



HAL
open science

Measuring Power Relations Among Locations From Mobility Data

Lucas Santos de Oliveira, Pedro Olmo Stancioli Vaz de Melo, Aline Carneiro
Viana

► **To cite this version:**

Lucas Santos de Oliveira, Pedro Olmo Stancioli Vaz de Melo, Aline Carneiro Viana. Measuring Power Relations Among Locations From Mobility Data. ACM MobiWac 2019 - 17th ACM International Symposium on Mobility Management and Wireless Access, Nov 2019, Miami, United States. hal-02302490

HAL Id: hal-02302490

<https://inria.hal.science/hal-02302490>

Submitted on 1 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Measuring Power Relations Among Locations From Mobility Data

Lucas Santos de Oliveira
lsoliveira@uesb.edu.br
Federal University of Minas Gerais
Belo Horizonte, Minas Gerais, Brazil

Pedro O. S. Vaz de Melo
olmo@dcc.ufmg.br
Federal University of Minas Gerais
Belo Horizonte, Minas Gerais, Brazil

Aline Carneiro Viana
aline.viana@inria.fr
Inria Saclay - Ile de France
Palaiseau, France

Abstract

Key location identification in cities is central in human mobility investigation as well as for societal problem comprehension. In this context, we propose a methodology to quantify the power of point-of-interests (POIs) in their vicinity, in terms of impact and independence – the first work in the literature (to the best of our knowledge). Different from literature, we consider the flow of people in our analysis, instead of the number of neighbor POIs or their structural locations in the city. Thus, we first modeled POI's visits using the multiflow graph model where each POI is a node and the transitions of users among POIs are a weighted direct edge. Using this multiflow graph model, we compute the *attract*, *support* and *independence powers*. The *attract power* and *support power* measure how many visits a POI gather from and disseminate over its neighborhood, respectively. Moreover, the *independence power* captures the capacity of POI to receive visitors independently from other POIs. Using a dataset describing the mobility of individuals in the Dartmouth College campus, we identify a slight dependence among buildings as well as the tendency of people to be mostly stationary in few buildings with short transit periods among them.

1 Introduction

The rise of mobile technologies and collective sensing in the last decade has contributed to the generation of large datasets that describe activity dynamics in cities, which created new opportunities for research [7, 24]. Many works address these opportunities by means of user mobility prediction and Point-of-Interest (POI) recommendation [6, 14].

However, another fundamental area of study in traditional urban literature is the investigation of the way in which city neighborhoods become popular and how movement of citizens impacts the number of visitations in POIs [7, 6, 4, 9].

In fact, identifying key locations, which are places where persons spend a considerable amount of time during the day or which they visit frequently [20], is central to understand human mobility and social patterns. Such understanding can, in turn, inform solutions to large-scale societal problems in fields as varied as telecommunications, ecology, epidemiology, and urban planning [15]. In the literature, many works address human mobility as a graph-based approach [4, 21] that creates an opportunity for identifying important locations by means of power relationships, where one has the capacity to influence others or events.

In his seminal work on power-dependence relations, Richard Emerson [10] claims that power is a property of the social relation: i.e., saying “X has power” is vacant, unless we specify “over whom”. Based on this claim, [3, 2] define power in the context of exchange networks, where the relationship in the network and their weights involve the transfer of valued items (i.e., information, time, money, energy). Different from centrality measures such as PageRank and closeness, which state that a node is central if it is connected with central nodes, Bozzo and Franceschet [3] consider that a node in a network is powerful if it is connected with powerless nodes. This notion of power leads to two assertions concerning the power of a node: (i) it is directly correlated with the number of its neighbors, and (ii) it is inversely correlated with the power of its neighbors [3]. In fact, the more ties a node has, the more powerful the node is. However, the second property, which is not habitual, characterizes power well: powerful nodes can impose their will on powerless nodes since the first has many other options to negotiate and the latter do not [2, 3].

Given the above context, in this paper, we propose the first graph-based methodology in the literature to quantify the power of POIs by means of user mobility trajectories. Different from [3], who infer the power of nodes from the sum reciprocal node degrees, we infer the power of a node from its visiting flows and according to the following three distinct approaches.

First, the *attract power*, which is the capacity of a POI to receive people from its vicinity. Second, the *support power*, which is the capacity of a POI to disseminate people over its vicinity. Given a large set of visits and mobility trajectories made by people, we calculate the potential impact [4], or influence, a POI has in its neighborhood. Imagine, for instance, a university campus and its impact in its nearby restaurants and bars when it is closed for summer vacations. Third and conversely, the *independence power*, which is the potential resilience a POI has to other POIs moving out from its neighborhood. Using the

same example, replacing (or shutting down) restaurants and bars in the vicinity of the campus will probably not affect its visitations much.

Unlike traditional centrality metrics that have limited utility for modeling flow in exchange networks and identifying places of power, our metrics tackle these problems and clearly distinguish among the three types of power a POI can assume: power to disseminate, to gather people and to be indifferent to the flow of people visiting the POIs in the vicinity.

In this work, we modeled the power quantification problem using the *multiflow graph model* (MGM) where each POI is a node and the transitions of users among POIs are a weighted direct edge. The weight of the edge is the number of transitions made by people who moved from one POI to the other. The *attract power* and *support power* are calculated using the user trajectories, which measures how many visits a POI attracts from and disseminate over its neighborhood, respectively. The *independence power* is calculated by measuring the number of visits a POI receives that are not influenced by any other POI. We applied this methodology on the mobility dataset collected in the Dartmouth College campus [18]. We found that the campus buildings have low *support power* and *attract power*; people tend to move over a few buildings and spend most of their time in the same building; and there is a slight dependence among buildings, even those with high *independence power* receive user visits from other buildings on campus.

Moreover, although we applied the methodology in the context of urban mobility, it can be used in other contexts, such as the analysis of the influence on social networks through the dissemination of information. Our metrics would help to identify who are the most influential people (*support power*) and who are the most passive people (*attract power*). Additionally, it can also be used to model message delivery problems in Mobile ad-hoc Networks (MANET). The effectiveness of message delivery could be tested since our metrics identify and quantify the most powerful places to spread and to attract people.

