

# 博士論文

道路変化イベントの把握のための機械学習  
を利用した統合的システムに関する研究

小 林 亘



# 目次

## 道路変化イベントの把握のための機械学習を利用した統合的システムに関する研究

|                                       |    |
|---------------------------------------|----|
| 第1章 序論                                | 1  |
| 1.1 はじめに                              | 1  |
| 1.2 研究の背景                             | 1  |
| 1.2.1 日本における道路の変化量                    | 1  |
| 1.2.2 道路変化イベントの情報を取り巻く現状              | 2  |
| 1.2.3 道路変化イベントの把握で得られる一般の道路利用者の便益とニーズ | 8  |
| 1.2.4 道路変化イベントの情報の収集の現状               | 10 |
| 1.2.5 研究の背景のまとめ                       | 12 |
| 1.3 先行研究                              | 12 |
| 1.3.1 先行研究に関する情報（原典情報）                | 12 |
| 1.3.2 公示情報を対象とした先行研究                  | 14 |
| 1.3.3 工事発注見通し情報を対象とした研究               | 15 |
| 1.3.4 工事入札公告を対象とした研究                  | 15 |
| 1.3.4 先行研究のまとめと明らかになった課題              | 16 |
| 1.4 研究の目的と取り組み                        | 17 |
| 1.5 論文の構成                             | 19 |
| 参考文献                                  | 22 |
| 第2章 原典情報の関連づけの手法とシステムデザイン             | 25 |
| 2.1 はじめに                              | 25 |
| 2.2 道路変化イベントに関する原典情報の構造化              | 25 |
| 2.2.1 供用開始情報                          | 25 |
| 2.2.2 道路区域の決定                         | 27 |
| 2.2.3 道路区域の変更                         | 29 |
| 2.2.4 道路開通情報                          | 31 |
| 2.2.5 工事情報                            | 33 |
| 2.3 原典情報の関連づけ手法                       | 40 |
| 2.3.1 原典情報の関連づけのためのキーの選定              | 40 |
| 2.3.2 場所による集約                         | 41 |
| 2.3.3 時間による整理                         | 47 |
| 2.3.4 場所と時間による関連づけ                    | 50 |
| 2.4 統合的システムのデザイン                      | 51 |
| 2.4.1 原典情報の投入から地図上での表示までの流れ           | 52 |
| 2.4.2 ユーザインタフェース                      | 52 |
| 参考文献                                  | 55 |

|  |    |
|--|----|
| 第3章 原典情報の関係の基礎分析.....                                | 57 |
| 3.1 公示情報の間の関係分析.....                                 | 57 |
| 3.1.1 使用した情報.....                                    | 57 |
| 3.1.2 調査結果.....                                      | 58 |
| 3.1.3 供用開始と道路区域情報の関係のまとめ.....                        | 62 |
| 3.2 道路の供用の開始と工事完了件数の関係の分析.....                       | 62 |
| 3.2.1 使用した情報.....                                    | 62 |
| 3.2.2 調査結果.....                                      | 63 |
| 3.2.3 供用開始と工事完了件数の関係のまとめ.....                        | 64 |
| 3.3 道路の開通と工事の関係の分析.....                              | 65 |
| 3.3.1 使用した情報.....                                    | 65 |
| 3.3.2 調査結果.....                                      | 65 |
| 3.3.3 道路開通と工事完了件数の関係のまとめ.....                        | 67 |
| 参考文献.....  | 68 |
| 第4章 工事情報からの道路変化イベントの簡易的な予測.....                      | 69 |
| 4.1 工事情報からの予測により得られる効果.....                          | 69 |
| 4.2 予測の評価方法.....                                     | 70 |
| 4.2.1 実際と予測結果の組み合わせの分類.....                          | 70 |
| 4.2.2 予測結果の評価.....                                   | 70 |
| 4.3 工事件数による簡易的な開通予測.....                             | 71 |
| 4.3.1 調査データ.....                                     | 72 |
| 4.3.2 開通の有無と工事件数に関する調査.....                          | 72 |
| 4.3.3 工事件数からの予測と評価.....                              | 76 |
| 参考文献.....  | 78 |
| 第5章 機械学習を用いた工事内容による予測.....                           | 79 |
| 5.1 機械学習を利用した予測の概要.....                              | 79 |
| 5.2 予測に用いる変数.....                                    | 80 |
| 5.2.1 工事内容から説明変数への変換.....                            | 81 |
| 5.3 教師データ.....                                       | 83 |
| 5.3.1 正例の選定.....                                     | 83 |
| 5.3.2 負例の選定.....                                     | 84 |
| 5.3.3 正例と負例の均衡化.....                                 | 85 |
| 5.4 使用するアルゴリズム.....                                  | 85 |
| 5.4.1 ナイーブベイズ分類器 (Naïve Bayes classification).....   | 86 |
| 5.4.2 AODE (Averaged One-Dependence Estimators)..... | 90 |
| 5.4.3 決定木.....                                       | 90 |
| 5.4.4 ランダムフォレスト.....                                 | 94 |

|   |     |
|---|-----|
| 5.5 機械学習を用いた工事内容による予測の実験.....           | 94  |
| 5.5.1 実験の方針.....                        | 94  |
| 5.5.2 実験方法.....                         | 95  |
| 5.5.3 道路開通に対する予測実験.....                 | 97  |
| 5.5.4 供用開始に対する予測実験.....                 | 102 |
| 5.6 予測の正確さの向上.....                      | 108 |
| 5.6.1 データ：量と質の向上.....                   | 108 |
| 5.6.2 説明変数の改良.....                      | 109 |
| 5.6.3 スパースデータに適したデータマイニングアルゴリズムの採用..... | 109 |
| 5.6.4 工事の順序関係（マルコフ連鎖）を利用する方法.....       | 115 |
| 5.7 件数による予測，機械学習による予測の利用の位置づけ.....      | 117 |
| 参考文献.....                               | 121 |
| 第6章 情報の流通とシステムの有用性に関する評価.....           | 124 |
| 6.1 情報の流通への取り組み方針.....                  | 124 |
| 6.2 流通の対象とする情報について.....                 | 124 |
| 6.3 道路変化イベントの公表について.....                | 125 |
| 6.4 工事情報の流通について.....                    | 126 |
| 6.4.1 統一的なフォーマットの提案と提供可能性.....          | 126 |
| 6.5 システムの有用性の評価.....                    | 127 |
| 第7章 総括.....                             | 129 |
| 7.1 本研究の成果のまとめ.....                     | 129 |
| 7.2 本研究の特徴.....                         | 131 |
| 7.2.1 本研究の独自性.....                      | 131 |
| 7.2.2 本研究の有用性.....                      | 132 |
| 7.3 課題.....                             | 133 |
| 7.4 今後の展望.....                          | 134 |
| 7.4.1 詳細調査との連動.....                     | 134 |
| 7.4.2 道路変化イベント以外の予測への拡張.....            | 135 |
| 7.5 おわりに.....                           | 135 |
| 付録.....                                 | 136 |
| 付録1 場所から地点・路線への関連づけ.....                | 136 |
| 付録2 次元圧縮の例.....                         | 137 |
| 付録3 負例の整理方法.....                        | 138 |
| 付録4 教師データの生成.....                       | 144 |
| 付録5 テストデータの生成.....                      | 145 |
| 付録6 データ関係 URL.....                      | 146 |
| 発表論文リスト.....                            | 148 |



## 第1章 序論

### 1.1 はじめに

日本においては多くの道路の変化があり、例えば、2013年から2014年には2,211kmの道路延長の増加が見られる<sup>4)</sup>。道路の利用に影響のある道路の変化を本研究では「道路変化イベント」と呼ぶこととする。道路を有効に利用するためには、道路変化イベントを把握することが必要である、そのためには、例えば、乗用車の7割に普及している<sup>30)</sup>カーナビゲーションの地図が更新されなければならない。実際のところ、道路の変化に対する情報の提供<sup>6)</sup>は道路管理者によって、情報の収集<sup>20)</sup>は地図調製者によって実施されている。しかし、地図の更新は道路の変化に追いついておらず、カーナビゲーション利用者に対するアンケート(2007年、2012年)では、カーナビゲーションが新たな道路に対応していないことが不満の第一位となっている<sup>21),31)</sup>。

このように、関係者による道路変化の提供と収集の努力はなされているにも関わらず、道路利用者が不満を示しているという状況に対し、その改善を図るための研究が行われてきた。先行する研究を調査した結果、本研究において主に3点の課題を見出した。第一章では、このような道路変化イベントを取り巻く現状と問題を研究の背景として概説し、先行研究から導かれた課題と、その解決に向けた本研究の取り組みを述べる。

### 1.2 研究の背景

#### 1.2.1 日本における道路の変化量

道路法第一条では同法の目的を、道路網の整備を図るため、道路に関して諸事項を定め、「・・・、もって交通の発達に寄与し、公共の福祉を増進すること」としている。このように交通の発達に寄与し、公共の福祉を増進するために整備されている道路網に対する予算規模は、国土交通省道路局・都市局から平成28年1月に発表された平成28年度道路関係予算概要によると37,784億円となっている。そのうちの約41%を占める直轄事業には内訳が示され、72%が改築その他、20%が維持修繕、8%が諸経費等に充てられることとなっている<sup>1)</sup>。予算が工事で執行された量は、例えば、建設工事受注動態統計調査で確認することができる。これによると公共機関の道路工事は2009年度には工事件数で56,280件、請負契約額で29,380億円であった<sup>2)</sup>。道路建設の最終段階である舗装工事について見たとき、一般社団法人日本道路建設業協会のまとめによると2009年度の実施量は2,884億円であった。そして、道路の全舗装費のうち45.2%が舗装新設に充てられていた<sup>3)</sup>。

このように道路を建設する過程に着目して変化量を求めることができる一方、道路の変化量そのものを国土交通省道路局が公表している道路統計年表から得ることができる。道路統計年表2013と同2014を比較すると、道路の実延長で道路統計年表2013の1,214,917kmに対して同2014では1,217,128kmと2,211kmの増加、道路と鉄道の立体交差箇所26,022箇所に対して26,388箇所と366箇所の増加、道路と鉄道の平面交差で30,150箇所から29,959

箇所と 191 箇所の減少を示している<sup>4)</sup>。

これらの統計量の他に、後述する道路の供用開始の告示の数などから、日本において道路の変化は多数発生していることが確かめられる。

### 1.2.2 道路変化イベントの情報を取り巻く現状

本研究は道路変化イベントを把握することを目指している。本研究では、道路変化イベントの定義を「道路の利用に影響のある道路の変化」とする。すなわち、道路のルートや形状が変化して、それらが実際に道路利用に供される（供用される）ものを対象とする。具体的には、除草や舗装の維持修繕などのように道路のルートや形状が変化しないものや、変化したものであっても通行できない工事中のものは除外され、次が対象となる。

- ① 異なる 2 点を結ぶ新たなルートが供用されるもの（例として、バイパス、橋梁、隧道）
- ② 既存の道路の形状が変更し供用されるもの（例として、拡幅、立体交差化、平面交差点の改良）

#### 1.2.2.1 道路変化イベントに関する道路管理者の提供情報

道路変化イベントを示す情報のうち、道路管理者から提供されるものを以下に示す。道路管理者とは道路の新設、改築、維持などの道路法上の管理行為を行うものであり、その数は、2014 年 4 月現在で国土交通大臣(1)、都道府県(47)、市町村(1718)の合計 1,766 である。

##### (a) 道路法に基づく道路の供用の開始の公示

道路法第十八条（道路の区域の決定及び供用の開始等）第二項では、「道路管理者は、道路の供用を開始し、又は廃止しようとする場合においては、国土交通省令で定めるところにより、その旨を公示し、・・・」と定められている（公示とは、広く知らせるための発表）。対象の規模の大小は問わないため、小規模なもの、例えば、延長距離の短いものや拡幅の幅の小さなものであっても全てが公示されることとなる。

図 1.2-1 に道路の供用開始の告示（告示という形態で公示されたもの）の例を示す。

このように新たに供用された道路が有効に利用に供されるよう、道路法ではその情報の公示を道路管理者に求めている。同法第一項はまた、道路の供用に先立って路線（道路の起点から終点）が定まった時点で、遅滞なく道路区域（道路を構成する土地の範囲）を決定して公示すること、道路区域を変更したときも同様とすることを定めている。本研究では同法に基づく道路の供用の開始の告示を「供用開始情報」、道路区域の決定または変更の告示を「道路区域情報」と呼ぶこととする。

| 新潟市告示第 2 1 8 号   |             | 掲示期間 4.1-4.14  |                  |
|--|-------------|--|------------------|
| 道路の供用開始について  |             |  |                  |
| <p>道路法（昭和 27 年法律第 180 号）第 1 8 条第 2 項の規定に基づき、道路の供用を次のように開始する。</p> <p>なお、関係図面は告示の日から 2 週間新潟市土木部土木総務課において、一般の縦覧に供する。</p> <p>平成 2 2 年 4 月 1 日</p> <p style="text-align: right;">新潟市長 篠 田 昭</p> |             |  |                  |
| 道路の種類  | 路線名         | 区 間  | 供用開始年月日          |
| 市道   | 北 2-115 号線  | 新潟市北区太郎代字水汲湯 2880 番 1 地先から<br>新潟市北区太郎代字水汲湯 2841 番 1 地先まで | 平成 2 2 年 4 月 1 日 |
| 市道   | 豊栄 1-820 号線 | 新潟市北区横土居字村前 304 番地先から<br>新潟市北区横土居字村前 291 番 1 地先まで        | 平成 2 2 年 4 月 1 日 |
| 一般国道   | 1 1 3 号     | 新潟市東区下山三丁目 1411 番 1 地先から<br>新潟市東区下山三丁目 1333 番 1 地先まで     | 平成 2 2 年 4 月 1 日 |

図 1.2-1 道路の供用開始の告示の例<sup>5)</sup>

(b) 一部の自治体が独自に公表している道路開通情報

一部の自治体では、法令に依らない独自の行政サービスとして供用開始情報のうち、重要なものを整理して道路開通情報として公表している。一部の自治体の独自のサービスであるため、対象の選択に関して統一的な定めは無い。図 1.2-2 に道路開通情報の例を示す。

表 1.2-1 は、供用開始情報と道路開通情報をインターネットで公開している自治体の一部について供用開始件数と道路開通情報の件数を調べたものである。

道路開通情報の区間長について、埼玉県、愛媛県、新潟市の 2007 年度から 2010 年度の 41 件の道路開通情報を調査したものを表 1.2-2 に示す。それぞれの自治体の道路開通の平均区間長は 1.0~1.3km、最短は 200m、最長は 7km であり、内訳は新たなルート（バイパス、小規模なショートカット）、橋梁の新設、架け替え、トンネル、既存の交差点の立体交差化であった。該当区間の長さを記している供用開始情

報と道路開通情報の例を表 1.2-3 に示した。

| 新潟市 Niigata City<br>みなとまち。みらいまち。新潟市  |           |              |            |              |
|--|-----------|--------------|------------|--------------|
| <a href="#">現在のページ</a> <a href="#">トップページ</a> <a href="#">くらし・手続き</a> <a href="#">道路と交通</a> <a href="#">道路整備・維持管理</a> <a href="#">道路開通情報</a> |           |              |            |              |
| 道路開通情報   |           |              |            |              |
| 道路開通情報一覧表  |           |              |            |              |
| 音声読み上げソフトをご利用される方へご案内: 開通情報の単位はキロメートルです。   |           |              |            |              |
| 平成24年度   |           |              |            |              |
| 現在、開通情報はありません。   |           |              |            |              |
| 平成23年度   |           |              |            |              |
| 開通路線一覧   |           |              |            |              |
| 路線名  | 開通箇所      | 開通延長(キロメートル) | 開通日        | 備考           |
| 国道460号大曾根バイパス  | 西蒲区今井～大曾根 | 0.2          | 平成23年6月17日 |              |
| 都市計画道路太平大淵線  | 東区岡山～東区江口 | 1.2          | 平成24年3月27日 | 都市計画決定幅員整備完了 |
| 国道113号横土居バイパス  | 北区横土居     | 1.2          | 平成24年3月29日 |              |

