# A heuristic for analyzing interactions between protectors and street animals

# Marcus Vinício de Oliveira Braga<sup>1</sup>, Lilian Berton<sup>1</sup>

<sup>1</sup>Instituto de Ciência e Tecnologia – Universidade Federal de São Paulo (UNIFESP) São José dos Campos – SP – Brazil

marcusoliveirabraga@gmail.com.br, lberton@unifesp.br

Abstract. This work proposes the construction of a cooperation network among animal welfare advocates and establishments involved in animal protection efforts. The network was modeled as a multi-layered architecture, considering interactions between individuals, establishments, and the animals assisted. The goal is to investigate the presence of communities and identify the most influential advocates in the provision of aid, using metrics such as modularity and degree centrality through the Louvain algorithm. Identifying these patterns aims to optimize the logistics and efficiency of support interactions, as well as to provide insights for planning preventive actions in public policies related to animal protection.

### 1. Introduction

The lack of control over the population of abandoned animals on the streets is a problem faced by many cities, especially in large cities. According to [Instituto Pet Brasil 2018], animals in vulnerable conditions are those that live under the care of families classified below the poverty line, or that live on the streets, but receive care from people other than those who care for them. The fate of these animals on the streets is uncertain, often resulting in death due to numerous adverse weather conditions, or euthanasia as a public policy control measure [Oliveira et al. 2008]. Animal protectors often feel overwhelmed by the high demand to which they are subjected daily, which generates a feeling of helplessness and frustration.

Nowadays, animal welfare has been receiving increasing attention and incentives [Horecka and Neal 2022]. Initiatives from various research fields are being proposed, from sociological studies that propose points to be observed for current adequacy to the needs of animal shelter and care, as seen in [Horecka and Neal 2022], creation of proposals for architecture of collaborative social platforms to assist in locating lost animals or adopting them [Bankar et al. 2024], and solutions for predicting the spread of animal diseases between farms [Kinsley et al. 2020] through the investigation of their interactions.

The difficult measurement, even if approximate, of the population of these animals [Horecka and Neal 2022] makes it impossible to better assess and efficiently use resources to provide care to those in need. The complex network of possible interactions and multiple agents in this context also limits the conventional approach of only one layer in the graph plane [Kinsley et al. 2020, Karimi et al. 2020], reinforcing the need to use networks with several layers to better represent and understand the real scenario.

This study proposes the construction of a cooperation network between animal welfare organizations and establishments involved in the animal cause, using data ob-

tained in real and simulated ways. The network was modeled as a multilayer architecture and the Louvain algorithm was applied, aiming to identify the presence of possible communities and most influential animal welfare organizations in providing aid, within a geographic radius based on modularity and degree centrality metrics. This identification seeks to improve logistics and efficiency in aid interactions, in addition to providing support for preventive public policy actions. In a preliminary analysis, it appears to be a cutting-edge network in the national and international scenario with a socioenvironmental impact bias.

The paper is organized into the following sections. Section 2 presents related work. Section 3 presents the network construction methodology and the applied heuristics. Section 4 presents the results and discussions of the obtained metrics. Finally, Section 5 presents the conclusions of the study and future work.

### 2. Related Works

In this section, we present work carried out by other authors that converge with this study: (i) techniques for detecting communities in complex networks; (ii) optimizations of the Louvain algorithm for graphs with large amounts of data and multilayer networks; (iii) use of Social Network Analysis (SNA) in favor of the animal cause. We will briefly discuss each of the initiatives and where our work intersects and complements it.

## 2.1. Multilayer networks and community detection

A complex network is a structure composed of nodes (or vertices) connected by edges (or links), whose organization presents non-trivial patterns of connectivity, typical of real systems, such as social, biological, technological or informational networks [Newman et al. 2011]. A complex network is represented as a graph G = (V, E), where V is the set of nodes and  $E \subseteq V \times V$  is the set of edges. A community in a network is a subset of nodes in which the internal density of connections is significantly greater than the density of connections with the rest of the network.

