

PLATEAU の建物情報を利用した地域犯罪の傾向分析 Analysis of Local Crime Trends using PLATEAU Building Information

向 直人¹⁾
Naoto Mukai

1 背景

平成 28 年に官民データ活用推進基本法が施行され、犯罪発生情報のオープンデータ化が進められている。都道府県警察は窃盗に関する情報をウェブサイト上で公開し、地域の犯罪情勢の把握や、防犯活動の促進に向けた利用を奨励している。本研究では、愛知県警察が公開している平成 30 年から令和 3 年までのオープンデータ [1] を利用して、犯罪発生と関連性の高い地理的な環境要因を抽出することを目的とする。地理的な環境要因を抽出することができれば、犯罪ホットスポットの推定が可能になり、将来の犯罪を予防することに繋がる。

環境要因を抽出する手法として RTM(Risk Terrain Modeling) が知られている [2, 3]。RTM では、環境要因として食料品店や集合住宅などを対象とし、犯罪発生地点における環境要因の近接性や密度を特徴量として、回帰モデルを生成する手法である。Kennedy らは、シカゴにおける暴力事件を対象に、RTM を利用して影響力の高い環境要因を推定した [4]。薬物活動、暴力団活動などに使用される建物、ギャングの活動が活発な地域、差し押さえされた物件などが、環境要因として特定された。また、犯罪発生地点の近接性を利用した手法として、カーネル密度推定 (KDE: Kernel Density Estimation) がある。中川らは、時空間に拡張したカーネル密度推定を利用して、犯罪ホットスポットを検出する手法を提案した [5]。カーネル密度推定のパラメータであるバンド幅を最適化することで、対象地域を効率的にカバーするパトロールを実現した。同様に、野らは RTM とカーネル密度推定を組み合わせ、埼玉県にある警察署管内の防犯パトロールのための経路を構築した [6]。近年は犯罪の時間的連続性に着目し、LSTM(Long Short-Term Memory) と呼ばれる再帰型ニューラルネットワークで犯罪発生を予測する手法も提案されている [7, 8]。LSTM は長期的な依存関係を把握できるため、時間の経過に伴う犯罪パターンを予測するのに適しているとされる。

本研究では、環境要因として 3D 都市モデル PLATEAU[9] の建物データを用いる。PLATEAU は、国土交通省が主導する 3D 都市のオープンデータであり、2023 年 4 月には愛知県の名古屋市を含む 117 の都市をカバーしている。建物の形状を LOD(Level of Details) と呼ばれる基準で表現していることが大きな特徴であり、LOD0 は建物の地表 (または屋上) の緯度・経度、LOD1 は建物を構成する複数の面の緯度・経度の情報が含まれる。また、「業務施設」や「商業施設」など 18 種類の建物の用途が属性として与えられている。これらの情報を定量化し、犯罪との関連性を抽出するために、犯罪発生地域の地図画像を生成するという方法を採用した。地図画像に用途で色分けされた建物の形状を含めることで、犯罪発生地域の周辺の様子を把握するこ

とが可能となる。この地図画像の建物ごとの画素数を特徴量として、同地域で生じた犯罪件数を回帰分析することで、建物の用途が犯罪発生に影響するかを明らかにする。

2 犯罪オープンデータ

愛知県警察が公開している犯罪オープンデータは、平成 30 年から令和 3 年までの 56,733 件の窃盗に関する犯罪データで構成されている。窃盗は「ひったくり」「車上ねらい」など 7 種類の犯罪種別 (手口) に分類され、発生年月日や発生時刻などの情報と共に記録されている。一方で、発生場所に関しては、個人情報の保護の観点から、「都道府県」「市区町村」「町丁目」の情報しか提供されておらず、正確な緯度・経度は不明である。そこで、情報試作室の住所ジオコーダー [10] を利用して、住所文字列から対応する緯度・経度に変換した。変換に際しては、住所の名寄せが不完全で、正確な緯度・経度を導出できないケースが存在する。そこで、大字までが一致した 13,319 件の住所のみを対象とすることにした。

PLATEAU は愛知県全域をカバーしていないため、名古屋市のデータのみを用いる。名古屋市の PLATEAU のデータは、一辺が約 1km の 3 次メッシュの単位で整備されており、栄などの中心部は LOD2、その他は LOD1 で提供される。ここでは、名古屋市の中心部を含む 2 次メッシュの 523657 と 523667 の一部を対象とする (2 次メッシュは 10×10 の 3 次メッシュを包含する)。図 1 にその対象地域を示す。この対象地域に含まれる住所は 13,319 件中の 2,100 件 (約 16%) であり、その範囲内で発生した窃盗は 56,733 件中の 17,506 件 (約 31%) であった。

