Erschienen in: EuroVis Workshop on Visual Analytics (EuroVA) / Turkay, Cagatay; von Landesberger, Tatiana (Hrsg.). - Genf : The Eurographics Association, 2019. - ISBN 978-3-03868-087-1
https://dx.doi.org/10.2312/eurova.20191122

Moving Together: Towards a Formalization of Collective Movement

J. Buchmüller[1] and E. Cakmak[1] and N. Andrienko[2] and G. Andrienko[2] and J. W. Jolles[1] and D. A. Keim[1]

[1]University of Konstanz, Germany
[2]Fraunhofer IAIS, Germany and City University, London, UK

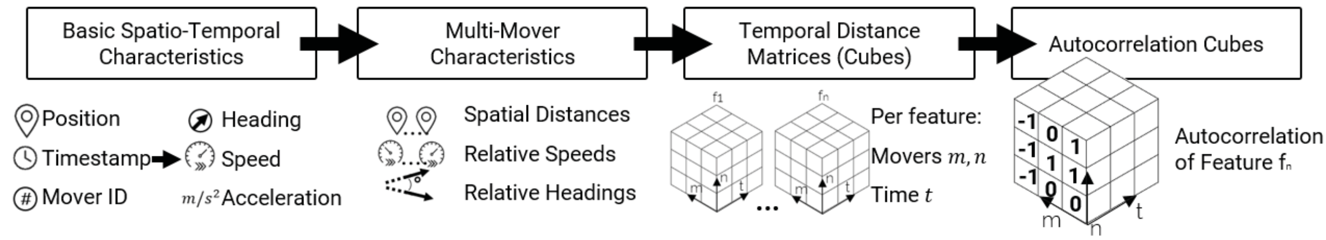

Figure 1: *We propose a data transformation process that enables a perspective shift in the exploration of coordinated movement. Basic spatiotemporal characteristics and derived features are transformed by the application of suitable distance functions to derive pairwise distances between the movers for all features. By calculating these distances for all time steps, 3D matrices or cubes can be extracted for each distance measure. By computing temporal autocorrelations in these cubes, coordination between two or more movers can be expressed.*

Abstract

While conventional applications for spatiotemporal datasets mostly focus on the relation between movers and environment, research questions in the analysis of collective movement typically focus more on relationships and dynamics between the moving entities themselves. Instead of concentrating on origin, destination and the way in between, this inter-mover perspective on spatiotemporal data allows to explain how moving groups are coordinating. Yet, only few visualization and Visual Analytics approaches focus on the relationships between movers. To illuminate this research gap, we propose initial steps towards a comprehensive formalization of coordination in collective movement based on temporal autocorrelation of distance matrices derived from basic movement characteristics. We exemplify how patterns can be encoded using autocorrelation cubes and outline the next steps towards an exhaustive formalization of coordination patterns.

1. Introduction

The field of spatiotemporal data analysis is a well-established research discipline with many real-world applications such as traffic analysis [AAR16], soccer analysis [SAMS*17] or animal movement analysis [SBJ*11]. In such applications, the movement of entities (movers) is typically influenced by their environment and other movers [WG09]. For instance, the movement of a soccer player depends on the motion of his team and the extent of the field. When it comes to the analysis of multiple movers, expert tasks frequently require a change of perspective by concentrating on the coordination between the movers instead of their environment (see Section 2.1). The coherent coordination between multiple proximate movers over time is also known as collective movement (moving cluster) [DWL08]. While manifold slightly divergent definitions for the term exist, a good general understanding of such self-organizing phenomena and the concomitant interactions can be taken for example from the work of Moussaid et al. [MGTH09] from a sociological or an ecological perspective by Westley et al. [WBTB18]. While various mechanisms for coordination exist, e.g. voice commands, these mechanisms are not reflected in movement data, and are difficult to track. Therefore, research on coordinated movement often relies on the detection of coordination by the analysis of spatio-temporal data only. Current approaches to detecting coordinated movements involve a prior definition of a particular pattern, such as "movers go in similar directions at similar speeds" or "movers keep within a certain distance from the group center." Such definitions do not cover all kinds of coordination, and it may be not known in advance what coordination patterns to expect. We propose an approach that does not assume any specific kind of coordination. The main concept of our approach is to estimate the *temporal autocorrelation* (afterwards referred to as autocorrelation for simplicity) in time series of pairwise feature distances (where the features are directions, speeds, etc.) between movers within a rolling time window. Coordinated movement manifests in multiple (but not necessarily all) pairs of movers having high positive autocorrelation values in one or more of the respective time series of distances/differences. After detection of such autocorrelations, interactive visual tools showing the respective movements can be applied for exploring and identifying the patterns. We consider this publication to be an initial step towards a full coordination-oriented formalization of collective movement.