Prior studies analyzed the impact that a new business has on the local market ecosystem [4, 6]. Others analyzed the resilience in terms of business survival characteristics [8, 26]. Different from these works, our work quantifies the impact of the POI in its vicinity individually, without category dependence. Besides, our work measures the POI resilience by means of visitation independence, i.e., it is invariant to the instability of vicinity visits.

In short, the main contributions of this work are threefold:

- We propose a graph-based methodology named *multiflow graph model* to identify key locations. From mobility user trajectories, we modeled the problem by means of Power relations among POIs.
- From this graph approach, we propose three different influence measures: the *attract power* and *support power*, which are, respectively, the capacity of a POI to gather and to disseminate people, and the *independence power*, which is the POI capacity of receiving visitors independently from others POIs.

- Different from other metrics of Power, the *independence power*, *attract power*, and *support power* ranked as powerful distinct POIs.

The rest of this paper is structured as follows. In Section 2, we describe the related work. In Section 3, we describe the Dartmouth dataset characteristics. Thereafter, in Section 4, we present a methodology to measure the POI impact power and independence. In Section 5, we apply this methodology in the Dartmouth dataset to identify the powerful POIs. Finally, in Section 6, we conclude our work.

2 Related Work

The literature is rich in solutions that aim to leverage human mobility. The rise of mobile technologies and collective sensing in the last decade has contributed to the generation of large datasets that describe activity dynamics in cities and has created new research opportunities [7, 24]. Researchers addressed this challenging task using different data sources, such as location-based social networks (LBSN) and mobility traces.

Such mobility data sources have been used on the study of POI recommendations [28, 27]. The general idea of such works is to exploit the social connections and the favourite POIs of users to recommend new places to be visited. Ye et al. [27], for instance, explored user preference, social influence and geographical influence to provide a POI recommendation service. Similarly, Zhang et al. [28] proposed a new approach called LORE to exploit the sequential influence of locations on users’ check-in behaviors for location recommendations.

There are also works that analyze mobility data to predict, as accurately as possible, the future location of individuals and their friends [23, 25]. D’Silva et al. [7], for instance, treated venue categories as proxies for urban activities to forecast the weekly popularity dynamics of a new venue establishment. Alternatively, Feng et. al. [11] proposed an attentional recurrent neural network model for predicting human mobility from lengthy and sparse trajectories. Moreover, Silveira et. al.[25] proposed a family of data-driven models, called MobHet, to predict human mobility using heterogeneous data sources. In a comparable way, Sadilek et. al. [23] explored the interplay between people’s location, interactions, and their social ties. Then, they proposed a system that predicts the location and social ties in online social networks.

Another relevant research problem is the identification of relevant (or important) locations in mobility traces, i.e., personal places of interest (PPOIs): home, work, or any place where persons spend a considerable amount of time during the day or which they visit frequently [20]. Based on this problem, Pavan et al.[20] proposed a mapping scheme of POIs onto a feature space to identify those important locations. Alternatively, Isaacman et al. [15] proposed techniques based on clustering and regression for analyzing anonymized cellular network data, and to discern semantically-meaningful locations. Similarly, Cambe et al. [4] proposed a framework to examine the role of new businesses in their respective local areas. Using urban activity, they measure the impact,

either positive or negative, that retail facilities have on each other. Different from this work, the method proposed in [4] is dependent on the POI category and uses a different impact percentage scale that varies around 1. Additionally, Cambe et al. only analyze the impact of a location in its vicinity and do not identify the most important places (see Section 4.2). Finally, the authors do not take into account the independence of one location in relation to its neighborhood (see Section 4.3).

Another way to infer the importance of POIs is by analyzing user transitions among them as an exchange network [3, 1]. Some studies addressed this problem as a relation of power, where the relationship in the network involves the transfer of valued items (i.e., information, time, money, energy)[3, 2, 1]. Moreover, it is advantageous to be connected to those who have few options. This type of relational power is endogenous with respect to the network structures, meaning that it is a function of the position of the node in the network [2].

Additionally, the study of power has a long history in economics (in its acceptance of bargaining power) [19, 22] and sociology (in its interpretation of social power)[5, 2, 12]. In this work, we use the definition of power found in [3, 2]. However, we tackle this problem in a different way, taking into account the flow of people to identify powerful locations, not only the number of neighbors or its structural location in the network. More importantly, *to the best of our knowledge, this is the first work in the literature that uses power relations to infer the importance of a POI in its vicinity.*

3 Dataset

In this work, we use the Dartmouth dataset [18] to evaluate our proposed methodology. According to Henderson et al. [13], Dartmouth College campus has over 190 buildings on 200 acres and more than 560 Cisco 802.11b access points (APs) were installed to cover the whole area. This dataset contains extensive trace information from the entire network from the Fall term of 2001 until Winter 2004. However, following Henderson et al.[13], we focus on Syslog messages collected during the Fall 2003 and Winter 2004 terms, a 17 week period from 2 November 2003 to 28 February 2004, inclusive. Dartmouth College had about 5500 students and 1200 faculty, and during the collection period, there were approximately 3200 under-graduates on campus, of which 75% owned a laptop.

Dartmouth data collection resulted in an extremely large dataset. Over the 17 weeks of the trace, they saw 7134 unique cards associated with an AP and received 32,742,757 Syslog messages. Then, the APs were classified by the type of building in which they are located: 221 residential, 147 academic, 72 administrative, 59 libraries, 45 social and 22 athletic. Residences include dormitories, fraternities, business school, and faculty housing. Social buildings include dining areas, an arts center, and a museum. Athletic facilities include skating rinks, football fields, boathouses, and a ski lodge [18].

According to Kotz et al. [18], the AP names are meaningless to people out-

side Dartmouth, so they have replaced the names with a descriptive name of the form Building#AP#; for example, ResBldg100AP2 is AP#2 in the Residential Building #100. In this work, we removed the AP number to eliminate cases where computers were associating and reassociating with several APs many times in succession. Additionally, we consider that a user is located in the building if he/she is connected to an AP that belongs to the building. For example, if a user is connected in ResBldg100AP2, then we consider that he/she is in ResBldg100.

