

空間データマイニングによる交通事故パターンの可視化

2007MI077 伊藤 公郎 2007MI230 鈴木 佑哉

指導教員 河野 浩之

1 はじめに

阪神高速道路での平成 21 年の交通事故件数は 1103 件と首都高速道路について多く発生している。阪神高速道路株式会社の阪神高速道路の交通安全対策アクションプログラム^{*1}より、平成 13 年から平成 17 年までの 5 年間における阪神高速道路で発生した事故を足し合わせて事故件数の多い区間を示したり、交通事故の発生件数を路面の乾燥時と湿潤時と比較すると、湿潤時には乾燥時の約 2 倍の事故が発生しており、事故原因を年度単位で分析すると雨天日数との相関が強いことがうかがえる。このように、交通事故多発地点の表示や交通事故と路面状況との関係性などが記されていた。しかし、路面状況ごとの事故発生場所やそれぞれの発生しやすい事故形態などは示されていない。

そこで本研究では、阪神高速道路の中でも交通量が多いと言われている環状線で路面状況の違いでの交通事故に着目し、路面の湿潤時と乾燥時の交通事故発生場所の違いやどの事故形態が発生しやすいのかなどを分析し、それらの状況が分かりやすいように R ではグラフ作成し、ArcGIS では阪神高速道路の環状線マップ上に交通事故データをマッピングにより可視化をする。R では主成分分析による解析も行なう。交通事故を分析するには、2003 年から 2005 年と 2006 年から 2008 年の間に阪神高速道路で実際に発生した交通事故データを 2 種類用いて検討する。前者の交通事故データは地図上へのマッピングに使用し、後者の交通事故データは R によるグラフ化と主成分分析に使用する。また、GIS での表示を路面状況や事故形態で分けて表示することでより正確な危険箇所を表示する。

2 交通事故の可視化に関連した先行研究

2.1 交通事故データの可視化に関する研究

Nada ら [1] は、直接スロベニア警察からと一般公開されている公的に利用可能な交通事故データベースから交通事故データを取得し、可視化を行なうために必要なデータを抽出、分析から交通事故と季節、曜日、時間帯、免許取得年数との相互関係やパターンを発見し可視化を行なっている。グラフ化に用いたデータは、1995 年から 2005 年はまでにスロベニアの 58 の州で起きた

453451 件の事故データを使用している。また GIS で用いたデータは、2006 年に首都リュブリャナとヴェレニエで起きた事故データを使用している。具体的にグラフ化では、各月と曜日、各時間と曜日で事故総数を表示し、どの月、曜日、時間に事故のピークがあるかを分析している。また、運転免許取得年数と交通事故件数、1995 年から 2005 年までの交通事故総数、交通事故死の件数のグラフを作成、分析しそれぞれの傾向を推測している。

表 1 は、Nada ら [1] の研究のほかに管野ら [2]、Linhua ら [3] の交通事故の可視化に関連した研究の論文で用いている事故データと可視化の方法をまとめたものである。

2.2 先行研究との相違点

先行研究では道路の路面状況を考慮されていないためこれらを考慮して分析、可視化をすればより正確なパターンや相互関係が発見できる。そこで本研究では、阪神高速道路の交通事故データを用いて、Nada ら [1] と同様にグラフを作成し、分析によって交通事故パターンや交通事故が多い時間帯や月などとの相互関係を可視化するのに加え、道路の路面状況と天候を考慮して分析し、地図上に事故発生場所の可視化をし、それぞれを比較することにより正確な事故パターンや相互関係を発見する。

