An intelligent tree planning approach using location-based social networks data

Staalduinen, J.H. van; Tetteroo, J.; Gawehns, D.; Baratchi, M.; Cao, L.; Kosters, W.A.; ... ; Takes, F.W.

An Intelligent Tree Planning Approach Using Location-Based Social Networks Data

Jan H. van Staalduinen(✉), Jaco Tetteroo, Daniela Gawehns, and Mitra Baratchi

Leiden Institute of Advanced Computer Science (LIACS), Leiden University, Leiden, The Netherlands
{j.h.van.staalduinen,j.tetteroo}@umail.leidenuniv.nl,
{d.gawehns,m.baratchi}@liacs.leidenuniv.nl

Abstract. How do we make sure that all citizens in a city have access to enough green space? An increasing part of the world's population lives in urban areas, where contact with nature is largely reduced to street trees and parks. As optional tree planting sites and financial resources are limited, determining the best planting site can be formulated as an optimization problem with constraints. Can we locate these sites based on the popularity of nearby venues? How can we ensure that we include groups of people who tend to spend time in tree deprived areas?

Currently, tree location sites are chosen based on criteria from spatial-visual, physical and biological, and functional categories. As these criteria do not give any insights into which citizens are benefiting from the tree placement, we propose new data-driven tree planting policies that take socio-cultural aspects as represented by the citizens' behavior into account. We combine a Location Based Social Network (LBSN) mobility data set with tree location data sets, both of New York City and Paris, as a case study. The effect of four different policies is evaluated on simulated movement data and assessed on the average, overall exposure to trees as well as on how much inequality in tree exposure is mitigated.

Keywords: Urban computing · Tree planning · Social network analysis · Community detection algorithms · Mobility data · Multi-objective optimization

1 Introduction

As of 2018, 55% of the world's population lives in urban areas, a number which is projected to grow to 68% by 2050.[1] The North-American continent stands

[1] United Nations Department of Economic and Social Affairs, World Urbanization Prospects 2018, https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf, p. xix, last visited 7 December 2020.

J. H. van Staalduinen and J. Tetteroo—Contributed equally to this project.

M. Baratchi et al. (Eds.): BNAIC/Benelearn 2020, CCIS 1398, pp. 157–171, 2021.
https://doi.org/10.1007/978-3-030-76640-5_10

Fig. 1. We describe four policies that combine data from multiple sources and produce a ranking of potential tree planting sites.

out in particular, where this number is already at 82%. While it is easy to point out the economical reasons for moving to the city – at least at first sight [5] – there are certainly downsides attached to urban life. One of them is the inescapable fact that cities, by definition [4], have a higher population density, leading to more built-up areas and thus a scarcer supply of nature than in rural areas. However, as Rohde and Kendle put it, "it is obvious from any casual observation that many human beings do not like to be dissociated from the natural world; as a nation we spend millions of pounds every year on garden and household plants" [15]. Indeed, contact with nature does seem to be linked to human well-being and positive emotional effects and is even said to strengthen urban communities [9,13]. Apart from socio-cultural benefits, urban greenery can help to mitigate two characteristically urban problems: air pollution due to traffic [10] and (extreme) warmth due to the urban heat island effect [12]. The inclusion of parks and street trees in city landscapes is, therefore, an important aspect of the urban planning process.

To date, socio-cultural arguments play a marginal if not non-existent role in formal frameworks describing criteria for selecting potential tree planting sites. They do not account for the amount of people that are accommodated by the newly planted trees. Following the established criteria, trees may end up in places where they are beneficial to some people, but its effects may not serve the majority of people, or may never reach the people yearning for them most.

We propose taking a data-driven approach based on available mobility data which allows considering additional tree planning criteria. Popular adoption of Location-Based Social Network (LBSN) applications has allowed the collection of valuable data representing the movement of people between venues. We identify policies that take people's movement into account when choosing potential tree planting sites. These are based on (1) site popularity, (2) existing tree density at potential planting sites, (3) existing tree density at other sites that are visited often by the same people and (4) a multi-objective combination of (1) and (3). Each of the policies takes another aspect of the data into account and provides a ranking for the potential planting sites, as schematically shown in Fig. 1. This ranking can be embedded within the criteria of established tree planning frameworks that currently lack this socio-cultural value and insight.