The first attempt to identify communities began with Girvan & Newman [Girvan and Newman 2002] who proposed the removal of edges with high *betweenness* in an interactive way. However, the high cost of this approach, especially in large networks, has driven the study of initiatives based on *modularity*, due to its lower number of iterations. In this context, one of the main algorithms is the Louvain algorithm, presented by [Blondel et al. 2008]. Much of the literature carries out studies on single-layer graphs, but there are situations where multilayer networks emerge. Thus, unlike traditional (monolayer) networks, where a community can be defined solely based on the density of connections, in *multiplex* networks a community can emerge from the combination of interactions in multiple layers simultaneously. Some proposals have been presented to improve this representation, for example that of [Karimi et al. 2020], which uses evolutionary algorithms in search of greater segregation of communities.

### 2.2. Louvain algorithm optimizations

The Louvain algorithm has high computational cost in dense graphs. [Zhang et al. 2021] introduced the concept of *Fast Louvain*, where after a pre-analysis of the graph, nodes with little significance to modularity are temporarily eliminated. This action reduces the

number of iterations and consequently the execution time. [Chaudhary and Singh 2019] proposes a hierarchical greedy approach for better community identification in large volumes of data. Recently, [Wang et al. 2021] proposes the use of weights in relation to the geographic distance between agents.

#### 2.3. Social networks in animal causes

[Brask et al. 2021] provides an introduction to animal social networks for complex systems scientists, highlighting areas of synergy and encouraging interdisciplinary collaborations. [Neethirajan and Kemp 2021] highlights the use of sensing technologies in investigating social network analysis of farm animals, such as cattle, to improve management and welfare. The work of [Finn et al. 2019] reviews how multilayer network analysis can be applied to the study of animal behavior, offering new insights into social structure and the positions of individuals within these structures.

To our knowledge, no previous study has explored the use of complex network science to model care networks for stray animals. Although existing literature applies social network analysis in contexts such as captive animal welfare, farm management, and behavioral ecology, the specific investigation of cooperation structures between protectors, shelters, and institutions related to the care of stray animals remains an open and underexplored research gap.

## 3. Materials and methods

## 3.1. Methodology

The development of the codes used in this study were carried out in Python and implemented in the Google Colab platform. For comparative studies and better future sharing, all collected data were saved in a sqllite database allocated within Google Drive. The codes and data captured for this study will be available at: https://github.com/.

Due to the lack of available data for the development of this study, simulations aligned with real data obtained from recognized sources for a specific city in Brazil were used. Demographic, population, geolocation and establishment data involved with the animal cause were obtained through the resources available from IBGE and Google tools, and guide the simulation of fictitious data such as protectors and animals distributed within the geographic field of the city following heuristics and proximity radius.

Access to the API provided by IBGE at https://cidades.ibge.gov.br allows access to the 2022 CENSUS with information on population and demographic density. This information will be essential for defining the heuristic for determining the number of animals per municipality, given the projection made by Instituto Pet Brasil and made available by the Ministry of Agriculture [Instituto Pet Brasil 2018] which reports that in 2018, 3.9 million animals are in a situation of allowance and of these, 69% are dogs and 31% are cats. This heuristic considers the population of street animals to be 5% of the human population.

$$Dogs = 0.69 \times Animals\_Street$$
 (2)

$$Cats = 0.31 \times Animals\_Street$$
 (3)

The protectors use a similar simulation methodology, considering that they are 1% of the population density of the municipality. In addition, they receive a real name with the help of the Faker library for greater realism.

$$Protectors = \frac{Population}{Demographic\_Density} \times (0.01 \times Demographic\_Density)$$
 (4)

Next, still for the same municipality, the Google Places API is consulted with the query text "petshops, veterinary clinics, animal NGOs in" concatenated with the city name and finally the Google GeoCode API to convert the address into latitude and longitude coordinates. With this, we obtained real and updated data on establishments involved with the veterinary theme and that can be potential agents of the action, as exemplified by Table 1.

