Can human mobility metrics tell us something about animal behavior?

M. Łoś[1], K. Siła-Nowicka[2], K. Smolak[1], W. Rohm[1], N. Van de Weghe[3]

[1]Wroclaw University of Environmental and Life Sciences, Wroclaw, Poland
Email: {marcelina.los; kamil.smolak; witold.rohm}@upwr.edu.pl

[2]University of Auckland, New Zealand
Email: katarzyna.sila-nowicka@auckland.ac.nz

[3]Ghent University, Belgium
Email: nico.vandeweghe@ugent.be

Abstract

Integrated Science of Movement is an interdisciplinary science that has been established to bridge methodological gaps between movement ecology and human mobility. This paper proposes a step towards it by transferring human mobility data mining techniques to the animal movement domain. We use well-know mobility study methods such as stop detection, spatial aggregation (DBSCAN), and two approaches of temporal aggregation (Next Time-Bin and Next Place) to process GPS tracking data of Herring Gulls. We also calculate a number of movement statistics (visitation frequency and distinct locations over time) from these data. The initial results indicate differences between breeding season and the remaining parts of the year.

1. Introduction

In recent decades, the availability and the evolution of sensors have allowed to record animal movement data with miniaturized tags, that are less invasive and available for different species (Kays *et al.* 2015). At the same time, the growing popularity of personal devices has enabled collecting high quality human movement data (Zheng *et al.* 2015). Furthermore, new technological solutions have increased processing power of computers. These advancements led to the rapid development of two individual disciplines: movement ecology and human mobility. Regardless of the similarities in data used, there is a methodological gap between these domains. To bridge it and improve understanding of animal and human movements Miller *et al.* (2019) and Demšar *et al.* (2021) argue to converge movement research domains into an Integrated Science of Movement.

Movement ecology focuses on understanding animal behaviors and responses to environmental changes (Nathan *et al.* 2008). A common way of studying animals' behavior is mapping their distribution through space and time. To describe areas used by the animal during their activities the home range estimation is defined (Burt 1943). Techniques for estimating these areas include geometric approaches such as minimum convex polygon (MCP) (Blair 1940) and probabilistic models, which calculate utilization distribution such as kernel density estimator (Péron 2019). Another approach in analyzing animal trajectory is the identification of movement behavior. Based on movement parameters (e.g. step length, turn angle, velocity) calculated from a trajectory, models search for statistically different characteristics to determine unique behaviors or processes (Miller *et al.* 2019). The most commonly used method to identify behavioral states from movement data is path segmentation (Edelhoff *et al.* 2016), which divides a trajectory into homogenous parts (segments) based on the time interval, spatial shape or semantic meaning.

Understanding human behaviors using mobility data often comes down to a process called semantic enrichment of movement trajectories (Parent *et al.* 2013). The process consists of three elements: trajectory segmentation, travel mode detection, and trip purpose identification. The modes and trip purposes can be transformed into sequences and have been used to understand spread of epidemic (Zhang *et al.* 2021, Xiong *et al.* 2020, Christaki 2015), predicting traffics (Kong *et al.* 2016), or developments in urban planning (Wang *et al.* 2019).

Data processing methods often used in human mobility studies may be useful to study animal behavior. In this paper we propose to adopt a combination of stop detection, spatial aggregation (DBSCAN), and two approaches of temporal aggregation (Next Time-Bin and Next Place), to see if it can contribute to a deeper understanding of animal behavior.

2. Methods

The steps of our research are presented as a scheme in Figure 1.