2. Collective Movement

Collective movement emerges from the coordination and information transfer between individual movers, as for example in two movers evading each other when on a collision course. Wood and Galton [WG09] proposed a taxonomy of collectives that describe spatial relationships between the collectives and individual movers. They define a spatial collective by having at least one form of spatial coherence between movers. Movers are classified according to the following five criteria: membership, coherence, spatial location, different roles, and depth. Dodge et al. [DWL08] proposed another definition of collective movement (moving cluster) as a group of movers that are close to each other and take a similar path over time. Galton and Wood [GW16] emphasized in their recent work that collective movement is more than just the sum individual movers motion and requires new analytical methods that highlight spatiotemporal coherence and correlation between movers over time. The scope of our approach encompasses fixed collectives which are moving in conjunction. We do not regard spatially and temporally disjoint movements such as road traffic over a large city. We assume the existence of a form of coordination between the movers, which does not necessarily imply similar movements but may involve any adjustments of movements of one entity to movements of others or to a common aim. From the data perspective, the analysis involves three types of attributes of movers: spatial (e.g., latitude, longitude), temporal (e.g., time), and thematic (e.g., the name of mover). The coherence between the spatiotemporal attributes of the movers has to be investigated to uncover the coordination processes in collective movement. Yet, several aspects of this task are insufficiently supported by current Visual Analytics approaches.

2.1. Expert feedback

To better understand the requirements of users in the field of collective movement, we collected expert feedback in informal interviews serving as a basis for our formalization efforts. We have interviewed two Ph.D. experts in Collective Animal Behavior, both on a postdoctoral level, covering both the global group level as well as the inter-individual perspective in their research. Both experts state similar issues concerning existing methods they apply: Research on Collective Animal Behavior is primarily driven by the idea of being able to model observed behavior by a minimal set of governing rules [Sum10]. Still, few techniques exist to come to an initial understanding of the observed behavior to derive abstractions for modeling. One of the primary tasks of our experts is to decode coordination between multiple movers, how it forms and breaks up again. Being a typical task of explorative visualization, they state that this task is not well supported by the tools available to them. Moreover, the experts also do not fully trust themselves to be able to extract all meaningful coordination patterns by hypothesizing about it. Thus, it is important to provide approaches which do not rely on previous assumptions. Further, the experts state that a variable temporal element is not regarded by most conventional approaches, which would enable the detection of "lagging" responses of other parts of the group. Thus, not only the pattern detection itself is an issue, but also extracting of temporally shifted pattern sequences.

2.2. Related Approaches

Tobler's first law of geography [Tob70] states that "everything is related to everything else, but near things are more related than distant things." In many domains, e.g., animal behaviour [Sum10] and soccer analysis [SSS*14], correlation between movers is expected to show mover coordination. For specific kinds of patterns, autocorrelation approaches have been applied in the behavioral ecology domain, for example by Bumann and Krause [BK93] or Jolles et al. [JBS*17]. Andrienko et al. [AAB*13a] exhaustively examined current movement analysis methods, with most methods focusing either on individual movers [HJS*08], events [DWL08], space [ASJ*10], or time [VLBA*12]. Few methods analyze movement data using temporal autocorrelation [GDGF10, KCO*12], yet disregard the coordination aspect. Laube et al. [LI02] analyze relative motion patterns of movers similarly to us. Yet, their approach deals with a single attribute with discrete values. Our approach employs multiple numeric attributes, which are not discretized to avoid possible information loss. As well, it is decentralized, meaning that it is not dependent on an aggregated feature derived from the entire group. For example, Andrienko et al. [AAB*13b] put the movers in relationship to a central trajectory derived from the whole group. Such centralized approaches are not able to uncover differently behaving subgroups existing within a group. In conclusion, an apparent gap can be identified in the Visual Analytics domain concerning methods for coordinated movement analysis.