Table 1. Example of Veterinary Clinics Collected

Name	Address	Type of Establish- ment	Latitude	Longitude
Hospital Veterinário	Av. Dr. Ademar de	veterinary_care,	-23.2008553	-45.8927735
Carinho de Bicho -	Barros, 1000 - Jardim	point_of_interest, es-		
Unidade Adhemar de	Sao Dimas, São José	tablishment		
Barros	dos Campos - SP,			
	12245-011, Brazil			
Veterinary Clinic Soft	R. Me. Paula de São	veterinary_care,	-23.2032522	-45.8995935
Affection - Unit Vila	José, 171 - Vila Ema,	health,		
Ema	São José dos Campos -	point_of_interest,		
	SP, 12243-010, Brazil	establishment		
Maple Vet 24 horas	R. Siqueira Campos,	pet_store, pharmacy,	-23.1800171	-45.8837732
<ul> <li>Clínica Veterinária</li> </ul>	314 - Centro, São	veterinary_care,		
Popular & Petshop	José dos Campos - SP,	store, health,		
	12210-250, Brazil	point_of_interest,		
		establishment		

Another simulation point that is extremely important for the development of this project is the interactions between the agents and the animals. A heuristic was applied that assigns an average number of aids that a protector performs and a desistance deflator that tries to simulate the fact that a protector may commit to, but not actually perform an aid.

$$Number\_Interactions = Protectors \times Number\_Help \times (1 - Coefficient\_Dropout)$$
 (5)

where,

- *Number\_Interactions*: Total number of interactions performed randomly.
- *Protectors*: Number of individuals or entities offering protection.
- *Number\_Help*: Amount of assistance provided.
- Coefficient Dropout: Proportion representing the dropout rate.

The municipality's boundary is given through the maximum latitude and longitude coordinates of the northeast and southwest extremities also obtained by the Google GeoCode API and all interactions and simulations are carried out within the blue rectangle as shown in Figure 1.



Figure 1. Bordering the city of São José dos Campos.

This study considers establishments and protectors as partner agents to interact with animals marked as abandoned by municipal boundaries, which are represented in the plot in blue and green, respectively. The types of interactions are parameterizable and the following were considered in the project: sponsorship, donation, temporary shelter and rescue. The animals were considered cats and dogs because the proportionality of division of population estimation studies contemplates only these species. The simulation process uses the real coordinates already obtained from the establishments found and respects the geographic limits of the city. They are created randomly according to the heuristics presented above within the coordinate range of the map. The interactions, although also random, respect the radius limit of 5 km of distance between the protector or establishment in relation to the animal, calculated by the coordinates through the haversine formula [Louwers 2022]. Plotting the interactions on the map using the Folium library allows for efficient visualization of the type of interactions performed, as shown in Figure 6.

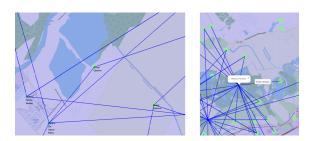


Figure 2. Simulated interactions between agents and animals.

Each edge received its weight taking into account the total distance between the vertices in question and the type of interaction performed between them, with the weight being the result of the product of these two attributes. The metric assigned to each type of interaction can be seen in Table 2.

Interaction Type	Weight	Description
Sponsorship	1.0	It represents the highest level of
		commitment to the animal.
Donation	0.8	Financial or material contribu-
		tion to help the animal.
Temporary Shelter	0.5	Availability to shelter the animal
		temporarily.
Rescue	0.3	Indicates less relative effort, but

Table 2. Weights Assigned to Interactions

The result of applying the various heuristics presented with the real data captured is a complex network that models the interactions between establishments, protectors and animals. In this network, these agents are considered vertices and the edges are weighted according to the relevance of the help and geographic proximity. To better understand the relationships in this network and identify possible communities in general and segregated into layers by type of interaction, the Louvain algorithm (from the python-louvain library) was applied. This algorithm is popular in community discovery, mainly due to its efficiency and scalability, and acts by maximizing modularity, which is a measure that highlights the internal connections between a set of vertices to be considerably larger than its external ones, as seen in Equation 6. These communities will allow finding groups that help animals and optimizing them from a logistical point of view, in addition to locating the main protectors within a group.

$$Q = \frac{1}{2W} \sum_{i,j} \left[ w_{ij} - \frac{k_i k_j}{2W} \right] \delta(c_i, c_j) \tag{6}$$

essential to save the animal.