図 1.2-2 道路開通情報の例<sup>6)</sup>

表 1.2-1 供用開始情報と道路開通情報の件数<sup>5),7),8),9),10),11)</sup>

| 自治体名 | 年度   | 供用開始件数 | 道路開通情報件数 |
|------|------|--------|----------|
| 新潟市  | 2013 | 158    | 4        |
| 新潟市  | 2014 | 249    | 3        |
| 島根県  | 2013 | 145    | 1        |
| 高知県  | 2014 | 68     | 3        |

表 1.2-2 道路開通情報に含まれる開通区間長(2007-2010)<sup>6),12),13),14),15),16)</sup>

| 自治体名 | 区間数 | 区間長(km) | 平均区間長(km) |
|------|-----|---------|-----------|
| 新潟市  | 14  | 0.2～2.5 | 1.1       |
| 愛媛県  | 11  | 0.6～1.8 | 1.0       |
| 埼玉県  | 16  | 0.2～7.0 | 1.3       |

表 1.2-3 供用開始情報と道路開通情報の例

(a) 供用開始情報

供用開始情報は延長を示していた道路管理者の自治体の公報より抜粋。なお、供用開始情報の全てに延長が記されるものではない。

| 道路種別 | 路線          | 区間  | 発生日<br>(供用日) | 道路管理者 | 概要     |
|------|-------------|---|--------------|-------|--------|
| 県道   | 294号奥の谷日比原線 | 吾川郡いの町清水上分字種々ヤシキ1571番1から吾川郡いの町清水上分字種々ヤシキ3427番まで | 平成27年2月17日   | 高知県   | 22メートル |
| 県道   | 三次江津線       | 江津市川平町南川上973番1地先から江津市川平町南川上973番1地先まで            | 平成26年7月11日   | 島根県   | 17メートル |

(b) 道路開通情報

| 道路種別  | 路線     | 区間                  | 発生日<br>(開通日) | 道路管理者 | 概要     |
|-------|--------|---------------------|--------------|-------|--------|
| 主要地方道 | 長岡栃尾巻線 | 新潟市南区清水～西蒲区六分(両郡橋)  | 平成21年3月29日   | 新潟市   | 1.9km  |
| 国道    | 299号   | 埼玉県飯能市長沢～井上(吾野トンネル) | 平成27年3月7日    | 埼玉県   | 約0.9km |

供用開始情報と道路開通情報の件数の差、供用開始情報と道路開通情報の内容の違いが示すように、道路開通情報は道路利用者の道路の選択に大きな影響を与える道路変化イベントのみを抽出して公表している。そして、道路開通情報として公表される供用の開始は、新たなルート、橋梁、隧道の開通、立体交差化といった機能、規模、地域に対する影響の大きさ(代替可能性や交通量、混雑度)などが総合的に判断されたものとなっている。

1.2.2.2 道路変化イベントに関する情報の収集者のリクワイヤメント

道路変化イベントに関する情報について、デジタル道路地図情報を収集し、カーナビゲーションシステムや電子地図に提供している<sup>17)</sup>一般財団法人日本デジタル道路地図協会(DRM)のリクワイヤメントを調査した。この機関を選定した理由は、①カーナビゲーション用の基盤となるデジタル道路地図データを共通化して民間企業に提供していること<sup>18)</sup>、②その提供の実績が平成27年6月までの累計で約82百万枚(DRMデータベースの提供枚数)に達し<sup>19)</sup>、十分な実績があると判断したためである。

同協会が道路管理者に情報の提供を要請している道路構造物・道路施設は、橋・高架(延長15m以上のもの)、トンネル、踏切、立体交差箇所、中央分離帯、横断歩道橋、料金所等の新設・増設、移設、廃止、名称の変更など多様である。そして、提供を求める時期は、都道府県道以上について供用の2年前となっている<sup>20)</sup>。このように早期に道路変化イベントの情報を必要とする理由は、情報を入手してから確認、データ作成等に時間(利用者の地図データが更新されるまでのリードタイム)が必要であるためである。リードタイムの内訳をカーナビゲーション地図の更新工程として図1.2-3に示す。

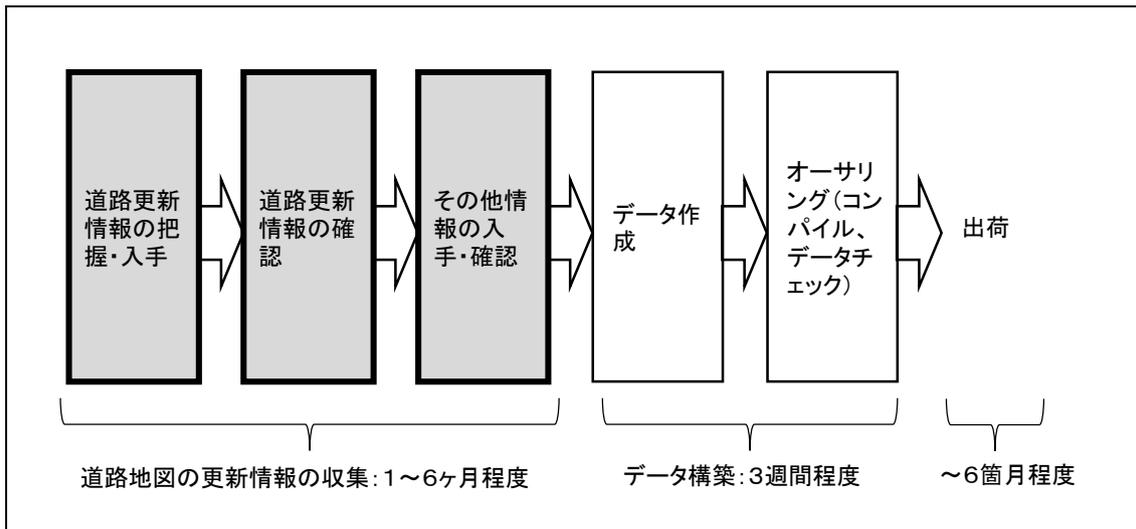


図 1.2-3 カーナビゲーション地図の更新工程

関本義秀，金澤文彦，松下博俊：次世代デジタル道路地図のあり方に関する研究<sup>21)</sup>，pp.48，図 4.2-2 情報作成方法を元に作成

また，中條氏（2012）により整理された道路構造変化の内容では，新規路線・バイパス等の新規供用，車線数の増加を伴う拡幅，道路中心線が概ね 15m 以上移動する線形改良などを流通を目指す情報として挙げている<sup>22)</sup>。

このほか，特定非営利活動法人 ITS Japan からは，種々の経路案内サービス，走行支援サービスに必要な道路情報の種別，内容，項目がまとめられ，これらの情報の流通が提言されている<sup>23)</sup>。

このように収集者のリクワイヤメントは多岐にわたる。

### 1.2.2.3 本研究で対象とする道路変化イベント

道路変化イベントの情報収集者である地図調製者等にとっては，自らのリクワイヤメントに基づいた情報が得られることが望ましいであろう。しかし，前項で述べたように収集者のリクワイヤメントはサービスによって異なるものがあるうえ，その情報を完全に把握することは，情報収集の専門部署を擁する地図調製者においても実現できておらず，現在は真値となるデータが得られていない。加えて，地図調製者が情報を収集する際に重視する点は，現在の地図と実態との差であり，既に変化した道路の変化の時期は問題ではないために，収集された情報には時期に関する情報が含まれない場合があるという問題がある。

これに対して，道路管理者の提供する道路変化イベントの情報には次の利点がある。

- ① 布施ら（2009）の研究<sup>24)</sup>で明らかにされたように，道路管理者の提供情報は地方公共団体，自動車・カーナビゲーション・地図関係の企業・団体の 15 組織へのアンケートで約 9 割の組織から有用との評価が得られている。
- ② 道路管理者の公表する情報は，行政機関であるために情報自体の精度が高く，収集の過程で不確定さが入り込まないため，真値として扱える。

③ いずれも場所と時期が示されている。

以上の理由から、本研究においては、対象とする道路変化イベントの真値として、道路管理者から公表されている次の(a), (b)を採用する。

(a) 道路法に基づく道路の供用の開始の公示（供用開始情報）

(b) 一部の自治体が独自に公表している道路開通情報

#### 1.2.2.4 道路変化イベントの情報の有用性

図 1.2-1, 図 1.2-2 に示した例のように、道路変化イベントは、道路の詳細な形状を示す情報では無いため、この情報だけでは道路地図を更新することはできない。しかし、詳細な情報を入手するために行う道路管理者に対する問い合わせや情報公開請求等において、その実施の前には、対象とする自治体、道路変化イベントの箇所や関係する工事などを絞り込むことが必要である。絞り込みができなければ、多くの問い合わせや多くの工事図面が必要となり、情報の請求者・提供者の双方に作業の負担を強いることとなる。さらに、図面等を入手した後に、地図を更新するために必要なものを選び出すことにも多大な労力が必要となる。

例えば、鹿児島県では、情報公開等に関する平成 21 年度の運用状況を公表しており、それによると請求内容の内訳の一位は道路位置図・平面図の 51 件であった<sup>25)</sup>。新潟県の情報公開開示請求は、政策研究大学院大学のデータベースから入手することができる<sup>26)</sup>。それによると新潟県に対する道路地図に関する情報公開請求は 2007 年度から 2009 年度の平均で年 87 件であった（警察と用地担当分を除く道路管理課と地域整備部対応分）。一部の例を表 1.2-4 に示す。このように、道路に関わる詳細な情報の収集は数多く行われ、情報公開請求に対する事務作業が請求者・公開者の双方に発生している。

これに対して、布施ら（2009）の研究<sup>24)</sup>で示されたように、道路変化イベントの情報によって対象を絞り込むことができれば、情報収集における図面などの提供依頼・開示請求の作業が効率化される。

表 1.2-4 情報公開請求の例（政策研究大学院大学、情報公開開示請求データベース）

| 請求日       | 文書名                | 請求件名               | 部課名           |
|-----------|--------------------|--------------------|---------------|
| 2009/2/26 | 道路の区域変更等に係る位置図等... | 道路の区域変更等に係る位置図等... | 道路管理課         |
| 2009/2/20 | 道路平面図              | 道路平面図              | 長岡地域振興局 地域整備部 |
| 2009/3/6  | 道路の区域変更等に係る位置図等... | 道路の区域変更等に係る位置図等... | 道路管理課         |
| 2009/3/18 | 道路の区域変更等に係る位置図等... | 道路の区域変更等に係る位置図等... | 道路管理課         |
| 2009/3/30 | 道路の区域変更等に係る位置図等... | 道路の区域変更等に係る位置図等... | 道路管理課         |
| 2009/3/30 | 道路の区域変更等に係る位置図等... | 道路の区域変更等に係る位置図等... | 道路管理課         |
| 2009/3/31 | 道路の区域変更等に係る位置図等... | 道路の区域変更等に係る位置図等... | 道路管理課         |
| 2009/3/24 | 道路の区域変更等に係る位置図等... | 道路の区域変更等に係る位置図等... | 道路管理課         |
| 2009/4/14 | 道路台帳附図             | 道路台帳附図             | 道路管理課         |

### 1.2.3 道路変化イベントの把握で得られる一般の道路利用者の便益とニーズ

#### 1.2.3.1 道路変化イベントの把握で得られる一般の道路利用者の便益

道路変化イベントを把握して最適な道路を選択することにより、移動時間の短縮、省エネ、事故の削減などの便益が得られる。例として、釜石市と遠野市を結ぶ国道 283 号線に対して平成 19 年 3 月に開通した仙人峠道路（図 1.2-5）の開通を把握し、これを利用することにより得られる便益の一部を表 1.2-5 に示す。



国土交通省三陸国道事務所釜石維持出張所の管理区間図

([http://www.thr.mlit.go.jp/sanriku/10\\_iji/kamaishi/kanrikukan/kanrikukan.html](http://www.thr.mlit.go.jp/sanriku/10_iji/kamaishi/kanrikukan/kanrikukan.html)) に一部加筆



一般国道 283 号線のトンネル



一般道 283 号「仙人峠道路」  
枯松沢橋と新仙人トンネル

写真は国土交通省三陸国道事務所「一般国道 283 号『仙人峠道路』とその開通効果」  
([http://www.thr.mlit.go.jp/sanriku/07\\_office/kouka/sennin/](http://www.thr.mlit.go.jp/sanriku/07_office/kouka/sennin/)) より

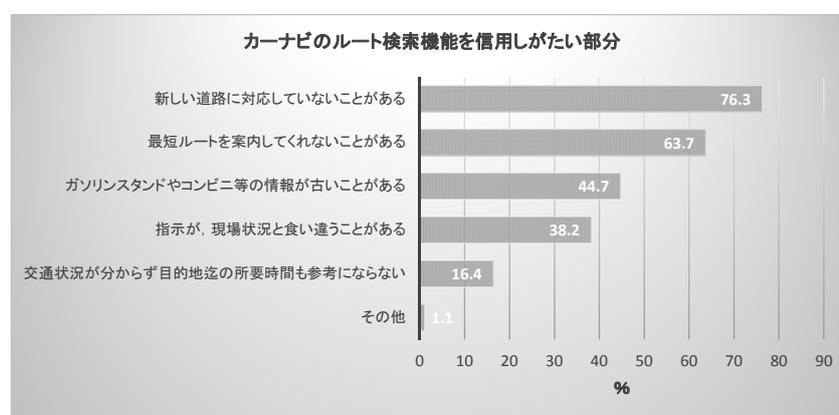
図 1.2-5 一般国道 283 号と仙人峠道路

表 1.2-5 仙人峠道路の開通を把握し利用する便益の例

| 効果の分類                           | 国道 283 号 (旧道) のみを利用できるとした場合 | 国道 283 号に加えて仙人峠道路が選択できる場合                                      | 備考                                     |
|---------------------------------|-----------------------------|--|--|
| 移動時間の短縮 <sup>29)</sup>          | 90 分<br>(救急車不在時間 3 時間)      | 40 分<br>(救急車不在時間 2 時間)   | 上有住地区の通報から釜石病院に収容されるまでの時間 (大船渡消防署住田分署) |
| 省エネ <sup>27)</sup>              | 34km,<br>2000cc 燃費 10km/l   | 29km,<br>2000cc 燃費 11km/l                                      | 釜石・遠野間の例                               |
| 交通事故発生件数 <sup>28)</sup>         | 16 件/年 (国道 283 号)           | 8 件/年 (国道 283 号と仙人峠道路の合計)                                      |  |
| 冬季 (12~2 月) 死傷事故 <sup>29)</sup> | 5 件/年 (平成 14-18 年平均)        | 1 件/年 (平成 19-20 年平均)   | 交通事故総合分析センター・釜石警察署・遠野警察署               |
| 観光・地域の活性化                       |                             | 大槌・池袋線の新設 <sup>27)</sup><br>青森からの新規バスツアー <sup>29)</sup> 24 本新設 |  |

### 1.2.3.2 道路変化イベントの把握に対する一般の道路利用者のニーズ

平成 26 年度の情報通信白書では、カーナビゲーションの 2013 年の出荷台数は 546.6 万台に達しており、7 割程度の乗用車に搭載されていると報告されている<sup>30)</sup>。このことから、現在、多くの道路利用者はカーナビゲーションを道路の選択に用いていると考えられ、道路変化イベントが確実に収集されているかどうか、また、そのニーズの有無についてはカーナビゲーションに対する利用者のアンケートから知ることができる。2007 年に行われた国土交通省の調査 (n=584)<sup>21)</sup>、2012 年に行われた民間企業の調査 (n=262)<sup>31)</sup>のいずれにおいても利用者の不満の第一位は新しい道路にカーナビゲーションが対応していないことであった。このことから、一般の道路利用者は道路変化イベントの把握に対してニーズがあり、しかし、現在の道路地図は道路の変化に十分に対応できていないことが伺える。民間企業による調査結果を図 1.2-6 に示す。



カーナビのルート検索に対して自己判断も織り交ぜる必要ありと回答した n = 262 名

図 1.2-6 運転とカーナビに関する意識調査 (株式会社ミックウェア, 2012)

道路変化イベントの把握は、このように現状の道路利用において必要性があるばかりでなく、実用化が待たれる新たなサービスにおいても必要とされている。先に挙げた ITS Japan による経路案内サービスや走行支援サービス<sup>23)</sup>のほか、準天頂衛星による高精度測位を利用した新たなサービスの創出においては高精度な測位に対応した地図の整備が全てのサービス分野にとって乗り越えるべき課題として整理されている<sup>32)</sup>。また、自動運転においては、距離センサーと地図のフュージョン技術の研究が行われ、現実世界と速やかに正確に一致する地図データが求められている<sup>33)</sup>。

