I. Introduction

This study aims to explore the urban planning process of a kindergarten location through different methods of accessibility estimation. Convenient pedestrian access and minimal car traffic are the main priorities of this accessibility exercise, but determining these is a difficult problem with many potential solution methods. While geographic data can be a useful tool for solving these issues, the potential drawbacks of this data must also be considered when designing solutions. This study considers two alternative methods of public amenity plan through mobility data and street network data, comparing their offered solutions and potential issues.

II. Mass Mobility Data Analysis Potential

Mobility data can be very useful for understanding and predicting human movement. Burke (2006) introduced the concept of participatory sensing, arguing that the use of social mobility data can increase personal and collective participation in projects such as urban planning initiatives, increasing data quality and credibility as well as the equity of resulting plans. The development of ML and deep learning, as well as the increased amount of locational data captured about individuals, has broadened the capabilities of data scientists to understand the movements of people through space.

This study uses recorded Gowalla data to plot the mobility patterns of several individuals through Cambridge in 2010. Functions were created to iterate through the Gowalla data, identifying instances where individuals recorded multiple check-in points on the same day and estimating their route with shortest path methods. The function first assumes that their travel mode was by foot and so uses the walking accessibility graph, but updates this estimation to bicycle using the bike graph or to driving using the driving graph depending on the estimated average speed of the user when travelling between the two checkpoints.

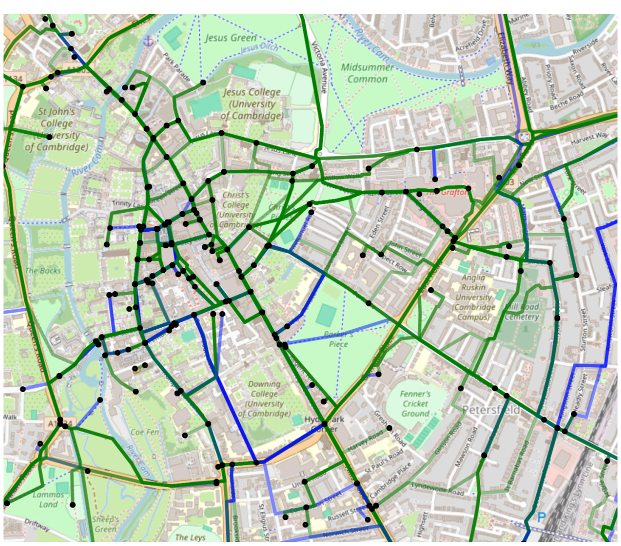


Figure 1: Predicted Walking and Cycling Journeys

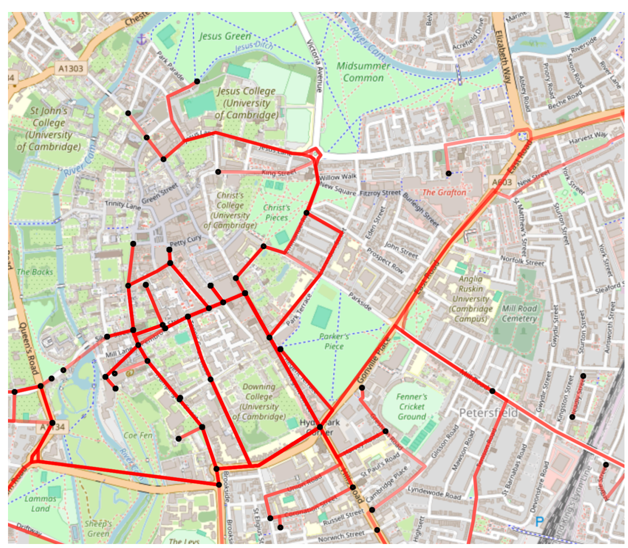
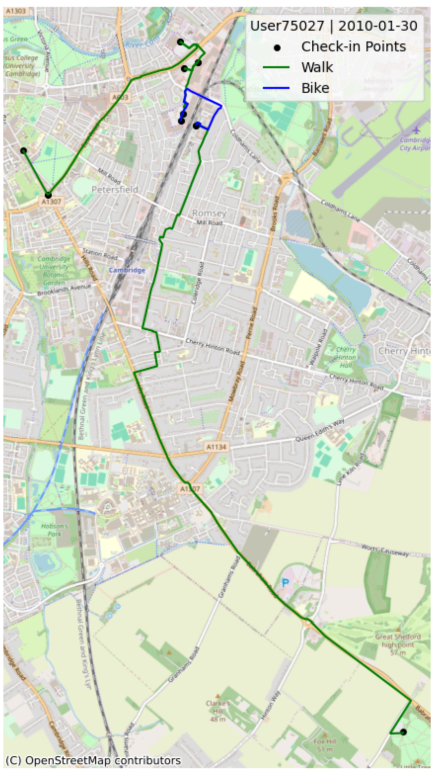