2.3. Open challenges

The analysis of collective movement characteristics requires the application of computational methods and models to identify coordination patterns of movers [BDFM18]. Without pre-defined pattern extraction, the characteristics indicative for coordination first have to be identified and extracted from raw data. The analysis of the underlying coordination highly depends on the spatial, temporal and quantitative scale of analyzed movers [LP11]. Relative to the number of movers, the potential number of coordination patterns is exponential, adding to an already challenging problem. New methods need to be developed that can detect coordination patterns across different scales. Such methods can help to examine how collective movement emerges from local coordination between movers or how global collective movement produces local coordination.

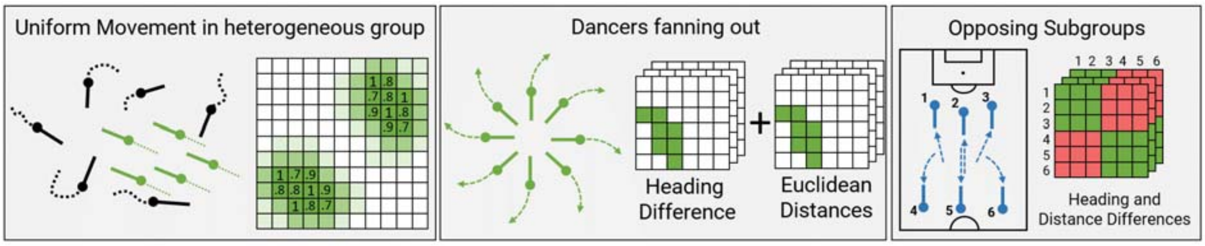

Figure 2: *Examples for patterns and how to detect them using autocorrelation matrices. On the left, five movers show coordinated movement, while others do not. They appear as a cluster of correlated entities in the autocorrelation of the heading distance feature. In the middle, a more complex pattern is shown with movers moving away from each other in unison. The pattern can be represented by a combination of change in heading, where the correlation stays positive over time, and uniform change in Euclidean distances between the movers over time. On the right, defence and midfielders of a soccer team regroup to their designated positions in opposing directions, resulting in groups of distinguishable positive and negative correlation values (green and red). This is an example for coordination despite opposing movements.*

3. Formalization

Fundamental to our approach is the idea that coordination can be expressed as correlation over time between movements of two or more movers. Consequently, instead of specifically characterizing patterns to extract, we expect that coordination manifests in quite strong temporal correlation, having a high probability of detection. Therefore, our approach relies on detecting temporal correlations between movers over a set of extracted features, which is equivalent to detecting *temporal autocorrelations* in time series of pairwise differences (distances) between the feature values of different movers. Autocorrelation [CO19, BJRL15] denotes the correlation between the values at time steps $t, t+1, ..., t+n$ and the values at time steps $t-\Delta, t+1-\Delta, ..., t+n-\Delta$, where Δ is a certain time lag. The most common is autocorrelation with $\Delta = 1$, which is called the first-order autocorrelation.

For detecting coordinated movements, the most relevant is positive first-order autocorrelation, which means that values in a time series are similar to the directly preceding values. In terms of pairwise distances or differences, it says that the latter does not change much from one time step to the next one. This refers not only to cases of keeping roughly the same values over time but also to gradual changes (decreases or increases). Thus, a gradual decrease of between-mover distances means that they concentrate in space and a gradual increase in the distances means that they spread. In the following parts, we first take a look at the necessary data preprocessing. Then, we will explain the autocorrelation concept and illustrate examples for the application of our framework.

3.1. Data Transformation

Figure 1 shows the pipeline of data transformations involved in our approach. The first step is to extract movement characteristics (features) from the raw data. Based on the positions of a mover in consecutive time steps, time-variant features are computed, such as heading, speed, and acceleration. In the second step, we shift the perspective from the features of individual movements to the between-movers perspective, which is reflected in the calculation of relative features encoding pairwise distances, in a general sense, between entities. These may include spatial distances, differences in speeds, headings, etc. Generally, any distance function applicable to the features can be used. By the pairwise calculation of distance functions for each time step, we create a 3D matrix, or cube, as shown in Figure 1, per feature. The cubes consist of 2D matrices corresponding to the time steps and containing the values returned by the distance functions chosen for the respective features.