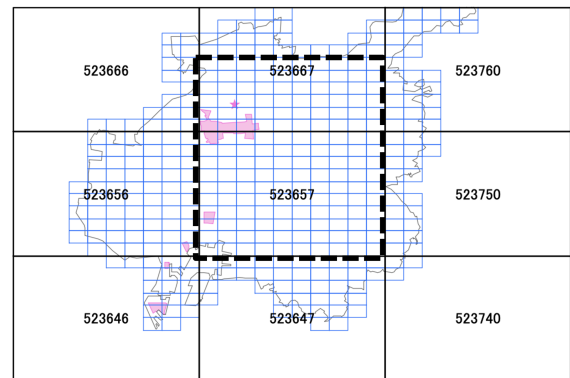


図 1 本研究で対象とする地域

表 1 は、この 17,506 件の窃盗における 7 種類の犯罪種別の発生件数と、各住所で発生した犯罪数の平均、標準偏差、変動係数を示している。発生件数は「自転車盗」が最も多く、全体の約 70% を占めている。次いで、車内の荷物を盗む「車上ねらい」、車に取り付けられた部品を盗む「部品ねらい」が続く。「ひったくり」は最も少なく、名古屋市の中心部であっても頻度は低いことがわかる。また、「ひったくり」の変動係数は約 4.5、「自

1) 椋山女学園大学 文化情報学部 文化情報学科

動販売機ねらい」の変動係数は約3.6と他に比べて大きく、これらの犯罪が発生しやすい地域と、そうでない地域の差が顕著であることが示唆される。本研究において抽出を目的とする地理的な環境要因が、この地域間の差を生む一因であると考えられる。

表1 犯罪種別ごとの発生件数に関する統計量

犯罪種別	件数	平均	標準偏差	変動係数
ひったくり	159	0.076	0.340	4.493
車上ねらい	2,194	1.045	1.695	1.622
部品ねらい	1,213	0.578	1.114	1.929
自動販売機ねらい	469	0.223	0.813	3.639
自動車盗	676	0.322	0.741	2.302
オートバイ盗	474	0.226	0.677	2.998
自転車盗	12,321	5.867	17.147	2.923

3 Plateau を利用した特徴量の生成

ジオコーディングにより住所文字列から変換した緯度・経度を中心とした地図画像を生成する。地図画像の生成には Mapbox の Static Tiles API[11] を利用した。生成される地図は、道路が白(255,255,255)、背景が黒(0,0,0)のバイナリ画像であり、サイズは256px×256pxである。このバイナリ画像に、表2に示す建物用途の配色に従って、PLATEAUの建物形状を重ね合わせて描画する。建物用途は「業務施設」や「商業施設」など18種類で構成され、それらの色差が大きくなるように配色を設定した。また、対象とする名古屋市の建物形状のデータはLOD1またはLOD2であり、建物の高さの情報を含むが、生成される画像は平面地図であることから、建物の地表面と接する緯度・経度の情報のみを用いて2次元のポリゴンとして描画した。

表2 建物用途と色

建物用途	色 (RGB)
業務施設	(255,0,0)
商業施設	(255,0,127)
宿泊施設	(255,0,255)
商業系複合施設	(127,0,255)
住宅	(0,0,255)
共同住宅	(0,127,255)
店舗等併用住宅	(0,255,255)
店舗等併用共同住宅	(0,255,127)
作業所併用住宅	(0,255,0)
官公庁施設	(127,255,0)
文教厚生施設	(255,255,0)
運輸倉庫施設	(255,127,0)
工場	(255,122,122)
農林漁業用施設	(255,122,188)
供給処理施設	(255,122,255)
防衛施設	(188,122,255)
その他	(122,122,255)
不明	(122,188,255)

ここで注目すべきは地図画像のズームレベル（表示倍率）である。低いズームレベルであれば、広い範囲の地域特徴を表現できるが、対象の犯罪とは無関係

の特徴まで含めてしまう可能性がある。一方、高いズームレベルであれば、狭い範囲の地域特徴しか表現できないが、対象の犯罪と関係が深い特徴のみを取り扱うことができると考えられる。そこで、本研究では、レベル15(1.830m/px)、レベル16(0.915m/px)、レベル17(0.457m/px)の3種類のズームレベルを比較することにした。各ズームレベルの地図画像の1辺の長さは、約468メートル、約234メートル、約117メートルに一致する。