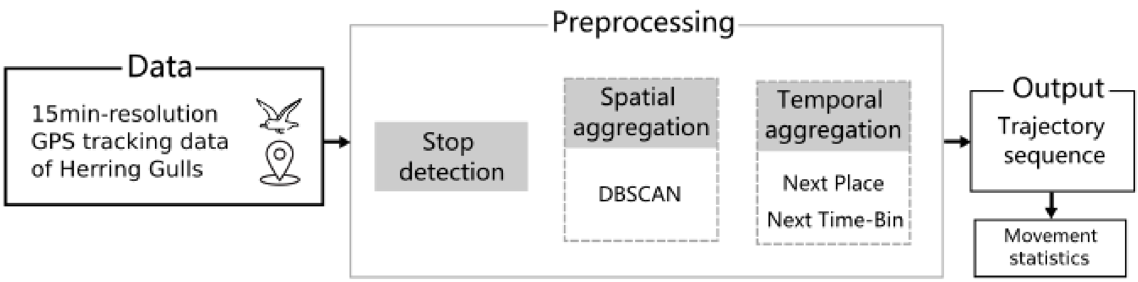

Figure 1. Illustration of the data preprocessing.

2.1 Data

In this project, we use GPS tracking data (1,073,945 points) of 11 adults Herring Gulls 4 females, 7 males) from a colony in Ostend, Belgium, covering a period from January to November 2017 (Stienen *et al*. 2016) (Figure 2). The birds were tracked with the University of Amsterdam Bird Tracking System (Bouten *et al.* 2013) collecting the data at 15-min intervals.

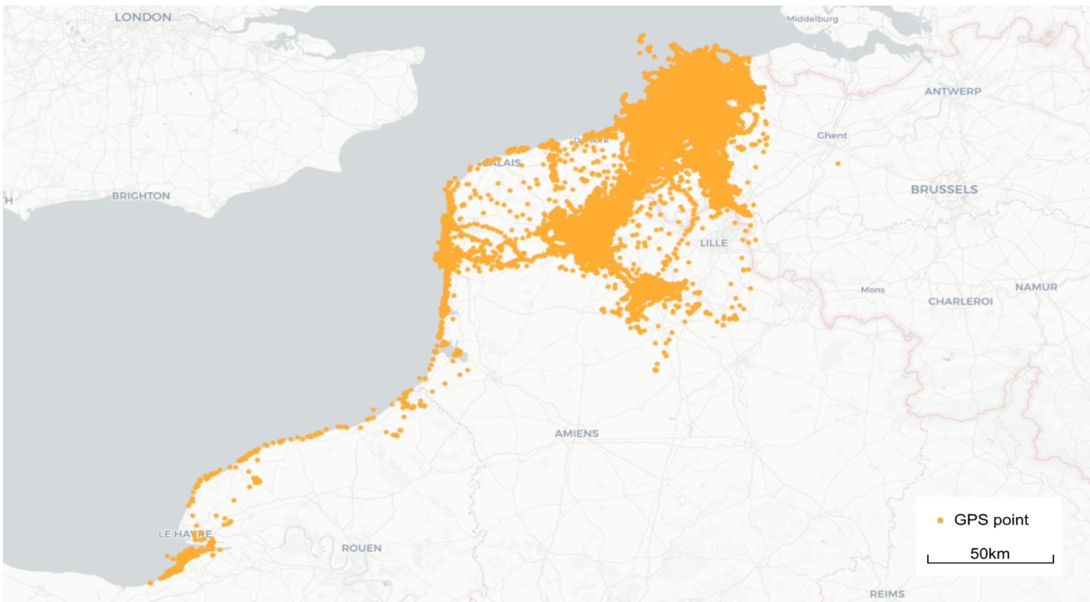

Figure 2. Map showing gull tracking data used in the study.

2.2 Spatial and temporal aggregation

To create a movement trajectory, the raw data was filtered and converted into a temporally ordered sequence of visited locations. This process is shown in Figure 3.