- Q: Total modularity of the graph.
- W: Total sum of the weights of the edges in the graph ( $W = \frac{1}{2} \sum_{i,j} w_{ij}$ ).
- $w_{ij}$ : Edge weight between nodes  $i \in j$ .
- $k_i$ : Weighted node degree i (sum of the weights of the edges connected to the node i).
- $c_i$ : Community to which the node i belongs.
- $\delta(c_i, c_j)$ : Kronecker delta function, which is 1 if  $c_i = c_j$  (the knots  $i \in j$  belong to the same community) and 0 otherwise.

#### 3.2. Dataset

For the application of this first phase of the study, the city of São José dos Campos, located in the state of São Paulo, was chosen. Considering the values obtained in the IBGE queries in relation to the 2022 CENSUS, we attributed a population of 697,054 people distributed at a density of 634 people per square kilometer. Heuristics were applied to obtain the values of animals, protectors and their interactions. However, for study purposes and to reduce computational costs, the quantities of animals and protectors were divided by 100. The values of establishments are real and obtained through geographic query.

Category	Value
Dogs	240
Cats	108
Protectors	69
Establishments	56
Interactions	11.849

**Table 3. Heuristic Values for Agents** 

All information is persisted and versioned in a database for control, economy and comparative studies with the same base, as illustrated in Tables 3, 4 and 5.

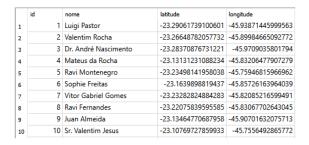


Figure 3. Database Users table.



Figure 4. Database Establishments table.

	id	especie	descricao	foto	latitude	longitude	data_reportada	id_usuario
1	1	Cachorro	Animi eius sed velit excepturi aliquam.	NULL	-23.16350504871157	-45.88755241328546	2023-08-31	47
2	2	Cachorro	Id officiis aliquid at ea.	NULL	-23.17119808233897	-45.85147423777401	2024-10-25	NULL
3	3	Cachorro	Mollitia fugit porro tempora.	NULL	-23.11786675412782	-45.88654310081596	2022-04-14	NULL
4	4	Cachorro	Vitae recusandae officiis voluptatibus unde accusantium ullam eaque.	NULL	-23.26913631980562	-45.76604887512912	2022-01-03	31
5	5	Cachorro	Fugit necessitatibus assumenda fugiat ut.	NULL	-23.21120594047402	-45.96357790039378	2024-06-22	NULL
6	6	Cachorro	Provident aperiam aliquam occaecati in iure voluptatibus.	NULL	-23.13860232975752	-45.81637382188983	2024-04-09	56
7	7	Cachorro	Officia cumque cupiditate reiciendis mollitia.	NULL	-23.27927937368133	-45.77150907938783	2024-09-06	NULL
8	8	Cachorro	Aut iste corrupti quo aperiam repudiandae ut.	NULL	-23.26874414534469	-45.8815704624169	2024-08-01	NULL
9	9	Cachorro	Vel autem pariatur pariatur.	NULL	-23.09498342475045	-45.83782237799479	2023-01-12	34
10	10	Cachorro	Laudantium perferendis distinctio similique.	NULL	-23.26290531861039	-45.86562372072651	2023-09-10	8

Figure 5. Database Animals table.