#### 1.2.4 道路変化イベントの情報の収集の現状

一般の利用者のアンケートでは不満が多いカーナビゲーション地図の更新であるが、地図調製者は多大な経費と労力を費やし、次の方法で道路の変化に関する情報を収集している。

##### ① 道路管理者への要請<sup>20),34)</sup>

日本デジタル道路地図協会のデジタル道路地図データベースの作成・更新方法<sup>20)</sup>、DRM ニュース<sup>34)</sup>中の事業計画により、同協会から道路管理者に対して道路変化イベントに先じて、情報の提供を要請していることを知ることができる。このように供用より前に情報の入手を目指す理由は、情報を入力してから道路変化イベントに対応するための準備（例えば、地図データの変更と製品化）に必要なリードタイムを確保するためである（図 1.2-5 を参照）。ただし、供用より 2 年前の工事図面を依頼しているのは都道府県以上の幹線道路であり<sup>20)</sup>、関係省庁から入手できなかった市町村道路、港湾道路、農林道等については前年度の変化分を後追いで入手するに留まっている<sup>34)</sup>。

この方法は、提供された情報が道路管理者によって選別されたものであるため、入手した情報の処理効率が良いという利点がある。しかし、法律に基づかない作業を道路管理者に要請しており、(a)道路管理者が必ずしも要請に応えるとは限らないこと、(b)厳密な審査が期待できず、作成担当者のミスやエラーが混入している可能性があること、という欠点があるため網羅性に課題がある。また、依頼者において道路管理者の数が 1700 以上と多いことから作業負担が大きいという欠点がある。

##### ② インターネットや報道の利用

インターネットや新聞地方紙等の報道に道路開通情報が掲載されることがある。日本デジタル道路地図協会や、地図調製企業では情報収集担当者が、これらを休むことなく情報収集している<sup>35)</sup>。

この方法の利点は、手軽であることであるが、欠点として、インターネットで入手できる道路変化イベントは一部に限られ、網羅性が確保されないほか、2 年後の開通のような長期的な情報が得られることは一般的にはなく、開通日の直前直後にしか情報が入手できないことが挙げられる。

以上のほか、道路変化の詳細を入手するために、次のような方法が利用されている。

③ 国土地理院の地図の利用<sup>20),34)</sup>

日本デジタル道路地図協会では、国土地理院の基盤地図情報及びこれに基づく電子国土基本図等を収集している。この方法は、網羅性は高いが、道路変化イベントの発生より前に情報を入手することが難しい。

④ 現地調査<sup>36),37)</sup>

現地を調査することにより情報を収集することが確実であることは言うまでもない。地図調製企業では、例えばモバイル・マッピング・システム（MMS）などの様々な手段で現地の調査を行っている。

この方法の利点は確実であることであるが、欠点として日本の道路の総延長は 2012 年 4 月 1 日時点で 1,273,620.4 km（国土交通省。林道，農道等は含まず）であり，全ての道路を捕捉することは多大な経費がかかることと，開通前に把握することが難しいことが挙げられる。現地調査に必要な経費については，関本らによる「次世代デジタル道路地図のあり方に関する研究」<sup>21)</sup>において，地図更新にかかる費用は各社とも社外秘としているが，推計として新規供用道路調査費用を約 4 億円，交差点調査費用を約 20 億円等とする値が示されている。

上記は，公開資料やヒアリングにより確認した情報の収集の方法である。前述のように道路管理者は道路法に基づいて供用開始を公示している（図 1.2-1 を参照）。しかし，公示をあまねく収集している地図調製者は確認できなかった。公示は道路法に基づくものであり，網羅性が高いという利点があるにも関わらず収集・利用が行われていない理由は以下と考えられる。(a)全国の道路管理者の公示をその都度入手しなければならない。(b)公示は紙媒体で行っている自治体が少なくない，(c)公示の内容を分析調査し，道路の供用に関する内容を抽出することが容易ではない。(d)道路の供用の公示は，その多くが実際の供用の開始と同時期に行われるために，リードタイムが確保できない。

表 1.2-6 は道路変化イベントの情報の収集方法を整理したものである。網羅性，リードタイム，効率の 3 つにおいてニーズを満たすものは無い。このため，情報収集者，例えば地図調製者は幾つかの方法を併用して道路変化イベントに関する情報収集を行っている。

表 1.2-6 道路変化イベントの把握方法の種類と特徴

| 把握方法            | 網羅性 | リードタイム | 効率 |
|-----------------|-----|--------|----|
| ① 道路管理者への要請     | △   | ○      | △  |
| ② インターネットや報道の利用 | ×   | ×      | ○  |
| ③ 国土地理院の地図の利用   | ○   | ×      | ○  |
| ④ 現地調査          | ○   | ×      | ×  |
| 供用開始の告示         | ○   | ×      | ×  |

### 1.2.5 研究の背景のまとめ

これまでに述べた研究の背景の要点は次の通りである。

- ✓ 道路には多くの変化が発生している。
- ✓ 道路の利用に影響のある道路の変化を本研究では「道路変化イベント」と定義した。
- ✓ 道路変化イベントを把握することは多くのステークホルダーにとって有益である。
- ✓ 道路管理者による供用開始の告示（供用開始情報）、道路開通情報を本研究では道路変化イベントの真値とする。
- ✓ 道路利用者の多くはカーナビゲーションを利用しているが、新たな道路へ対応していないことに不満が多い。
- ✓ 道路管理者は法に基づき供用開始の情報を提供している。
- ✓ 地図調製者は、道路管理者の情報提供に依らず、独自の方法で情報を収集している。
- ✓ このように、情報の提供と収集はかみ合っておらず、道路利用者は道路変化イベントの情報を十分に利用できていない。

道路変化イベントを取り巻く現状を図 1.2-7 に示した。

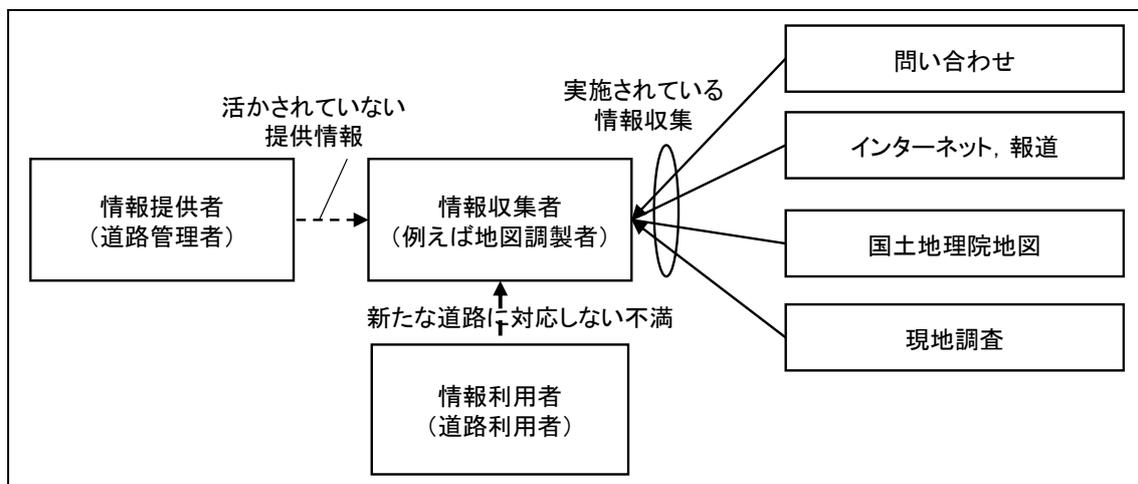


図 1.2-7 道路変化イベントを取り巻く現状

## 1.3 先行研究

### 1.3.1 先行研究に関する情報（原典情報）

先行研究では、道路変化イベントを直接的に示す情報だけでなく、間接的に道路変化イベントを示すと考えられた工事に関する情報も扱っている。ここでは、個々の先行研究について述べる前に、先行研究が扱っている情報を俯瞰し整理する。

道路変化イベントに関して道路管理者等が管理する加工前の一次的な情報（以下、「原

典情報」と呼ぶ)として、道路変化イベントを直接に表す「道路の供用の開始」の告示(供用開始情報)、道路管理者から公表される「道路開通情報」があることを既に述べた。そのほかには、道路変化イベントに間接的に関係する「道路区域」に関する告示(道路区域情報)、道路工事の情報などがある。なお、全ての道路工事が道路変化イベントに関係するとは限らないが、道路工事情報を収集する時点では関係するものだけを取捨選択することができず、取捨選択のルールを見出すこと自体が本研究で取り扱うべき課題の一部である。本研究で取り扱う道路工事の情報は、(a)工事の目的を道路そのものの建設や維持管理であるものであり、かつ、(b)施工場所が道路である(道路の予定地を含む)ものとした。(a)のスクリーニングは、工事を発注する主体が道路管理者であって、道路予算で執行されるものを選ぶことにより抽出できる。(b)のスクリーニングは、施工場所が路線認定されている場所(例えば、山上の道路気象観測施設や道路管理事務所を除く)ことによって実施することができる。このことは、第2章に詳しく述べる。なお、本研究において、公共工事を民間工事と区別する必要がある場合には「公共工事」とし、公共工事あるいはその中の道路工事であることが明らかな場合には単に「工事」とする。

表 1.3-1 原典情報の一覧

| 事業      | 道路区域の決定・変更    | 道路工事                      |                              |                          |                          | 供用の開始                   |                | 管理               |
|---------|---------------|---------------------------|------------------------------|--------------------------|--------------------------|-------------------------|----------------|------------------|
|         | 種別            | 道路変化イベントを直接的に示さない情報(間接情報) |                              |                          |                          | 道路変化イベントを直接的に示す情報(直接情報) |                | 道路変化イベントの後に整理される |
| 原典情報の種類 | 道路区域の決定・変更の告示 | 工事発注見通しの公表                | 工事入札公告                       | 工事実績情報(一部公表)             | 契約内容の公表                  | 道路開通情報(広報)              | 道路の供用の開始・廃止の告示 | 道路台帳の更新(閲覧)      |
| 標準的な時期  |               | 工事開始前(年度当初)               | 工事契約前                        | 工事開始直後                   | 契約後                      | 開通時                     | 供用時            |                  |
| 目的      | 行為制限する道路区域示す  | 入札契約の透明性の確保               | 契約の適正化                       | 工事品質の確保、契約の適正化           | 入札契約の透明性の確保              | 道路利用者へのサービス             | 道路の供用の宣言       | 道路管理             |
| 根拠法令等   | 道路法           | 公共工事の入札及び契約の適正化の促進に関する法律  | 会計法、公共工事の入札及び契約の適正化の促進に関する法律 | 中央建設業審議会、契約図書(土木工事共通仕様書) | 公共工事の入札及び契約の適正化の促進に関する法律 | 無し                      | 道路法            | 道路法              |

先行研究及び本研究に関係する原典情報を表 1.3-1 に示す。路線の認定を終えた道路予定地について、道路変化イベントの発生までには、道路区域の決定・変更、道路工事の施工を経て、道路の供用に至るという事業の流れがある。一般的には表 1.3-1 の事業は左から右へ順に進行することを想定している。このため、事業の進行に連動して左から右への順に情報が発生する。これらのうち、告示、公表、公告は公開される情報であるが、法令では必ずしも電子的な手段による公開を求めているため、その手段の選択は実施する行政機関に任されている。

### 1.3.2 公示情報を対象とした先行研究

道路法第十八条に基づく公示（供用開始情報、道路区域情報）を対象とした研究には、布施らによる「公示情報に基づく道路更新情報のクローリングシステムの検討」（土木情報利用技術論文集 vol18,2009）<sup>24)</sup>がある。クローリングとは、インターネットで公開されているサイトから自動的に情報を取得することである。この研究では次のことが示されている。

#### （1）インターネットにおける公示情報の提供状況（網羅性）

公示は公報に掲載されるため、公報のインターネット上の公開状況を調査した結果、全ての都道府県と 12 の政令指定市では確認できたものの、6 つの都道府県下の市町村に対する調査では約 3%にとどまったことが報告されている。ここから、インターネットによる公示情報の提供が大規模な行政機関に偏っていることが伺える。

#### （2）クローリングによる情報の構造化の困難さ

クローリングシステムによって収集できた件数は 175 件中 157 件であったこと、情報の抽出（スクレイピング）に当たって形態素解析とテキスト構文解析を行っていることから、公示を構造的なデータとすることが容易ではないことが伺える。

#### （3）情報によるリードタイムの確保の困難さ

供用開始時期と公報掲載時期の関係についての調査によって、全ての道路管理者において公示情報は供用開始の数日前から 1 箇月後までの間に公報に掲載されているという結果が報告されている。ここから、公示情報では、実際の道路変化イベントに対してリードタイムが確保されないとと言える。

#### （4）公示情報の有用性

上記研究において試行的に公開した公示情報のクローリングシステムに対する有用性の確認において、「・・・情報収集における図面などの提供依頼・開示請求においては、現状のシステムでの十分に実利用が可能であるとの回答を得ている。」との報告がなされており、15 組織に対するアンケートで 9 割から有用との回答を得たという報告から、道路変化イベントの把握に対する公示情報の有用性は高いと考えることができる。

このように公示の情報は道路変化イベントの把握に有用であるが、インターネットで公開している道路管理者は限定されており、構造的なデータとすることは容易では無く、リードタイムが確保できないという問題が指摘されていた。

### 1.3.3 工事発注見通し情報を対象とした研究

「工事発注見通し情報」を利用する研究も行われている。関本らによる「工事発注見通し情報を用いた全国における道路更新情報の自動抽出に向けた試み」(土木学会論文集 D3 (土木計画学), Vol.68, No.3, 2012)<sup>38)</sup>がそれである。この研究では次のことが報告されている。

#### (1) 網羅性

工事発注見通し情報は、首都圏の自治体の 576 のサイトの 9 割以上で案件名称、概要、工事発注日、町域名・地先等が掲載されているとの報告がなされている。

#### (2) 構造的なデータとするための困難さ

構造的なデータを得るために、(a)電子ファイルを取得する、(b)電子ファイルを HTML やエクセルフォーマットであっても全て PDF 化する、(c) OCR 機能を用いて CSV 形式に読み取る、(d)道路更新に関する工事を抽出する、という手順が採用されている。このような手順に加えて、100 自治体に対する検証ではデータの正規化(構造化)に至った自治体数が 72 であったと報告されており、また、道路更新に関する工事を抽出に用いる文字列を手動で作成したとの報告もあることから、構造化は容易ではないことが伺える。

このように工事発注見通し情報を構造的なデータとすることは容易ではないが、多くの自治体でその公表はなされている。また、工事発注見通しは、「公共工事の入札及び契約の適正化の促進に関する法律施行令」により基本的に毎年度 4 月 1 日に公表することとされている。公表から入札契約が行われるまでの期間、工事契約から工事完了までの期間は案件によって異なるので、リードタイムの長さを一律に示すことはできないが、一般的には数箇月以上のリードタイムが得られる。

### 1.3.4 工事入札公告を対象とした研究

自治体が工事発注見通し情報に続いて公表する情報が工事入札公告であり、これを用いて道路変化イベントの把握を行うための研究が中條による「工事入札公告を用いた道路更新情報の自動収集・推定に関する研究」(2012)<sup>22)</sup>である。この研究では次のことが報告されている。

#### (1) 道路管理者が有する原典情報の有用性

この研究において工事入札公告を対象とする前に、工事入札公告に加えて、道路管理者が有する原典情報である①公示情報、②開通情報、③電子納品(工事成果の完成図書の電子ファイル)に対して、道路更新情報の利用価値の認識について「地理空間情報流通実験コンソーシアム」会員への調査を行い、いずれも 6 から 8 割程度 (n=60) の利用価値があるとの結果が報告されている。

#### (2) 工事入札公告のインターネットでの公表の状況(網羅性)

工事入札公告については、全国の自治体の電子的手段による提供が約 900 あったことが示されている。

### (3) リードタイム

工事入札公告は工事の前に提供される情報であるから、入札契約までの期間と工事完了までの期間がリードタイムとして得られることが示されている。

### (4) 構造化の困難さと精度

工事入札公告から道路更新に関係するものを選ぶため、案件名と概要から形態素解析によりキーワードを抽出して、道路管理者毎、道路レベル毎（都道府県道、市町村道）に学習モデルを構築している。その結果、精度についての的中率（正確度、Accuracy と同義）が 65～87%、網羅率が（真陽性率、再現率と同義）71～79%であり、道路管理者毎、道路レベル毎（都道府県道、市町村道）にモデルを構築すれば運用できる可能性があると報告されている。ここでは、道路変化の真値として、主として三重県の道路台帳附図を利用している。