Figure 2: Predicted Car Journeys

From plotting the check-in points and most likely routes of all users, as well as potential transportation methods, this data shows us where the most well traversed regions of Cambridge might be. Placing a kindergarten in the city centre near Trinity Lane would make it conveniently accessible for many people who travel through there, without exposing children to dangerous and polluting car traffic.

However, there are several potential errors that might arise from this decision. Because the data only records check-in locations, significant time lag between check in and departure to the next location could result in misclassification of travel type. For example, user102829's second journey is classified as walking because of their travel speed, but their distance covered alongside a major road - even though it is on the OSM's walking graph - might indicate that this is incorrect. More of the walking and biking journeys might actually be driving journeys, increasing car traffic in areas that this analysis deemed to be pedestrian-friendly. Furthermore, Gowalla's data is limited to those who signed up for and regularly used the app, which is likely a small subset of the population not representative of the larger whole. Relying on too much mobility data for urban planning risks excluding those who don't have constant access to cell phones from accessibility to essential social amenities (Van Dijk, 2020).

III. Mobility Data Privacy Risks



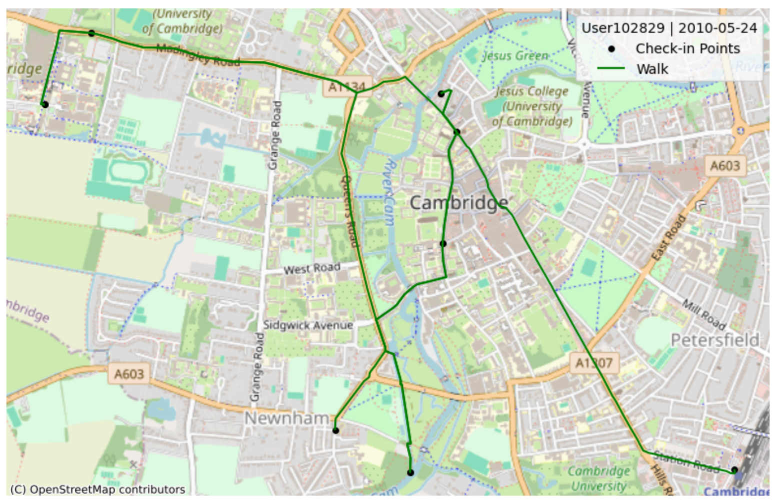


Figure 3: User 75027 Predicted Routes

Figure 4: User 203829 Predicted Routes

User ID	Date	Longest Route (m)	Average Route (m)	Total Route (m)
75027	2010-01-30	7169.50	1521.77	12174.15
102829	2010-05-24	4547.45	1757.71	12303.99

Figure 4: Selected Users' Mobility Summary Table

From this data on users 75027 and 102829 we have learned what stops they made on the selected dates, and their predicted routes and travel modes. Using Google Maps to search up locations, it can be ascertained that User 75027 visited a park in the afternoon, then an apartment, made several stops in a shopping centre including potentially a Costa Coffee, an Asda, a Sports Direct, a Pizza Hut, then a Tesco Superstore late that evening. User 102829 started out by the Cambridge Rail Station around noon, then visited several areas around a university hub, a nature reserve, and central Cambridge college campuses. Through locational knowledge alone, ideas of who the users might be already begin to emerge; perhaps a parent doing the shopping, or an academic attending on-campus duties. By cross referencing mobility datasets with other location-based datasets, researchers have shown that it is not only possible but quite easy to identify anonymous location-logging individuals from 10 check-ins (Rossi & Musolesi, 2014), or by knowing their three top locations (Zang & Bolot, 2011).

