

都市空間の評価における異種グラフの応用に関する検討 Heterogeneous Graph for Urban Space Evaluation

増田 亮[†] 永吉 勤[†] 小林 悠一[†]
Ryo Masuda Tsutomu Nagayoshi Yuichi Kobayashi

1. はじめに

異種グラフ(Heterogeneous graph)は、複数の種類のノードとエッジで構成される特殊なグラフである。通常のグラフ(Homogeneous graph)が単一のデータ種類の関係性のみを表現するのに対し、異種グラフは複数の種類のデータ間の関係も明示的に構造化する。このため、異種グラフを用いることで、通常のグラフでは難しい、任意のデータ間の関係を柔軟に記述することができる。多くの実社会の現象は、異なる要素が相互に関連し、複雑なネットワークを形成するため、異種グラフはインターネット、共著ネットワーク、地図情報など[1]、幅広い実社会のドメインでの応用が期待されている。

機械学習の一手法であるグラフ表現学習の分野では、DeepWalk[2]、node2vec[3]、LINE[4]、metapath2vec[5]など、グラフ構造からノードの埋め込み表現(グラフを構成するノードの潜在的な特徴を低次元のベクトル空間に埋め込んだもの)を自動的に学習する手法が提案されている。グラフ表現学習の手法により異種グラフから埋め込み表現を取得することでクラスタリングやエッジ分類等の下流タスクを実行でき、Web ページのカテゴリ分類や、論文と著者の執筆の関係性分析[6]などが実現できる(図 1)。

このように、データを異種グラフにより適切に表現できれば、グラフ表現学習を適用し、様々な下流タスクを実施することができる。しかし、異種グラフの既存の適用例の多くは、ノードを Web ページ、エッジをハイパーリンクで表現できる Web に代表されるように、ノード・エッジを予め明確に定義でき、異種グラフの設計が比較的容易なドメインに限られている。

異種グラフの応用が期待されるドメインの一つに、都市空間の評価がある。近年の都市計画では、データ駆動型まちづくりが注目されており、機動的な都市空間活用が重視されている [7]。具体的な事例として、にぎわい創出をめざし、道路にウッドデッキ等を一時設置するパークレットなどがある。このような空間活用は、地点に配置する都市機能を柔軟に変化させることで実現されるため、その効果を事前評価するには、都市データを用いて地点と都市機能の関係性を定量化することが重要である。異種要素間の関連を記述可能な異種グラフは、この関係性分析において有望な手段であるといえる。しかし、異種グラフの結果の妥当性は、入力グラフの設計に大きく依存するため、都市空間の複雑な相互関係を捉え、都市データを適切に構造化できる異種グラフを設計することが肝要である。

そこで本研究では、都市空間全体の評価を目的とし、都市空間を構成する要素を表現可能な異種グラフの設計を検討する。また、オープンデータを用いて設計した異種グラフの妥当性を確認する。

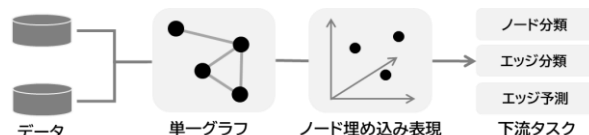


図 1 異種グラフを用いた分析の流れ

2. 都市空間を評価可能なグラフ設計の提案

2.1 着目点

前述の通り、都市空間の評価には、ある地点とある都市機能の親和性など、地点・都市機能の関係性を定量化した情報が必要である。そのため、設計する異種グラフは、地点・都市機能の 2 つの要素を含む単一グラフであることが求められる。しかし、地点と、そこに立地する都市機能を単純に接続しただけでは、地点間や都市機能間の相互作用を表現できず、また単一グラフであることも保証されない。

本研究では、都市空間の要素が空間上で相互に関係することに着目し、地点・都市機能の階層関係に加えて、地点間も空間的な関係に基づいて接続した異種グラフの設計を提案する(図 2)。このグラフ設計に従うことで、都市データを適切に構造化し、すべての都市機能・地点の空間関係を表現した単一グラフに統合できる。

2.2 グラフ設計

提案するグラフ設計手法を適用し、都市データから異種グラフを構築する流れを説明する。前提として、地点の位置データと、各地点の都市機能(レストラン、カフェ等)の立地データが入手されている状況を考える。はじめに、「地点」タイプのグラフを構築する。地点タイプのグラフは、空間上の地点をノードとし、エッジが関係性の強い地点間を接続する。このとき、地点間のエッジの有無は、地点ノード間の地理的な近接性か、道路網などのネットワークデータに基づき設定する。つぎに、地点タイプのノードと、「都市機能」タイプのノードを接続する。地点と都市機能タイプのノード間にエッジを接続することで、ある都市機能が特定の地点に位置する状態を表現する。これにより、位置データ・立地データから単一グラフを構築できる。