このように、入札公告情報はリードタイムを得られる情報である。これを道路変化の予測に用いるためには自治体毎に学習モデルを生成する必要がある。先行研究では道路台帳附図を調査してモデルを作成したが、大きな労力が必要であると考えられる。

#### 1.3.4 先行研究のまとめと明らかになった課題

これらの先行研究のまとめを表 1.3-2 に示す。

表 1.3-2 先行研究のまとめ

| 道路変化イベントの把握 | 使用原典情報       | 先行研究                                 | 先行研究調査の視点       |                  |                 |  |
|-------------|--------------|--------------------------------------|-----------------|------------------|-----------------|--|
|             |              |                                      | リードタイム確保能力      | 道路変化イベントの把握      | インターネットからの取得    | データの構造化                                      |
| 直接          | 公示情報（道路の供用等） | 公示情報に基づく道路更新情報のクロールリングシステムの検討        | ×<br>供用開始と同時に公表 | ○<br>有用性は高い評価    | △<br>主として大規模自治体 | 可能だが容易ではない（多様なファイル形式、多様なフォーマット、多様な語句への対応が必要） |
| 間接          | 工事発注見通し      | 工事発注見通し情報を用いた全国における道路更新情報の自動抽出に向けた試み | ○<br>工事発注準備段階   | －<br>対比していない     | ○<br>9割以上       |  |
| 間接          | 工事入札公告       | 工事入札公告を用いた道路更新情報の自動収集・推定に関する研究       | ○<br>工事契約前      | △<br>自治体毎のモデルが必要 | △～○<br>約半数      |  |

先行研究を調査した結果によって示された問題点から、本研究で課題を以下の3点と設定した。

### 課題1：道路変化イベントと工事情報の関連づけにおける課題

道路変化イベントを直接に示す供用開始の告示は、早いものでも供用開始の数日以内以降となっている。リードタイムを確保するために、道路変化イベントに先立って行われる工事の情報を利用する研究が行われてきた。工事入札公告を用いた先行研究では、その研究で定義した道路変化の内容を工事から予測し、実際の道路変化と突き合わせて評価を行っていた。道路変化の情報は、予測の評価だけでなく、自治体別に学習モデルを生成するためにも必要である。そこでは道路変化の抽出に道路台帳附図を用いていた。変化の内容を正確に捉えるには道路台帳附図は適した手段である。しかし、図面から該当する情報を抽出していくことは情報の閲覧、変化の抽出などを手作業で行わなければならない。これを改善するために、道路変化イベントと工事情報とを関連づける方法を確立することを一つ目の課題に設定した。

### 課題2：工事情報からの道路変化イベントの予測における課題

「1.2.4 道路変化イベントの情報の収集の現状」に述べたように、情報収集者が望む道路変化の種類は多く、また、把握する期間として2年前というリクワイヤメントもある。したがって、多様なリクワイヤメントに応える予測方法の開発が求められる。

また、工事入札公告から道路変化を予測した先行研究では、自治体毎に予測の為の学習モデルが必要と述べられている。しかし、自治体毎に学習モデルを生成するには、情報の入手や該当する情報の抽出において大変な労力が必要となる。この問題を解決するため、より簡易に利用できる予測手法の開発も望まれる。

これらの予測方法の開発を二つ目の課題に設定した。

### 課題3：情報の流通における課題

先行研究は、いずれも自動的な情報の収集に取り組んでいる。そこでの最も大きな問題は、インターネットから電子的に取得できる情報が限られていることである。さらに、取得とその利用には、提供される情報形態と内容が多様性に富むものであるため、自動化しにくいという問題もある。例えば、提供システムへのアクセス方法の違い（HTML、ファイル、検索システムを経由するアクセス）によって多様な取得方法への対応が必要になること、電子的なフォーマットの違い（HTML、PDF、エクセルなど）によって多様なテキスト化への対応が必要になること、さらに表記のゆれがあることなどが具体的な例として挙げられる。

情報から生成される構造的なデータが、本研究を支える基盤である。このため、利用しやすい（マシンリーダブルな）情報を流通させるための環境づくりを三つ目の課題に設定した。

## 1.4 研究の目的と取り組み

実際に利用されている情報収集方法に関する調査では、公示の情報や工事情報といった原典情報を網羅的に使用しているものは確認できなかった。しかし、これらの情報の有用性は先行研究で示されているところである。本研究は、これらの情報を活用して道路変化イベ

ントの把握を早期かつ網羅的に行う仕組みの実用化を目指すものである。

この仕組みの実用化によって道路利用者はこれまで以上に変化した道路の情報を速やかに得ることができる。そして、道路管理者に対する問い合わせや情報公開請求の効率化により、道路管理者と地図製作者双方の事務作業の効率化が期待できる。また、このような仕組みは、道路に関する情報の公表を定めた道路法の趣旨に適うものであり、オープンデータ・オープンガバメントの事例の一つともなるものである。

本研究では、先行研究で示された課題に対して、次のように取り組む。

#### 課題1への取り組み：道路変化イベントと工事情報の関連づけ

道路変化イベントと工事情報とを対照するため、工事情報と道路変化イベントを関連づける方法を提案する。この結果生成されるデータセットを投入できる統合的なシステムを構築し、この手法の実現性を確認する。データセットを予測の改善に利用し、システムを情報の流通の促進に活かすこととする。

#### 課題2への取り組み：工事情報からの道路変化イベントの予測の改善

課題1で生成されたデータセットを利用して、情報収集者の多様なリクワイヤメントに応える予測方法、簡易に利用できる予測方法を開発する。多様な予測を行うために、機械学習の適用を検討する。

#### 課題3への取り組み：情報の流通の促進

課題1で構築した統合的なシステム、課題2で取り組んだ予測手法を情報の提供者に示すことで、情報の流通への理解を促し、情報の流通の促進を目指すものとする。

なお、本研究では「地方自治体」を道路管理者の対象として選んだ。その理由は次の通りである。全国的あるいは広域的な組織である国土交通省と高速道路会社については数も少なく道路変化イベントが把握しやすいのに比べて、地方自治体は、その数が多く情報提供への要請への対応がまちまちであるうえ全国に点在しているので情報収集の効率が悪い。このことから、地方自治体に対する情報収集のニーズが高いと考えたためである。

本研究において設定した課題と取り組みの要点を図1.4-1に示す。

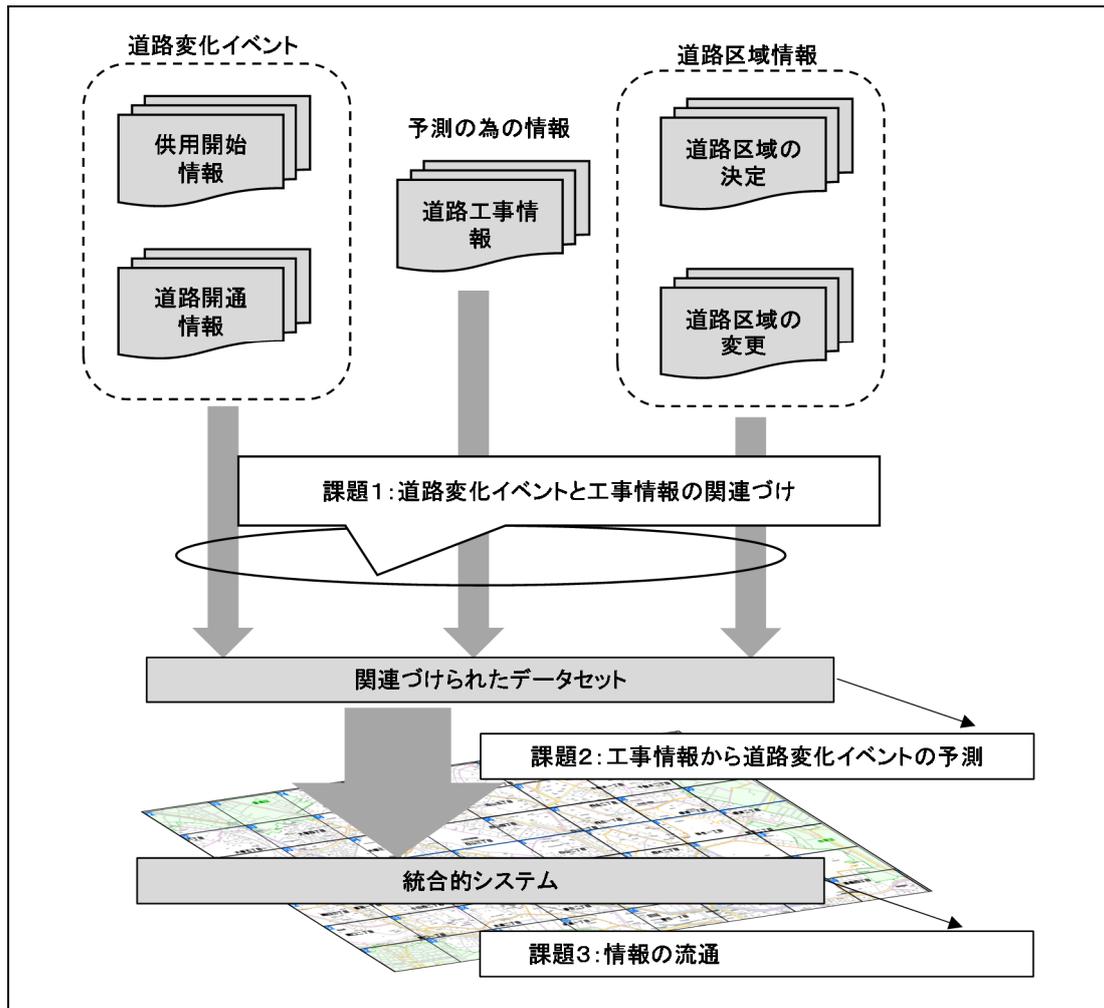


図 1.4-1 本研究で設定した課題と取り組み

## 1.5 論文の構成

本論文は次の7章により構成される。

第1章はこの章であり、研究の背景として道路変化イベントの把握に対する現状と問題点を述べた。そして、公示の情報と工事情報の取得と活用に関するこれまでの先行研究、ここで示された問題点を述べた。最後に、問題点から導かれた課題に対して本研究の目的と取り組みを述べた。

第2章では、道路変化イベントを把握するために利用できると考えた原典情報、具体的には供用開始情報、道路開通情報、工事情報、道路区域情報を関連づけて利用するための手法を提案する。そして、提案した手法によって生成されたデータセットを可視化できる統合的なシステムのデザインについて述べる。本章は主として課題1への取り組みについて述べるものである。

第3章では、関連づけた原典情報に対して、統合的システムを使用して基礎的な分析を行

う。道路区域と供用開始の関係、道路変化イベントと工事とが、どのような時間の間隔で、どのような割合で発生しているかについて明らかにする。本章は課題2の予測方法の開発の基礎となる分析である。

第4章では、工事情報のデータから道路変化イベントを予測することの利点と評価方法について述べる。そして、簡易に工事完了件数から道路開通を予測する手法を提案し、実際のデータにより検証する。本章は主として課題2のうち、簡易な予測方法の開発について述べるものである。

第5章では、工事の内容を用いて、供用開始と道路開通の2種類の道路変化イベントを数箇月前と2年前に予測する手法を提案する。この予測手法で使用する機械学習の手法について簡単に解説する。実際のデータにより予測結果を検証する。そして、予測手法全般に関するまとめを述べる。本章は主として課題2のうち、多様な予測方法の開発について述べるものである。

第6章では、本研究全体の基盤となる情報の流通について述べるものである。本研究では道路管理者へ統合的システムを提示することによってこの課題へ取り組んでおり、その調査結果を述べる。本章は主として課題3への取り組みについて述べるものである。