3.2. Calculation of Autocorrelation

Computing the temporal autocorrelation of a full-time series of pairwise differences means checking if the respective movers coordinated their movements over the whole period covered by the dataset. Apart from such permanent coordination, researchers may be interested in detecting coordination that existed during shorter intervals. Depending on the character of the studied movers and their movement, as well as on the temporal resolution of the data (i.e., recording frequency), a researcher needs to estimate the minimal length of a time interval (i.e., the number of consecutive time steps) L_{min} in which coordinated movements can be reflected in the data. Then, cases of temporary coordination are detected by calculating the autocorrelation over the rolling time window of the length L_{min} (see Figure 3). Hence, instead of (or in addition to) a single "global" autocorrelation value for each time series, we obtain a sequence of interval-based autocorrelation values associated with the time steps 0, 1, ..., $N - L_{min}$, where N is the length of the full-time series.

With this approach, L_{min} sets the lower bound for the expected duration of movement coordination, but does not limit the upper bound: higher duration manifests in the preservation of high autocorrelation values in multiple consecutive time steps. The value for L_{min} should not only be meaningful from the domain perspective but also permit valid computation of autocorrelation from the statistical perspective. Statistical literature points out that the probability of detecting autocorrelation in series with less than 20 steps is small [SMS10]. However, this primarily refers to weak autocorrelations, which are mostly not interesting for detecting coordinated movement. The statistical literature also suggests that, among numerous autocorrelation measures that have been proposed, some may work better than others for short time series [DT93]. One of such measures is the C-statistic [You41], which is an estimator based on the sum of squared differences between successive observations Y over n time steps:

$$C = 1 - \frac{\sum_{t=1}^{n-1} (Y_t - Y_{t+1})^2}{2\sum_{t=1}^{n} (Y_t - \overline{Y})^2}$$

While the C-statistic performed well in a comparison by Solanas et al. [SMS10], the authors state as well that depending on the length of the time series and negative or positive sign of the autocorrelation, other estimators might work better. Consequently, the choice of the estimator is use-case dependent and needs to be investigated in more detail and subject to different data types. Having chosen an appropriate L_{min}, the task of detecting coordination patterns transforms to a search for high autocorrelation values within individual cubes. Visual approaches can then be applied to display the relative movement patterns corresponding to the detected autocorrelations.

3.3. Examples

For illustration purposes, we provide three examples of how mover coordination can be read from temporal autocorrelation matrices in Figure 2. On the left, a uniform movement of several movers results in a dense cluster moving over time. The resulting autocorrelation matrix shows a cluster, which can be identified easily. The second example shows a more complicated situation to encode, which other approaches have difficulties in capturing: We assume dancers have been arranged in a small circle and are now moving outward in unison. In this case, coordination is expressed in the movers moving away from each other in a uniform way. We can detect this pattern by examining the relative heading between the movers, which stays stable over time. Yet, this does not suffice to distinguish this specific pattern, so we also observe correlations in the development of relative distances between the movers, which increase pairwise proportionally. This can be captured by the autocorrelation function, and we can now compare the two cubes, checking which cells coincide in the correlation between the two cubes. If movers correlate in both cubes, we can confirm a fanning pattern. In the third example, defence and midfielders of a soccer team regroup after the ball has left their area of responsibility. The two subgroups running towards different areas in different directions create two distinguishable clusters of positive correlation values between the members of the subgroups and negative values between the two groups, together forming a more complex coordination pattern.

4. Discussion

We propose a novel formalization approach towards the analysis of coordination in spatio-temporal data. While established methods predominantly focus on a task- or environment-oriented perspective, we advocate a pairwise study based on autocorrelation of relative movement features, thus observing coordination instead of environmental interaction. Main characteristic is the transformation of spatio-temporal data into autocorrelation cubes. By exploring pairwise correlations over time, coordination patterns of different duration can be extracted. Also, lagging responses and pattern sequences can be observed by combining appropriate distance matrices. In comparison to current approaches, our method takes up an intrinsic perspective on coordination by observing relative inter-mover characteristics instead of absolute trajectories. At the same time, we do not make assumptions on group features to detect patterns, as