4 Methodology

In this section, we describe the *multiflow graph model* where each location is a node and transitions of users between locations define the weighted directed edges. From this graph, we compute the *attract power*, *support power*, and *independence power* metrics.

4.1 Multiflow Graph Model

The input necessary to construct the *multiflow graph model* is a set of visits \mathcal{C} over a set of POIs \mathcal{P} made by a set of users \mathcal{U} . For simplicity, we organize the set \mathcal{C} into disjoint ordered sets $C_1, \dots, C_{|\mathcal{U}|}$, where $C_i = \{c^1, c^2, \dots\}$ corresponds to all the visits c^j made by user u_i , in chronological order. A visit c^j is a tuple $\langle t, p \rangle$, where t is the time the visit started and p is the POI where it took place. We denote by $t(c^j)$ and $p(c^j)$ the timestamp and the POI of *visit* c^j , respectively.

Then, from each ordered set C_i , we construct the corresponding set of *trajectories* S_i of user u_i . S_i contains all trajectories user u_i made during their social days, i.e., the distinct places u_i visited in chronological order. More formally, we transform every ordered set $C_i = \{c^1, c^2, \dots\}$ into a set $S_i = \{S_i^1, S_i^2, \dots\}$, where $S_i^k \in \mathcal{S}$ is a sequence of chronologically ordered POIs visited by user u_i , i.e., $S_i^k = \{p^1, p^2, \dots\}$. This transformation is done by adding, in chronological order, the POI $p(c^j)$ of each visit $c^j \in C_i$ to a sequence $S_i^k \in S_i$.

This process starts by making $k = 1$ and by adding $p(c^1)$ to S_i^k . Then, for each of the following visits c^j , we do the following:

- If $t(c^j) - t(c^{j-1}) \geq 6$ hours, make $k = k + 1$ and add $p(c^j)$ to S_i^k ;
- Else if $t(c^j) \geq 6:00$ and $t(c^{j-1}) < 6:00$, make $k = k + 1$ and add $p(c^j)$ to S_i^k ;
- Else if $p(c^j) = p(c^{j-1})$, do nothing and process POI $p(c^{j+1})$;
- Else if $p(c^j) \in S_i^k$, make $k = k + 1$ and add $p(c^j)$ to S_i^k ;
- Else, add $p(c^j)$ to S_i^k ;

This guarantees that all trajectories have the following properties: (i) all visits occur on the same social day, which begins at 6:00 and ends at 5:59 the next day; (ii) all transitions between consecutive POIs happened in less than 6 hours; (iii) all POIs of a trajectory are distinct, i.e., each user trajectory S_i^k is a simple directed path with no cycles.

We restricted $t(c^j) - t(c^{j-1}) < 6$ hours because Kotz et al. [17] suggested that around 90% of user session durations are less than 6 hours. So, for differences $t(c^j) - t(c^{j-1})$ greater than or equal to $6h$, we consider that the user has started a new *trajectory*. Also, if in a sequence of visits the user returns to a POI already present in this sequence, we break this sequence and start a new trajectory from this latter visit to prevent cycles to be formed.

Some trajectories are dense, and clearly represent what the user did in that day. On the other hand, some trajectories are composed by a single visit. In order to give the sense of flow to every trajectory, we added an artificial special location \mathbf{H} at the beginning and at the end of each user *trajectory* S_i^k , simply called as *Home*. Thus, each *trajectory* $S_i^k = \{p^1, \dots, p^n\}$ is transformed to $S_i^{k*} = \{p^0, p^1, \dots, p^n, p^{n+1}\}$, where $p^0 \equiv \mathbf{H}$ and $p^n \equiv \mathbf{H}$. For simplicity, we will keep denoting each trajectory by $S_i^k = \{p^1, \dots, p^n\}$, but with $p^1 = \mathbf{H}$ and $p^n = \mathbf{H}$ for all trajectories.

As the last step before constructing the *MGM*, we assign a *signature* κ to each trajectory $S_i^k = \{p^1, p^2, \dots, p^n\}$, which is simply defined by the ordered sequence of POIs present in S_i^k , i.e., $\kappa(S_i^k) = p^1 p^2 (\dots) p^n$. Note that trajectories with the same signature κ may exist within a set S_i and across different sets of users' trajectories S_i and S_j . All trajectories that have the same signature κ denote sequences of visits, or trajectories, over the exact same POIs. Additionally, we denote by $n(\kappa)$ the number of trajectories with signature κ . We also denote by S_κ the ordered sequence of POIs p^1, p^2, \dots, p^n defined by signature κ , which from now on we call a *signature trajectory* S_κ . The size of S_κ is denoted by $|S_\kappa|$. Finally, we denote by \mathcal{S} the set of all *signature trajectories* S_κ .

From the *trajectories* in \mathcal{S} , we assemble the *multiflow graph model* using the process described in Algorithm 1. For each trajectory $S_\kappa \in \mathcal{S}$, we traverse the whole trajectory S_κ and, for each location p^i , we create the edge $e(< p^i, p^j, \kappa >)$ between the POIs p^i and p^j . Additionally, we assign to the edge weight $w(< p^i, p^j, \kappa >)$ the value from $n(\kappa)$.

Algorithm 1: Multiflow Graph Model - (MGM)

```

Data: Trajectory set  $\mathcal{S}$ 
for  $S_\kappa \in \mathcal{S}$  do
  for  $i \leftarrow 1$  to  $\|S_\kappa\| - 1$  do
     $e(< p^i, p^j, \kappa >)$ ;
     $w(< p^i, p^j, \kappa >) = n(\kappa)$ ;
  end
end

```

Figure 1 shows a figurative example of the *multiflow graph model* that we represent as a MultiDiGraph $G(V, E, W)$, where: (i) vertices $v \in V$ represent

POIs; (ii) the multi-edge arrows $e \in E$ represent the user transitions between POIs; and (iii) the weights $w(v_i, v_j, \kappa) \in W$ are represented on the arrows from v_i to v_j . The trajectory signature κ is represented by its color, e.g., the edge between POI **A** and **R** is colored in yellow. Note, for instance, that 10 user trajectories start in **A** (red arrow), visit **L** and finish in **R**. Thus, the weight $w(v_i, v_j, \kappa) \in W$ represents the total number of trajectories from users who visited v_j just after visiting v_i , using the trajectory κ . Nodes Academic **A**, Library **L**, and Restaurant **R** represent campus buildings and nodes **H** represent the single artificial node *Home*. Although we represent **H** as multiple nodes in the figures, this is done only to ease the visualization.