This leads mobility data to be useful not only for public good but also for more nefarious purposes. Loss of geographical privacy opens the possibility of customer tracking, where corporations could profile individuals based on their visited locations, classifying them into micro markets in order to manipulate consumer behaviour through personalised marketing; making individuals locatable also opens them to risks from bad actors seeking to scam, blackmail, extort, or repress (Clarke & Wigan, 2011). Truly anonymizing the data to safeguard individual locations requires aggregating away the spatial or temporal granularity that makes it useful for tasks such as urban flow modelling or urban planning. In order to analyse accessibility without risking locational privacy, more methods and data must be explored.

IV. Geographic Analysis Alternatives

Studies researching street networks and urban vitality such as Yue and Zhu (2019) and Lan et al (2022) have found significant correlations between urban morphology and mobility patterns. By transforming urban street maps into graph networks, it is possible to measure how connected locations are to the rest of the city, and in what ways, through centrality measures. This analysis method comes with limitations; planned grid cities are less likely than organic ones to have a strong correlation between street centrality and high foot or vehicle traffic. However, there are also more ways to supplement these methods, such as with increased land use analysis (Maheshwari, Jain & Chopra, 2024) or integrating machine learning methods into graphs with graph neural networks (Lan et al, 2022).

There are several varieties of centrality measures. This study focuses on degree centrality, or the number of connections a street has, and betweenness centrality, which measures number of shortest paths that pass through a street and can to some degree predict traffic flow (). Cambridge's street networks were analysed to find the streets with the highest walking degree centrality and driving betweenness centrality. These measures were chosen on the basis that a kindergarten should be well-connected to pedestrians in order to serve local families, but should also avoid busy roads for safety reasons. A land use base map was added to identify residential areas to target, and to find potential kindergarten locations near green spaces, which have been shown to correlate with positive outcomes (Dadvand et al, 2015).

Figure 5 shows an area in the northern side of town near Orchard Park with large amounts of accessible pedestrian streets in a residential area, near green spaces and further from major roads. The analysis suggests this could be a good location for a kindergarten, but that further in-person study should be conducted to confirm these findings.