Since animals do not interact and at this stage of the study we are not considering direct actions between establishments and protectors, the network (Fig. 6) was implemented as undirected and has 474 vertices and 1231 edges. The difference between the final network and the input data occurs mainly because the generation of interactions is random and when greater than 5 km in relation to the agent it is ignored, in addition to the possible repetition of an interaction. In this case, regardless of the number of times it occurs, it will be considered unique.

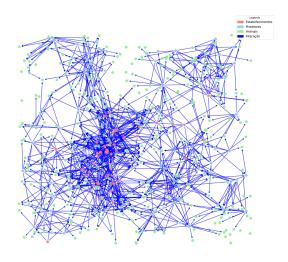


Figure 6. Interaction graph.

## 3.3. Algorithm and environment settings

The Google Colab infrastructure was used in the standard version of the platform on a Linux system with x86\_64 architecture, 12Gb of RAM and a 2 core processor. The study did not include a GPU and all data persistence used a sqllite file within Google Drive. The algorithms used by Louvain were not modified, however the amount of data was reduced, as reported in the data acquisition section, for computational cost purposes.

## 3.4. Analysis criteria

The analysis criteria defined were the detection of communities through the Louvain algorithm using modularity to determine the quality of these clusters. The most important vertices were discovered through the degree centrality property, highlighting those with the greatest number of connections and edges that reflect their distance and interaction.

## 4. Results and Discussions

By separating the graph into layers within the interactions possible by the system, we can analyze a distinct behavior between them. As seen in Table 4 and Figure 7, the high modularity rates calculated by the Louvain algorithm suggest a high coupling and quality of relationships.

Regardless of the type of interaction, all communities appear to develop around the actual establishments in the network. When plotting this plot, the geographic position

of the agents was maintained in search of patterns. The rescue layer has a more dispersed core compared to the others, while the sponsorship layer is the most concise and practically does not stray from the core.

Table 4. Modularity by interaction layer

Layer	Modularity
rescue	0.8904
temporary shelter	0.8436
donation	0.8440
sponsorship	0.8729

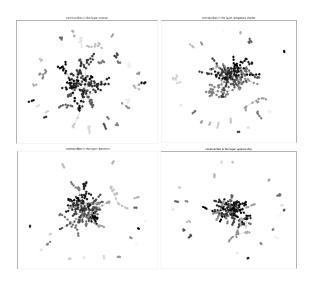


Figure 7. Communities by layers of interactions.

When applying the Louvain algorithm to the graph as a whole, modularity drops drastically to 0.5972. This suggests that, although the quality of the community structure is considerable when observed by only one type of interaction, it loses strength when considering unconnected vertices that orbit around the center. This type of behavior corroborates that the evaluation by type of interaction is more efficient than in general terms.

Within the communities, or in general, it was possible to identify through degree centrality the protectors and establishments within the network that collaborate most with aid within their radius of action, considering the number of connections and the weight of each one of them, according to the interaction carried out and the distance between the agents (Table 5 and Figure 8). With this approach, we can apply incentives in a targeted manner and identify possible bottlenecks.

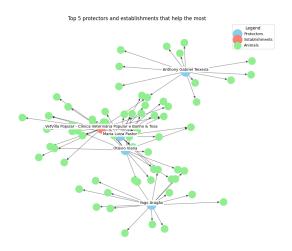


Figure 8. Top 5 establishments that help the most.

Table 5. Main Protectors and Establishments.

Name	Help
Otávio Viana	37.72
Maria Luiza Pastor	35.85
Yago Aragão	33.59
VetVilla Popular - Clínica Veterinária Popular e Banho & Tosa	32.62
Anthony Gabriel Teixeira	31.87

## 5. Conclusion and Future Work

Measuring the stray animal population is a challenging task due to several objective and subjective factors. This information is crucial for decision-making by public bodies and for NGOs and independent animal protection organizations.

The first critical point is obtaining data in an assertive and direct manner, as there are no recent studies with an adequate level of reliability. The dynamic characteristics of the animal population and the positioning radius of individuals are also challenging and require more in-depth studies and techniques. To overcome this obstacle, projection heuristics based on population or demographic indicators are applied.