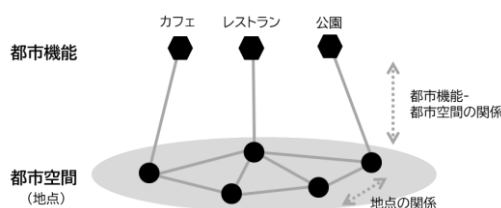


図 2 都市空間を評価可能なグラフ設計

[†] (株) 日立製作所 研究開発グループ
Hitachi, Ltd., Research & Development Group

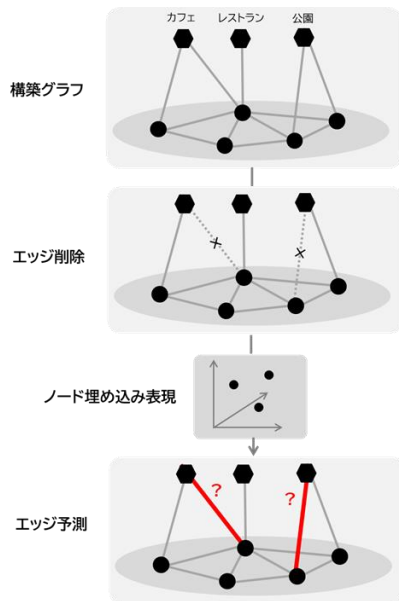


図 3 性能評価の流れ

2.3 都市評価に関する情報生成

構築したグラフを入力とし、グラフ表現学習の手法を適用することで、都市機能と地点の関係性を踏まえたノード埋め込み表現が獲得される。各ノードの埋め込み表現のベクトルに基づき、ノード分類、エッジ分類、エッジ予測など、グラフ表現学習の一般的な下流タスクを実施することで、ある地点に適した都市機能の予測など、都市評価に必要な情報を出力できる。

3. オープンデータによる性能評価

3.1 評価の概要と手順

オープンデータを用いて、提案グラフの妥当性を評価する。前述の通り、都市空間の評価の目的の一つは、地点と都市機能の関係性を定量化することである。そこで、提案グラフに基づくノード埋め込み表現が、既存の都市空間の関係性を適切に反映できているかを検証する。

評価の手順を説明する(図 3)。まず、地点の位置データと都市機能の立地データを用いて、提案手法に基づくグラフを構築する。つぎに、地点タイプと都市機能タイプを接続するエッジの一定割合をランダムに削除する。最後に、ノードの埋め込み表現を獲得した上で、削除したエッジの存在を予測するタスクを実行し、予測性能を評価する。このとき、エッジ有無の予測には、両端ノードの埋め込み表現に基づくロジスティック回帰を用いる。

3.2 分析対象地域・使用データ

本評価では、日本国内の自治体の一つを分析対象地域に選定し、OpenStreetMap[8]が提供する地図データを用いて分析を実施した。使用データには 9,284 地点の座標と 32 種類の都市機能の立地が記録されている。

3.3 評価の設定

本評価では、グラフ表現学習の手法に node2vec[3]を利用し、128 次元のノード埋め込み表現を生成した。なお、削

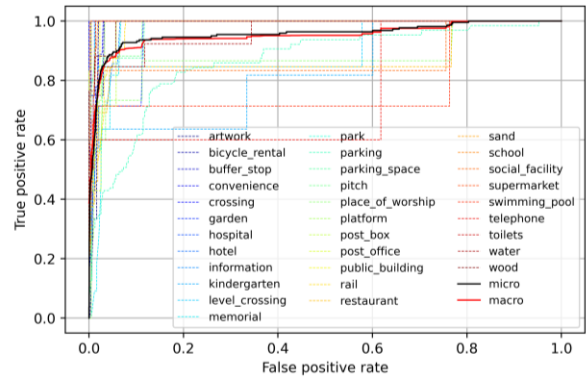


図 4 ROC 曲線による予測性能の評価結果

除エッジの割合は 20%、1 ノードあたりのサンプリング数は 5 回、サンプリングあたり移動回数は 50 回、ウィンドウ幅は 7、node2vec のパラメータ $p \cdot q$ はともに 1 とした。得られた埋め込み表現から、32 種類の都市機能ノードに対して、削除したエッジの存在を予測するタスクを実施した。

3.4 予測性能の評価結果

32 種類の都市機能の予測性能を図 4 の ROC 曲線に示す。結果から、提案するグラフ設計に基づく予測は、ほぼすべての都市機能で、ランダムな予測より高い精度を示したことがわかる。ここから、設計した異種グラフが、異種都市データ間の相互作用を一定程度表現できたと考えられる。

4. おわりに

本研究は、都市空間の評価への異種グラフの応用を検討し、地点・都市機能の関係性を推定可能とする異種グラフ設計を提案した。性能評価の結果、提案グラフは、ランダムな予測と比較し、優れた精度を示すことを確認した。

参考文献

- [1] Sun, Y., Han, J., "Mining heterogeneous information networks: a structural analysis approach", ACM SIGKDD explorations newsletter, Vol.14, No.2, pp.20-28 (2013).
- [2] Perozzi, B., Al-Rfou, R., Skiena, S., "DeepWalk: Online learning of social representations", In Proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining, pp. 701-710 (2014).
- [3] Grover, A., Leskovec, J., "node2vec: Scalable feature learning for networks", In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp.855-864 (2016).
- [4] Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J., Mei, Q., "LINE: Large-scale information network embedding", In Proceedings of the 24th international conference on world wide web (2015).
- [5] Dong, Y., Chawla, N., Swami, A., "metapath2vec: Scalable Representation Learning for Heterogeneous Networks", In Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining, pp.135-144 (2017).
- [6] Sun, Y., Han, J., Yan, X., Yu, P. S., Wu, T., "Pathsim: Meta path-based top-k similarity search in heterogeneous information networks", Proceedings of the VLDB Endowment, Vol. 4, No. 11 (2011).
- [7] 伊藤孝紀, 岩崎翔太, 鈴木篤也, 西田智裕, "道路空間再編に向けた社会実験の効果検証—岡崎市康生通りのパークレットを対象として", 日本建築学会論文集, Vol.86, No.779, pp.197-207 (2021).
- [8] OpenStreetMap Foundation, "OpenStreetMap", licensed under Open Data Commons Open Database License.