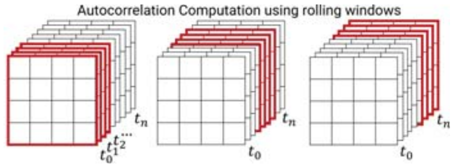

Figure 3: *Calculation of the autocorrelation features on a given distance matrix using a rolling window approach from front to back.*

task- and event-oriented approaches often do. This way, it is also possible to display relations between different subsets of movers. A meaningful ordering of entities in the matrices has to be established to enable visual detection of subgroups and clusters. Since coordination between spatially close entities is usually higher than between more distant ones, an ordering according to spatial proximity appears sensible. We intend to investigate spatial linearization functions such as space-filling curves to derive a 1D linear ordering of entities, as for example Buchmüller et al. use [BJC*19]. One of the main challenges is the computational complexity of the involved matrix computations, as the matrix size increases exponentially with growing numbers of movers, time steps and features. As stated in Section 2, our approach is intended for fixed collectives rather than independent movers, implying spatial and temporal proximity of involved movers. Consequently, scalability to very large numbers of movers is not our primary concern. Rather, a very long time period and/or a very high temporal resolution of the datas can pose a challenge. Yet, we expect that we can leverage the similarities inherent to collective movement to apply trajectory compression techniques and time step clustering according to movement features to reduce the complexity. We intend to explore these approaches for larger real-world datasets in a followup publication. While so far, we can only hypothesize that our approach is able to cover most coordination patterns in collective movement data, we intend to derive a thorough classification of encoded patterns. Finally, choosing the right rolling window size for the autocorrelation poses a challenge, since it needs to be adapted to the observed movers' behavior.

In summary, learning about the coordination of movers can reveal behaviors, and thus, motivations of the observed entities. Recent works and the emerging field of collective movement research pose new questions towards the coordination of movers: Instead of asking where the movers go, the primary question is how the entities are coordinating on their way. Consequently, the domain of spatio-temporal data analysis has to provide methods to reflect this change of perspective. With this paper, we propose an initial step towards a concept for a formal, coordination-oriented foundation for processing collective movement data. Furthermore, having access to domain experts and datasets in the areas of collective animal behavior (e.g., bird, fish and locust swarms) and team sport analysis (e.g., soccer or rugby data), we intend to test and expand our formalization in a practical implementation on collective movement datasets as direct followup to this publication. Thereby, we want to explore the coordination patterns encoded in basic movement as well as higher-order autocorrelations like acceleration or turn patterns. By applying our formalization to datasets with varying parameters, e.g., in sampling rate, time series length, or nature of the observed behavior, we also want to explore the application of more complex features.

References

[AAB*13a] ANDRIENKO G., ANDRIENKO N., BAK P., KEIM D., WROBEL S.: *Visual analytics of movement*. Springer Science & Business Media, 2013. 2

[AAB*13b] ANDRIENKO N., ANDRIENKO G., BARRETT L., DOSTIE M., HENZI P.: Space transformation for understanding group movement. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2169–2178. 2

[AAR16] ANDRIENKO N., ANDRIENKO G., RINZIVILLO S.: Leveraging spatial abstraction in traffic analysis and forecasting with visual analytics. *Information Systems 57* (2016), 172–194. 1

[ASJ*10] AHAS R., SILM S., JÄRV O., SALUVEER E., TIRU M.: Using mobile positioning data to model locations meaningful to users of mobile phones. *Journal of urban technology 17*, 1 (2010), 3–27. 2

[BDFM18] BIRKIN M., DODGE S., FASY B. T., MANN R. P.: From observations to prediction of movement (dagstuhl seminar 18282). In *Dagstuhl Reports* (2018), vol. 7, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik. 2

[BJC*19] BUCHMÜLLER J., JÄCKLE D., CAKMAK E., BRANDES U., KEIM D. A.: Motionrugs: Visualizing collective trends in space and time. *IEEE transactions on visualization and computer graphics 25*, 1 (2019), 76–86. 4

[BJRL15] BOX G. E., JENKINS G. M., REINSEL G. C., LJUNG G. M.: *Time series analysis: forecasting and control*. John Wiley & Sons, 2015. 3

[BK93] BUMANN D., KRAUSE J.: Front individuals lead in shoals of three-spined sticklebacks (gasterosteus aculeatus) and juvenile roach (rutilus rutilus). *Behaviour 125*, 3-4 (1993), 189–198. 2

[CO19] CLIFF A., ORD K.: Spatial autocorrelation / a. d. cliff and j. k. ord. *SERBIULA (sistema Librum 2.0)* (03 2019). 3