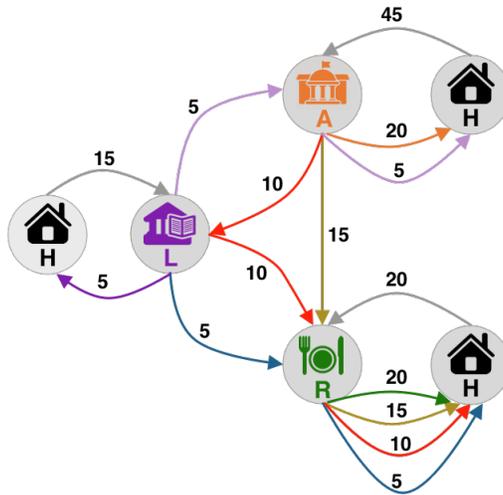


Figure 1: Mobility graph example.

4.2 Support and Attract Power

In this section, we show how to compute the *support power* and *attract power* using the *multiflow graph model* previously introduced.

Attract power. First, to assemble the *multiflow graph model* as shown in Figure 2, we compute the out-degree proportion for each weight $w(\langle p^i, p^j, \kappa \rangle) \in W$ associated to each edge $e \in E$ as shown in Algorithm 2. For each edge weight $w(\langle p^i, p^j, \kappa \rangle) \in W$, we compute the out-degree proportion by dividing its weight by the total sum of the edges weights w from the outgoing neighbors of node p^i , given by the function N_{out} . We denote by $N_{out}(p^i)$ the set of outgoing neighbors of node p^i . Similarly, we denote by $N_{in}(p^i)$ the set of incoming neighbors of node p^i .

Then, after computing the out-degree proportions, we discard the Home **H** edge values as shown in Figure 3. Thereafter, for compute the *attract power* of each node p^i , we use Algorithm 3 described as follows. First, for each node $v_i \in V$, we initialize its *attract power* value with 0. Then, for each *trajectory* $S_\kappa \in \mathcal{S}$,

Algorithm 2: W_{out} edge weights out-degree proportion

Data: Edge weight set W , Outgoing neighbors function N_{out}

for $w(< p^i, p^j, \kappa >) \in W$ **do**

$w_{out}(< p^i, p^j, \kappa >) \leftarrow \frac{w(< p^i, p^j, \kappa >)}{\sum_{p^j \in N_{out}(p^i)} w(< p^i, p^j, - >)};$

end

we traverse all nodes p^i of this trajectory and sum cumulatively the proportions of edge weight $w_{out}(< p^{i-1}, p_i, \kappa >)$ associated to the edge $e(< p^{i-1}, p^i >)$. At the same time, for each node p^i that we traverse, we add the cumulative total to its *attract power*.

The intuition of this algorithm is that POI p^i attracts a portion of the people who left the POI p^{i-1} and, throughout the trajectory, each POI p^i is indirectly responsible for this portion of visits in its predecessors. The reason why we use the cumulative sum.

Algorithm 3: Attract Power

Data: out-degree edge weight set W_{out} , Trajectory set \mathcal{S} , Vertices set V

for $v_i \in V$ **do**

$attract[v_i] \leftarrow 0;$

end

for $S_\kappa \in \mathcal{S}$ **do**

$tot \leftarrow 0;$

for $i \leftarrow 2$ **to** $\|S_\kappa\| - 1$ **do**

$tot+ = w_{out}(< p^{i-1}, p^i, \kappa >);$

$attract[p^i]+ = tot;$

end

end

For example, as shown in Figure 3, to compute the *attract power* of the restaurant **R**, we sum the weights from incoming edges from **A** and **L**: $0.3+0.4+0.2+0.2 = 1.1$. Note that we cumulatively sum the weights of the red edges from POI **A** to **R**. Similarly, to compute the *attract power* of the Academic building **A**, we use the unique income edge with value 0.2 (arrow purple). Finally, for library **L**, we use the unique value 0.2 (arrow red). In short, the *attract power* of POI **A** is the cumulative sum of importance from each POI in its incoming neighborhood.

It is important to mention that the *attract power* has two substantial meanings. First, it shows the capacity of the restaurant **R** to gather people and, consequently, how powerful POI **R** is. Second, if location **R** closes its doors in this period, probably, 1.1 places will be impacted, because POI **R** will stop receiving visitors from these places.

Support power: First, to assemble the *multiflow graph model* as shown in Figure 4, we compute the in-degree proportion for each weight $w(< p^i, p^j, \kappa >) \in W$ associated to each edge $e \in E$ as shown in Algorithm 4. For each edge weight $w(< p^i, p^j, \kappa >) \in W$, we compute the in-degree proportion by dividing its weight by the total sum of the edges weights w from the incoming neighbors of node p^j , given by the function N_{in} .