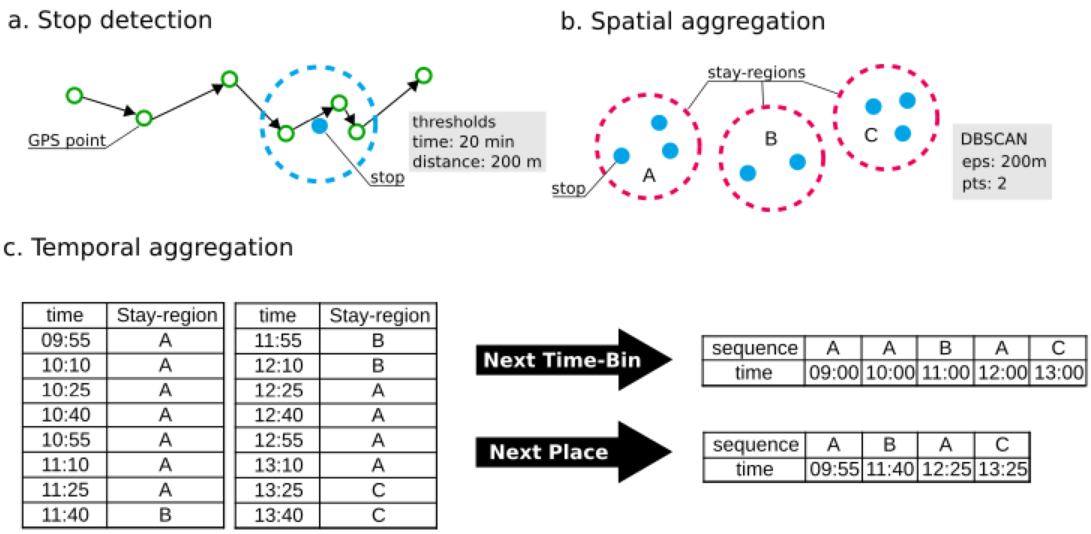

Figure 3. Three steps of preprocessing. a. In the first step stops (blue dots) are identified as centroids of selected green GPS points. b. Identified stops (blue dots) are aggregated spatially to stay-regions (pink dashed-line circles) using DBSCAN algorithm. c. The last step of preprocessing consists of two different methods of temporal aggregation: Next Time-Bin and Next Place. In the Next Time-Bin hours in the sequence – represent time interval (e.g 9:00 indicates 'from 9:00 to 10:00')

As a first step in this process, we identified all the stops. The GPS points were grouped into a stop as long as they were within 200m from each other and no more than 20 minutes apart (Figure 3a). Next, the stops were aggregated spatially and temporally. Using the DBSCAN algorithm (2pts, 200m) the stops were clustered into stay-regions (Figure 3b). After the spatial aggregation part, the stops were ordered temporally into sequences using two approaches: Next Time-Bin (NTB) and Next Place (NP) (Cuttone *et al.* 2018) (Figure 3c). In NTB, the location is periodically detected at a selected time interval (one hour) which means that there could be so-called self-transitions between sequences (Figure 3c). In NP, the sequence contains only transitions between places.

2.3 Movement statistics

After extracting movement sequences into NTB and NP, we aggregated them into monthly bins and calculated two statistics:

1. Visitation frequency – using a one-hour temporal resolution, the visitation frequency f corresponds to the ratio of visits in a stay-region to the number of visits in all detected stay-regions during the observation period (month). In human mobility, the frequency f of the r^{th} most visited stay-region follows Zipf's law:

$$f_r \sim r^{-\zeta} \tag{1}$$

where ζ=1.2 ± 0.1 (Song *et al.* 2010).

2. Distinct locations over time - $S(t)$ shows the total number of locations visited within one hour. In human mobility $S(t)$ is expected to follow:

$$S(t) \sim t^{\mu} \tag{2}$$

where $\mu = 0.6 \pm 0.02$ (Song *et al.* 2010)

3. Preliminary results and Discussion

The visitation frequency f was calculated for all stay-regions detected using the methodology presented in Figure 3. Figure 4 shows the average f for the first two most popular stay-regions. For both approaches, we can see a peak in the breeding season (eggs are laid from the beginning of May, the chicks fledge in early August) for the first place in the rank. From April to June, the average f for the first location (based on map analysis, we assume that it may be their colony) is higher than 70% using NTB and 40% using NP. The differences between NTB and NP are caused by self-transitions. The NP trajectory sequence (by definition) consists only of changing locations, whereas in NTB the same location may occur several times resulting in self-transitions. Despite these differences, both charts (Figure 4a and 4b) show a tendency towards increasing the f from April to June for the first place in the rank. The studied birds reduce the number of visited places and spend more time in one location during a breeding season. Furthermore, variations in movement patterns over the months can be influenced by seasonal climate change.