[DT93] DECARLO L. T., TRYON W. W.: Estimating and testing autocorrelation with small samples: A comparison of the c-statistic to a modified estimator. *Behaviour research and therapy 31*, 8 (1993), 781–788. 3

[DWL08] DODGE S., WEIBEL R., LAUTENSCHÜTZ A.-K.: Towards a taxonomy of movement patterns. *Information visualization 7*, 3-4 (2008), 240–252. 1, 2

[GDGF10] GELFAND A. E., DIGGLE P., GUTTORP P., FUENTES M.: *Handbook of spatial statistics*. CRC press, 2010. 2

[GW16] GALTON A., WOOD Z.: Extensional and intensional collectives and the de re/de dicto distinction. *Applied Ontology 11*, 3 (2016), 205–226. 2

[HJS*08] HUYS R., JIRSA V. K., STUDENKA B. E., RHEAUME N., ZELAZNIK H. N.: Human trajectory formation: Taxonomy of movement based on phase flow topology. In *Coordination: Neural, behavioral and social dynamics*. Springer, 2008, pp. 77–92. 2

[JBS*17] JOLLES J. W., BOOGERT N. J., SRIDHAR V. H., COUZIN I. D., MANICA A.: Consistent individual differences drive collective behavior and group functioning of schooling fish. *Current Biology 27*, 18 (2017), 2862–2868. 2

[KCO*12] KAWALE J., CHATTERJEE S., ORMSBY D., STEINHAEUSER K., LIESS S., KUMAR V.: Testing the significance of spatio-temporal teleconnection patterns. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining* (2012), ACM, pp. 642–650. 2

[LI02] LAUBE P., IMFELD S.: Analyzing relative motion within groups of trackable moving point objects. In *International Conference on Geographic Information Science* (2002), Springer, pp. 132–144. 2

[LP11] LAUBE P., PURVES R. S.: How fast is a cow? cross-scale analysis of movement data. *Transactions in GIS 15*, 3 (2011), 401–418. 2

[MGTH09] MOUSSAID M., GARNIER S., THERAULAZ G., HELBING D.: Collective information processing and pattern formation in swarms, flocks, and crowds. *Topics in Cognitive Science 1*, 3 (2009), 469–497. 1

[SAMS*17] SACHA D., AL-MASOUDI F., STEIN M., SCHRECK T., KEIM D. A., ANDRIENKO G., JANETZKO H.: Dynamic visual abstraction of soccer movement. In *Computer Graphics Forum* (2017), vol. 36, Wiley Online Library, pp. 305–315. 1

[SBJ*11] SPRETKE D., BAK P., JANETZKO H., KRANSTAUBER B., MANSMANN F., DAVIDSON S.: Exploration through enrichment: a visual analytics approach for animal movement. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (2011), ACM, pp. 421–424. 1

[SMS10] SOLANAS A., MANOLOV R., SIERRA V.: Lag-one autocorrelation in short series: Estimation and hypotheses testing. *Psicologica: International Journal of Methodology and Experimental Psychology 31*, 2 (2010), 357–381. 3, 4

[SSS*14] SACHA D., STEIN M., SCHRECK T., KEIM D. A., DEUSSEN O., ET AL.: Feature-driven visual analytics of soccer data. In *2014 IEEE conference on visual analytics science and technology (VAST)* (2014), IEEE, pp. 13–22. 2

[Sum10] SUMPTER D. J.: *Collective animal behavior*. Princeton University Press, 2010. 2

[Tob70] TOBLER W. R.: A computer movie simulating urban growth in the detroit region. *Economic geography 46*, sup1 (1970), 234–240. 2

[VLBA*12] VON LANDESBERGER T., BREMM S., ANDRIENKO N., ANDRIENKO G., TEKUŠOVÁ M.: Visual analytics methods for categoric spatio-temporal data. In *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)* (2012), IEEE, pp. 183–192. 2

[WBTB18] WESTLEY P. A. H., BERDAHL A. M., TORNEY C. J., BIRO D.: Collective movement in ecology: from emerging technologies to conservation and management. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences 373 1746* (2018). 1

[WG09] WOOD Z., GALTON A.: A taxonomy of collective phenomena. *Applied Ontology 4*, 3-4 (2009), 267–292. 1, 2

[You41] YOUNG L.: On randomness in ordered sequences. *The Annals of Mathematical Statistics 12*, 3 (1941), 293–300. 3