3 阪神高速道路データと研究の流れ

3.1 交通事故データ可視化の流れ

まず、阪神高速道路の交通事故データから R によるグラフ化と解析、ArcGIS による地図上への表示に必要な項目だけを選択して今回必要ない項目は予め削除しておく。R では、必要な項目だけにしたデータを取り込み交通事故との関係性が分かるグラフの作成と解析を行う。次に、ArcGIS にデータを取り込む前に阪神高速道路の交通事故データには座標は記録されていないが kp(キロポスト) が記録されていたので、他に kp と座標が対応しているデータを用意する。交通事故データと kp と座標の対応データを PostgreSQL に格納する。そして、kp により結合しファイルを作成する。作成したファイルを地図データとともに ArcGIS に取り込み路面状態、事故の形態ごとに地図上に事故データを表示させることで可視化を行う。最後に、R で作成したグラフと解析結果、ArcGIS による地図上への表示結果から何らかの交通事故との相互関係やパターンを発見する。

^{*1} <http://www.hanshin-exp.co.jp/company/torikumi/jutai/actionprogram.html>

表 1 事故データを可視化する方法の比較

論文名	事故データ	可視化の方法
Mining Spatio-temporal Data of Traffic Accidents and Spatial Pattern Visualization[1]	スロベニアで 1995 年から 2005 年までに発生した事故	グラフと GIS による可視化
時空間 GIS を用いた交通障害情報の視覚化 [2]	阪神高速道路で 2006 年の 2 月から 7 月までに発生した事故	GIS による可視化
A GIS-based Bayesian approach for analyzing spatial-temporal patterns of intra-city motor vehicle crashes[3]	アメリカのテキサス州ハリス郡で 1996 年から 2000 年で発生した事故	グラフと GIS による可視化

3.2 使用するデータの紹介

3.2.1 R で使用する阪神高速の交通事故データ

本研究で R によるグラフ化, 解析を行うのに使用する交通事故データは, 2006 年から 2008 年までの 3 年間の阪神高速道路で実際に発生した交通事故データである。まず交通事故データは, 3 年間で 7135 件記録されており, 交通事故データの中身としては, 交通事故の発生日時(年, 月, 日, 曜日, 時間)や天候, 路面状態, 場所(路線名(上下), 車線), 事故の形態など 64 項目が記録されている。使用する項目は, グラフでは発生日時 月, 発生日時 曜日, 時間 発生 時, 路面状態, 事故の形態を使用し, 主成分分析では路面状態, 事故形態, 事故原因を使用する。

3.2.2 ArcGIS で使用する阪神高速の交通事故データ

次に本研究で ArcGIS による地図上への表示で使用する交通事故データは, 2003 年から 2005 年までの 3 年間の阪神高速道路で実際に発生した交通事故データである。まず交通事故データは, 堂島, 阿波座カーブ, 空港集約, 池田線 0.5KP カーブ, 伊丹 TN15.3-16.0, 池田線 6.2-9.6KP, 加島カーブ, 守口下り 0.0-1.0, 扇町カーブ, 環守渡りカーブ, 守口上下 4.7-6.5, 東大阪 0.5-0.6, 堺 1.2-1.4, 港線 5.7-5.8 の 14 のエリアに分けられていて 3 年間で 1463 件記録されており, 交通事故データの中身としては, 交通事故の発生日時や路面状態, 路線名, 事故の形態, kp(キロポスト) など 74 項目が記録されている。ArcGIS では, 発生月, 発生日, 発生時, 曜日, 路面状態, 事故形態, kp を使用する。

3.2.3 阪神高速道路周辺の空間データ

本研究では, ArcGIS で地図を表示するのに国土地理院の基盤地図情報ダウンロードサービスから地図データを取得する。ダウンロードするファイルは基盤地図情報縮尺レベル 2500 の JPGIS 形式で符号化した XML ファイルである。しかし, ArcGIS へ取り込むにはファイル形式を変換する必要がある。ファイル変換には, 基盤地図情報閲覧コンバータソフトを使う。基盤地図情報閲覧コンバータソフトを使用して XML 文書形式のファ