Our paper makes the following contributions:[2]

- We describe novel data-driven criteria for potential tree planting site selection based on information on people's movement from a venue interaction network;
- We analyse the impact of these policies in a way that uncovers inequalities between groups of citizens and shows which policies decrease these inequalities;
- We apply this method to rank venues as potential tree planting sites in New York City and Paris.

This paper is organized as follows. Section 2 presents the related work. We give the problem definition in Sect. 3, for which we present, as potential solutions, the tree-planning policies in Sect. 4. In Sect. 5 we describe our experiments, applying our methods to New York City and Paris. The results are discussed in Sect. 6. Finally, Sect. 7 presents our concluding remarks.

2 Related Work

Most of the work in the field of tree planning revolves around selecting appropriate tree species for predetermined planting sites [17,18]. This reflects the observations by Spellerberg and Given [18] and Pauleit [14] that tree planning is often an afterthought in the urban design process and characterised by pragmatism. While the visual aesthetic of trees and socio-cultural function of green spaces in the city seem to be important motives for planting trees, the first motive seems to play only a small role in the tree planning process [16] and the second motive is not reflected in the sparse body of site selection criteria that we could find.

The work by Amir and Misgav [1], in which they aim to describe a complete tree planning decision framework, does incorporate criteria on site selection. They define three useful categories of criteria, which are *spatial-visual*, *physical and biological* and *functional*. Criteria relating to the *socio-cultural* function of green spaces, however, are missing. We observed several works describing site selection criteria [7,14], but those fall within the category of *physical and biological* criteria that are essential for the survival of the tree. Moriani et al. [10] did use population density in a planting priority index, but as they focused on the air pollution-reducing quality of trees, this still falls within the category of *functional* criteria. We believe then, that the body of site selection criteria is still incomplete and that we can contribute to this framework by introducing new socio-cultural criteria that take people's movement into account.

As a way to capture the general movement patterns of people within cities, we utilize data collected by LBSNs. As defined by Zheng and Zhou [21], social networks are social structures that consist of individuals connected to each other via specific types of interdependencies. In LBSNs these individuals are connected

[2] This work earlier participated and was selected for the Future Cities Challenge co-organised by Foursquare at NetMob 2019. The work has not been published elsewhere.

through their shared experience, interacting with the locations in the network. Oftentimes, in LBSNs users announce their visit to venues through a so-called check-in option. The check-in data can provide information about the movement of people between a network of venues. The structure of such a network can be explored to find underlying patterns. For instance, locations can be grouped based on the similarity between user profiles [8]. Hung et al. [6] use these user profile similarities to find user communities. Girvan and Newman [3], however, use clustering algorithms on the full network to detect communities, eliminating the need for individual trajectories.

Most of these approaches have considered studying the network properties of LBSN data without considering how such information can be used in improving urban aspects. Recently, Arp et al. [2] have shown how such data can be used in optimising the state of traffic within the city. A recent trend in which private companies make their data available through various "Data for Good" programs helps to advance research in the field. In this paper, we aim to study whether such data can be used for improving decision making regarding the optimal allocation of resources, in this case trees, throughout the city.

3 Problem Definition

To find solutions to the planting site selection problem, we combine urban data consisting of venue locations and movements between them with tree location data. Given the undirected network graph $G = (V, E, W, T)$, where nodes $v \in V$ represent venues and weighted edges $e = (v_1, v_2), \quad e \in E$ represent movements of people between a pair of venues v_1 and v_2, with weight $w_e \in W$ denoting the number of movements between the pair of venues, as well as the tree density $td_i \in T$ value for each node v_i, each policy creates a ranking of planting sites; the goal is to find a policy that satisfies a certain objective. The objectives we selected are (1) the best absolute increase in number of tree encounters among citizens and (2) the largest decrease in inequality in allocation of resources for citizens. These are further explained in Sect. 5.4.

4 Methods

In this section, we describe the four different planting site selection policies: degree (Sect. 4.1), tree density (Sect. 4.2), community tree density (Sect. 4.3) and a combination in the form of a Pareto ranking (Sect. 4.4).

4.1 Policy 1 – Degree

A first possible approach to maximize the impact of planting a tree, is to plant it near a place where many people pass by. From this perspective, the goal is to find the venues that are maximally popular among visitors. To find these locations we maximize the degree of all nodes v in the undirected network graph, defined as the sum of the weights of the edges that are connected to it.

