

大阪府内における住宅系不動産の 間取りのグラフデータベースの自動作成と分析

川上 春子

1. はじめに

古くから住宅の間取りは、住宅の質を決定するうえで非常に重要であると考えられてきた。事実、大学の住宅の設計課題は、間取りを中心として展開されることが一般的であるし、間取りに関する書物も多数出版されている¹⁾。ある程度経験を積んだ人間であれば、間取りの良しあしを主観的・直感的に判断できる。しかし計算機などで間取りを定量的に評価することを目的とした場合、間取りを何らかの手段でモデル化する必要がある。このためにグラフ理論を用いたアプローチが用いられることが多いが、この先駆けとなったのは、ロンドン大学で始まった Space Syntax²⁾である。日本ではこれとは別に山中らの研究³⁾が初期の研究として存在する。部屋や空間をノード(点)、それらの動線としてのつながりをエッジ(線)として表現してアクセスグラフを構築し、ネットワーク分析を適用することで、空間のつながりからみた間取りの性質を定量的に評価することができる。またグラフマイニングといった、グラフ構造に対応した比較的新しい機械学習の手法を用いた研究も行われている⁴⁾。

こうした間取りの研究でボトルネックになるのが、如何に間取りのデータを収集し、グラフ化するかということである。過去の研究では、間取りを図面や書籍等から読み取って、手作業でグラフ化することが一般的に行われてきた。しかしこの方法では多数のデータを作成するには限界があった。しかし近年、国立情報学研究所が、不動産ビッグデータとして、全国の約 533 万件の物件データを収録した LIFULL HOME'S データセットを研究者向けに公開している⁵⁾。このデータには、賃料や築年数といった数量的な物件のデータ以外に、間取りや物件の画像も収録されているため、量としては問題のない規模を誇っている。しかし、肝心の間取りのデータは画像(ラスター)データであるため、そのままでは上記の分析に用いることができない。このため、深層学習の技術を用いて、ラスターデータを変換する取り組みが最近行われており、中でも山田ら⁶⁾は、深層学習の手法を用いて、アクセスグラフを抽出する方法を開発している。

以上の背景から本研究では、LIFULL HOME'S データセ

ットと、山田らの方法を用いて、大阪府内の住宅系不動産の間取りをアクセスグラフとして抽出し、どの程度の精度でグラフが生成できるのかを検証する。次に、生成されたアクセスグラフを、グラフの同型性によって判定し、異なるアクセスグラフをデータベース(カタログ)化するとともに、それらについてネットワーク分析を行い、特徴を把握する。

2. 研究方法

本研究は以下の手順で行う。

2.1 データの準備

LIFULL HOME'S データセットに収録されているデータの中で、物件の基本情報が掲載されている「賃貸物件スナップショットデータ」と、解像度の高い間取り画像が含まれる「高精細度間取り図画像データ」を物件固有の ID で紐づけ、所在値と間取りのタイプから、必要な画像を抽出する。さらに、収録されている画像は間取りだけでなく、内観や外観の画像も含まれているので、それらのチェックを行う。また、モノクロ画像と RGB 画像が混在しているので、2.2 の手法を適用するために、モノクロ画像は RGB の色空間に変換する。

2.2 アクセスグラフの生成

使用するアクセスグラフ化のツールは、東京大学山崎俊彦研究室で開発されたものなので、それを借用して利用する。このツールは、図 1 の手順で間取り画像をアクセスグラフに変換する。

まず、Step 1 として入力された間取り画像を、深層学習の手法であるセマンティック・セグメンテーション(SS)により、各建築構成要素に色分けする。SS とは、画素に対し、車、人、道路などの決められた複数の領域に画像を塗り分けることである。この手法では、間取り画像に対して表 1 に示す計 14 種類の建築構成要素に色分けする。これにより、間取り画像が画素レベルで認識され、画像中のどこにどのような建築構成要素があるのかを推定できる。