Another characteristic is the size of the database and its computational cost to perform operations. In this study, for example, the data was reduced by 100 times for applicability. However, even for cities with low population density, this value can be heavy for home computers and/or low processing power, requiring a processing service that is compatible with the size of the databases.

The results in identifying communities within the graph created were promising. The modularity value obtained shows a relationship above random between protectors, establishments and animals. Identifying the main protectors was also important for understanding where to support actions and inputs in the expansion of the aid network, as well as understanding the main bottlenecks.

As next steps, we intend to improve the heuristics for generating animals and

protectors, adding new indicators such as the Human Development Index (HDI) and the Urbanized Area range, assigning different weights to each quadrant of the area and not on an average basis as is currently the case. In addition, we intend to use this database in a more reliable way through a mobile application, where users can interact and feed this information in a community and in real time. The use of GPU parallelism in higher performance programming languages with an optimized version of Louvain also shows promise for overcoming the problems of database size and the applicability of simulation and evaluation algorithms in an improved version.

# 6. Ackowledgement

We thanks Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes).

#### References

- Bankar, S., Bhagwat, H., Karande, A., Sargar, P., and Dabhade, S. A. (2024). Enhanced stray animal care using machine learning. *International Research Journal of Modernization in Engineering, Technology and Science*, 6(2):102–105. e-ISSN 2582-5208.
- Blondel, V., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics*, page P10008.
- Brask, J. B., Ellis, S., and Croft, D. P. (2021). Animal social networks: an introduction for complex systems scientists. *Journal of Complex Networks*, 9(2):cnab001.
- Chaudhary, L. and Singh, B. (2019). Community detection using an enhanced louvain method in complex networks. In *15th International Conference on Distributed Computing and Internet Technology (ICDCIT)*, volume 11319 of *Lecture Notes in Computer Science*, pages 243–250. Springer.
- Finn, K. R., Silk, M. J., Porter, M. A., and Pinter-Wollman, N. (2019). The use of multi-layer network analysis in animal behaviour. *Animal behaviour*, 149:7–22.
- Girvan, M. and Newman, M. (2002). Community structure in social and biological networks. *Proc. National Academy of Sciences*, 99:7821–7826.
- Horecka, K. and Neal, S. (2022). Critical problems for research in animal sheltering, a conceptual analysis. *Frontiers in Veterinary Science*, 9:804154.
- Instituto Pet Brasil (2018). Inteligência de Mercado: Um Olhar Sobre o Setor Pet no Brasil.
- Karimi, F., Lotfi, S., and Izadkhah, H. (2020). Multiplex community detection in complex networks using an evolutionary approach. *Expert Systems with Applications*, 146:113184.
- Kinsley, A. C., Rossi, G., Silk, M. J., and VanderWaal, K. (2020). Multilayer and multiplex networks: An introduction to their use in veterinary epidemiology. *Frontiers in Veterinary Science*, 7:596.
- Louwers, J. (2022). Calculate geographic distances in python with the haversine method. *Medium*.

- Neethirajan, S. and Kemp, B. (2021). Social network analysis in farm animals: Sensorbased approaches. *Animals*, 11(2):434.
- Newman, M., Barabási, A.-L., and Watts, D. J. (2011). *The structure and dynamics of networks*. Princeton university press.
- Oliveira, E. T., Nascimento, M. L., Silva, R. F., and Costa, L. A. (2008). A questão dos animais abandonados e o papel do médico veterinário. *Revista Brasileira de Higiene e Sanidade Animal*, 2(1):56–79.
- Wang, C., Wang, F., and Onega, T. (2021). Network optimization approach to delineating health care service areas: Spatially constrained louvain and leiden algorithms. *Transactions in GIS*, 25(2):1065–1081. PMCID: PMC8140423.
- Zhang, J., Fei, J., Song, X., and Feng, J. (2021). An improved louvain algorithm for community detection. *Mathematical Problems in Engineering*, 2021:1–14.