Then, after computing the in-degree proportions, we discard the Home **H**

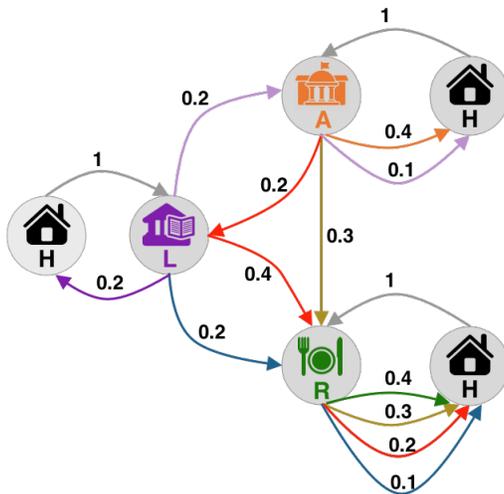


Figure 2: out-degree distribution to compute Attract Power

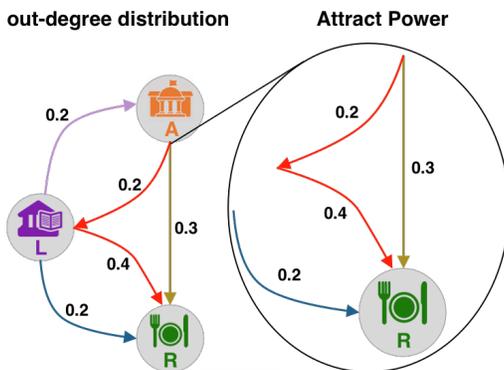


Figure 3: Attract Power

edge values as shown in Figure 5. Thereafter, for compute the *support power* of each node p^i , we use Algorithm 5 described as follows. First, for each node $v_i \in V$, we initialize its *support power* value with 0. Then, for each *trajectory* $S_\kappa \in \mathcal{S}$, we traverse all nodes p^i of this trajectory and sum cumulatively the proportions of edge weight $w_{in}(\langle p^{i-1}, p^i, \kappa \rangle)$ associated to the edge $e(\langle p^{i-1}, p^i \rangle)$. At the same time, for each node p^{i-1} that we traverse, we add the cumulative total to its *support power*.

Note that different from *attract power*, to calculate the *support power* cumulatively, we traverse the path in reverse order, i.e. from the destination to the starting POI. The intuition of this algorithm is that POI p^{i-1} disseminate a portion of the people who arrives in the POI p^i and, throughout the trajectory, each POI p^{i-1} is indirectly responsible for this portion of visits in its successors.

Algorithm 4: W_{in} edge weights in-degree proportion

Data: Edge weight set W , Incoming neighbors function N_{in}

for $w(\langle p^i, p^j, \kappa \rangle) \in W$ **do**

$w_{in}(\langle p^i, p^j, \kappa \rangle) \leftarrow \frac{w(\langle p^i, p^j, \kappa \rangle)}{\sum_{p^i \in N_{in}(p^j)} w(\langle p^i, p^j, - \rangle)}$;

end

The reason why we use the cumulative sum.

Algorithm 5: Support Power

Data: In-degree edge weight set W_{in} , Trajectory set \mathcal{S} , Vertices set V

for $v_i \in V$ **do**

$support[v_i] \leftarrow 0$;

end

for $S_\kappa \in \mathcal{S}$ **do**

$tot \leftarrow 0$;

for $i \leftarrow \|S_\kappa\| - 1$ **to** 2 **do**

$tot+ = w_{in}(\langle p^{i-1}, p^i, \kappa \rangle)$;

$support[p^{i-1}] + = tot$;

end

end

For example, as shown in Figure 5 to compute the *support power* of academic building **A**, we sum the weight of all its outgoing edges: $0.2 + 0.4 + 0.3 = 0.9$. Note that we cumulatively sum the weights of the red edges from POI **A** to **R**. Similarly, to compute the *support power* of library building **L**, we sum the outgoing edges' weights $0.1 + 0.2 + 0.1 = 0.4$ and for restaurant **R**, the sum is 0, because there is no outgoing edge. In short, the *support power* of POI A is the cumulative sum of importance given by each POI in its outgoing neighborhood.

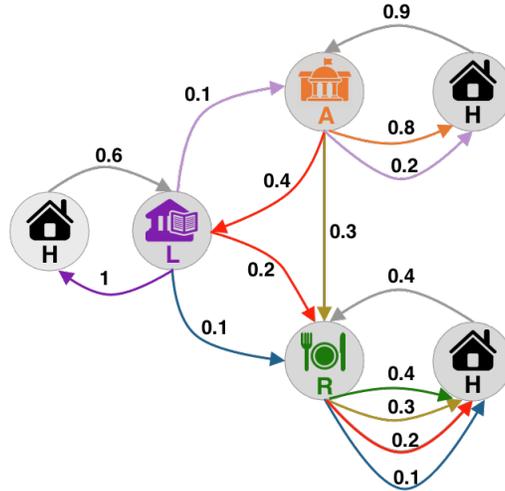


Figure 4: in-degree distribution to compute Support Power

Similar to *attract power*, *support power* also has two substantial meanings.

First, it shows the capacity of academic building **A** to disseminate people and, consequently, how powerful POI **A** is. Second, if location **A** closes its doors in this period, probably, 0.9 places will be impacted, because **A** will stop sending visitors to these places.

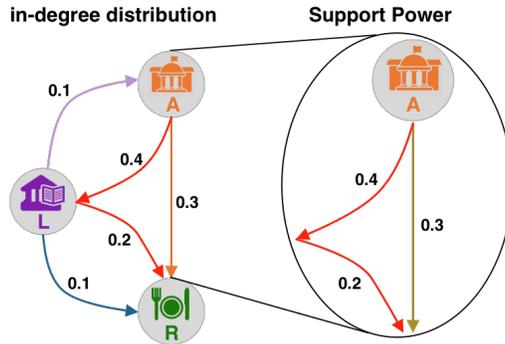


Figure 5: Support Power

4.3 Independence Power

The *multiflow graph model* also enables to compute another important metric that we call *independence power*, which shows the capacity of a POI to receive visitors independently from others POIs. Using the special **H** location associated with each POI in the *MGM*, we can infer how many POIs from incoming mobility flow have this building as a starting point, in *support power* case, and how many POIs from outgoing mobility flow have this same building as an ending point, in the case of *attract power*.

More formally, Algorithm 6 shows how to compute the *independence power* of a POI. For each node $p^i \in V$, we sum to the variable *in* the proportions of the incoming edges' weights $w(< H, p^i, \kappa >) \in W$, computed by dividing its weights by the total sum of the edges weights w from the incoming neighbors of node p^i . Also, we sum to the variable *out* the proportions of the outgoing edges' weights $w(< p^i, H, \kappa >) \in W$, computed by dividing its weights by the total sum of the edges' weights w from the outgoing neighbors of node p^i . Finally, the node p^i 's *independence power* is given by the harmonic mean between the variables *in* and *out* values. Similar to F1 Score, harmonic mean is appropriate for situations when the average of rates is desired.

