From networks to flows: Using flow maps to understand mobility patterns in cattle trade

Sima Farokhnejad[1], Eraldo Ribeiro[2] and Ronaldo Menezes[1,*]
[1]Department of Computer Science, University of Exeter, UK
[2]Department of Computer Science, Florida Institute of Technology, USA
sf503@exeter.ac.uk, eribeiro@fit.edu, *r.menezes@exeter.ac.uk

Keywords: cattle trade networks, flow maps, cattle movement

The understanding of mobility patterns is very important in a wide range of research areas [1, 2], and it has a direct impact on policymaking in urban environments, disease spreading analyses, and control of epidemics [3, 4], to name a few. Despite the power of networks in modelling mobility, alternative methods may be more useful or complementary in capturing idiosyncrasies of mobility patterns, depending on the datasets, or areas being investigated. Mobility datasets are often large and usually contain plenty of individual records. This feature adds too many details to the network generated from the dataset, making the modelling itself often confusing; simulation becomes time-consuming, and it is difficult to have a global analysis of mobility. Furthermore, to analyse global patterns of mobility during a period of time, it is critical to be able to look beyond microscopic details and able to see global patterns. To add to these challenges, mobility datasets are always subject to uncertainty [5]. Many datasets, including human mobility and livestock trade, show evidence of uncertainty. As a result of the forgotten or unreported records, or the scale of the datasets, the networks generated from the dataset are anything but flawless, which could affect traceability or reliability of the results.

The dataset of cattle movement in Brazil is an example of scale and uncertainty; the state we deal in this work (Minas Gerais) produces more cattle than many countries around the world. Economic pressures in poorer nations coupled with poor tracking infrastructure make it easier for people to trade cattle without reporting it. According to a recent report in a political magazine [1], these unreported trades occur frequently in Brazil as a mechanism akin to money laundering. The Amazon's vast national forest, which cannot be cultivated, has been ravaged over the last two decades by miners and thieves of public lands looking for hardwoods and new pastures for cattle raising. Animal Transit Guides (ATGs) indicate that there are over 91,000 cattle heads raised on stolen land, leading to a number of cases where cattle are transported clandestinely, without registration in any ATG.

As a result, there is missing information in these datasets. Due to its imperfection, the dataset loses the ability to trace the path of trades at some point. However, regardless of these microscopic details, analysing the path of the trades is extremely relevant, as well as recognising the global mobility patterns throughout time. The use of flow maps as an alternative representation of trade is a novel method for understanding trade patterns over time. It is a general approach that could be applied to any type of mobility dataset, but here we focus on cattle movement. Flow maps are constructed using all the details of individual trades. However, they do not need to include details in each trade to get to a model of global behaviour, which makes it much more compatible for the use in large and complex mobility datasets.

The mobility dataset is used to generate a vector field (on a map) in order to produce a flow map. There are so-called critical points that need to be analysed. Observing the vector field behaviour near these points gives us insights into its characteristics. It is our opinion that *sinks* and *sources* play a major role in epidemiological research as destinations and starting points of infection, respectively. An analysis of sink-source dynamics in mobility trades is a valuable approach that is captured in moving from network methods to flow methods. In order to understand the global behaviour of the trade patterns, it is essential to analyse the dynamics of the critical points on the flow maps of trades of different months. Monitoring the changes in sinks and sources on maps within time windows (monthly, for instance) can be effective in finding the level of similarity between trade patterns in two consecutive months (or other periods of time).

[1] `https://piaui.folha.uol.com.br/materia/lavagem-da-boiada/` *The Laundering of Cattle (In Portuguese)* (last accessed: 1 Nov 2022).