Step 2 では、SS を行った画像の画素を走査して接続関係を調べ、ルールベースでアクセスグラフの作成を行う。ノードの作成は、一定以上の面積を持つ領域を抽出する

ことで行う。エッジの作成は、同じドアに隣接している部屋間、あるいは直接隣接している部屋間にも行う。これにより、実際の部屋の接続関係を正しく反映したグラフ構造を作成することができる。

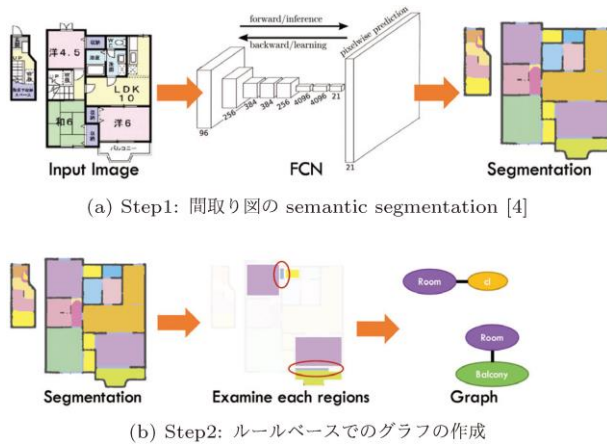


図1 間取り画像からアクセスグラフを作成する手順⁶⁾

表1 建築構成要素の定義

ラベル	説明
wall	外壁、内壁
tatami	和室
west	洋室
dk	ダイニング、キッチン、リビング
WC	トイレ
bath	浴槽
washing	洗面所、脱衣所
entrance	玄関
balcony	バルコニー、ベランダ、テラス
rouka	廊下
stairs	階段
cl	クローゼット、押入れ、下駄箱
doors	開き戸、引き戸、折戸、窓
unknown	記載のない箇所、不明箇所

2.3 アクセスグラフの精度評価

生成されたアクセスグラフを 100 個ランダムサンプリングし、間取りが正しくアクセスグラフに変換されているか、されていない場合はどのような間違いの傾向があるのかなどを、目視で調査・集計する。

2.4 アクセスグラフの同型性判定

生成したアクセスグラフがどの程度のバリエーションを有しているかを調べるために、グラフの同型性判定を行う。Python のネットワーク分析ライブラリの Network X を用いて、ノードラベル付きの同型性判定を行う。

2.5 アクセスグラフの分析

抽出されたユニークなグラフに対して、度数などの基本統計量をまとめるとともに、間取りの分析でよく行われるネットワーク分析を行い、各アクセスグラフの特徴

を分析する。

3. 結果

本研究の結果と考察を述べる。

3.1 アクセスグラフの生成

2.1 については、533 万件程度の物件データがあるため、大阪府かつ 3LDK という条件で、Python で該当物件を抽出した結果、48497 件の間取り画像が抽出された。

2.2 により、これらの画像からアクセスグラフを生成した。

3.2 アクセスグラフの精度評価

2.3 により、生成されたアクセスグラフの精度を確かめたところ、正解率は 52% であった。間違いのパターンとして、セマンティックセグメンテーションの時点で面積が小さい部屋、文字表記がない領域であるドアや壁、廊下の精度が低いことなどが挙げられる (図 2)。また、セマンティックセグメンテーションができていても、グラフ化の時点においてノードにつながるエッジがないグラフになっていることが多いことが分かった。ここでは、このようにグラフ上の任意の 2 頂点間のパスが存在しないグラフは、連結していないと表現する。



図2 アクセスグラフの誤変換の例

3.3 物件の選定

2.4 の同型性判定を行うにあたって、正確にグラフ化ができていない物件に絞る必要がある。上記の連結していないグラフは除いたうえで、表 2 に示した妥当な間取りの物件に絞った結果、48496 件から 18497 件になった。