Furthermore, the *independence power* value of a node ranges from 0 to 1, i.e., from totally dependent to totally independent from other POIs.

5 Results

In this section, we apply our methodology in the Dartmouth dataset to identify powerful and independent buildings.

Algorithm 6: Independence Power

Data: Edge weight set W , Vertices set V

```
for  $p^i \in V$  do
  independence[ $p^i$ ]  $\leftarrow$  0;
  in, out  $\leftarrow$  0;
  for  $w(\langle H, p^i, \kappa \rangle) \in W$  do
    in +=  $\frac{w(\langle H, p^i, \kappa \rangle)}{\sum_{p^h \in N_{in}(p^i)} w(\langle p^h, p^i, - \rangle)}$ ;
  end
  for  $w(\langle p^i, H, \kappa \rangle) \in W$  do
    out +=  $\frac{w(\langle p^i, H, \kappa \rangle)}{\sum_{p^j \in N_{out}(p^i)} w(\langle p^i, p^j, - \rangle)}$ ;
  end
  independence[ $p^i$ ] =  $2 * (in * out) / (in + out)$ 
end
```

5.1 Comparing the Campus Routine

Before showing the results for our proposed metrics, we perform a sanity check by examining the initial and final locations of the trajectories in the Dartmouth dataset. If the trajectories represent daily routines on campus, then we expect a substantial amount to start and end at residential buildings. Table 1 shows the percentage of times each type of building started and ended a trajectory and its most ranked buildings. Observe that, as expected, most of the trajectories start and end at residential buildings. Also, it is expected that a large number of trajectories do not start and end at residential buildings, as some students do not live on campus or do not turn on their computers before leaving or after arriving home by the end of the day. Furthermore, in a similar analysis, Henderson et al.[13] and Kotz et al. [17] ranked most of these top buildings as the busiest on campus, given that these are communal areas visited by many, if not most students. These results lead us to conclude that our methodology captured the daily routine on campus.

Table 1: Buildings with highest starting and ending number of trajectories

Building type	start(%)	end(%)	top start	top end
Academic	0.31	0.33	AcadBldg18	AcadBldg18
Administrative	0.03	0.04	AdmBldg20	AdmBldg20
Athletic	0.02	0.02	AthlBldg8	AthlBldg8
Library	0.13	0.17	LibBldg2	LibBldg2
Residential	0.43	0.36	ResBldg30	ResBldg83
Social	0.08	0.08	SocBldg1	SocBldg1

5.2 Identifying the Powerful Buildings

Given that our approach captured the student transitions routine on campus, we compute the *support power* and *attract power* for the Dartmouth dataset, as described in Section 4.2. Furthermore, we compare our approach with the following metrics of power: in-degree centrality (in_dg_cen), out-degree centrality (out_dg_cen), eigenvector centrality in (eigcen.in), eigenvector centrality out

(eigcen_out) and Bozzo and Franceschet power [3] (bf_power). Results are shown in Table 2. Note that the most powerful buildings are the social ones, such as libraries, academics, and restaurants. These buildings are hubs, i.e., they receive and disseminate students from and to all over campus, and are, in general, the busiest areas on campus [13].

Regarding the differences among the metrics, observe that the *attract power* rank has two buildings that do not appear in the baselines’ ranks: AcadBldg18 and AthlBldg8. Intuitively, having a high *attract power* means that these buildings are responsible for receiving a large fraction of students from other campus buildings. Similarly, the *support power* rank has two buildings that do not appear in the baselines’ ranks: AcadBldg18 and AthlBldg8. Having high *support power* means that these buildings are responsible for a substantial fraction of people arriving at other campus buildings. Furthermore, the *independence power* rank has all buildings that do not appear in the baselines’ ranks. Having high *independence power* means that a substantial fraction of people have these buildings as unique destinations on campus. The differences seen in our proposed power metrics are the consequence of an aspect that is not captured by the other baseline metrics: *the flow of people and the inter-dependence of this flow among the buildings*. In this direction, note how the bf_power has a contrasting rank in relation to our power metrics. Recall that the bf_power is computed using only the number of neighbors of the node, i.e., it ranks as powerful the buildings connected to the ones with the lowest degree. In the context of mobility transitions, it is not necessarily related to the inter-dependence among neighboring locations.

Table 2: Top 5 Dartmouth powerful buildings

Power	1st	2nd	3rd	4th	5th
attract	LibBldg1	LibBldg2	AcadBldg18	AthlBldg8	SocBldg4
support	LibBldg1	LibBldg2	AcadBldg18	SocBldg4	AthlBldg8
independence	AdmBldg9	AthlBldg5	ResBldg72	ResBldg76	ResBldg69
in_dg_cen	LibBldg1	SocBldg4	LibBldg2	SocBldg1	AcadBldg19
out_dg_cen	LibBldg1	SocBldg4	LibBldg2	SocBldg1	AcadBldg19
eigcen_in	LibBldg1	SocBldg4	LibBldg2	AcadBldg19	SocBldg1
eigcen_out	LibBldg1	SocBldg4	LibBldg2	AcadBldg19	SocBldg1
bf_power	AcadBldg9	SocBldg3	ResBldg22	LibBldg1	LibBldg2

To deeper investigate the differences among these ranks, we computed Kendall’s tau rank correlation [16] between the whole ranks produced by the metrics. The correlation ranges from -1 to 1 , and it is significantly greater than 0 when the ranks are similar (or identical for a correlation of 1). Conversely, it is significantly lower than 0 when the ranks are dissimilar (or fully different for a correlation of -1). If the correlation is close to 0 , then the ranks are independent.

Figure 6 shows the Kendall’s tau correlation results. Note that all of our proposed metrics have almost negligible correlations among themselves and with the other metrics. The highest, in absolute values, is between out_dg_cen and in_dg_cen with a positive correlation of 0.15 . These results show that our proposed metrics are different from traditional complex network metrics and from

the usual power metric [3].