イルをシェープファイル形式へ変換する。本研究では, 阪神高速道路周辺の西区, 中央区, 天王寺区, 北区, 浪速区の地図データを使用する。

4 事故データの可視化

4.1 R でのグラフ化

図 1, 図 2 は交通事故総数のグラフであり, 図 1 は曜日と月ごと, 図 2 は曜日と時間ごとのグラフである。色は白が交通事故件数が少なく色が濃くなっていくごとに交通事故件数が増加しているのを表している。図 2 からは 1 月や 2 月のように冬は少なく 6 月, 7 月, 8 月の夏に多く交通事故が発生しているのが分かる。また, 6 月, 7 月, 8 月でも木曜日, 金曜日, 土曜日, 日曜日のように週末により交通事故が発生しているのが分かる。図 3 からは夜や明け方は少なく, 朝や昼, 夕方に多く特に 7 時から 11 時の午前中が交通事故の発生が多いのが分かる。また 7 時から 11 時でも月の時を同様に週末により交通事故が発生しているのが分かる。

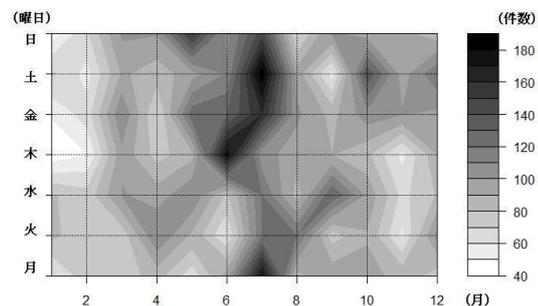


図 1 曜日と月ごとの交通事故総数

図 3, 図 4 は事故形態別交通事故発生件数のグラフである。図 3 は乾燥時の時間ごと, 図 4 は湿潤時の時間ごとのグラフである。縦軸が交通事故件数で横軸が時間であり, 線の色ごとに緑が車両接触, 赤が追突, 黒が施

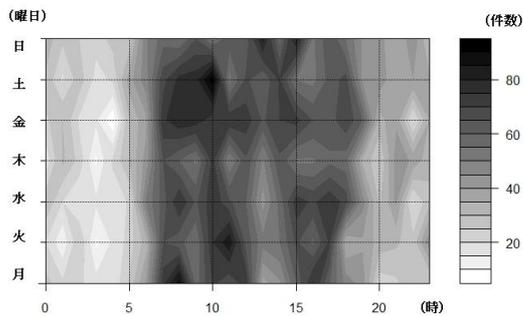


図2 曜日と時間ごとの交通事故総数

施設接触を表している。路面が乾燥時には車両接触と追突は図3から分かるように朝や昼、夕方の交通事故が多く発生する時間帯にどちらも増加しているのが分かる。しかし、施設接触だけはどの時間帯でも同じくらいの発生件数で変化が見られないのが分かる。また車両接触も追突も最も交通事故件数が多いのが10時から11時の間にあるのが分かる。路面が湿潤時には車両接触と追突は朝や昼、夕方の時間帯にすこし増加しているがほとんど変化が見られないのが分かる。しかし、施設接触は図4から分かるように特に交通事故件数が多かった7時から11時の時間帯には減少していて夜や明け方などの交通事故件数が少ない時間帯にピークが来ているのが分かる。

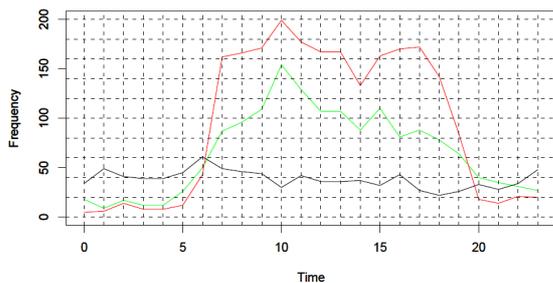


図3 乾燥時の時間ごとの事故形態別交通事故発生件数

4.2 Rによる解析

ここでは、阪神高速道路の交通事故データを使用してRによる解析を行なう。Rで行なう解析は、多変量のデータの合成変量を考へて、少ない次元でデータを把握する統計的方法である主成分分析を行なう。