4.2 Policy 2 – Tree Density

A second possible approach to maximize the impact of planting a tree, is to identify locations that are visited by people who do not regularly come across trees. In this case, the number of people who will come across the newly planted tree may be lower than in the previous case, where the location would be frequented by many people. Nevertheless, the people who do encounter the tree may gain more from the encounter because of their lack of earlier encounters. For this policy, we find locations by minimizing the sum of trees in the direct vicinity of venues, which we define as a radius of 25 m around each venue. We call this sum of trees the tree density td_i of a venue v_i.

4.3 Policy 3 – Community Tree Density

If we just prioritize venues with few trees in their immediate vicinity, we would discard the reality that people move about and that people are thus prone to visit multiple venues. A single venue that has few trees in its vicinity might not be a major problem if the usual crowd for this venue also regularly visits other venues that do have more trees in the neighbourhood.

Using LBSNs, we can use this observation in our objective. To this end, we introduce a measure we call the *community tree density*. This measure intends to highlight *groups* of *related venues* that have a low tree density, instead of *single* venues that have a low tree density. A relation between venues, in this sense, is determined by people travelling often between those venues. Using this policy, we aim to minimize the community tree density.

Using graph theory parlance, these related venues can be discovered through the task called *community detection*. A community is a group of nodes of which the nodes are densely connected with each other, but much less with the rest of the network [3].

To detect the communities, we use the Leiden community detection algorithm [20]: a fast algorithm that is able to find communities with high quality. It optimises modularity, a measure that compares the density of connections within a community with the density between communities [11].

As it is computationally heavy to compute the modularity of a community, the algorithm uses heuristics to approximate it. Therefore, it does not necessarily return the best community layout. To gain confidence in the robustness of our communities, we run the algorithms N times to find different community partitions. In Sect. 5.2, we find $N = 50$ to be reasonable.

We compute the community tree density ctd_i^n in community detection iteration n for a venue v_i by averaging the td_j – which is computed as in Sect. 4.2 – of all venues v_j that are in the same community C_k^n as venue v_i. As we run the community detection algorithm N times we obtain the overall community tree density ctd^i of a venue v_i by averaging over each of its computed community tree densities, as shown in Eq. 1.

$$ctd_i = \frac{1}{N}\sum_{n=1}^{N}\Big(\frac{1}{|C_k^n|}\sum_{v_j \in C_k^n} td_j\Big), \quad v_i \in C_k^n \tag{1}$$

4.4 Policy 4 – Pareto Ranking

The policies discussed above, venue degree and community tree density, could both be important in discovering the most suitable location(s) for one or more new trees. Indeed, a venue with a low tree density coefficient could have only one visitor, whereas other venues in the same community that have a similarly low tree density coefficient could have many visitors. In this case, the latter venue(s) would be more appropriate as a tree planting site. It is therefore important to take both objectives into account. To achieve this, we borrow a method from multi-objective optimization theory, the Pareto front [19].

We combine the venue degrees, i.e., the popularity of venues, with community-based tree density coefficients by detecting the set of venues that are Pareto efficient, i.e., the venues that are found by minimizing the tree density coefficient and maximizing the influence of the venue: the optimal trade-offs between the two measures. For these venues it is impossible to improve for one objective, without impairing the other objective. Also called the Pareto front, the venues in this set could meet our criterion of helping most people needing trees. To rank all tree planting sites on both objectives, we first compute the Pareto front and assign the appropriate rank to the locations in this set, and then remove the Pareto front from the set of locations. In this manner we iteratively compute Pareto fronts, rank venues on this front and then remove these from the set until all sites are ranked.

5 Experimental Set-Up

In this section, we describe the experimental set-up we will use to compare the tree planting policies for two cities, New York City and Paris. This section is structured as follows. In Sect. 5.1, we list the properties of our data sets. In Sect. 5.2, we conduct an experiment to find a suitable hyperparameter setting for the community detection algorithm. In Sect. 5.3, we describe how we simulate the movement of citizens through the city, since we don't have precise trajectory information. Finally, in Sect. 5.4, we explain what we want to measure with the experiments and how we evaluate the results.

5.1 Data Sources

Two Case Studies. We conducted two case studies to investigate the implementation and workings of our criteria using real data. For this, we chose to focus on New York City and Paris, as for both cities data sets describing venue interactions and tree locations were available, which are both needed for computing the rankings according to the policies. These data sets are described below.

Venue Interaction Data. Foursquare City Guide is a mobile app that recommends places to its users based on their likes or check-ins. The venue interaction data set provided by Foursquare comprises of two parts: venues and movements between them. Venues in this set are locations people can visit. Venue coordinates are recorded, as well as their name and a category. Movements are recorded when individuals make consecutive check-ins at different locations.