第7章では、総括として本論の成果をまとめ、加えて課題と展望について述べる。

下に論文の構成を示す。

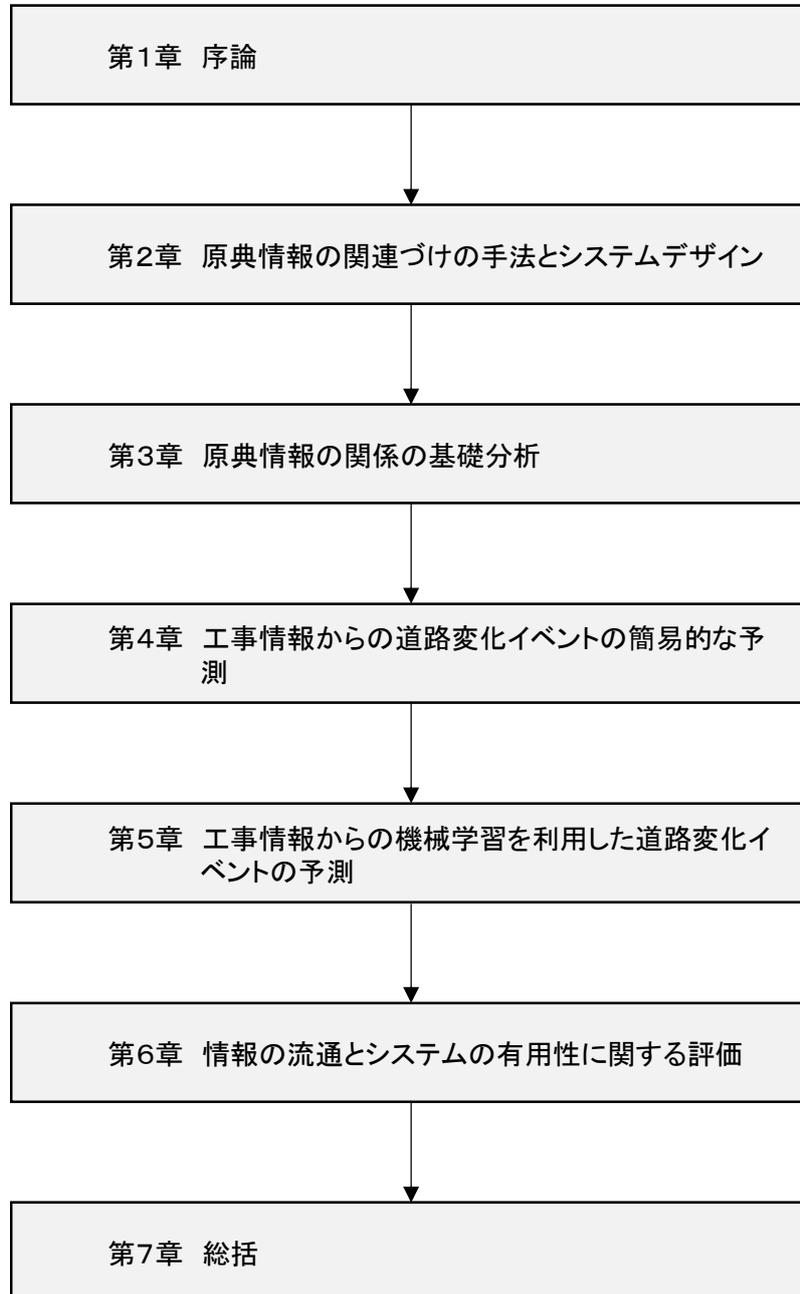


図 1.4-1 論文の構成

## 参考文献

### 1.2 研究の背景

- 1) 国土交通省道路局，都市局：平成 28 年度道路関係予算概要，<http://www.mlit.go.jp/comm on/001116627.pdf>, pp.2, pp.33-34, 平成 28 年 1 月. 2016.2.10.アクセス.
- 2) 国土交通省：道路統計年報（2014，2013），道路現況総括表，<http://www.mlit.go.jp/road/ir/ir-data/tokei-nen/2014tokei-nen.html>, <http://www.mlit.go.jp/road/ir/ir-data/tokei-nen/2013tokei-nen.html>. 2015.10.1 アクセス
- 3) 総務省統計局：日本統計年鑑第 9 章建設業 9 - 3 発注者及び目的別工事分類別公共機関からの受注工事件数及び請負契約額，<http://www.stat.go.jp/data/nenkan/zuhyou/y0903000.xls>, 2012.10.13.アクセス.
- 4) 一般社団法人日本道路建設業協会：統計資料道路事業費と舗装費 10 年間の推移（データ），[http://www.dohkenkyo.com/shiryu/g\\_pdf/costs/h02.pdf](http://www.dohkenkyo.com/shiryu/g_pdf/costs/h02.pdf), 2012.10.13.アクセス.
- 5) 新潟市総務部：行政文書情報の提供・公開窓口（公示文書検索），<https://info-navi.city.niigata.lg.jp/iDiscPub/noticeSearch.do>, 2015.8.15.アクセス.
- 6) 新潟市道路計画課：道路開通情報一覧表，<https://www.city.niigata.lg.jp/kurashi/doro/road/doroseibi/dorokaitsuinfo/index.html>, 2012.8.15.アクセス.
- 7) 新潟市道路計画課：道路開通情報一覧表，<https://www.city.niigata.lg.jp/kurashi/doro/road/doroseibi/dorokaitsuinfo/index.html>, 2015.11.26.アクセス.
- 8) 島根県総務部総務課：県報（供用開始），<http://www.pref.shimane.lg.jp/admin/pref/info/kenpou/>, 2015.8.1.アクセス.
- 9) 島根県道路建設課：道路開通情報（地区別にリンク），[http://www.pref.shimane.lg.jp/infra/infos/douro\\_kaitu/kaitu/](http://www.pref.shimane.lg.jp/infra/infos/douro_kaitu/kaitu/), 2015.11.3.アクセス.
- 10) 高知県法務課：高知県公報発行月別閲覧（供用開始），<http://www.reikisyuutou.pref.kochi.lg.jp/kakokouhou/tsukisentaku.htm>, 2015.8.1.アクセス
- 11) 高知県道路課：平成 26 年度道路法関連告示一覧，<http://www.pref.kochi.lg.jp/soshiki/170701/kokuj-i-itan26.html>, 公開日 2015 年 3 月 31 日, 2015.8.1.アクセス.
- 12) 愛媛県土木部道路建設課：主要道路開通情報，<http://www.pref.ehime.jp/h40400/5744/kaituu/kaituu.html>, 2010.3.19.更新, 2013.9.29.アクセス.
- 13) 埼玉県道路街路課：道路開通情報，<http://www.pref.saitama.lg.jp/site/kaituszyouhou/>, 2012.10.7.アクセス.
- 14) 埼玉県道路政策課：道路開通情報（国道），<http://www.pref.saitama.lg.jp/site/dousei1006/905-20091224-51.html>, 2010.3.19.更新, 2010.8.15.アクセス.
- 15) 埼玉県道路政策課：道路開通情報（県道），<http://www.pref.saitama.lg.jp/site/dousei1006/dousei035.html>, 2012.2.3.更新, 2012.8.15.アクセス.
- 16) 埼玉県道路政策課：道路開通情報（その他），<http://www.pref.saitama.lg.jp/site/dousei1006>

- /dousei059.html, 2012.1.17.更新, 2012.8.15.アクセス.
- 17) 一般財団法人日本デジタル道路地図協会：利活用事例, <http://www.drm.jp/database/use.html>, 2012.9.10.アクセス.
  - 18) 池田, 小林, 平野：いかにしてカーナビゲーションシステムは実用化されたか, *Synthesiology*, Vol.3 No.4, pp295-296, 国立研究開発法人産業技術総合研究所, 2010.
  - 19) 一般財団法人日本デジタル道路地図協会：DRM ニュース（平成 27 年度 7-9 月期までの DRM データベース提供実績について）, No.51, pp.6, [http://www.drm.jp/introduction/news/drm\\_news51.pdf](http://www.drm.jp/introduction/news/drm_news51.pdf), 平成 28 年 1 月 1 日発行, 2016.1.2.アクセス.
  - 20) 一般財団法人日本デジタル道路地図協会：作成・更新方法, <http://www.drm.jp/database/structure.html>, 2014.1.1 アクセス
  - 21) 関本義秀, 金澤文彦, 松下博俊：次世代デジタル道路地図のあり方に関する研究, 国土技術政策総合研究所資料, ISSN136・6-7328, No. 372, pp.2, pp.15, pp.48, 2007.
  - 22) 中條覚：工事入札公告を用いた道路更新情報の自動収集・推定に関する研究, 東京大学博士論文, pp.10, 2012.
  - 23) 特定非営利活動法人 ITS Japan：安全・環境に資する走行支援サービス実現のための道路情報整備と流通へ向けた提言, <http://www.its-jp.org/wp-content/uploads/2010/09/2d241ee556cc6bcfb250a3130e642658.pdf>, pp.13, 2008.4, 2012.10.13.アクセス.
  - 24) 布施, 松林, 中條, 高橋, 脇島, 山口：公示情報に基づく道路更新情報のクローリングシステムの検討, 土木情報利用技術論文集, vol.18, pp.281-290, 土木学会, 2009.
  - 25) 鹿児島県総務部学事法制課：鹿児島県の情報公開・個人情報保護制度（平成 21 年度の運用状況）, 平成 22 年 10 月, <https://www.pref.kagoshima.jp/ab04/kensei/joho/unyoudocuments/unyokyo21v1.pdf>, 2015.8.4 アクセス
  - 26) 政策研究大学院大学：情報公開開示請求データベース, <http://gclip1.grips.ac.jp/disclosure/>, 2015.7.4.アクセス
  - 27) 国土交通省三陸国道事務所：仙人峠道路開通半年後の効果, [http://www.thr.mlit.go.jp/sa-nriku/07\\_office/kouka/sennin/kouka\\_6months/index.html](http://www.thr.mlit.go.jp/sa-nriku/07_office/kouka/sennin/kouka_6months/index.html), 2015.12.12.アクセス.
  - 28) 国土交通省三陸国道事務所：仙人峠道路開通 1 年後の効果, [http://www.thr.mlit.go.jp/sa-nriku/07\\_office/kouka/sennin/kouka\\_1year/index.html](http://www.thr.mlit.go.jp/sa-nriku/07_office/kouka/sennin/kouka_1year/index.html), 2015.12.12.アクセス.
  - 29) 国土交通省三陸国道事務所：仙人峠道路開通 2 年後の効果, [http://www.thr.mlit.go.jp/sa-nriku/07\\_office/kouka/sennin/kouka\\_2year/index.html](http://www.thr.mlit.go.jp/sa-nriku/07_office/kouka/sennin/kouka_2year/index.html), 2015.12.12.アクセス.
  - 30) 総務省：平成 26 年版情報通信白書第 4 章第 1 節 3(2)車と ICT, <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h26/html/nc141320.html>, 2015.6.4 アクセス
  - 31) 株式会社ミックウェア：運転とカーナビに関する意識調査, 報道機関向け報告, 2012.8.8
  - 32) 準天頂衛星を利用した新産業創出研究会：準天頂衛星を利用した新産業創出研究会報告書（案）, 第 3 章第 8 節横断的な課題, 平成 24 年 9 月, [http://www.meti.go.jp/committee/kenkyukai/seisan/juntenchousei/006\\_03\\_00.pdf](http://www.meti.go.jp/committee/kenkyukai/seisan/juntenchousei/006_03_00.pdf), 2015.8.4 アクセス

- 33) 須田, 青木 : 自動運転技術の開発動向と技術課題, 情報管理, Vol.57, No.11, pp.809-817, 2015.
- 34) 一般財団法人日本デジタル道路地図協会 : DRM ニュース, No.48, 平成 27 年度事業計画, [http://www.drm.jp/introduction/news/drm\\_news48.pdf](http://www.drm.jp/introduction/news/drm_news48.pdf), 2015.9.1.アクセス
- 35) INTERNET Watch : カーナビの地図データはこうして作られていた! 点と線の地図「道路ネットワーク」とは(地図会社を訪ねて~インクリメント P 株式会社), [http://internet.watch.impress.co.jp/docs/column/chizu/20150331\\_695365.html](http://internet.watch.impress.co.jp/docs/column/chizu/20150331_695365.html), 2015/3/31, 2015.10.9.アクセス.
- 36) 内田宗治 : ゼンリン住宅地図と最新ネット地図の秘密, 実業之日本社, ISBN 978-4-408-11063-9, 2014.
- 37) トヨタマップマスター : 情報・資料の収集, <http://www.mapmaster.co.jp/mapdb/gather.html>, 2015.10.8.アクセス.

### 1.3 先行研究

- 38) 関本義秀, 中條覚, 南佳孝, 山口章平, 山田晴利, 布施孝志 : 工事発注見通し情報を用いた全国における道路更新情報の自動抽出に向けた試み, 土木学会論文集 D3 (土木計画学), Vol. 68, No. 3, pp.117-128, 2012.  
(他の先行研究は 22), 24)は, 1.2 研究の背景に既出)

## 第2章 原典情報の関連づけの手法とシステムデザイン

### 2.1 はじめに

道路変化イベントを直接的に示す供用開始情報、道路開通情報は、道路変化イベントと同時に公表されることが多く、情報を利用できる形に整える準備時間（リードタイム）を得ることができない。これに対して、間接的な情報である工事情報の利用が研究されてきた。道路管理者等が管理するこれらの一次的な情報を、本研究では原典情報と総称している（表1.3-1参照）。原典情報は、道路の供用の開始に向けた一連の事業の各段階で発生するものであるが、先行研究の道路管理台帳と工事情報の結び付け（ただし、手作業による突合）以外に原典情報を相互に関連づけた資料は見当たらなかった。

関連づけが行われていない理由は、(a)情報の目的が異なるため関係づけの必要性が無いと考えられている、(b)原典情報の種類に応じて担当が異なることが少なくなく、原典情報での情報のやりとりがない、(c)原典情報が一対一に対応するとは限らず関係づけは容易には行えない、といったものではないかと推察する。

しかし、関連づけによってデータセットを生成すれば、次の効果が見込める。

- ① 2つ目の課題である工事情報からの道路変化イベントの予測に利用できる。
- ② データセットを可視化するシステムは3つ目の課題である情報の流通に利用できる。

そのために本章において原典情報の関連づけの手法とシステム化のためのデザインに取り組む。本章では、最初に原典情報を構造的なデータに変換する方法を整理する。対象とする原典情報は、(a)供用開始情報、(b)道路区域の決定、(c)道路区域の変更、(d)道路開通情報、(e)工事情報、である。次に、構造化されたデータを関連づける方法を提案する。関連づけには場所と時間を利用した。本研究では、関連づけられた情報を可視化する統合的なシステムを構築した。本章の最後には、そのシステムデザインについて述べる。

### 2.2 道路変化イベントに関係する原典情報の構造化

#### 2.2.1 供用開始情報

供用開始情報（道路の供用の開始の公示）が道路法第十八条第二項に基づくものであることについては「1.2.2.1 道路変化イベントに関する道路管理者の提供情報」に述べた。道路法施行規則（最終改正平成27年1月23日国土交通省令第四号）第三条には、道路の供用の開始又は廃止の公示において掲げる事項として、次の4点を挙げている。

- 一 路線名
- 二 供用開始又は廃止の区間
- 三 供用開始又は廃止の期日
- 四 供用開始又は廃止の区間を表示した図面を縦覧する場所及び期間

自治体において実際に発出されている供用開始情報の例を図2.2-1に示す。

●金沢市告示第99号  
 道路法（昭和27年法律第180号）第18条第2項の規定により、次の道路の供用を開始します。  
 なお、その区間を表示した図面は、金沢市土木局道路管理課において平成25年4月1日から同月15日まで一般の縦覧に供します。  
 平成25年4月1日  
 金沢市長 山野之義

| 路線名                  | 区間                                   | 供用開始日     |
|----------------------|--------------------------------------|-----------|
| 1 級幹線 119号<br>金沢駅通り線 | 本町1丁目 576番 先から<br>本町1丁目 572番 先まで     | 平成25年4月1日 |
| 十一屋町線 10号            | 十一屋町 310番 6先から<br>十一屋町 310番 1先まで     | 平成25年4月1日 |
| 金石西2丁目線 26号          | 金石西2丁目 559番 1先から<br>金石西2丁目 535番 1先まで | 平成25年4月1日 |
| 二塚9号<br>専光寺町西線 36号   | 専光寺町ツ 17番 7先から<br>専光寺町ツ 17番 9先まで     | 平成25年4月1日 |

(a) 金沢市公報第2757号の3（平成25年4月1日）（一部）

高知県告示第234号  
 道路法（昭和27年法律第180号）第18条第2項の規定により、道路の供用を次のとおり開始する。  
 その関係図面は、平成26年4月11日から2週間高知県土木部道路課及び高知県中央西土木事務所において一般の縦覧に供する。  
 平成26年4月11日  
 高知県知事 尾崎 正直

- 道路の種類 県道
- 路線名 谷地日下停車場
- 道路の区域

| 供用開始区間                                    | 延長<br>(メートル) | 供用開始年月日    |
|---|--------------|------------|
| 高岡郡日高村沖名字長崎126番5から<br>高岡郡日高村沖名字落合ノ上93番1まで | 190          | 平成26年4月11日 |

(b) 高知県公報第9631号（平成26年4月11日）（一部）

号外第68号 島根県報 平成26年4月4日

島根県告示第229号  
 道路の供用を次のように開始するので、道路法（昭和27年法律第180号）第18条第2項の規定により告示する。  
 その関係図面は、告示の日から15日間島根県土木部道路維持課及び当該道路を管轄する地方機関において一般の縦覧に供する。  
 平成26年4月4日  
 島根県知事 溝口 晋兵衛

| 道路の種類 | 路線名      | 供用開始の区間                        | 延長<br>(メートル) | 供用開始年月日   | 管轄する地方機関の名称 | 備考 |
|-------|----------|--------------------------------|--------------|-----------|-------------|----|
| 県道    | 松江鹿島美保関線 | 松江市鹿島町佐陀本郷35番1地先から同番3地先まで      | 25.37        | 平成26年4月4日 | 松江県土整備事務所   |    |
| "     | 大田桜江線    | 邑智郡川本町大字南佐木167番1地先から同225番1地先まで | 324.20       | 平成26年4月4日 | 県央県土整備事務所   |    |
| "     | 美郷飯南線    | 邑智郡美郷町酒谷489番地先から同494番1地先まで     | 112.00       | 平成26年4月4日 |             |    |

(c) 島根県報号外第68号（平成26年4月4日）（一部）

図 2.2-1 道路の供用開始の告示の例

原典情報を構造的なデータに変換する際の方針として、法令・規則等で定められた項目<sup>1),2)</sup>、実際に流通している公示情報の項目を調査し、それらの内容を損なうことなく格納できるデータテーブル構造を作成することとした。

(a)基本的な構造

道路の供用の開始と廃止の1件は、対象とする区間ごとに告示され、1件の供用開始には1区間だけが含まれる（ただし、告示には数件以上の供用開始が含まれることが少なくない）。このため、1区間の供用の開始又は廃止を1レコード（1件）として管理することができる。

(b)フィールド（項目）

フィールド（項目）は、道路法施行規則に定められた項目のみとしている自治体（例、図 2.2-1(a)）もあるが、供用開始する区間の延長を付記している自治体（例、図 2.2-1(b),同(c)）もある。延長は道路変化イベントの大きさを示す有益な情報である。このため、標準的に延長を設けることとした。その他、公示に関わる情報を格納できるものとした。

これらを踏まえて設計したデータテーブルの構造を表 2.2-1 に示す。ID はデータベースに格納する際の識別番号である。

表 2.2-1 道路の供用の開始および廃止のためのデータテーブルの構造

| ID          | 道路管理者 | 供用開始日    | 道路種類 | 路線名        | 区間（地名1）                | 区間（地名2）                   |
|-------------|-------|----------|------|------------|------------------------|---------------------------|
| 15100-4-001 | 新潟市   | 2014/4/1 | 市道   | 西南 5-30 号線 | 新潟市西区内野西二丁目 2932 番地先から | 新潟市西区内野西二丁目 3217 番 1 地先まで |

| 告示日      | 告示番号         | 管轄 | 延長(m) | 備考 |
|----------|--------------|----|-------|----|
| 2014/4/1 | 新潟市告示第 346 号 |    |       |    |



2.2.2 道路区域の決定

道路区域の決定（道路の供用の開始の公示）が道路法第十八条第一項に基づくものであることについては「1.2.2.1 道路変化イベントに関する道路管理者の提供情報」に述べた。道路法施行規則（最終改正平成 27 年 1 月 23 日国土交通省令第四号）第二条には、道路の区域の決定又は変更の公示において掲げる事項として、次の 4 点を挙げている。