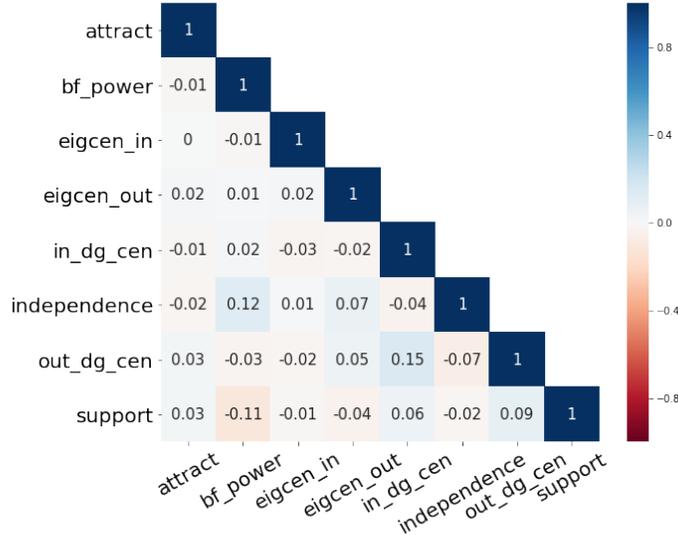


Figure 6: Kendall’s tau correlation among metrics

5.3 Powerful Buildings

From the Dartmouth mobility graph, we ranked the buildings according to the three power metrics introduced in Section 5. From these ranks, we analyze how our metrics of power relate to each other for the most powerful buildings. Figure 7 shows the scatter plot of the top 10 ranked buildings according to *attract power*, *support power*, and *independence power*. While the union set among the top 10 buildings for each metric totalizes 22 buildings, the intersection set results in 8 buildings. This shows that these metrics complement themselves, although, as expected, some natural overlap exists, especially between *attract power* and *support power*.

First, observe in Figure 7a that *LibBldg1*, *LibBldg2* and *AcadBldg18* stand out from the other buildings as the three most powerful according with *attract power* and *support power*, with values ranging from ≈ 3 to ≈ 5 . Intuitively, this means that these buildings have the potential to affect the visitation activity corresponding to the total amount of visits received by ≈ 3 to ≈ 5 other buildings. Furthermore, note that the highest-ranked buildings are, in the majority, social and library buildings, what is expected. According to [13], users spent less time in social and library than other buildings. Moreover, since these are communal area, many users have those buildings in their routine trajectories, which explains their high power of disseminating and of gathering users.

Regarding *independence power*, as we see in Figures 7b and 7c, the most powerful building is *AdmBldg9*, with *independence power* of 0.921. Intuitively,

tract power and *support power* measure how many visits a POI gather from and disseminate over its neighborhood, respectively. Moreover, the *independence power* is calculated by measuring the number of visits a POI receives that are not influenced by any other POI.

We applied our methodology in the Dartmouth dataset, which contains mobility traces of individuals over more than 190 buildings in the Dartmouth College campus. The results show that, in general, Dartmouth buildings have moderate to low *support power* and *attract power*, what is explained by the tendency of people to move over only a few buildings and to spend most of their time in the location. Nevertheless, we identified a slight dependence among buildings, even those buildings with high *independence power* receive user visits from other buildings on campus.

The novelty of our approach in methodological terms stems from the use of a graph-based approach, combined with the theory of power relations in exchange networks to tackle human mobility challenges. Moreover, our approach differs from traditional metrics of centrality in some circumstances such as exchange networks where the relationship in the network involves the transfer of valued items (i.e., information, time, money, energy). In this scenario, traditional metrics have limited utility in predicting powerful places. One limitation of our work is that we only consider the homogeneous impact of POIs in their vicinity, which does not take into consideration the heterogeneity of POIs categories and spatio-temporal bias effects. In addition, the impact brought by cycles in trajectories at the computation of the metrics is not considered. In future work, we will address these limitations and we will expand our analysis in other different size and nature traces.

References

- [1] Mohsen Bayati et al. “Bargaining dynamics in exchange networks”. In: *Journal of Economic Theory* 156 (2015), pp. 417–454.
- [2] Phillip Bonacich. “Power and centrality: A family of measures”. In: *American journal of sociology* 92.5 (1987), pp. 1170–1182.
- [3] Enrico Bozzo and Massimo Franceschet. “A theory on power in networks”. In: *Communications of the ACM* 59.11 (Oct. 2016), pp. 75–83. ISSN: 00010782. DOI: 10.1145/2934665. arXiv: arXiv:1510.08332v2. URL: <https://arxiv.org/pdf/1510.08332.pdf><http://dl.acm.org/citation.cfm?doid=3013530.2934665>.
- [4] Jordan Cambe et al. “Modelling Cooperation and Competition in Urban Retail Ecosystems with Complex Network Metrics”. In: (2019).
- [5] Karen S Cook and Toshio Yamagishi. “Power in exchange networks: A power-dependence formulation”. In: *Social networks* 14.3-4 (1992), pp. 245–265.