The data set contains check-in information from between April 2017 and March 2019 of ten different cities around the world, from which we picked Paris and New York City as examples. Note that other mobility data can be used to replace or augment the Foursquare data where available (e.g. traffic data or WiFi scans), as long as we know which venues are connected by people's movement.

Table 1. Description of venue interaction data set (Foursquare).

	New York City		Paris	
	Original	After pre-processing	Original	After pre-processing
# venues (nodes)	17,975	15,610	7,133	6,291
# interactions (edges)	7,920,000 (directed, parallel)	246,605 (undirected)	7,920,000 (directed, parallel)	182,187 (undirected)

We pre-processed this data set by creating a network where nodes represent venues and edges represent movements, and removing small unconnected 'islands' of up to 3 nodes that were not connected to the large *connected component*. We also removed the venues for which no location information was known. Finally, we flattened bidirectional edges into a single undirected edge for which the edge weight denotes the summed number of interactions between two given venues. Table 1 shows the number of nodes and edges before and after pre-processing.

Tree Location Data: the tree census data set of New York[3] contains information on street trees in New York City and surrounding cities. It contains information on among others the *species* and *health* of the trees, as well as their *longitude* and *latitude*. As only street trees were counted, trees in parks were not taken into account in the tree survey and are therefore not present in the data set.

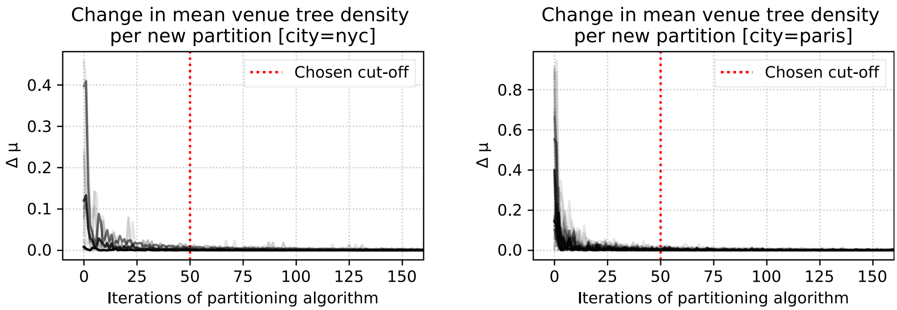
The Parisian tree census data set[4] contains similar information on its trees, most notably the locations using *longitude* and *latitude*. It should be noted that for the Parisian tree data set only trees in the city center were recorded.

[3] 'TreesCount! 2015 Street Tree Census', data set provided by the NYC Department of Parks & Recreation, https://data.cityofnewyork.us/Environment/2015-Street-Tree-Census-Tree-Data/uvpi-gqnh, last visited 7 December 2020.

[4] 'Les arbres', data set provided by the *Direction des Espaces Verts et de l'Environnement* of the city of Paris, https://opendata.paris.fr/explore/dataset/les-arbres/, last visited 7 December 2020.

5.2 The Number of Iterations of the Community Detection Algorithm

As described in Sect. 4.3, to obtain a stable value for the community tree density for each venue, we run the community detection for N iterations and for each venue compute the mean community tree density over these iterations. To find a proper value for N, we tracked how much the computed mean venue tree density value per venue changed after each iteration of the community detection algorithm. We show this change as the difference between two consecutive mean values ($\Delta\mu$) for each of the venues in Fig. 2. We observe that after 50 iterations, the mean value for each venue is approximately stable.



(a) New York City (b) Paris

Fig. 2. The mean values stabilise with more iterations of the algorithm.

5.3 Simulation

We want to compare the effect of different policies on individual citizens' exposure to trees. The individualized trajectory data needed for such a comparison was not available from the real life data set for privacy reasons. We therefore simulated trajectories by generating random walks over the movement graph extracted from our LBSN. Each random walk represents one citizen, visiting five venues in their city on one day. For each venue v_i in the random walk, the next venue $u_i \in \text{adj}(v_i)$ was randomly chosen from its neighbouring nodes, where the probability of a node to be chosen corresponds to the weight of the connecting edge. For each of the cities, we sampled 1% of the population size, resulting in 85,510 and 21,483 simulated trajectories for respectively New York and Paris.