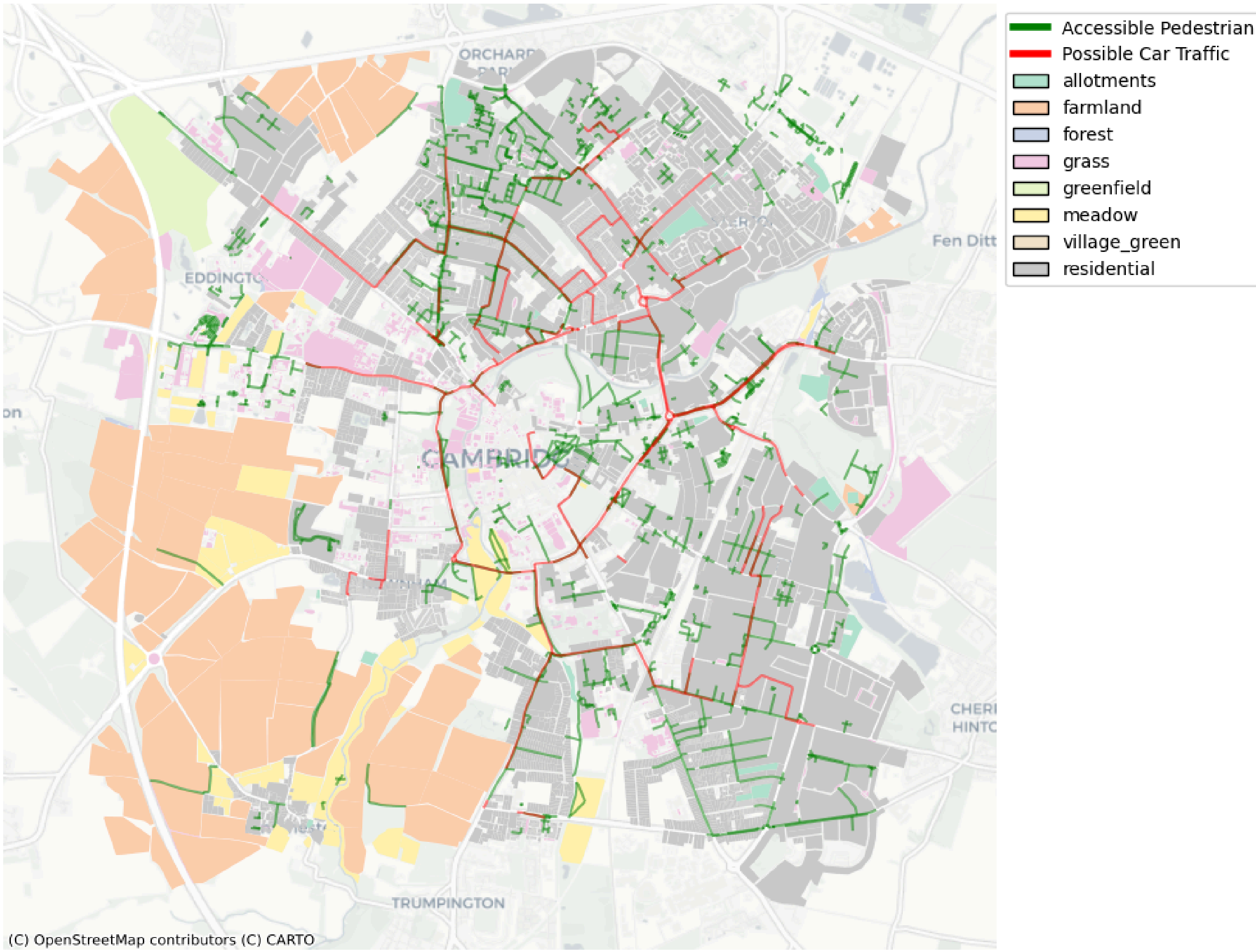


Figure 5: Kindergarten Geographical Factor Map

References:

Burke, J. A., Estrin, D., Hansen, M., Parker, A., Ramanathan, N., Reddy, S., & Srivastava, M. B. (2006). Participatory sensing.

Clarke, R., & Wigan, M. (2011). "You are where you've been: The privacy implications of location and tracking technologies." *Journal of Location Based Services*, *5*(3-4), 138-155.

Dadvand, P., Nieuwenhuijsen, M. J., Esnaola, M., Forns, J., Basagaña, X., Alvarez-Pedrerol, M., ... & Sunyer, J. (2015). Green spaces and cognitive development in primary schoolchildren. *Proceedings of the National Academy of Sciences*, *112*(26), 7937-7942.

Kazerani, A., & Winter, S. (2009, June). Can betweenness centrality explain traffic flow. In 12th AGILE international conference on geographic information science (pp. 1-9). Germany: Leibniz Universität Hannover.

Lan, T., Cheng, H., Wang, Y., & Wen, B. (2022). Site selection via learning graph convolutional neural networks: A case study of Singapore. *Remote Sensing*, *14*(15), 3579.

Maheshwari, A., Jain, Y., & Chopra, S. S. (2024). Spatial inequity index for essential facility access at building-level in highly dense urban areas. *Cities*, *149*, 104926.

Rossi, L., & Musolesi, M. (2014). It's the way you check-in: Identifying users in location-based social networks. In Proceedings of the second ACM conference on Online social networks (pp. 215-226).

Van Dijk, J. (2020). *The digital divide*. John Wiley & Sons.

Yue, H., & Zhu, X. (2019). Exploring the relationship between urban vitality and street centrality based on social network review data in Wuhan, China. *Sustainability*, *11*(16), 4356

Zang, H., & Bolot, J. (2011, September). Anonymization of location data does not work: A large-scale measurement study. In *Proceedings of the 17th annual international conference on Mobile computing and networking* (pp. 145-156).