- 一 道路の種類
- 二 路線名
- 三 次のイ、ロ又はハに掲げる場合の区分に応じそれぞれイ、ロ又はハに定める事項
  - イ 区域の決定の場合（ロに掲げる場合を除く。） 敷地の幅員及びその延

長

ロ 法第四十七条の七の規定により立体的区域とする区域の決定の場合  
イに掲げる事項並びに当該立体的区域とする区間及びその延長

ハ 区域の変更の場合 変更の区間並びに当該区間に係る変更前の敷地の  
幅員及びその延長並びに変更後の敷地の幅員及びその延長

四 区域を表示した図面を縦覧する場所及び期間

幾つかの異なる自治体において実際に発出されている道路区域決定の例を図 2.2-2 に示す。

| 新潟市告示第 53 号  |            | 掲示期間 1.20-2.2                                      |           |        |
|--|------------|--|-----------|--------|
| 道路区域の決定について  |            |  |           |        |
| <p>道路法（昭和 27 年法律第 180 号）第 18 条第 1 項の規定に基づき、道路区域を次のように決定した。</p> <p>なお、関係図面は告示の日から 2 週間新潟市土木部土木総務課及び新潟市北区役所建設課において、一般の縦覧に供する。</p> <p>平成 26 年 1 月 20 日</p> <p style="text-align: right;">新潟市長 篠 田 昭</p> |            |  |           |        |
| 道路の種類  | 路線名        | 区 間  | 敷 地 の     |        |
|  |            |  | 幅員 (m)    | 延長 (m) |
| 市道   | 北 3-120 号線 | 新潟市北区西名目所 3890 番 1 地先から<br>新潟市北区西名目所 3890 番 4 地先まで | 6.0~10.0  | 149.0  |
| 市道   | 北 3-121 号線 | 新潟市北区西名目所 3702 番 2 地先から<br>新潟市北区西名目所 3702 番 3 地先まで | 12.5~19.6 | 37.9   |
| 市道   | 北 3-122 号線 | 新潟市北区西名目所 3702 番 2 地先から<br>新潟市北区西名目所 2646 番地先まで    | 6.0~10.0  | 186.3  |
|  |            | 新潟市北区西名目所 2617 番 4 地先から                            |           |        |

(a) 新潟市告示第 53 号（平成 26 年 1 月 20 日）

| <p>●金沢市告示第131号</p> <p>道路法（昭和27年法律第180号）第18条第1項の規定により、市道の区域を次のように決定します。</p> <p>なお、その区域を表示した図面は、金沢市土木局道路管理課において平成26年4月1日から同月15日まで一般の縦覧に供します。</p> <p>平成26年4月1日</p> <p style="text-align: right;">金沢市長 山 野 之 義</p> |               |         |         |  |
|--|---------------|---------|---------|--|
| 道路の種類  | 路線名           | 幅 員 (m) | 延 長 (m) |  |
| 一般市道   | 浅野本町 2 丁目線22号 | 6.0     | 54      |  |
| 一般市道   | 駅西本町 6 丁目線12号 | 6.0     | 36      |  |
| 一般市道   | 天神町 1 丁目線19号  | 6.0     | 48      |  |
| 一般市道   | 弓取12号割出町線42号  | 6.0     | 94      |  |

(b) 金沢市公報第 2792 号の 3（平成 26 年 4 月 1 日）

図 2.2-2 道路の区域の決定の告示の例

道路区域に関する情報の目的は道路を作る区域であることを宣言するものであり、そのため、区域の大きさが示されている。そして、図 2.2-2 のように供用日は記載されず、日付に関しては告示日のみである。道路区域決定は、対象とする区間ごとに告示される。このため、1区間の道路区域決定を1レコード(1件)として管理することができる。フィールド(項目)は、道路法施行規則に定められた項目と実際に発出されている告示内容は一致している。データベースに格納するための構造を表 2.2-2 に示す。ID はデータベースに格納する際の識別番号である。

表 2.2-2 道路区域の決定のためのデータテーブルの構造

| ID  | 道路管理者 | 告示日        | 道路種類 | 路線名          | 区間(地名1)        | 区間(地名2)          |
|-----|-------|------------|------|--------------|----------------|------------------|
| 148 | 金沢市   | 2014-04-01 | 一般市道 | 弓取12号割出町線42号 | 石川県金沢市割出町54番5先 | 石川県金沢市割出町38番3先まで |

| 幅員(m) | 延長(m) | 告示番号 | 管轄          | 備考                    |
|-------|-------|------|-------------|-----------------------|
| 6.0   | 94    | 131号 | 金沢市土木局道路管理課 | 告示133号の区間を準用。公報2792-3 |

### 2.2.3 道路区域の変更

道路区域の変更(道路の供用の開始の公示)が道路法第十八条第一項に基づくものであることについては「1.2.2.1 道路変化イベントに関する道路管理者の提供情報」に述べた。道路法施行規則(最終改正平成27年1月23日国土交通省令第四号)第二条の道路の区域の決定又は変更の公示において掲げられた4点については前項を参照されたい。

実際の道路区域の変更の公示の例を図 2.2-3 に示す。前項で述べたように道路区域に関する情報の目的は道路を作る区域であることを宣言するものであり、区域の大きさが示されているが供用日は記載されず、日付に関しては告示日のみである。

## 告 示

## 島根県告示第228号

道路の区域を次のように変更したので、道路法（昭和27年法律第180号）第18条第1項の規定により告示する。

その関係図面は、告示の日から15日間島根県土木部道路維持課及び当該道路を管轄する地方機関において一般の縦覧に供する。

平成26年4月4日

島根県知事 溝 口 善兵衛

| 道路の種類 | 路線名      | 道路の区域                              |        |                         | 管轄する地方機関の名称     | 備考                                 |   |
|-------|----------|------------------------------------|--------|-------------------------|-----------------|------------------------------------|---|
|       |          | 区 間                                | 変更前後の別 | 敷地の幅員<br>延 長            |                 |                                    |   |
| 県 道   | 松江鹿島美保関線 | 松江市鹿島町佐陀本郷35番1地先から同番3地先まで          | 前      | メートル<br>15.67～<br>18.84 | メートル<br>25.37   | 松江県土整備事務所<br>道路改良工事<br>拡幅<br>仮設道設置 |   |
|       |          |                                    | 後      | 15.67～<br>73.81         | 25.37           |                                    |   |
| #     | 松江島根線    | 松江市島根町大芦3123番1地先から同地先まで            | 前      | 15.00～<br>28.00         | 24.00           | 減幅<br>不用物件発生<br>事業用地と交換            |   |
|       |          |                                    | 後      | 15.00～<br>22.00         | 24.00           |                                    |   |
| #     | 川本波多線    | 飯石郡飯南町志津見636番地先から同817番地先まで         | 前      | 47.00～<br>108.00        | 85.00           | 雲南県土整備事務所<br>減幅<br>町道移管            |   |
|       |          |                                    | 後      | 12.00                   | 85.00           |                                    |   |
| #     | 美郷大森線    | 大田市水上町福原字旗ヶ坪383番4地先から同386番4地先まで    | 前      | A                       | 4.00～<br>6.00   | 232.00                             | 左記のA及びBは、関係図面に表示する敷地の区分をいう。<br>ダブルウェイ解消<br>市道移管 |
|       |          |                                    |        | B                       | 13.00～<br>19.00 |                                    |   |
|       |          |                                    | 後 B    | 13.00～<br>50.00         | 130.00          |                                    |   |
|       |          | 大田市水上町福原字吉田原756番5地先から同町福原字クベ755番地先 | 前      | A                       | 3.50～<br>6.50   | 70.50                              | 左記のA及びBは、関係図面に表示する敷地の区分をいう。                     |

島根県報号外第 8 号（平成 26 年 2 月 7 日）

島根県告示第 61 号の一部

図 2.2-3 道路区域の変更の告示の例

(a)基本的な構造

供用開始と道路区域決定は、1件に含まれる区間は1であった。これに対して、区域の変更では、道路法施行規則第二条3のハにあるように「変更の区間並びに当該区間に係る変更前の敷地の幅員及びその延長並びに変更後の敷地の幅員及びその延長」、すなわち、変更前と変更後の情報が示される。さらに変更前、変更後のそれぞれにダブルウ

エイ、トリプルウェイと呼ばれる複数の区間が含まれる場合がある。このように1件の情報の中に複数の区間が含まれることが道路区域の変更の情報の特徴である。データテーブルの設計においては、複数の区間への対応と同時に1件の告示であることを表現する必要があると考えた。そのため、リレーションの正規化を行い、道路区域の変更の1件を示す要素と、個々の区間に関する要素に分離したデータテーブルの構造とした。

(b)フィールド（項目）

公示に関する基本的なフィールドはメタデータテーブルに設け、各区間の情報と道路区域の変更における位置づけ（前後の別と、前後夫々の識別）を区間のテーブルに設けた。

表 2.2-3 にデータテーブルの構造を示す。

表 2.2-3 道路区域の変更のためのデータテーブルの構造

(道路区域の変更の各件を示すメタデータテーブル)

| 道路管理者 | メタデータ番号 | 公示日        | 道路の種類 | 路線名    | 告示番号 | 管轄        | 備考  |
|-------|---------|------------|-------|--------|------|-----------|---|
| 島根県   | 177     | 2013/12/27 | 県道    | 出雲三刀屋線 | 845  | 出雲県土整備事務所 | 左記のA, B, C及びDは、関係図面に表示する敷地の区分をいう。ダブルウェイ区間B1の延伸によるB3へ... |

(道路区域変更の1件に対応する複数の区間のテーブル)

| 道路管理者 | メタデータ番号 | 詳細番号 | 区間（地名1（から））        | 区間（地名2（まで））          | 前後の別 | 前後内の識別 | 幅員(m)       | 延長(m)  |
|-------|---------|------|--------------------|----------------------|------|--------|-------------|--------|
| 島根県   | 177     | 1    | 出雲市大津町 525 番3地先    | 出雲市上島町 2617 番4地先まで   | 前    | A      | 6.00～23.00  | 5900   |
| 島根県   | 177     | 2    | 出雲市上塩冶町 1686 番1地先  | 出雲市上塩冶町 1629 番1地先まで  | 前    | B 1    | 20.10～33.50 | 246.1  |
| 島根県   | 177     | 3    | 出雲市船津町 381 番3地先    | 出雲市上島町 2617 番4地先まで   | 前    | B 2    | 12.00～80.00 | 2755   |
| 島根県   | 177     | 4    | 出雲市上塩冶町 2869 番29地先 | 出雲市上塩冶町 2866 番11地先まで | 前    | B 3    | 12.50～59.00 | 470    |
| 島根県   | 177     | 5    | 出雲市大津町 525 番3地先    | 出雲市上島町 2617 番4地先まで   | 後    | A      | 6.00～23.00  | 5900   |
| 島根県   | 177     | 6    | 出雲市上塩冶町 1686 番1地先  | 出雲市上塩冶町 2866 番11地先まで | 後    | B 1    | 12.50～71.20 | 1838.1 |
| 島根県   | 177     | 7    | 出雲市船津町 381 番3地先    | 出雲市上島町 2617 番4地先まで   | 後    | B 2    | 12.00～80.00 | 2755   |

2.2.4 道路開通情報

道路開通情報は、供用開始情報のうち、大規模で道路交通に与える影響が大きいと道路管

理者が判断したもの（例えば、新たな道路、橋梁、隧道、交差点改良等）である。このため、データフォーマットは供用開始情報と類似している。基本的な構造は1件を1レコードとして管理できる。相違点は、道路開通情報は公示ではないので公示日や公示番号が付与されないこと、供用開始の延長はメートルで示されるのに対して道路開通の延長はキロメートルであることが多いことである。実際の例を図 2.2-4 に、設計したデータテーブルの構造を表 2.2-4 に示す。

**一般国道378号 俵津玉津トンネル(俵津バイパス一部区間)**

---

**開通日**

- 2014年7月12日

**対象区間**

- 西予市明浜町俵津 ~ 宇和島市吉田町深浦

**工事事務所**

- 西予土木事務所

**整備内容**

- 延長943m (トンネル部697m)

**関連リンク**

- [地理院地図 俵津玉津トンネル付近\(国土地理院\)\(外部サイトへリンク\)](#)

**関連ファイル**

- [一般国道378号 俵津玉津トンネル事業箇所説明資料\(PDF:1,108KB\)](#)
- [一般国道378号 俵津バイパスの概要について\(PDF:1,423KB\)](#)

**写真**

【整備後状況】



【整備前状況】



愛媛県：主要道路開通，<http://www.pref.ehime.jp/h40400/5744/kaituu/kaituu.html>

図 2.2-4 道路開通情報の例

表 2.2-4 開通情報のデータテーブルの構造

| ID | 道路<br>管理者名 | 開通日        | 道路<br>種類 | 路線名      | 区間（地名1） | 区間（地名2） | 延長<br>(km) | 備考         |
|----|------------|------------|----------|----------|---------|---------|------------|------------|
| 10 | 高知県        | 2015-01-14 | 県道       | 高知南インター線 | 高知市高須新町 | 高知市五台山  | 1.6        | 4車線化(高知新聞) |
| 3  | 新潟市        | 2013-07-24 |          | 小新亀貝線    | 新潟市西区亀貝 | 新潟市西区小新 | 0.8        |            |

## 2.2.5 工事情報

1 件の工事からは複数の情報が生成される。多くの自治体が生成している工事の情報を発生する時期の順に並べると、(a)工事発注見通しの公表（四半期ごと，入札公告より前），(b)工事入札公告（工事の入札の前），(c)工事实績情報（契約後 10 日以内），(d)契約内容の公表（契約後）となる。これらは同一の工事に対してそれぞれ発生するものであり，時間が進むにつれて情報の確実さは増す。仮に，1 件の工事に対して，別種の情報（例えば，工事入札公告と契約内容）が得られた場合には，精査して 1 つの情報とする必要がある。早期の段階の情報をを用いることと，情報の確実さはトレードオフとなる。

### 2.2.5.1 本研究で対象とする工事

公共工事には道路に関係の無いものが含まれているほか，道路における工事であっても道路自体の変更を工事の目的とするのではなく，道路に埋設されたインフラ（例えば，上下水道や電力，通信のための管路やケーブル）が工事の目的物であって，道路には影響を与えないこと（原形復旧）を目指す工事がある。公共工事のうち，本研究で対象とする工事を抽出するための第一段階のスクリーニングは，道路を目的物とする工事の選別となる。これらの工事は道路の予算で執行され，道路管理者が実施する工事であることから，工事の発注者を道路管理者とすることで抽出できる。道路管理者としての発注を選ぶためには，道路建設課発注といった名称だけでは判断できる場合はそれを利用する。土木建設課のように河川管理者との機能も持つ場合には工事件名や工事場所を利用する。工事件名や工事内容から抽出する方法については先行研究<sup>22), 38)</sup>で紹介されている。本研究で調査したところ，いわゆる道路や路線名を示すもののほか，「トンネル」，「隧道」，「橋梁」，「橋」が件名に含まれているものは道路工事として抽出することが適当であった。第二段階のスクリーニングとして，施工場所が道路あるいは道路予定地である場所であるものを抽出することにより，山上の道路用気象観測施設のための工事や道路管理事務所の営繕工事などの工事を除外することができる。このようなスクリーニングによって抽出した工事を本研究では道路工事とする。

しかし，道路が工事の目的物であっても，除草や清掃といった維持保守であって，道路の変化に寄与しない工事が含まれる。これらは，道路変化イベントの予測には不要な工事であるが，維持保守工事において軽微な改修を行って供用開始が告示されることや，維持保守と明確に件名に記載されていないものが無いとは言えないと考えた。このため，これらの除去は行わず，これらが含まれた道路工事の中から道路変化イベントを適切に予測することを，後述する道路変化イベントの予測において取り扱う。図 3.1-3 に目的と場所によって道路工事を抽出するためのベン図を示す。