5.4 Evaluation

For the evaluation of our framework we have the following goal: we want to show the different ways in which each planting policy improves the city. We measure this improvement in two ways. First, we want to investigate which policy is best at increasing the overall number of tree encounters for all citizens. Second, we want to investigate which policies are best for targeting specifically the trajectories that are lacking most in tree encounters and are therefore more suitable to decrease the level of inequality among citizens in this regard.

The evaluation should thus provide answers to the following questions:

1. Overall, which of the proposed policies increase(s) the number of tree encounters by the citizens the most?
2. Which of the policies is/are best suited for removing inequalities in the number of tree encounters between citizens?

Ideally, we would have compared our policies with existing tree planting policies as baselines, but these are not (well) described yet. We therefore compare the performance of our different proposed policies and consider a random assignment as a baseline.

6 Results

In this section, we describe the results of our experiments, in which we apply the four different tree planting policies to the two cities and evaluate them using simulated random walks. This section is structured as follows. In Sect. 6.1, we describe the situation in each city before we plant any new trees. In Sect. 6.2, we show the distribution of the values the policies use to rank the venues. In Sect. 6.3, we use the policies to plant new trees and analyse the result.

6.1 Initial Situation

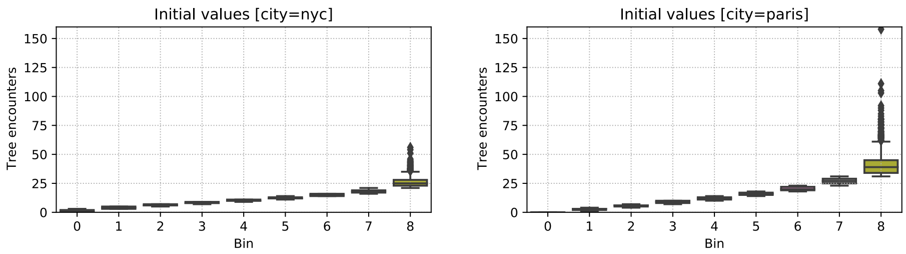
To define the initial situation, we counted the number of trees encountered along each trajectory, simulated by a random walk, and grouped these trajectories into nine bins of equal size, ordered in ascending order according to the number of tree encounters, in order to be able to compare intuitively with the new situation later (see Fig. 3). The two cities have a similar distribution of tree encounters.

6.2 Ranking the Venues

By applying our framework to the data, we generated four rankings: one for each planting policy as discussed in Sect. 4. We show the distribution of the values of each planting site in Fig. 4 for New York City and Fig. 5 for Paris.

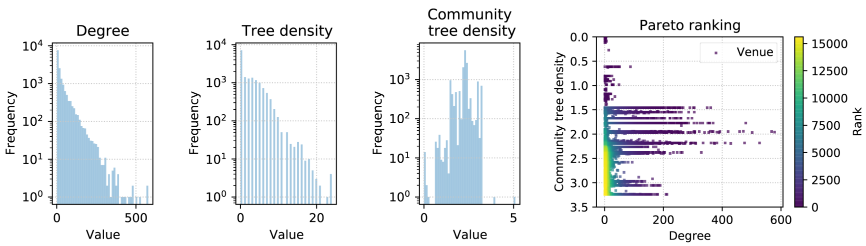
Figures 4a and 5a show the degree distribution. For both cities, this follows a power law, where most venues have a low degree, and only some venues have a high degree. When using this policy, venues with a high degree are chosen as desirable planting site. Figures 4b and 5b show the distribution of trees. They are similarly distributed, following a power law.

In Figs. 4c and 5c we show the community tree density. For both cities, we see some venues where the communities are especially tree-sparse, but most venues have on average a small number of trees in their community. When using either the tree density policy or the community tree density policy, locations with a lower density are given priority as desirable planting sites.



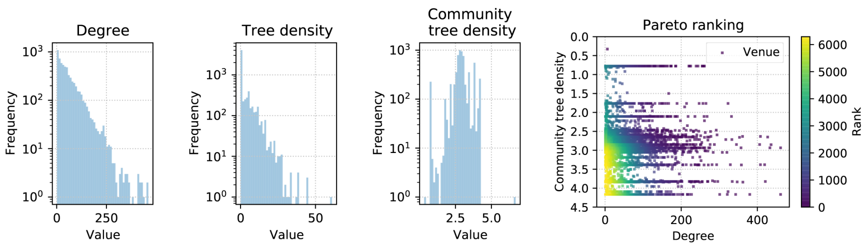
(a) New York City (b) Paris

Fig. 3. The distribution of tree encounters per random walk in the initial situation grouped in ordered bins of ascending numbers of tree encounters.