Rにより主成分分析を行なった結果以下のような結果が得られる。

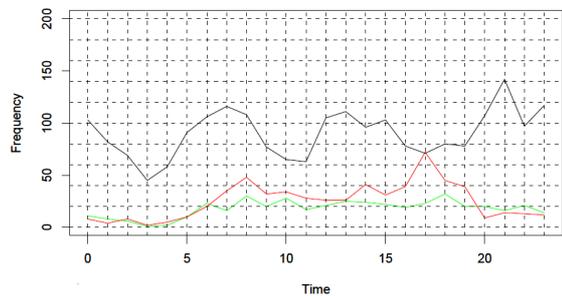


図4 湿潤時の時間ごとの事故形態別交通事故発生件数

	Comp.1	Comp.2
乾燥	0.327	-0.622
湿潤	-0.330	0.617
施設接触	-0.581	-0.363
車両接触	0.231	0.153
追突	0.342	0.204
ハンドルブレーキ操作不適當	-0.341	
前後方不注意	0.397	0.166

第2主成分までの累積寄与率は83.3%となり、ここまでを考察する。第1主成分は乾燥、車両接触、追突、前後方不注意が正で湿潤、施設接触、ハンドルブレーキ操作不適當が負である。したがって正は乾燥時の安全不確認であり、負は湿潤時のスピードの出しすぎであるといえる。第2主成分は湿潤、車両接触、追突、前後方不注意が正で乾燥、施設接触、ハンドルブレーキ操作不適當が負である。したがって正は湿潤時の安全不確認であり、負は乾燥時のスピードの出しすぎであるといえる。

4.3 ArcGISによる可視化

ここでは、ArcGISにより阪神高速道路の交通事故を乾燥時と湿潤時また、それぞれについて施設接触、車両接触、追突の3つについて地図上に表示した結果を示す。発生件数が1-4件、5-10件、11-20件、21-30件、41-79件と発生件数を色の濃さと円の大きさで分けて表示している。表示した結果のうち、特徴がよく見られた結果を図5、図6に示す。図5は乾燥時の施設接触、図6は湿潤時の施設接触の表示結果である。

5 実行結果の考察

5.1 Rによるグラフ化と主成分分析の考察

阪神高速道路での事故形態は路面状況や時間によって変化が見られた。図3と図4から車両接触と追突に関しては乾燥時に多く発生しており、事故件数が時間によって事故件数が大きく変動している。日中に事故が増

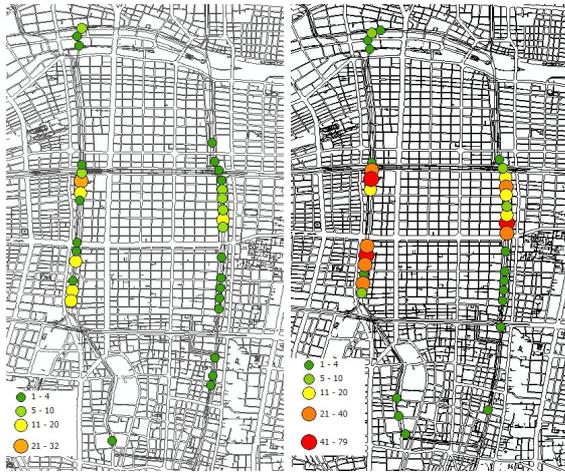


図5 乾燥時の施設接触

図6 湿潤時の施設接触

加しているため交通量と大きく関係しているということが考えられる。施設接触は湿潤時に多くどの時間も湿潤時の方が事故件数が多かった。これは路面状況が事故原因と関係があるためだと考えられる。主成分分析した結果は、第1主成分を見ると正の方には直線の長い路線が多く、負の方には路線の出入り口付近が多いという傾向がある。第2主成分の結果を見ると正の方には急カーブのある路線が多く、負の方には合流、分流地点が多いという傾向がある。