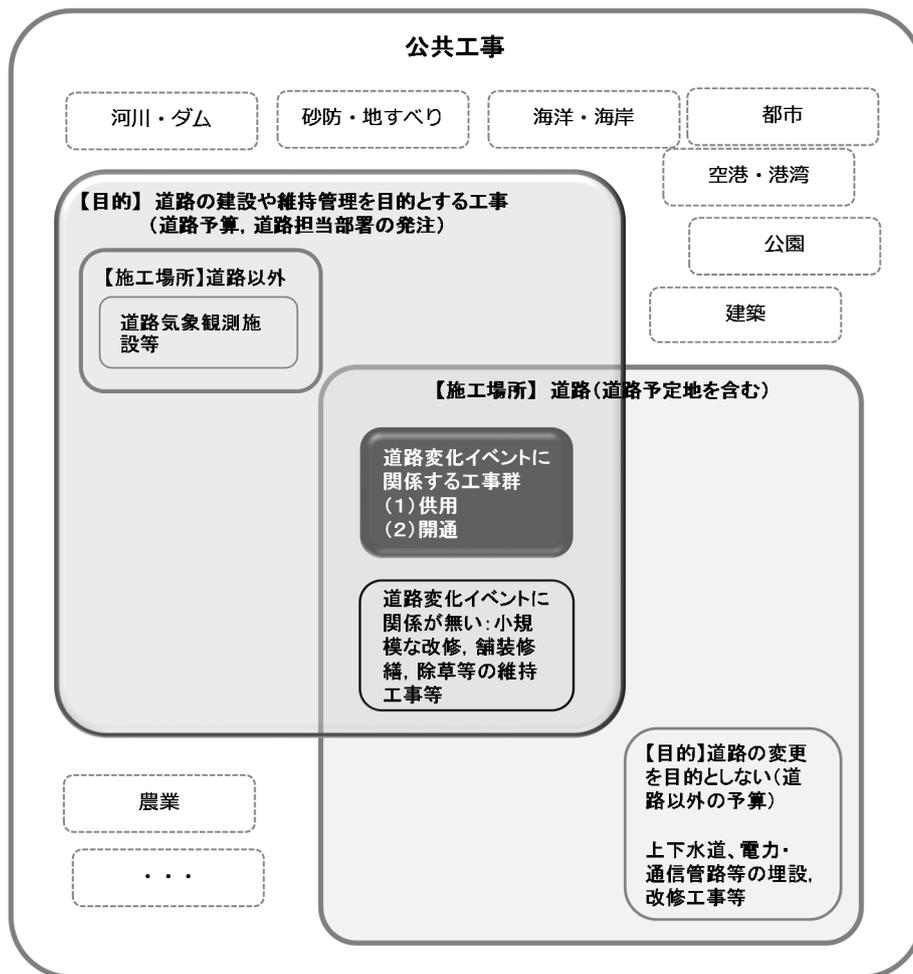


図 2.2-5 目的と場所による道路工事の分類

### 2.2.5.2 法律に定めのある公共工事の情報

前述の(a)工事発注見通しの公表（四半期ごと，入札公告より前），(b)工事入札公告（工事の入札の前），(c)工事実績情報（契約後 10 日以内），(d)契約内容の公表（契約後）のうち，(c)工事実績情報以外は，「公共工事の入札及び契約の適正化の促進に関する法律」に情報の作成と公表の根拠を見出すことができる。

同法では第二章「情報の公表」において，各公共工事の発注者である国，特殊法人等，地方公共団体に，公共工事の発注の見通し，入札及び入札及び契約の過程，並びに公共工事の契約の内容に関する事項の公表を求めている。同法の施行令では，公表の具体的な方法を「官報又は時事に関する事項を掲載する日刊新聞紙に掲載する方法」あるいは「公衆の見やすい場所に掲示し，又は公衆の閲覧に供する方法」のいずれかとし，公衆の閲覧の場合は「閲覧所を設け，又はインターネットを利用して閲覧に供する方法によらなければならない。」としている。契約の内容に関する事項については，「公共工事の名称，場所，種別及び概要」，「工事着手の時期及び工事完成の時期」などが示されている。

国土交通省、総務省、財務省は、国、特殊法人等、地方公共団体の同法に対する取組状況を調査、公表している。平成23年9月1日の調査によると、発注見通しの公表は、国、特殊法人等、地方公共団体全体で98.2%（機関数1900）、公共工事の名称、場所、種別及び概要の公表は98.0%（同1895）、工事着手の時期及び工事完成の時期の公表は94.7%（同1832）であった。そして、インターネット上で入札公告等の情報を公表しているとした機関は77.3%（1495）、していない機関は23.0%（444）であった<sup>3)</sup>。

このように、8割弱の機関についてインターネットによる入札公告等が行われており、電子的な手段でデータを収集することが可能となっている。ただし、先行研究が示したように、発注者毎に様々なフォーマットで提供されているため、情報の集約には高度なシステムと相当の労力が必要であることから、これら入札情報を集約し配信することをビジネスとして行う者も存在している。これらの情報サービス事業の主な顧客は、入札を行う業者、すなわち、工事業業者や調査・設計・測量・用地等の業者である。同法に基づいて公表された発注見通し（図2.2-6(a)）、入札公告（図2.2-6(b)）、入札結果（図2.2-6(c)）の例を示す。

| 公表<br>番号 | (1)名称                 |                 | (2)工事場所                     | (3)工事種別 | (4)工事規模 | (5)工事概要                    | (6)発注時期 | (7)期間 | (8)発注事務所・課 |
|----------|-----------------------|-----------------|-----------------------------|---------|---------|----------------------------|---------|-------|------------|
|          | 路線名等                  | 工事名             |                             |         |         |                            |         |       |            |
| 1        | 一般国道423号<br>(止々呂美吉川線) | 直結Gランプ橋梁上部工事    | 箕面市下止々呂美地内                  | 鋼橋上部工   | ★       | 鋼単純箱桁橋L=39m                | 第4四半期   | 10ヶ月  | 池田土木事務所    |
| 2        | 一般府道 余野茨木線            | 道路改良工事          | 豊能郡豊能町木代地内                  | 土木一式    | ★-4     | 排水構築物工 1式                  | 第4四半期   | 5ヶ月   | 池田土木事務所    |
| 3        | 一般国道423号              | 道路防災工事          | 池田市伏尾町地内                    | 土木一式    | ★-3     | 道路防災工事 1式                  | 第4四半期   | 10ヶ月  | 池田土木事務所    |
| 4        | 主要地方道茨木摂津線            | (安元)道路拡幅工事      | 茨木市大字安元地内                   | 土木一式    | ★-2     | L=150m、擁壁工L=150m、側溝工L=150m | 第4四半期   | 12ヶ月  | 茨木土木事務所    |
| 5        | 都市計画道路十三高槻線           | 街路築造工事その1(正雀工区) | 吹田市南正雀3丁目内外                 | 土木一式    | ★-3     | 街路築造工 1式                   | 第4四半期   | 9ヶ月   | 茨木土木事務所    |
| 6        | 主要地方道大阪中央環状線          | 道路改良工事(鳥飼工区)    | 摂津市鳥飼和道2丁目地内                | 土木一式    | ★-2     | 道路改良工 1式<br>透音壁設置工 1式      | 第4四半期   | 24ヶ月  | 茨木土木事務所    |
| 7        | 都市計画道路 堺松原線           | 電線共同溝整備工事       | 松原市天美東3丁目地内<br>～松原市天美東4丁目地内 | 土木一式    | ★-3     | 推進工 L=70m                  | 第3四半期   | 3ヶ月   | 富田土木事務所    |
| 8        | 都市計画道路 大和川線           | 道路構造物設置工事       | 松原市三宅西7丁目内外                 | 土木一式    | ★-3     | コンクリート壁工 1式                | 第4四半期   | 4ヶ月   | 富田土木事務所    |
| 9        | 一般国道 (新)309号          | 道路付属設備工事        | 南河内郡河南町中地内外                 | 電気工事    | ★-4     | 付属設備工 1式                   | 第4四半期   | 6ヶ月   | 富田土木事務所    |
| 10       | 都市計画道路 堺港大堀線          | 近鉄跨線橋 上部工事(その5) | 松原市天美南5丁目地内 外               | 鋼橋上部工   | ★       | 3径間連続鋼床版箱桁橋 L=141m         | 第4四半期   | 20ヶ月  | 富田土木事務所    |
| 11       | 一般府道 大阪羽曳野線           | 舗装道補修工事(さくら町工区) | 藤井寺市さくら町地内 外                | 舗装      | ★-3     | 舗装補修 1式                    | 第3四半期   | 3ヶ月   | 富田土木事務所    |

(a)発注見通しの例

(大阪府都市整備部：http://www.pref.osaka.lg.jp/attach/4423/00014353/4kai(tosei)kouji.pdf)

(一部のみ抜粋)

入札公告(建設工事)

次のとおり一般競争入札に付します。

申請等の受付は、土曜日、日曜日及び祝日等(行政機関の休日に関する法律(昭和63年法律第91号)第1条に規定する行政機関の休日)を除く、午前9時から午後6時(電子入札の場合)、又は、午前9時15分から午後6時(紙入札の場合(下記4.(1)の担当部局の受付時間))とする。ただし、申請期限等の最終日の受付時間は、電子・紙入札ともに別表1のとおりとする。

平成26年12月10日

分任支出負担行為担当官

東北地方整備局

仙台河川国道事務所長 牧 哲



1. 工事概要

- (1) 工事名 吉野沢道路改良工事(入札保証金納付対象案件)
- (2) 工事場所 宮城県本吉郡南三陸町歌津字吉野沢 地内
- (3) 工事内容
  - 掘削工(切土工) 50,000 m<sup>3</sup>
  - 盛土工 50,000 m<sup>3</sup>
  - 函渠工 B4.7×H4.7 52.5 m
  - 排水構造物工 1式
- (4) 工期 契約締結日の翌日から平成27年10月30日まで
- (5) 工事実施形態

本工事における工事実施形態は下記のとおりとする。

- ① 本工事は、総価契約単価合意方式の対象工事である。
- ② 本工事は、価格以外の要素と価格を総合的に評価して落札者を決定する総合評価落札方式(施工能力評価型(I型))の適用工事のうち、品質確保のための体制その他の施工体制の確保状況を確認し、施工内容を確実に実現できるかどうかについて審査し、評価を行う施工体制確認型総合評価落札方式の試行工事である。
- ③ 本工事は、建設工事に係る資材の再資源化等に関する法律に基づき、分別解体等及び特定建設資材廃棄物の再資源化等の実施が義務付けられた工事である。
- ④ 本工事は、契約締結後に施工方法等の提案を受け付ける契約後VE方式の試行工事である。
- ⑤ 本工事は、現場経験の少ない技術者の技術力向上を図るため、主任技術者又は監理技術者を専

(b) 入札公告の例(東北地方整備局仙台河川国道事務所)(一部のみ抜粋)

| 平成27年度工事に係る入札及び契約に関する公表 |  |                  |                |
|-------------------------|--|------------------|----------------|
| 公表日: 27.10.27           |  | 更新日: 当初          |                |
| 名称                      | H27年度高石市橋梁長寿命化対策工事(新北新橋他)  | 施工年度             | 平成27年度         |
| 場所                      | 高石市内   | 工事担当課            | 街路河川課          |
| 種別                      | 土木一式   | 工期 契約締結日~28.3.18 | 契約者            |
| 概要                      | 【新北新橋】ひびわれ補修工1式 表面含浸工1式<br>【緑園9号橋】断面補修工1式 表面含浸工1式<br>【西取石2号橋】ひびわれ補修工1式 断面補修工1式 表面含浸工1式<br>【新仇浪橋】床版防水工1式<br>【新伽羅橋】床版防水工1式 | 住所               | 東羽衣7-3-17      |
|                         |  | 契約日              | 27.10.27       |
| 入札・契約の方法                | 指名競争入札   | 入札日時             | 27.10.26 10:00 |
| 指名・随意契約理由               | 施工能力の現状・不誠実な行為の有無・工事施工実績・経営及び信用の状況・既発注工事の受注及び進捗の状況・地理的条件   |                  |                |
| 予定価格(税抜)                | ¥8,072,000   | 最低制限価格(税抜)       | ¥6,426,000     |
| 落札金額(税抜)                | ¥6,426,000   | 契約金額(税込)         | ¥6,940,080     |
| 入札結果                    |  |                  |                |
| 業者名                     | 第1回入札金額(税抜)  | くじによる抽選の結果       | 備考             |
|                         |  | 辞退               |                |
|                         | ¥6,426,000   |                  |                |
|                         | ¥6,426,000   |                  |                |
|                         | ¥6,426,000   |                  |                |
|                         | ¥6,426,000   | 落札               |                |

(c) 入札結果の公表の例

(高石市総務部契約検査課: <http://www.city.takaishi.lg.jp/ikkrwebBrowse/material/files/group/10/nyuusatukekka2710.pdf>) (一部のみ抜粋, 業者名は公表されているが本論文の主旨に関係が無いので筆者にてマスクした)

図 2.2-6 公共工事の入札及び契約の適正化の促進に関する法律に基づく工事情報の公表例

### 2.2.5.3 契約に定めのある公共工事の情報

次に、公共工事の契約関係図書に定めのある情報について述べる。1993年、建設大臣（当時）の諮問機関である中央建設業審議会において建議された「公共工事に関する入札・契約制度の改革について」に基づき、発注者が客観的な基準により建設業者を選定することを目的として「各公共発注機関が共同で利用できる」、「自らが発注した工事施工実績情報のみならず、他機関が発注した情報も得ることができる」データベースが整備されることとなった<sup>4)</sup>。このデータベースは工事实績情報システム（コリンズ）と呼ばれ、全国的な公共工事の実績が工事实績情報として登録され、発注者は応札者の実績を他の発注機関のものであっても確認することができるようになっている。前記の「公共工事の入札及び契約の適正化の促進に関する法律」に関する国交省等の調査において、技術審査を実施している機関のうち、「建設業者の同種・類似工事の経験」を技術審査の実施項目とする発注機関は88.7%（548）となっている<sup>3)</sup>。このことから、このデータベースは発注機関において広く利用されていることが伺える。1994年に工事实績情報システムの運用が開始されてから2011年3月までの登録件数は約345万件、利用機関は、国土交通省、47都道府県、19政令市、791市区町村など合計1030（2011年3月）となっている<sup>5)</sup>。

このデータベースへの工事实績情報の登録は、国土交通省、都道府県をはじめとする多くの発注者において、契約図書（土木工事共通仕様書等）によって受注者への登録を義務付けている。その一般的な方法は、「請負金額500万円以上の工事について、契約、契約変更、工事完成のそれぞれの時点で10日以内に発注機関の確認を受け、工事实績情報を工事实績情報システムへ登録すること」となっている<sup>6)</sup>。これにより、①工事毎に一つの工事实績情報が登録されること、②契約から10日以内に工事完成予定日（工期）が登録されること、③発注機関の確認を受けていること、という特徴がある。

工事实績情報の公共工事における捕捉状況を確認するため、統計法に定められた基幹統計である「建設工事施工統計調査」との比較を行った<sup>7)8)</sup>。2006年度から2010年度までの建設工事施工統計調査における元受完成工事高に対して、工事实績情報システムに登録された工事实績情報のうち、建設工事施工統計調査と同じ発注機関を対象とした工事の請負金額（各年度に完了登録されたもの）の比率は、107.4%～98.7%であった。このことから、多くの公共工事は工事实績情報システムに登録されていると考えられる。

工事实績情報の内容は、工事件名、工事開始日・完了日、発注機関、コリンズ情報確認機関、工事場所、工種、工法・型式などである。その例（一部の項目を抜粋したもの）を表2.2-5に示す。工事实績情報は単一の事業者の基準があり、一つのデータベースに格納されているため、データの品質、データ加工性が良好である。工事实績情報のうち、発注者の承諾を得た簡易版は工事实績情報公開（一般公開）として公開されている（図2.2-7）<sup>9)</sup>。また、有料サービスであるJACIC NETにおいてコリンズ・テクリス一般検索サービスとして工事实績情報の一部が公開されている（2016年1月6日現在）<sup>10)</sup>。