表2 物件の絞り込み条件

部屋	数
ダイニングキッチン	1
トイレ	1
浴室	1
洗面	1
バルコニー	0,1,2,3
玄関	1
廊下	1
階段	0
クローゼット	0,1
洋室+和室	3

3. 4 同型性判定

同型性判定に用いる手法とその結果について述べる。

3. 4. 1 NetworkX の概要

前節で選定した物件について NetworkX を用いて同型性判定をする。NetworkX は、Python ベースのネットワーク分析ツールである。複雑ネットワークの生成・操作・分析などに用いられる。NetworkX は有向グラフやマルチグラフなど、多様なネットワークを表現できるデータ構造を有している。また、ネットワーク分析のための、数多くの関数やアルゴリズムを実装している。

3. 4. 2 同型性判定の結果

選定された 18497 件の物件について同型性判定をした結果、5141 通りの間取りパターンが出てきた。ただしクローゼットを考慮すると、間取りのバリエーションが著しく増えることが分かったため、独立したクローゼット以外はアクセスグラフからは除いている。

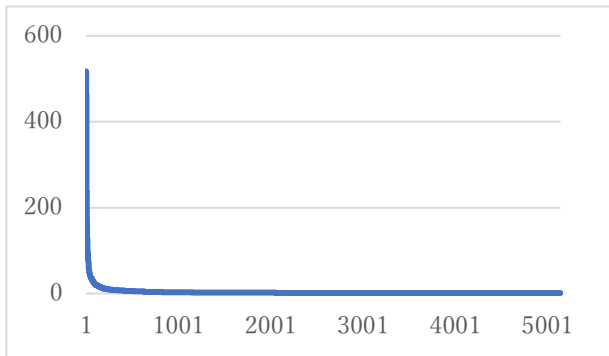


図3 同型なアクセスグラフの降順の度数分布

一番多いと判定された間取りと、1 パターンのみと判定された間取りの一例を、以下に示す。



図3 最多パターンの間取り (件数: 517 件)



図4 1 パターンのみの間取りの例

3. 5 アクセスグラフのネットワーク分析

2.5 により、抽出された 18497 件の物件に対して、ネットワーク分析を行った。本研究では間取りに関連したグ

ラフ特徴量として、次数中心性と玄関から各部屋までのネットワーク距離を求めた。

3. 5. 1 次数中心性

次数中心性とは多くの頂点と隣接している頂点が中心点とするものである。ノード i の次数中心性 DC_i はノード i の次数 k_i そのもの。つまり、ノードのリンク数が多いほど中心性が高くなる。NetworkX では、次数を自分以外のノードの数 $(n - 1)$ で割って正規化したものを使用している。式としては以下で示される。

$$DC_i = \frac{k_i}{n-1}$$

各部屋の次数中心性の平均値を見ると、廊下、dk が高いことが分かる。他の部屋は大きく差をつけて低いことが分かる。

表4 次数中心性の平均値

room	average
rouka	0.560
dk	0.428
washing	0.272
tatami	0.202
west	0.196
balcony	0.195
wc	0.145
entrance	0.135
bath	0.131
cl	0.021

3. 5. 2 玄関から各部屋までのネットワーク距離

経路探索において、実際にある辺をたどって行って最短距離を求める幅優先探索を用いて、玄関から各部屋まで何ステップで到達できるかを求めた。Networkx の関数 `bfs_edges` では引数として、対象ネットワーク・探索開始点・探索の深さの上限を与えることができる。本研究では、対象ネットワークを 2 章で生成したアクセスグラフ、探索開始点を玄関とする。

幅優先探索の結果、3 と 4 が多かった。玄関から 3 部屋または 4 部屋通って、一番遠い部屋に行けるということである。この幅優先探索を使えば、リビングを通らずに洋室に行きたいなど、動線を指定して物件を選ぶことができる。

3. 6 築年

最後に、築年数とアクセスグラフの関係を分析した。まず、全体の竣工年のグラフを見ると、1989 年までは着工数は右肩上がりだったが、1989 年～1992 年にかけて減っている(図5)。バブル期で着工数が増え、1991 年にバブル

