figure 1: Map of Munich illustrating the borders between non-free and free return regions (free inside the polygons) with the city of Munich on the left and the free return region satellite of TU Munich on the right.[4]

In order to demonstrate the effectiveness of distribution-based group anomaly detection for revealing optimization potential in the scope of public rental bike rides for which there is usually no ground truth information about actual group anomalies available the following research questions arise:

- RQ1 How is it possible to detect abnormal groups of bike rides by distribution-based GAD with unsupervised machine learning methods?
- RQ2 How is it possible to define meaningful group structures for distribution-based GAD on bike rides?
- RQ3 How to model normal and abnormal group behavior for distribution-based GAD on bike rides?
- RQ4 How can GAD model predictions of abnormal bike ride groups be used to optimize urban planning?

For investigation in these research questions the study uses the continuously updated and publicly available real-world datasets provided at MVG Rad website[2]. These MVG Rad datasets contain bike ride information for rental start and end like time, GPS location (longitude and latitude), rental and return bike station name (e.g. Josephsplatz) and flag (GPS location matches station location yes or no) for the years 2015 to 2021 - in total 3342060 bike rides. All information provided ensures the anonymity of the bicyclists.

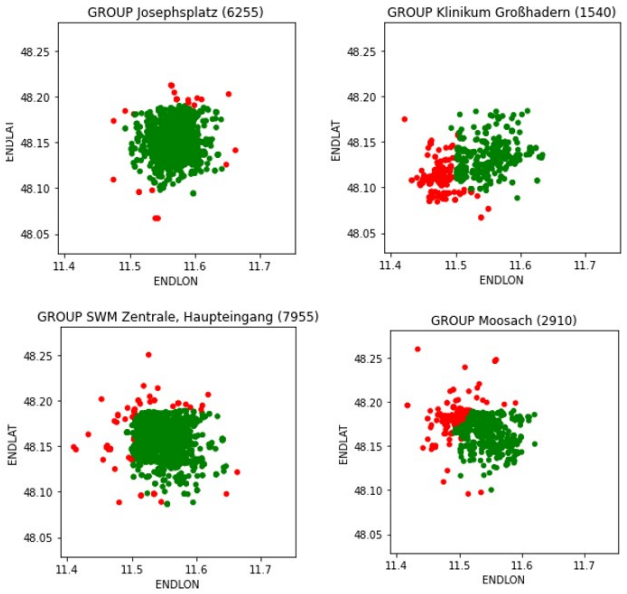

Figure 2: Illustration of normal return region ride distributions on the left and abnormal distributions on the right. (considering only returns of red non-free return region distributions and excluding green free return regions)

In this domain (group) behavior is represented by transportation activities of GPS equipped rental bikes providing its location only from the start and end location of a rental bike ride together with the related date and time. For this study this information is the foundation to learn patterns of spatio-temporal behavior.

3.2 Methodological Approach

Before distribution-based GAD can be applied in the context of urban planning the GAD problem is required to be defined accordingly to the given problem domain of rental bike rides.

Definition of groups and group members: Empirical analysis revealed that there are differences in the distributions of bike ride end locations (cf. Figure 2) of bike rides with specific rental bike stations as start location with station name as GID. Thus transportation activity groups can be represented by rental bike stations as start location whereas the corresponding bike rides shape the group as group members. The behavior of a member of a group $\mathcal{G}_m$ is described by $x_m \in \mathbb{R}^2$ with $x_m = (x_{EndLon}, x_{EndLat})$ as feature vector. In order to represent the collective behavior of all group members the covariance matrix Σ_m is calculated for all group members x_m and transformed as group feature vector $x'_m = (\sigma_{11_m}, \sigma_{12_m}, \sigma_{21_m}, \sigma_{22_m})$ with $x'_m \in \mathbb{R}^4$ representing the covariance of the distributions of GPS locations where the bike got returned. (RQ2)

Definition of normal and abnormal group behavior: Although different covariances are likely to get distinguished by distribution-based GAD methods they have semantically no meaning for optimizations in urban planning. Considering distributions of rental

[3]https://www.mvg.de/services/mvg-rad/mvg-rad-agb.html
[4]https://www.mvg.de/dam/jcr:d8d39828-995f-4d6a-892b-7466d041ab3b/geschaeftsgebiet-mvg-rad.pdf

stations (start locations) of free and non-free return regions one recognizes that the majority of rentals from free return regions also end in free return regions (cf. Josephsplatz or SWM Zentrale in Figure 2). According to the domain guidelines that in non-free regions bike returns are only allowed at bike stations the variance of return coordinates is expected to be very low. Nevertheless there are rental bike stations in non-free return areas which show frequent bike rides with high variance return coordinates of frequently visited points of interest (cf. Figure 2 with e.g. Moosach, Klinikum Großhadern), which is normal behavior from user perspective but abnormal behavior according to the domain guidelines. Rental stations with high variance bike rides to rarely visited points of interest remain considered as normal. (RQ3)