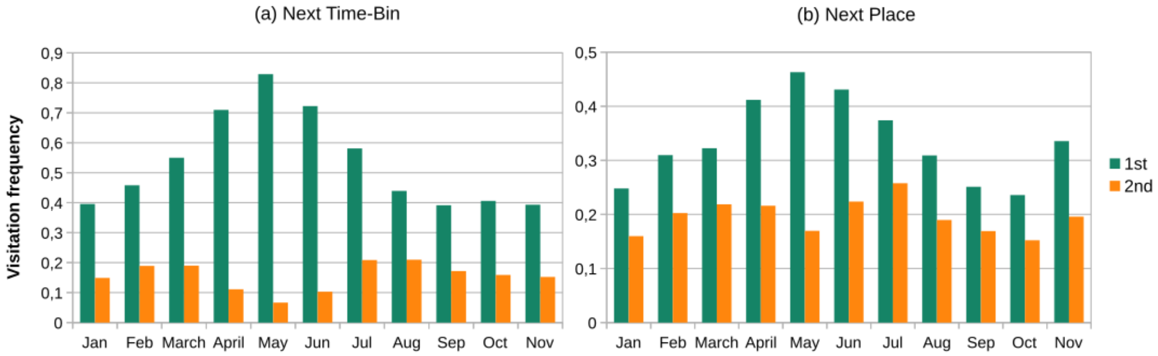

Figure 4. The average visitation frequency for first (green) and second (orange) location during selected months for NTB and NP approach.

In human mobility f and *S(t)* are empirical observations, which indicate that trajectories follow reproducible scaling laws (1) and (2). These have not been applied for animal data before. To evaluate the potential of these statistics for movement ecology we calculated ζ (1) and μ (2) factors for the study period for the seagull data using NTB and NP approaches (Table 1). The value of ζ for NTB starts to decrease in March reaching the minimum (-3.42) in May. For NP ζ is lower from April to June in comparison to the rest of the year. This suggests that the movement patterns of seagulls change during the breeding season. In terms of *S(t)*, we can see that μ factor decreases from March to July in NP and from April to May in NTB.

Table 1. The values of ζ and μ factor for Jan-Nov 2017.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov
ζ (NTB)	-1.31	-1.38	-1.67	-2.40	-3.42	-2.63	-1.76	-1.41	-1.32	-1.38	-1.29
ζ (NP)	-0.93	-1.00	-1.04	-1.24	-1.38	-1.31	-1.19	-1.09	-0.99	-0.94	-0.87
μ (NTB)	0.56	0.55	0.55	0.25	0.37	0.77	0.63	0.54	0.57	0.57	0.54
μ (NP)	0.57	0.50	0.35	0.38	0.40	0.38	0.36	0.6	0.58	0.6	0.53

Preliminary results produced in this paper show that commonly used methods in human mobility have the potential to be used in animal movement research, but there are some limitations to overcome. All calculations were made by calculating statistics in monthly bins, however, it is an arbitrary division that does not correspond to animals' mobility patterns. Moreover, people rarely change the locations of their significant places, whereas, seagulls do not have the same main places throughout the year. In future work, we will split the data based on significant mobility changes instead of using one month as a splitting point. Applying this more adjusted division of data could improve the results and better reflect the mobility of gulls.

Acknowledgements

Authors are supported by the National Science Centre Grant (UMO-2019/35/O/ST6/04127). The datasets analyzed during the current study are available in the Movebank Data Repository and Zenodo, https://doi.org/10.5281/zenodo.3541811, (Stienen, 2019).

References

Bekoff, M., Mech, L.D., 1984, Simulation analyses of space use: home range estimates, variability, and sample size. *Behavior Research Methods, Instruments, and Computers,* 16, 32–37.