が崩壊したことを反映していると思われる。

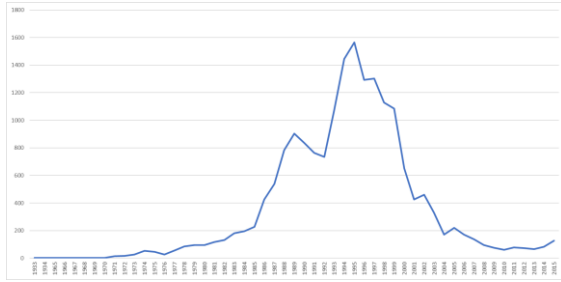


図5 竣工年の分布

次に年代別に間取りの特徴を見ていく。まず、本研究で1980年代にしか見られなかった間取りは和室が2部屋であることが多く、かつldkの中心性が高かった(図6)。1990年代にしか見られない間取りは、和室が1部屋であることが多かった(図7)。2000年代以降にしか見られない間取りは和室がなく、3部屋とも洋室であることが多かった(図8)。これらのことから、和室の数が減っていることが分かる。時代が進むにつれて和室の人气がなくなっていったものとみられる。

一方、古くから現代まで続く間取りとして、和室が1部屋、洋室が2部屋の間取りが多く見られた。和室の人气がなくなっているとはいえ、一定の需要があることが分かる。(図9)



図6 1980年代にしか見られない間取りの例(15件)



図7 1990年代にしかない間取りの例(17件)



図8 2000年以降にしかない間取りの例(17件)



図9 古くから現代まで続く間取りの例(125件)

4. まとめ

本研究では、大阪府内の住宅系不動産の間取りをアクセスグラフとして抽出し、どの程度の精度でグラフが生成できるのかを検証し、生成されたアクセスグラフを、グラフの同型性によって判定し、異なるアクセスグラフをデータベース(カタログ)化するとともに、それらについてネットワーク分析を行い、特徴を把握した。

中心性においては、廊下やダイニングキッチンが高く、より多くの部屋とつながっていることが分かった。中心性を見ることで、ldkが独立しているほうがいいなどの要望にも応えられる。

玄関から各部屋までのネットワーク距離においては、今後リモートワークが進み、家で仕事をする場合などにかせると考える。仕事部屋を始点にして、生活音のするリビングが2ステップ目に出てこないような間取り、隣り合わない間取りなどが分かる。

築年においては、全体を見ると、バブルやバブル崩壊、震災、増税など、社会の変化を反映していることが分かった。年代別に間取りを見ると、特に和室の数において特徴が表れていた。

本研究は大阪府下の3LDKの賃貸物件を対象に行ったが、より広範囲で、他の間取りについても分析をすることで地域や間取りごとの特徴がより明確になるとと思われる。

参考文献

1. 飯塚 豊, 間取りの方程式, エクスナレッジ, 2014
2. Hillier, B. Hanson, J., The Social logic of space. Cambridge University Press, 1984
3. 山中知彦, 原広司, 藤井明, 渡辺健一, 計画学におけるグラフ理論適用に関する研究 その2 応用編, 日本建築学会論文報告集, 342, 62-72, 1984
4. 瀧澤重志, 吉田一馬, 加藤直樹, グラフマイニングを用いた室配置を考慮した賃料分析 京都市郊外の3LDKを中心とした賃貸マンションを対象として, 日本建築学会環境系論文集, 73(623), 139-146, 2008
5. 国立情報学研究所, LIFULL HOME'S データセット, <https://www.nii.ac.jp/dsc/idr/lifull/> (参照日 2020.9.27)
6. 山田万太郎, 汪雪婷, 山崎俊彦, 相澤清晴, 深層学習を用いた不動産間取り図のグラフ化と物件検索への応用, 人工知能学会第33回全国大会論文集, 2019