- [6] Krittika D’Silva et al. “If I Build It, Will They Come?: Predicting New Venue Visitation Patterns Through Mobility Data”. In: *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. SIGSPATIAL ’17. Redondo Beach, CA, USA: ACM, 2017, 54:1–54:4. ISBN: 978-1-4503-5490-5. DOI: 10.1145/3139958.3140035. URL: <http://doi.acm.org/10.1145/3139958.3140035>.
- [7] Krittika D’Silva et al. “Predicting the temporal activity patterns of new venues”. In: *EPJ Data Science* 7.1 (2018), p. 13. ISSN: 2193-1127. DOI: 10.1140/epjds/s13688-018-0142-z. URL: <https://epjdatascience.springeropen.com/articles/10.1140/epjds/s13688-018-0142-z>.
- [8] Krittika D’Silva et al. “The Role of Urban Mobility in Retail Business Survival”. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2.3 (2018), pp. 1–22. DOI: 10.1145/3264910. URL: https://www.krittikadsilva.com/pdfs/DSilva%7B%5C_%7DUBicomp%7B%5C_%7D2018.pdf.
- [9] Marco De Nadai et al. “The Death and Life of Great Italian Cities: A Mobile Phone Data Perspective”. In: *Proceedings of the 25th International Conference on World Wide Web*. WWW ’16. Montrécal, Québec, Canada: International World Wide Web Conferences Steering Committee, 2016, pp. 413–423. ISBN: 978-1-4503-4143-1. DOI: 10.1145/2872427.2883084. URL: <https://doi.org/10.1145/2872427.2883084>.
- [10] Richard M. Emerson. “Power-Dependence Relations”. In: *American Sociological Review* 27.1 (1962), pp. 31–41. ISSN: 00031224. URL: <http://www.jstor.org/stable/2089716>.
- [11] Jie Feng et al. “DeepMove: Predicting Human Mobility with Attentional Recurrent Networks”. In: *Www* 2 (2018), p. 10. DOI: 10.1145/3178876.3186058.
- [12] Linton C Freeman. “Centrality in social networks conceptual clarification”. In: *Social networks* 1.3 (1978), pp. 215–239.
- [13] Tristan Henderson, David Kotz, and Ilya Abyzov. “The Changing Usage of a Mature Campus-wide Wireless Network”. In: *Proceedings of the 10th Annual International Conference on Mobile Computing and Networking*. MobiCom ’04. Philadelphia, PA, USA: ACM, 2004, pp. 187–201. ISBN: 1-58113-868-7. DOI: 10.1145/1023720.1023739. URL: <http://doi.acm.org/10.1145/1023720.1023739>.
- [14] Ramon Hermoso, Jürgen Dunkel, and Florian Rückauf. “Should I Stay or Should I Go?: Exploiting Visitor Movements to Derive Individualized Recommendations in Museums”. In: *Proceedings of the 21st ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*. MSWIM ’18. Montreal, QC, Canada: ACM, 2018, pp. 343–346. ISBN: 978-1-4503-5960-3. DOI: 10.1145/3242102.3242140. URL: <http://doi.acm.org/10.1145/3242102.3242140>.

- [15] Sibren Isaacman et al. “Identifying Important Places in People’s Lives from Cellular Network Data”. In: *Proceedings of the 9th International Conference on Pervasive Computing*. Pervasive’11. San Francisco, USA: Springer-Verlag, 2011, pp. 133–151. ISBN: 978-3-642-21725-8. URL: <http://dl.acm.org/citation.cfm?id=2021975.2021988>.
- [16] William R Knight. “A computer method for calculating Kendall’s tau with ungrouped data”. In: *Journal of the American Statistical Association* 61.314 (1966), pp. 436–439.
- [17] David Kotz and Kobby Essien. “Analysis of a Campus-wide Wireless Network”. In: *Wireless Networks* 11.1–2 (Jan. 2005), pp. 115–133. DOI: 10.1007/s11276-004-4750-0. URL: <http://www.cs.dartmouth.edu/~dfk/papers/kotz-jcampus.pdf>.
- [18] David Kotz et al. *CRAWDAD dataset dartmouth/campus (v. 2004-11-09)*. Downloaded from <https://crawdad.org/dartmouth/campus/20041109>. Nov. 2004. DOI: 10.15783/C71593.
- [19] John F. Nash. “The Bargaining Problem”. In: *Econometrica* 18.2 (1950), pp. 155–162. ISSN: 00129682, 14680262. URL: <http://www.jstor.org/stable/1907266>.
- [20] Marco Pavan et al. “Finding Important Locations: A Feature-Based Approach”. In: *Proceedings - IEEE International Conference on Mobile Data Management* 1 (2015), pp. 110–115. ISSN: 15516245. DOI: 10.1109/MDM.2015.11.
- [21] Sancheng Peng et al. “Social influence modeling using information theory in mobile social networks”. In: *Information Sciences* 379 (2017), pp. 146–159. ISSN: 00200255. DOI: 10.1016/j.ins.2016.08.023. URL: <http://dx.doi.org/10.1016/j.ins.2016.08.023>.
- [22] Sharon C Rochford. “Symmetrically pairwise-bargained allocations in an assignment market”. In: *Journal of Economic Theory* 34.2 (1984), pp. 262–281.
- [23] Adam Sadilek, Henry Kautz, and Jeffrey P. Bigham. “Finding Your Friends and Following Them to Where You Are”. In: *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining*. WSDM ’12. Seattle, Washington, USA: ACM, 2012, pp. 723–732. ISBN: 978-1-4503-0747-5. DOI: 10.1145/2124295.2124380. URL: <http://doi.acm.org/10.1145/2124295.2124380>.
- [24] Thiago H. Silva et al. “Urban Computing Leveraging Location-Based Social Network Data: A Survey”. In: *ACM Comput. Surv.* 52.1 (Feb. 2019), 17:1–17:39. ISSN: 0360-0300. DOI: 10.1145/3301284. URL: <http://doi.acm.org/10.1145/3301284>.
- [25] Lucas M. Silveira et al. “MobHet: Predicting human mobility using heterogeneous data sources”. In: *Computer Communications* 95 (2016), pp. 54–68. ISSN: 01403664. DOI: 10.1016/j.comcom.2016.04.013. URL: <http://dx.doi.org/10.1016/j.comcom.2016.04.013>.

- [26] Lei Wang et al. “On the brink: Predicting business failure with mobile location-based checkins”. In: *Decision Support Systems* 76 (2015), pp. 3–13. ISSN: 01679236. DOI: 10.1016/j.dss.2015.04.010. URL: <http://dx.doi.org/10.1016/j.dss.2015.04.010>.
- [27] Mao Ye et al. “Exploiting Geographical Influence for Collaborative Point-of-interest Recommendation”. In: *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR ’11. Beijing, China: ACM, 2011, pp. 325–334. ISBN: 978-1-4503-0757-4. DOI: 10.1145/2009916.2009962. URL: <http://doi.acm.org/10.1145/2009916.2009962>.
- [28] Jia-Dong Zhang, Chi-Yin Chow, and Yanhua Li. “LORE: Exploiting Sequential Influence for Location Recommendations”. In: *Proceedings of the 22Nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. SIGSPATIAL ’14. Dallas, Texas: ACM, 2014, pp. 103–112. ISBN: 978-1-4503-3131-9. DOI: 10.1145/2666310.2666400. URL: <http://doi.acm.org/10.1145/2666310.2666400>.