図2は名古屋市中区3丁目、図3は名古屋市中区新栄1丁目、図4は名古屋市中区千種区不老町の地図画像である。左からレベル15、レベル16、レベル17の地図画像を表している。各住所の犯罪発生件数を表3に示す。名古屋市中区栄3丁目は、名古屋市の繁華街の中心であり、489件の犯罪が発生している。これは、2,100件の住所の中で最多である。商業施設(255,0,127)が目立って存在しており、「車上ねらい」「オートバイ盗」「自転車盗」がワースト1位である。名古屋市中区新栄1丁目は、繁華街から少し外れた住宅街であり、住宅(0,0,255)や共同住宅(0,127,255)が多く存在している。名古屋市中区3丁目と比べて総犯罪件数は少ないが、「ひったくり」がワースト2位、「部品ねらい」がワースト2位、「自動販売機ねらい」がワースト1位である。名古屋市中区千種区不老町は、大学が設置された郊外の地域であり、文教厚生施設(255,255,0)が大部分を占めている。他地域に比べ犯罪件数は少ないが「自転車盗」は頻繁に発生していることがわかる。このように、建物用途で表された地域の特徴量と、同地域で発生する犯罪種別には関係があるように思われる。そこで、地図画像に占める建物用途の画素数をカウントすることで、特定の住所に対応する地域の特徴量とする。

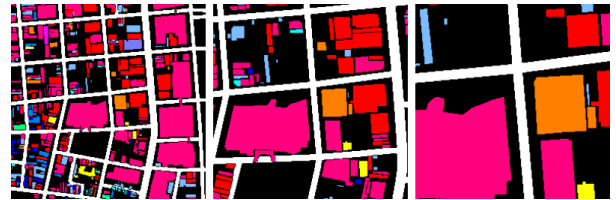


図2 地図画像(名古屋市中区栄3丁目)

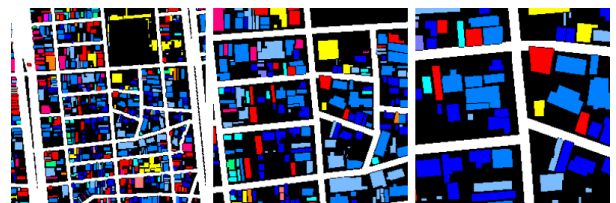


図3 地図画像(名古屋市中区新栄1丁目)



図4 地図画像(名古屋市中区千種区不老町)

表 3 犯罪件数の比較

犯罪種別	栄 3 丁目	新栄 1 丁目	千種区不老町
ひったくり	4	4	0
車上ねらい	23	17	2
部品ねらい	8	10	1
自動販売機	10	11	0
自動車盗	3	7	0
オートバイ	13	6	0
自転車盗	427	140	74
合計	489	195	77

3 種類のズームレベルにおける建物種別の画素数の平均を表 4, 画素数の変動係数を表 5 に示す (対象地域に存在しなかった「商業系複合施設」「作業用併用住宅」「防衛施設」を除く)。名古屋市の対象地域には「住宅」や「共同住宅」が多く存在し、「農林漁業施設」や「宿泊施設」は殆ど存在しないことがわかる。また、ズームレベルが高くなると、各建物の画素数も増加することが確認できる。これは地図画像に占める建物の面積が相対的に大きくなるのが理由と考えられる。同様にズームレベルが高くなると、地図画像に含まれる範囲が狭くなり、建物の用途が限定されることから、変動係数も上昇することが確認できる。そもそも存在が少ない「農林漁業施設」や「宿泊施設」は、変動係数が高くなる傾向があるが、犯罪を誘引する要因にはなりにくいと考えられる。一方で、「業務施設」や「商業施設」は一定数存在することに加え、変動係数が比較的高い傾向がある。これらの建物用途は、生じやすい犯罪種別を判断するための特徴量となりえると考えられる。また、「住宅」や「集合住宅」は変動係数が低く、住所間での差異が小さいことから、特徴量としては価値が低いことが予想される。

表 4 建物用途ごとの画素数の平均

建物用途	レベル 15	レベル 16	レベル 17
業務施設	808.1	867.2	925.3
商業施設	890.8	894.3	1036.8
宿泊施設	26.0	27.4	43.10
住宅	5661.7	7019.6	7895.5
共同住宅	2734.6	3237.1	3543.3
店舗等併用住宅	537.1	637.0	688.6
店舗等併用共同住宅	246.0	273.0	283.4
官公庁施設	98.7	105.1	117.7
文教厚生施設	1238.8	1359.0	1503.3
運輸倉庫施設	481.6	525.6	682.0
工場	715.5	796.5	894.7
農林漁業施設	2.1	2.7	1.8
供給処理施設	45.9	52.1	42.2