(a) (b) (c) (d)

Fig. 4. The distribution of values for site selection policies (city: NYC).



(a) (b) (c) (d)

Fig. 5. The distribution of values for the site selection policies (city: Paris).

Finally, Figs. 4d and 5d show the Pareto rankings of the cities. Here we set the community tree density objective against the degree objective. As the degree is maximized and the community tree density minimized, priority is given to those venues that are closest to the top right corner.

6.3 Planting the Trees and Analyzing the Result

After defining the initial situation, for each city, we selected the 10% most suitable locations according to each of the planting policies. This amounted to 1,561 locations for New York and 629 locations for Paris. We then analysed the results as follows.

Table 2. Increase of tree encounters along random walks.

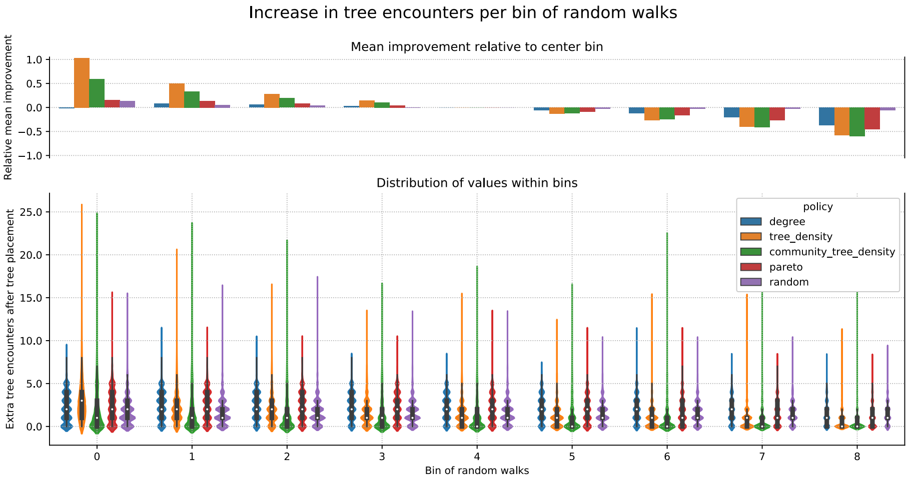
	New York City		Paris	
	Mean	Std	Mean	Std
Degree	**2.124**	1.591	2.220	1.288
Tree density	1.617	1.841	**2.972**	1.376
Community tree density	1.242	2.003	2.775	1.547
Pareto	1.991	1.685	2.471	1.329
Random	1.632	1.394	2.930	1.421

Q1 – Best Overall Performance. The mean increase of tree encounters per random walk, re-counted after applying the policies, is shown in Table 2. The standard deviation of the increase is also relevant for comparison of the results. While a small standard deviation shows the improvements were reached over the entire range of trajectories, a higher standard deviation shows a focus towards a subgroup of the trajectories. Both could be desirable.

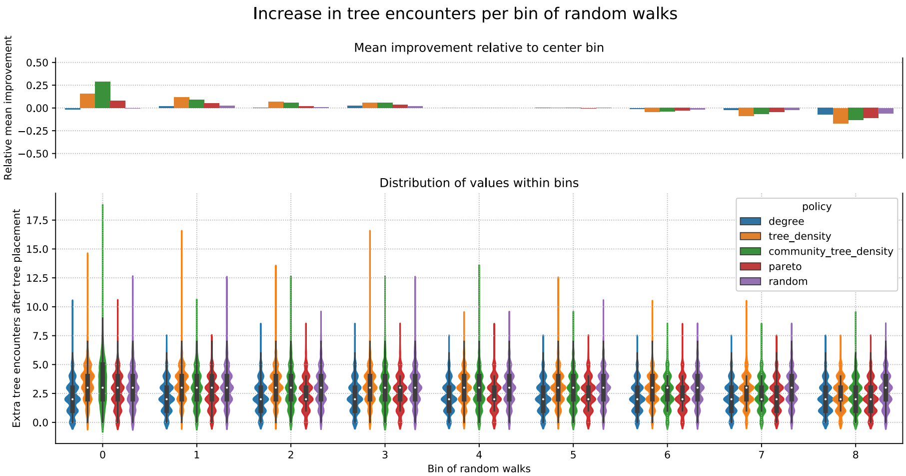
We see that there are differences between the two cities.