Blair W.F. 1940, Notes on Home Ranges and Populations of the Short-Tailed Shrew., Ecology, 21(2), 284–288.

Bouten, W., Baaij, E.W., Shamoun-Baranes, J., Camphuysen, K.C.J., 2013, A flexible GPS tracking system for studying bird behaviour at multiple scales. *J. Ornithol.*154, 571–580.

Burt, W.H., 1943. Territoriality and home range concepts as applied to mammals. *Journal of Mammalogy*, 24, 346–352.

Christaki, E, 2015 New technologies in predicting, preventing and controlling emerging infectious diseases., *Virulence*, 6(6), 558–565.

Cuttone A., Lehmann S., González M.C., 2018, Understanding Predictability and Exploration in Human Mobility., *PJ Data Sci.* 7, 2.

Demšar, U., Long, S.J., Benitez-Paez, F., Brum Bastos, V., Marion, S.,Martin, G., Sekulić, S., Smolak, K., Zein, B., Siła-Nowicka, K., 2021, Establishing the integrated science of movement: bringing together concepts and methods from animal and human movement analysis, I*nternational Journal of Geographical Information Science,* 35, 1-36.

Edelhoff H., Signer J., Balkenhol N., 2016, Path segmentation for beginners: An overview of current methods for detecting changes in animal movement patterns., Movement Ecology, 4(1).

Kays, R., Crofoot, M.C., Jetz, W., Wikelski, M., 2015, Terrestrial animal tracking as an eye on life and planet. *Science.* 348. 2478.

Kong, X., Xu, Z., Shen, G., Wang, J., Yang, Q., Zhang, B., 2016, Urban traffic congestion estimation and prediction based on floating car trajectory data. Future Generation Computer Systems, 61, 97–107.

Miller, H., Dodge, S., Miller, J., Bohrer, G., 2019, Towards an Integrated Science of Movement: Converging Research on Animal Movement Ecology and Human Mobility Science. *International Journal of Geographical Information Science*, 33, 1-22.

Nathan, R., Getz, W.M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., Smouse, P.E., 2008, A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences*, 105 (49), 19052–19059.

Parent, C., Spaccapietra, S., Renso, C., Andrienko, G., Andrienko, N., Bogorny, V., Gkoulalas-Divanis, A. Macedo, J., Pelekis, N., Theodoridis, Y., Yan, Z., 2013, Semantic Trajectories Modeling and Analysis. *ACM Computing Surveys,* 45(4), 42

Péron G., 2019 Modified home range kernel density estimators that take environmental interactions into account., Movement Ecology; 7(1), 1–8.

Song, C., Koren, T., Wang, P., Barabási, A.-L., 2010, Modelling the scaling properties of human mobility. *Nat Phys,* 6(10), 818–823.

Stienen, E.W.M., Desmet P., Milotic, T., Hernandez, F., Deneudt, K., Matheve, H., Bouten, W., 2019, HG_OOSTENDE - Herring gulls (Larus argentatus, Laridae) breeding at the southern North Sea coast (Belgium). Dataset. https://doi.org/10.5281/zenodo.3541811

Wang, J., Kong, X., Xia, F., Sun, L., 2019, Urban Human Mobility: Data-Driven Modeling and Prediction. *ACM SIGKDD Explorations Newsletter*, 21, 1–19.

Xiong, Ch., Hu, S.,Yang, M., Luo, W., Zhang, L., 2020, Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections. *Proceedings of the National Academy of Sciences of the United States of America*, 117(44), 27087–27089.

Zhang, J., Feng, B., Wu, Y., Xu, P., Ke, R., Dong, N., 2021, The effect of human mobility and control measures on traffic safety during COVID-19 pandemic. *PLoSONE*, 16, 493–497.

Zheng, Y., 2015, Trajectory Data Mining: An Overview. *ACM Transactions on Intelligent Systems and Technology,* 6, 1-41.