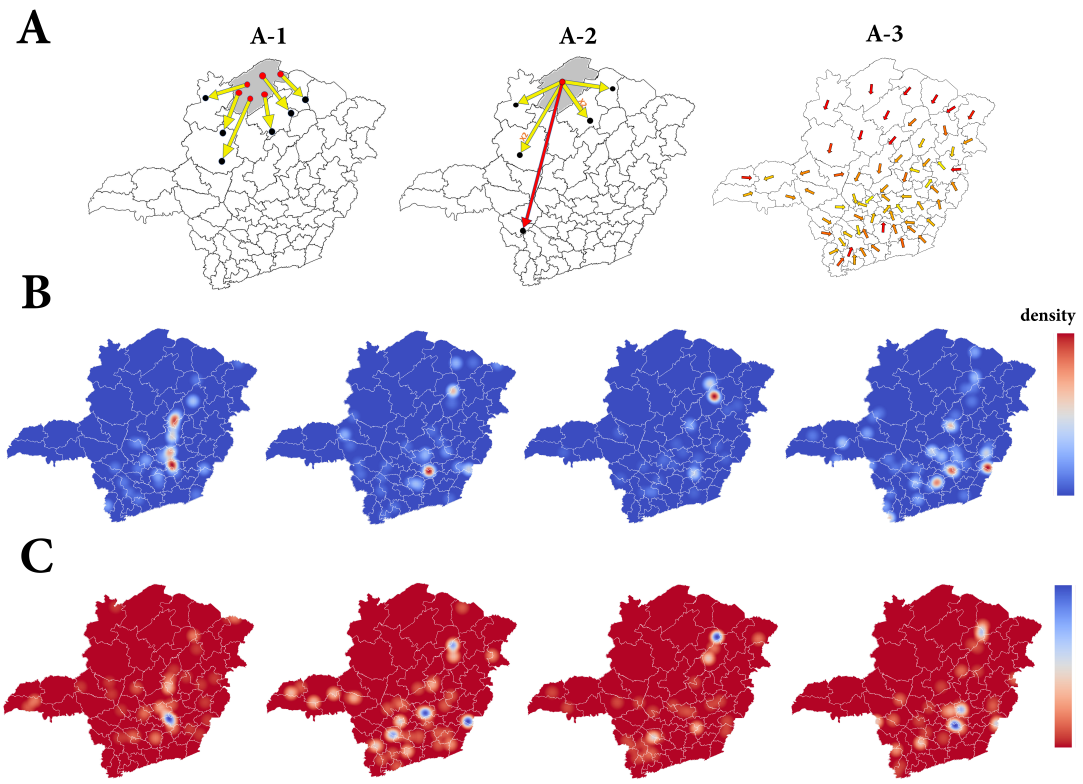

Figure 1: **A) An illustration of how to generate a flow map from a cattle mobility dataset. A-1**) Take the grey part as an example and concentrate on the trades happening between it and other parts. **A-2**) The trades are captured with a vector starting from the centre of the location (in grey) and ending at the centre of the destinations (yellow vectors). All vectors are then aggregated using a relevant method. Here, we add the vectors into a final vector (red vector), which we assign to the grey location. **A-3**) Visualisation of what we call the flow map. For a better visualisation, all vectors are displayed with the same length. The colour of each vector corresponds to a range that shows the size of the vector. **B) Sinks and C) Sources are visualised in flow maps for four consecutive months.** Initial flow maps are generated based on trades happening within specific months. A triangle-based interpolation method is used to create a vector field from the initial flow map. We found critical points in this vector field and identified sinks and sources. Sink and source areas are shown on monthly trade flow maps. In **B**, red areas indicate locations with a high density of sinks in the vector fields. Dark blue shows the areas with zero density of sink points. In **C**, the emphasis is on sources. Dark blue areas indicate high-density of sources, and dark red areas are the ones empty of source points.

In this work, we examine a dataset related to cattle movement in the state of Minas Gerais in Brazil. Figure 1-A is a toy example of how trades can be used to produce an output vector for each part (here each part is a micro-region in the state, although the regions be for a higher or lower granularity if needed). Trades within one part are first removed from the records; we do not consider trades internal to one area. As an example, consider the grey part in Figure 1-A-1, each trade originating from this part is captured as a vector. In Figure 1-A-2 for each trade between a location in the grey part and a destination in another part, a vector with the start point in the centre of the grey part and the end point in the centre of a destination part is assigned. By employing a method for the combination of vectors (here the average of the vectors is determined), the vectors of the grey part are merged into a final vector (red vector in Figure 1-A-2).

The result is a flow map for the trades that happen over a given period of time. Figure 1-A-3 shows a flow map with vectors in the centres of micro-regions, which could be regarded as a scatter vector field. Having an approximation of the vector value at each point of the map allows us to easily analyse the global pattern of the vector field. The interpolation methods involve constructing vectors for new points based on known scatter vectors. A triangle-based interpolation [6] is the technique we used to estimate the value of vectors at points

inside the map that have not been calculated.

Points with vector sizes equal to zero are critical points in a vector field. Classification of these critical points as sink, source, and saddle points is according to the sign of the eigenvalues of the Jacobian matrix [7]. Understanding how sink and source patterns change from month to month can help us understand the level of dynamics we have in the phenomenon being captured and help policymakers reach better decisions. Similarities or differences in the patterns of sinks and sources in different months of the year determine how predictable a year is in terms of mobility. Heatmaps in Figure 1-B and 1-C show the density of sinks and sources in flow maps generated from monthly trades. Over a period of four consecutive months, the sinks and sources area are recognised.

We developed an approach that can be applied to any type of mobility dataset which contains sources and destinations in trajectories. It provides an alternative/complementary tool to network methods for analysing the dynamic patterns of mobility, and it is independent of intermediate steps reached between source and destination of moving entities. Availability of numerous techniques to combine trades originated from each part to produce the final vector for that part is a strong aspect of this approach. As a result, it becomes more relevant to a particular type of mobility dataset and analysis goals; for instance, one could consider the number of cattle heads being transported as a factor in the value of the vectors. Additionally, any kind of dataset that contains origin-destination values can be converted to a flow map using this approach. The dynamic behaviour of the network during the time could then be evaluated using flow maps.

References

[1] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, "Understanding individual human mobility patterns," *nature*, vol. 453, no. 7196, pp. 779–782, 2008.

[2] H. Barbosa, M. Barthelemy, G. Ghoshal, C. R. James, M. Lenormand, T. Louail, R. Menezes, J. J. Ramasco, F. Simini, and M. Tomasini, "Human mobility: Models and applications," *Physics Reports*, vol. 734, pp. 1–74, 2018.

[3] D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J. J. Ramasco, and A. Vespignani, "Multiscale mobility networks and the spatial spreading of infectious diseases," *Proceedings of the National Academy of Sciences*, vol. 106, no. 51, pp. 21 484–21 489, 2009.

[4] A. Arenas, W. Cota, J. Gómez-Gardeñes, S. Gómez, C. Granell, J. T. Matamalas, D. Soriano-Paños, and B. Steinegger, "Modeling the spatiotemporal epidemic spreading of covid-19 and the impact of mobility and social distancing interventions," *Physical Review X*, vol. 10, no. 4, p. 041055, 2020.

[5] Y. Zheng, "Trajectory data mining: an overview," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 6, no. 3, pp. 1–41, 2015.

[6] D. Watson and G. Philip, "Triangle based interpolation," *Journal of the International Association for Mathematical Geology*, vol. 16, no. 8, pp. 779–795, 1984.

[7] M. Smolik and V. Skala, "Vector field interpolation with radial basis functions," in *Proceedings of SIGRAD 2016, May 23rd and 24th, Visby, Sweden*, no. 127. Linköping University Electronic Press, 2016, pp. 15–21.