For New York, the degree policy on average increased the number of tree encounters most and is thus best suited for increasing the overall number of tree encounters. The tree density policy and especially the community tree density policy had a high standard deviation, indicating that while on average they did not increase the tree encounters as much as the degree policy, they did target specific trajectories more than others. The Pareto policy achieved a mean and standard deviation in between that of the degree and community tree density policies, as expected as it is comprised of both policies.

For Paris the degree policy increased the number of tree encounters least. In this city the tree density policy outperformed all others and is thus best suited for increasing the overall number of tree encounters. Again, the community tree density policy yielded the highest standard deviation, indicating site selection near specific trajectories. The Pareto policy is outperformed by the random baseline, due to its dependency on the degree policy.



(a) New York City



(b) Paris

Fig. 6. Mean increase of tree encounters on sorted random walks. From left to right the bins hold trajectories that were increasingly tree-dense before site selection and tree planting (see Fig. 3). The top figures show the mean values of the violin plots underneath, but present them relative to the center bin.

To answer our first question from Sect. 5.4, then, we have to make a distinction between the two cities. For New York City, the degree policy seems to be the best at increasing the number of tree encounters in general for New York. It might however be prone to a green-get-greener phenomenon, which means that established venues may be solidified in the new situation. For Paris, the tree density policy has the best overall increase in tree encounters.

Q2 – Best Policy for Targeting Inequality. To answer the second question in Sect. 5.4, we need a more detailed analysis. This is shown in Fig. 6, where the random walks are grouped into the same ordered bins as in Fig. 3.

For both cities, the figures at the bottom of Fig. 6 show the distribution of the mean increase of tree encounters per walk of each bin. The top figures show the same values, but present them relative to the center bin. This allows us to easily spot whether a policy preferences sites in tree-sparse over tree-dense trajectories, which will create an equalizing effect.

When we look at the results for both New York (Fig. 6a) and Paris (Fig. 6b), we see that the tree density and community tree density policies consistently have the biggest difference between the center bin, left-most bins and right-most bins and therefore have the most focus towards creating equality. In New York, the tree density policy performs considerably better than the community tree density policy, as the violin of the community tree density policy is quite wide at the bottom and quickly grows slim. This effect is less visible for Paris, where both policies tend to perform equally well, but the community tree density does have the edge in the left-most bin.

As was also the case when evaluating which policy performed best under Q1 above, the performance of the Pareto policy tends to rank between the degree policy and both tree density policies. This could mean that in certain use cases, this policy could prove to be a valuable compromise between the policies regarding both objectives outlined in Sect. 5.4.

To answer the second question, then, the tree density and community tree density policies have, for both cities, the most 'equalizing' effect, because their improvements are targeted specifically at the tree-sparse trajectories.

7 Conclusion

In this paper, we propose novel criteria that can be used when selecting potential tree planting sites. The nature of these criteria is socio-cultural, capturing people movement between venues. Having implemented them as policies for a case study on New York City and Paris, we show that they are applicable in the field and can be used to support decision-makers by providing them with the ranking policy most appropriate for their goals.

From our experiments, we observe that there is no single policy that outperforms all others. Depending on the goal of urban planners, one may select the degree policy to increase the average tree encounters, or the community tree density policy to target sites in tree-sparse trajectories. When faced with the challenge of selecting tree planting sites, there are policy choices to be made, and it is important to analyse the situation on a detailed level.

We want to note that while tree survey data was available for the case studies, these data sets were far from perfect. City planners who have the intention of adopting a data-driven method will not only have to decide which site-selection approach to use, but also make sure that the underlying data is sufficiently complete and high in quality. On a similar note, we want to mention that Foursquare data represents only a fraction of citizen's movement, possibly not representing all groups of citizens. Other mobility data such as traffic data or WiFi scans can be used to paint a more representative picture.

We conclude that the newly introduced data-driven socio-cultural approach to finding a tree planting site that benefits different communities of city dwellers is feasible and can easily be implemented by urban planning organizations. Integration of this approach depends on the availability of detailed records of existing trees and movement data of city inhabitants.

In the future, we think it would be quite possible to extend this work towards other site selection applications, such as communal waste bins. There is however also more work to be done in verifying the results, both by extending it to other cities, and also by implementing it with different data sources to enrich the analysis.