5.2 ArcGISによる可視化の考察

施設接触は乾燥時と湿潤時の交通事故発生場所は似ていて堺線本線合流点や信濃橋出路、長堀入路付近に集中して発生しているが湿潤時の方が多く発生している。次に車両接触は堺線本線合流点など似かよった場所で交通事故が発生しているが乾燥時では空港線合流地点に多く発生していたが湿潤時にはあまり発生していない。最後に追突は乾燥時と湿潤時で多く発生した場所では似ているが乾燥時には湿潤時とまたそれぞれの施設接触と車両接触とも違い堺線本線合流点や信濃橋出路、長堀入路付近以外にも夕陽丘入出路や松原線分岐、合流点、えびす町入路、なんば出路付近など環状線kp7-9付近など全体的に交通事故が発生していた。乾燥時の事故件数のが湿潤時よりも多かったが、実際には雨の降った日数は晴れの日よりも少ない。そこで年間の天気との割合と事故件数の乾燥と湿潤の割合を調べた。大阪府は年間降水日数が2000年は93日、2005年は92日、2007年は98日であった。また、雪日数は2000年は14日、2005年は10日、2007年は5日であった。晴れと曇りの日数は明確なデータがないため、365日から雨と雪の降った日数を引いた日数とする。すると2000年は258日、2005年は

263日、2007年は262日である。そのため晴れと曇りの日数と雨と雪の日の日数の割合は各年とも約7:3である。そのため晴れと曇りの日の路面状況を乾燥、雨と雪の日の路面状況を湿潤とすると、乾燥時と湿潤時で事故発生率が同じなら事故件数は約2.3倍となる。施設接触の事故件数は乾燥時に179件、湿潤時に464件である。天候の割合を考慮すると湿潤時は乾燥時に比べ5.9倍事故が起きやすいと考えられる。車両接触の事故件数は乾燥時に127件、湿潤時に55件である。天候を考慮すると1.0倍であるため車両接触は路面状況は関係がないと考えられる。追突の事故件数は乾燥時に342件、湿潤時に86件である。天候を考慮すると乾燥時は湿潤時に比べ1.7倍事故が起きやすいと考えられる。

6 おわりに

阪神高速道路環状線において路面状況と事故の形態ごとに可視化を行った結果、施設接触は湿潤時に乾燥時の5.9倍事故が起きやすく、車両接触は路面状況にあまり関係なく、追突は乾燥時に湿潤時の1.7倍事故が起きやすいという特徴がわかった。また事故発生場所において、施設接触は堺線本線合流地点や信濃橋出路、長堀入路付近、車両接触は空港線分岐地点付近、追突は堺線本線合流地点付近で比較的多く発生していることがわかった。しかし、本研究で扱った阪神高速道路の交通事故データは、土佐堀出路や北浜出口、渡辺橋分岐、なんば出口付近のデータが不足していたため最も多くの事故が発生している場所を正確に特定することができなかった。そのため、不足したデータを追加して分析や地図上への表示が必要となる。また、交通事故発生場所の詳細な周辺情報も含めて正確に交通事故が多く発生する原因の更なる検討が今後の課題となる。

参考文献

- [1] Nada Lavrac, Domen Jesenovec, Nejc Trdin, Neza Mramor Kosta, "Mining Spatio-temporal Data of Traffic Accidents and Spatial Pattern Visualization," *Metodoloski zvezki* Vol.5, No.1, pp.45-63, 2008.
- [2] 菅野寿美朗, 武藤心吾, "時空間 GIS を用いた交通障害情報の視覚化," *南山大学数理情報学部情報通信学科卒業論文*, pp.184-187, 2006.
- [3] Linhua Li, Li Zhu, Daniel Z Sui, "A GIS-based Bayesian approach for analyzing spatial-temporal patterns of intra-city motor vehicle crashes," *Journal of Transport Geography*, Vol.15, pp.274-285, 2007.