4 実験

建物の用途ごとの画素数を説明変数、犯罪種別ごとの件数を目的変数として、線形回帰分析を適用する。説明変数として、表 4 に示されている 13 種類の建物用途に、道路と背景の 2 種類を加えた 15 種類の画素数を用いる。説明変数と目的変数は、平均 0, 標準偏差 1 に標準化する。また、説明変数の重みとなる偏回帰係数は正の値となるように制約した。表 6 に犯罪種別ごとの決定

表 5 建物用途ごとの画素数の変動係数

建物用途	レベル 15	レベル 16	レベル 17
業務施設	1.366	1.608	2.164
商業施設	1.280	2.046	3.244
宿泊施設	5.784	9.856	13.089
住宅	0.519	0.597	0.726
共同住宅	0.517	0.683	0.953
店舗等併用住宅	0.709	0.920	1.314
店舗等併用共同住宅	1.471	2.153	3.330
官公庁施設	3.132	5.526	7.579
文教厚生施設	0.969	1.509	2.203
運輸倉庫施設	1.743	2.287	3.287
工場	2.370	2.909	3.485
農林漁業施設	15.548	20.380	18.636
供給処理施設	3.994	6.474	8.346

係数 R^2 を示す。3 種類のズームレベルを比較すると、広い範囲を地図画像に含むレベル 15(1.830m/px) の決定係数 R^2 が最も高い値となった。このことから、犯罪と関わりがある地域の特徴を表現するには、一定の広さが必要になることがわかる。また、7 種類の犯罪種別を比較すると、「自転車盗」の決定係数 R^2 が最も高くなった。これは、そもそも「自転車盗」の犯罪件数が、他と比べ圧倒的に数多く、学習に適していたことが理由と考えられる。それでも、決定係数 R^2 は 0.273 と小さく、当てはまりが十分とは言えない結果となった。また、表 1 で変動係数の高かった「ひったくり」と「自動販売機ねらい」は決定係数 R^2 が 0.12 以上となっており、他の犯罪種別に比べて高い値を示した。発生件数は少なくとも、住所間で差異が大きいことが、影響したと考えられる。

表 6 犯罪種別ごとの決定係数 R^2

犯罪種別	レベル 15	レベル 16	レベル 17
ひったくり	0.130	0.074	0.050
車上ねらい	0.124	0.087	0.049
部品ねらい	0.071	0.044	0.024
自動販売機ねらい	0.124	0.092	0.052
自動車盗	0.046	0.026	0.012
オートバイ盗	0.089	0.062	0.037
自転車盗	0.273	0.193	0.112

偏回帰係数が上位の建物用途を表 7 に示す。「ひったくり」は業務施設が多くを占める地域で発生していることがわかる。例に上げた名古屋市中区 3 丁目のようにオフィスビル街などが該当していると考えられる。「車上ねらい」「部品ねらい」「自動車盗」は、いずれも車を狙った犯罪であるが、異なる特徴の地域で発生することがわかる。「車上ねらい」と「自動車盗」は業務施設や商業施設が多い市街地で発生する。一方、「部品ねらい」は背景が多くを占めていることから、建物が少ない郊外の地域で発生する。「自動販売機ねらい」は、「ひったくり」と同様に、業務施設の近くで発生している。これは自動販売機が、業務施設の近くに設置されることが多いからだと考えられる。「自転車盗」は業務施設と商業施設が多い地域で発生していることから、駅周辺などの繁華街が該当していると考えられる。

表 7 偏回帰係数が上位の建物用途 (レベル 15)

犯罪種別	1 位	2 位	3 位
ひったくり	業務施設 (.35)	背景 (.18)	道路 (.13)
車上ねらい	業務施設 (.20)	商業施設 (.20)	背景 (.19)
部品ねらい	背景 (.37)	道路 (.24)	業務施設 (.17)
自動販売機	業務施設 (.34)	背景 (.20)	道路 (.10)
自動車盗	業務施設 (.22)	背景 (.18)	道路 (.12)
オートバイ	背景 (.25)	商業施設 (.23)	業務施設 (.15)
自転車盗	業務施設 (.32)	商業施設 (.27)	背景 (.20)