Acknowledgments. We thank Foursquare and the organisation of Netmob for organising the Future Cities Challenge and providing us with access to Foursquare's data set.

References

1. Amir, S., Misgav, A.: A framework for street tree planning in urban areas in Israel. Landscape Urban Plan. **19**(3), 203–212 (1990)
2. Arp, L., van Vreumingen, D., Gawehns, D., Baratchi, M.: Dynamic macro scale traffic flow optimisation using crowd-sourced urban movement data. In: 2020 21st IEEE International Conference on Mobile Data Management (MDM), pp. 168–177 (2020)
3. Girvan, M., Newman, M.E.: Community structure in social and biological networks. Proc. Natl. Acad. Sci. **99**(12), 7821–7826 (2002)
4. Hall, S.A., Kaufman, J.S., Ricketts, T.C.: Defining urban and rural areas in us epidemiologic studies. J. Urban Health **83**(2), 162–175 (2006)
5. Harris, J.R., Todaro, M.P.: Migration, unemployment and development: a two-sector analysis. Am. Econ. Rev. **60**(1), 126–142 (1970)
6. Hung, C.C., Chang, C.W., Peng, W.C.: Mining trajectory profiles for discovering user communities. In: Proceedings of the 2009 International Workshop on Location Based Social Networks, pp. 1–8. LBSN 2009 (2009). https://doi.org/10.1145/1629890.1629892
7. Jim, C.Y.: A planning strategy to augment the diversity and biomass of roadside trees in urban Hong Kong. Landscape Urban Plan. **44**(1), 13–32 (1999)
8. Li, Q., Zheng, Y., Xie, X., Chen, Y., Liu, W., Ma, W.Y.: Mining user similarity based on location history. In: Proceedings of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. GIS 2008 (2008). https://doi.org/10.1145/1463434.1463477
9. Maller, C., Townsend, M., Pryor, A., Brown, P., St Leger, L.: Healthy nature healthy people: 'contact with nature' as an upstream health promotion intervention for populations. Health Promotion Int. **21**(1), 45–54 (2006)

10. Morani, A., Nowak, D.J., Hirabayashi, S., Calfapietra, C.: How to select the best tree planting locations to enhance air pollution removal in the MillionTreesNYC initiative. Environ. Pollut. **159**(5), 1040–1047 (2011)
11. Newman, M.E.J.: Analysis of weighted networks. Phys. Rev. E **705**, 056131 (2004). https://doi.org/10.1103/PhysRevE.70.056131
12. Norton, B.A., Coutts, A.M., Livesley, S.J., Harris, R.J., Hunter, A.M., Williams, N.S.: Planning for cooler cities: a framework to prioritise green infrastructure to mitigate high temperatures in urban landscapes. Landscape Urban Plan. **134**, 127–138 (2015)
13. Parry-Jones, W.L.: Natural landscape, psychological well-being and mental health. Landscape Res. **15**(2), 7–11 (1990)
14. Pauleit, S.: Urban street tree plantings: identifying the key requirements. In: Proceedings of the Institution of Civil Engineers-Municipal Engineer, vol. 156, pp. 43–50. Thomas Telford Ltd (2003)
15. Rohde, C.L.E., Kendle, A.D.: Human Well-being, natural landscapes and wildlife in urban areas: a review. English Nat. Sci. (1994)
16. Roy, S., Davison, A., Östberg, J.: Pragmatic factors outweigh ecosystem service goals in street tree selection and planting in South-East Queensland cities. Urban Forest. Urban Greening **21**, 166–174 (2017)
17. Sjöman, H., Nielsen, A.B.: Selecting trees for urban paved sites in Scandinavia. Urban Forest. Urban Greening **9**(4), 281–293 (2010)
18. Spellerberg, I.F., Given, D.R.: Trees in urban and city environments: a review of the selection criteria with particular reference to nature conservation in New Zealand cities. Landscape Rev. **12**(2), 19–31 (2008)
19. Stadler, W.: A survey of multicriteria optimization or the vector maximum problem, part I: 1776–1960. J. Optimization Theory Appl. **29**(1), 1–52 (1979)
20. Traag, V., Waltman, L., van Eck, N.J.: From Louvain to Leiden: guaranteeing well-connected communities. Sci. Rep. **9**, 5233 (2019). https://doi.org/10.1038/s41598-019-41695-z
21. Zheng, Y., Zhou, X.: Computing with Spatial Trajectories. Springer Publishing Company, Incorporated (2011)