Model Selection for distribution-based GAD: Due to the non-availability of verified labeled group samples we choose an unsupervised machine learning approach to distinguish between normal and abnormal group distributions. Related but different to the One-class support measure machine (OCSMM)[6], introduced as generalization of One-Class Support Vector Machine (OCSVM)[7] for distribution-based GAD, we formulate the group behavior by the covariance matrix Σ_m for group member feature vectors x as described in the RQ2 paragraph above. The group feature vectors x'_m act as reshaped covariance matrices as training samples representing group distributions by following the top-down approach as well. In order to detect abnormal bike ride group distributions with high variant but frequently appearing end location coordinates our method of choice is OCSVM in combination with our covariance-based group feature vector x'_m and with the radial base function (RBF) as non-linear kernel method. (RQ1)

4 EXPERIMENTS

4.1 Experimental Setup and Evaluation

For evaluation of our approach we use the MVG Rad dataset introduced in section 3.1 for the years 2015 to 2018 (1247057 samples) and pseudo labels based on an empirically selected threshold (1e-5) for group size weighted absolute covariance sums. As stated in section 3.2 for RQ3 only bike rides of rental stations from non-free return regions are relevant to revealing optimization potential for urban planning. This and further filtering of invalid zero coordinates reduces the data from 1247057 rides from 329 groups to 14183 rides from 157 groups. For the hyperparameter search of the critical OCSVM parameter γ a grid search has been applied resulting in an ideal choice of $\gamma = 5.55$. Our approach has been evaluated based on the metrics Area under the ROC curve (AUROC) and Recall.

4.2 Result Discussion

Differently to the work of [6] which used group means as training samples for OCSVM with less success in their evaluation, we could achieve reasonable results (ROC: 0.886 Recall: 1.0) with OCSVM in combination with our covariance-based group feature vector representations as training samples for distribution-based GAD.(RQ1) The application of distribution-based GAD in combination with OCSVM[7] as unsupervised machine learning method allows to detect frequently visited non-free return regions as potential future free return region satellites to further improve the accessibility for bike rentals in the domain of urban planning. Such free return region satellites can be public points of interest like universities (e.g. the already established free return region satellite for TU Munich - cf. Figure 1), schools, sport or shopping centers (Fürstenried West), locations of large company areas (Am Hart), stadiums (Allianz Arena), fairground areas (Messestadt West) but also churches, hospital areas (Klinikum Großhadern), settlements or other area segments which are frequently visited but without directly visible bike rental station. The improved accessibility in free return region satellites could save time and effort, and motivate for a more frequent usage of bikes instead of combustion engine based vehicles to sustainably contribute to emission reduction (RQ4).

5 CONCLUSION

In conclusion we introduced GAD as machine learning approach for spatio-temporal collective behavior detection in smart cities. In scope of a case study we demonstrated the effectiveness of distribution-based GAD as optimization approach for urban planning revealing frequently visited areas without directly visible bike stations in non-free return regions and propose for these free region satellites as optimization potential to improve the accessibility of rental bikes to the population. For future work an additional investigation in temporal group anomalies evaluating if timely restricted free region satellites are reasonable and if distribution-based GAD is also effective for other transportation services, e.g. eScooters, or for other bike datasets like in [4], might be of interest. Furthermore, the comparison of the proposed methodology with other distribution-based GAD methods might be interesting as well as the investigation in further possibilities to optimize characterization, aggregation and scoring functions for effective GAD.

ACKNOWLEDGMENTS

This work has been funded by the German Federal Ministry of Education and Research (BMBF) under Grant No. 01IS18036A. The authors of this work take full responsibility for its content.

REFERENCES

[1] Raghavendra Chalapathy, Edward Toth, and Sanjay Chawla. 2018. Group anomaly detection using deep generative models. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 173–189.
[2] O.Y. Ercoskun. 2011. *Green and Ecological Technologies for Urban Planning: Creating Smart Cities: Creating Smart Cities.* Information Science Reference.
[3] Ralph Foorthuis. 2021. On the nature and types of anomalies: A review of deviations in data. *International Journal of Data Science and Analytics* 12, 4 (2021), 297–331.
[4] Thomas Koch and Elenna R Dugundji. 2021. Taste variation in environmental features of bicycle routes. In *Proceedings of the 14th ACM SIGSPATIAL International Workshop on Computational Transportation Science*. 1–10.
[5] Aditya Kuppa, Slawomir Grzonkowski, Muhammad Rizwan Asghar, and Nhien-An Le-Khac. 2019. Finding Rats in Cats: Detecting Stealthy Attacks using Group Anomaly Detection. *2019 18th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/13th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)* (2019), 442–449.
[6] Krikamol Muandet and Bernhard Schölkopf. 2013. One-class support measure machines for group anomaly detection. *arXiv preprint arXiv:1303.0309* (2013).
[7] Bernhard Schölkopf, John C. Platt, John Shawe-Taylor, Alex Smola, and Robert C. Williamson. 2001. Estimating the Support of a High-Dimensional Distribution. *Neural Computation* 13 (2001), 1443–1471.
[8] Edward Toth and Sanjay Chawla. 2018. Group deviation detection methods: a survey. *ACM Computing Surveys (CSUR)* 51, 4 (2018), 1–38.
[9] Weng-Keen Wong, Andrew Moore, Gregory Cooper, and Michael Wagner. 2003. WSARE: what's strange about recent events? *Journal of Urban Health* 80, 1 (2003), i66–i75.
[10] Liang Xiong, Barnabás Póczos, and Jeff G. Schneider. 2011. Group Anomaly Detection using Flexible Genre Models. In *NIPS*.