表 5 で変動係数が低かった住宅と共同住宅の偏回帰係数を表 8 に示す。偏回帰係数は小さく、犯罪への影響は殆どないと言える。しかしながら、住宅は「自動車盗」に対してはわずかに影響している。これは、住宅街で自動車盗が発生する可能性があることを示唆している。また、共同住宅は「部品ねらい」「オートバイ盗」への影響が強い。共同住宅に駐車されている車やオートバイが狙われることを示唆しているように思われる。

表 8 住宅と共同住宅の偏回帰係数

犯罪種別	住宅	共同住宅
ひったくり	0.000	0.075
車上ねらい	0.000	0.056
部品ねらい	0.000	0.106
自動販売機ねらい	0.000	0.070
自動車盗	0.002	0.080
オートバイ盗	0.000	0.105
自転車盗	0.000	0.023

5 まとめ

本研究では、愛知県の犯罪オープンデータを対象に、地理的な環境要因を推定する新たな手法を提案した。環境要因として 3D 都市モデルの PLATEAU に含まれる建物情報を用いて、犯罪発生地点の周辺の地図画像を生成し、建物用途ごとに画素数をカウントして地域の特徴量とした。線形回帰分析の結果から、建物用途ごとの画素数という特徴量は、犯罪種別ごとの件数を回帰するには不十分であるが、影響力の強い建物用途を抽出できる可能性を示した。特に「自動販売機ねらい」に関しては、サンプル数が多いことから決定係数 R^2 も比較的高く、業務施設や商業施設が多くを占める地域で発生することが明らかとなった。十分な決定係数 R^2 を得られなかった原因として、サンプル数が不十分であったことが挙げられる。PLATEAU がカバーしている地域だけを対象としたことから、「ひったくり」などのサンプル数が極端に少なくなってしまう。愛知県だけでなく、他の地域のデータも加えることで、サンプル数を増やすことが必要となる。また、特徴量として画素数を用いたことから、地図内に含まれる建物数などの特徴が失われてしまった。RTM で採用されている近接性や密度などの特徴量を加えることで、さらなる改善が期待できると考えられる。さらに、PLATEAU に含まれる建物の高さのデータを加えた 3 次元の地図から特徴量を抽出するなどの方法を検討したい。

参考文献

- [1] 愛知県警察 愛知県警察犯罪オープンデータサイト. <https://www.pref.aichi.jp/police/anzen/toukei/opendata/seian-s/crimeopendata.html>.
- [2] Leslie Kennedy, Joel Caplan, and Eric Piza. Risk clusters, hotspots, and spatial intelligence: Risk terrain modeling as an algorithm for police resource allocation strategies. *Journal of Quantitative Criminology*, Vol. 27, pp. 339–362, 09 2010.
- [3] Joel M. Caplan, Leslie W. Kennedy, Jeremy D. Barnum, and Eric L. Piza. Risk terrain modeling for spatial risk assessment. *Cityscape*, Vol. 17, No. 1, pp. 7–16, 2015.
- [4] Leslie Kennedy, Joel Caplan, Eric Piza, and Henri Buccine-Schraeder. Vulnerability and exposure to crime: Applying risk terrain modeling to the study of assault in chicago. *Applied Spatial Analysis and Policy*, Vol. 9, , 08 2015.
- [5] 中川淳子, 西村祥治, 宮野博義. 犯罪発生履歴データの機械学習による時空間カーネル密度推定型犯罪予測の最適化. 情報処理学会論文誌, Vol. 61, No. 7, pp. 1255–1266, jul 2020.
- [6] 野貴泰, 糸井川栄一. 犯罪多発地点の予測に基づく防犯パトロール経路に関する提案. 地域安全学会論文集, Vol. 31, No. 0, pp. 195–204, 2017.
- [7] Yong Zhuang, Matthew Almeida, Melissa Morabito, and Wei Ding. Crime hot spot forecasting: A recurrent model with spatial and temporal information. *2017 IEEE International Conference on Big Knowledge (ICBK)*, pp. 143–150, 2017.
- [8] Tsion Eshetu Meskela, Yidnekachew Kibru Afework, Nigus Asres Ayele, Muluken Wendwosen Teferi, and Tagele Berihun Mengist. Designing Time Series Crime Prediction Model using Long Short-Term Memory Recurrent Neural Network. *International Journal of Recent Technology and Engineering (IJRTE)*, Vol. 9, No. 4, pp. 402–405, November 2020.
- [9] PLATEAU by MLIT. <https://www.mlit.go.jp/plateau/>.
- [10] 住所ジオコーダー. <https://www.info-proto.com/jageocoder/>.
- [11] Mapbox. <https://www.mapbox.com/>.