図 2.2-7 工事実績情報公開の入り口画面

表 2.2-5 工事実績情報の例（一部の項目のみ）

| 項目                   | 内容  |
|----------------------|---|
| 自社が請け負った一体的先行契約工事の有無 | 一体的先行契約工事無し   |
| 工事件名                 | (一) 三刀屋佐田線吉野工区社会資本整備総合交付金(改良)工事   |
| 路線・水系名等              | 一般県道三刀屋佐田線  |
| 請負金額                 | XX,XXX,XXX (円)  |
| 工期                   | 2010年08月05日～2011年02月28日   |
| 発注機関                 | 島根県出雲県土整備事務所  |
| 受注形態                 | 単独  |
| VE対象工事               | VE対象工事ではない  |
| ISO対象工事              | ISO対象工事ではない   |
| 工事分野                 | 道路  |
| 工事の業種                | 土木一式工事  |
| 工事種別                 | 一般土木工事  |
| 工種, 工法・型式            | 土工事/掘削または切土工<br>水路・管路工事/水路・管路工<br>法面工事/法面工<br>アスファルト舗装工事/アスファルト舗装工<br>軟弱地盤処理工事/表層混合処理工法 |
| 施工場所                 | 島根県出雲市佐田町吉野地内   |
| 施工地域                 | その他の地域  |
| 夜間工事の有無              | 夜間工事無し  |
| 交通規制の有無              | 交通規制有り  |
| 規制道路の種別              | 都道府県道   |
| 規制道路の交通量(台/日)        | 5,000未満   |
| 規制車線数                | 規制車線数1/全車線数1  |
| 交通の確保手段              | 通行止,片側交互通行  |
| 近接施工の有無              | 近接施工無し  |

#### 2.2.5.4 工事情報のデータテーブルの構造化

工事情報のデータテーブルは、工事の1件を1レコードとする。

工事情報のデータ項目は、公共工事の入札及び契約の適正化の促進に関する法律等により公表されている工事情報や公開されている工事实績情報を参考にして整理した。これに加えて、後述する工事内容から道路変化イベントを予測するために、工事内容を個々の変数として扱えるようカラムを追加している。

ただし、工事の情報には個人情報と工事の契約に関する項目が含まれるものがある。これらの項目については、道路管理者の一部に、情報が道路変化イベントの把握以外に利用されることを懸念する意見があった（このことについては第6章に示す）。これまでの公示と開通の情報では、得られた情報を全て格納する方針であった、しかし、工事情報では、予測に利用しない項目を含むことによって、情報の流通に支障となることを避けるため、フィールドは道路変化イベントの把握に利用できると考えた必要最小限のものに留めた。

表 2.2-6 工事情報のデータテーブルの構造

| No. | 項目名                | 入力必要                        | 記述方法                            | データ例                  | データ生成方法  |
|-----|--------------------|-----------------------------|---------------------------------|-----------------------|--|
| 1   | データベース ID          | 入力不要                        | 自治体番号+西暦年<br>下二桁+当該年度の<br>工事の連番 | 389093005             | システム   |
| 2   | 道路管理者              | 必須                          | 自治体名                            | 〇市                    | 工事情報から直接   |
| 3   | 工事完了日              | 必須                          | 西暦, 和暦可                         | 2013/3/10             | 工事情報から直接   |
| 4   | 工事件名               |                             |                                 | 市道〇号改良工事              | 工事情報から直接   |
| 5   | 請負額                |                             | 単位円                             | 12,345,678            | 工事情報から直接   |
| 6   | 路線種別               |                             |                                 | 市道                    | 直接又は件名から   |
| 7   | 路線名                |                             |                                 | 〇号線, 市東線              | 直接又は件名から   |
| 8   | 工事起点               | 必須                          | 地名                              | 〇市〇町〇                 | 工事情報から直接   |
| 9   | 起点別名               |                             | 通称名等                            | 曙橋                    | 記載があれば   |
| 10  | 工事終点               |                             | 地名                              | 〇市〇町△                 | 工事情報から直接   |
| 11  | 終点別名               |                             | 通称名等                            |                       | 記載があれば   |
| 12  | 起点緯度               |                             | 100 進法                          | 35.1234               | ジオコーディング   |
| 13  | 起点経度               |                             | 100 進法                          | 135.1234              | ジオコーディング   |
| 14  | 終点緯度               |                             | 100 進法                          |                       | ジオコーディング   |
| 15  | 終点経度               |                             | 100 進法                          |                       | ジオコーディング   |
| 16  | 工事概要               | 16 または<br>17~33 のい<br>ずれか必須 | テキスト                            | アスファルト舗装, 路<br>盤 100m | 工事概要から   |
| 17  | アスファルト舗装           |                             | 有るときのみ yes                      | yes                   | 工事概要を空白な<br>どの区切り文字ま<br>たは形態素解析<br>で, 工種や工法を<br>示す用語に分解す<br>る。用語は, 工事工 |
| 18  | 路盤路床               |                             | 有るときのみ yes                      | yes                   |  |
| 19  | 土木一式               |                             | 有るときのみ yes                      |                       |  |
| 20  | トンネル, コンクリ<br>ート舗装 |                             | 有るときのみ yes                      |                       |  |
| 21  | 橋梁上部工              |                             | 有るときのみ yes                      |                       |  |
| 22  | 橋梁下部工              |                             | 有るときのみ yes                      |                       |  |

|    |            |  |                           |           |   |
|----|------------|--|---------------------------|-----------|---|
| 23 | 専門工事(とび等)  |  | 有るときのみ yes                |           | 種の体系ツリー, 土木用語辞典とのマッチングによる. これらを17の項目に集約する |
| 24 | 基礎工事       |  | 有るときのみ yes                |           |   |
| 25 | 土工         |  | 有るときのみ yes                |           |   |
| 26 | 道路付属物      |  | 有るときのみ yes                |           |   |
| 27 | 歩道, 情報ボックス |  | 有るときのみ yes                |           |   |
| 28 | 法面, 擁壁     |  | 有るときのみ yes                |           |   |
| 29 | 防水         |  | 有るときのみ yes                |           |   |
| 30 | 水路, 管路     |  | 有るときのみ yes                |           |   |
| 31 | 撤去         |  | 有るときのみ yes                |           |   |
| 32 | 維持保守       |  | 有るときのみ yes                |           |   |
| 33 | 緑化         |  | 有るときのみ yes                |           |   |
| 34 | 工事開始日      |  | 西暦, 和暦可                   | 2012.11.3 | 工事情報から                                    |
| 35 | 備考1        |  | 任意テキスト                    |           |   |
| 36 | 備考2        |  | 任意テキスト                    |           |   |
| 37 | 夜間工事の有無    |  | yes/no                    | yes       | 工事の詳細又は工事実績情報から                           |
| 38 | 交通規制の有無    |  | yes/no                    | yes       | 工事の詳細又は工事実績情報から                           |
| 39 | 規制道路の種別    |  |                           | 市道        | 工事の詳細又は工事実績情報から                           |
| 40 | 規制道路の交通量   |  | 1:5000 台未満,<br>2:5000 台以上 | 1         | 工事の詳細又は工事実績情報から                           |
| 41 | 規制車線数      |  | 数値                        | 2         | 工事の詳細又は工事実績情報から                           |
| 42 | 工事中の交通確保手段 |  | テキスト                      | 切り回し      | 工事の詳細又は工事実績情報から                           |
| 43 | 近接構造物      |  | テキスト                      | 中央駅       | 工事の詳細又は工事実績情報から                           |

## 2.3 原典情報の関連づけ手法

### 2.3.1 原典情報の関連づけのためのキーの選定

本研究では、原典情報を関連づけるためのキーとして、場所（空間）と時間を利用する手法を提案する。他の項目、例えば、工事件名にはその命名規則には統一的なルールが無く、独自の省略形（例えば、簡A総など。図 2.3-3 参照）も使用されている。これと比べて、場所と時間はいずれも自治体による揺らぎがなく、普遍的な指標であると言える。また、全ての原典情報には場所と時間を扱う項目がある。ただし、工事情報においては、項目があっても十分な正確性を有しない場所が登録されていることがあることに注意を要する。重要なことは、原典情報は場所に固定された事柄であり、そして、原典情報は順序のある事柄である点である。 これらの理由から、場所と時間を関連づけのキーに利用する。

## 2.3.2 場所による集約

### 2.3.2.1 原典情報における場所の表記

原典情報における場所の表記の状況を以下に述べる。

#### (1) 供用開始情報と道路区域情報（いずれも道路法に基づく情報）の場所

調査した2県2市では、告示情報の場所（区間の始点から区間の終点まで）は、日本の一般的な住所表記に準じて示されていた。最も長い形で表記すると、街区方式の住居表示を実施している区域においては、

- ✓ 都道府県 - (郡) - 市町村・特別区 - (政令指定都市の区) - 町・字 - 街区符号 - 住居番号

といった表記方法であり、それ以外の区域においては地番を用いて以下のようなになる。

- ✓ 都道府県 - (郡) - 市町村・特別区 - (政令指定都市の区) - 町・字 - 地番 - (支号)

ここでの「町・字」の町は地方自治法第260条において「市町村区域内の町若しくは字の区域」として表現される「町」である。供用開始情報と道路区域情報は道路法に基づく法律事務であり、字以下のレベルまでの場所が正確に記述されている。ただし、都道府県や市町村が明らかな場合には省略されている。また、道路は地番が付されていないため、その地点から最も近い地番を用いて「XX番地先」などの表記がされている。

#### (2) 道路開通情報の場所

住居表示に準じた表記であるが、番地までは記載されていないことが多い。

#### (3) 工事情報の場所

工事情報も住居表示に準じた表記となっている。工事によっては、管内等の広域を対象とするものがあり、その場合には、字レベルでの場所の特定はできない。また、字レベルでの記述の無いものも散見される。場所への変換の割合については後述する。工事実績情報では項目「施工場所」中に次の4種類の工事場所が登録できるが、起点住所の登録以外は任意となっている。登録方法を示すドキュメントには、「都道府県名 - 市区町村名 - 丁目番地等」との例示がされている。

- ① 起点もしくは代表地点の住所（以下、「起点住所」）
- ② 終点の住所（以下、「終点住所」）
- ③ 起点もしくは代表地点の座標（緯度経度）（以下、「起点座標」）
- ④ 終点の座標（緯度経度）（以下、「終点座標」）

### 2.3.2.2 場所を取り扱う単位（大きさ）

いずれの原典情報においても場所は日本の一般的な住所表記に準じて示されており、町丁目・字のレベルであれば関連づけられる可能性が高いと考えた。そこで、①町丁目・字のレベルにより関係する原典情報同士を有効に関連づけられるか、②町丁目・字のレベルでの関連づけは可能であるか、の2点を確認する。

#### (1) 町丁目・字のレベルの関連づけの有効性

場所の識別に町丁目・字のレベルを使用したときに道路変化イベントのあった町丁目・字

に道路変化イベントに関連の無い原典情報が混入する期待値は、原典情報の発生場所が町丁目・字に対して一様な確率分布を持つと仮定すると、式(2.1)によって求められる。ここで、Nは原典情報の発生数であり、Mは町丁目・字の数である。

表 2.3-1 に実際のデータでの計算例を示す。Mには自治体の全ての町丁目・字の数ではなく、管理する道路が明らかに存在する町丁目・字として、5年間に実際に工事が行われた町丁目・字を採用した。式(2.1)から算出した年間に工事が観測される期待値は0.2件未満であった。このように無関係な工事が対象とする区画に年間1件を超えて観測される可能性は低く、町丁目・字のレベルで識別された場所が同じであれば関連づけても問題は無いと考えた。

$$\text{町丁目・字に原典情報が観測される期待値} = \frac{N}{M} \quad (2.1)$$

表 2.3-1 5年間に道路工事が行われた町丁目・字の工事数の年間期待値

| 地方自治体名 | 年間平均道路工事件数 | 5年間に工事が行われた町丁目・字の数 | 町丁目・字で工事が発生する期待値(年間) |
|--------|------------|--------------------|----------------------|
| 埼玉県    | 4,405件/5年  | 6,787              | 0.13                 |
| 新潟市    | 1,642件/5年  | 1,957              | 0.17                 |

## (2) 町丁目・字のレベルでの集約の可能性

次に、原典情報を統一された場所名で取り扱える可能性を、地名辞典を使用して調査した。本研究では町丁目・字を扱うために次の3種類の地名辞典を使用し比較した。

### ① 郵便番号データ(日本郵便株式会社; 2012年8月時点での登録数123,314)<sup>11)</sup>

郵便番号データは、町名から〇丁目を除く部分、および大字(これを日本郵便株式会社では「町域」と呼んでいる)に設定されており、小字または通称には原則として設定されていない。

### ② 大字・町丁目位置参照情報(国土交通省; 2012年8月時点での登録数253,818)<sup>12)</sup>

字・町丁目レベル位置参照情報は、日本における住所体系のうち、市、町、村、区、特別区の直下に属す行政区である「大字」、「町丁目」、自治体によっては「町字」を示す住所代表点を集めたものである。

### ③ 街区レベル位置参照情報の大字・町丁目(国土交通省; 2012年8月時点での登録数220,992)<sup>12)</sup>

街区レベル位置参照情報は全国の都市計画区域相当範囲を対象に、街区単位(「〇〇町△丁目□番」)の位置座標(代表点の緯度・経度、平面直角座標)を整備したデータである。

これらの地名辞典の一単位が識別できる大きさを、単純に面積をデータ数で割って求めた。日本全国で見ると、最もデータ数の少ない①郵便番号データを使用した場合には、その区域を正方形としたときの一边は1.75 km (3.065 km<sup>2</sup>) となり、最もデータ数の多い②大字・町丁目位置参照情報を使用した場合には1.22 kmとなる。都市部ではこれより小さく、都市部以外ではこれより大きなものになる。例として埼玉県、新潟市において郵便番号を使用した場合の一つの郵便番号が示す町域を正方形としたときの一边の長さを表2.3-2に示す<sup>13),14)</sup>。

表 2.3-2 郵便番号が示す町丁目・字レベルの大きさ（埼玉県、新潟市）

| 自治体名 | 面積                       | 郵便番号数 | 郵便番号あたりの面積            | 一边の長さ    |
|------|--------------------------|-------|-----------------------|----------|
| 埼玉県  | 3,767.92 km <sup>2</sup> | 2,927 | 1.287 km <sup>2</sup> | 1.134 km |
| 新潟市  | 726.10 km <sup>2</sup>   | 987   | 0.736 km <sup>2</sup> | 0.858 km |

郵便番号数（ビル等に割り当てられたものも含む）を用いて自治体面積を正方形に分割した場合

3種類の地名辞典はデータの整備年月や網羅状況が異なる。いずれを用いるかべきかを判断するため、実際の工事データとの適合性の評価を行った。2006年度から2010年度に完了した道路工事の工事場所（工事实績情報）に対してこれら3種類の地名辞典を適用し、地名辞典で場所を特定（照合）できた割合で適合性の評価を行った。照合できたとする判断は、地名辞典の地名が原典情報の場所に含まれたときとした。原典情報の場所には番地などが示されている場合があり、地名辞典の地名よりも長いことがあるためである。その結果を工事实績情報における場所の登録状況と併せて表2.3-3に示す。

表 2.3-3 工事实績情報の工事場所の登録状況と地名辞典との適合性

| 年度        |                            | 2006   | 2007   | 2008   | 2009   | 2010   |
|-----------|----------------------------|--------|--------|--------|--------|--------|
| 工事件数(件)   |                            | 85,984 | 82,285 | 77,894 | 87,077 | 83,872 |
| 工事場所の登録状況 | 起点住所に登録あり(工事件数比 %)         | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |
|           | 終点住所に登録あり (同)              | 17.26  | 18.52  | 18.78  | 36.35  | 43.17  |
|           | 路線に登録あり(同)                 | 28.86  | 28.28  | 27.43  | 52.95  | 63.88  |
|           | 始点に緯度経度登録あり (同)            | 0.02   | 0.03   | 0.09   | 31.80  | 42.55  |
|           | 終点に緯度経度登録あり (同)            | 0.02   | 0.02   | 0.06   | 20.59  | 28.50  |
| 地名辞典      | 郵便番号データにて整合(工事件数比 %)       | 71.78  | 74.37  | 75.02  | 81.46  | 84.00  |
|           | 大字・町丁目レベル位置参照情報にて整合(同)     | 56.17  | 58.20  | 58.85  | 63.83  | 66.67  |
|           | 街区レベル位置参照情報(大字・町丁目)にて整合(同) | 45.52  | 48.29  | 49.23  | 52.80  | 54.28  |

場所が特定できないものは、工事場所が「管内一円」や複数の都道府県に及ぶ広い範囲の工事であって町丁目・字が示されていない工事、工事場所に示された地名が地名辞典にないものであった。表 2.3-3 に示した結果から、最も多くの工事場所が変換された郵便番号データを地名辞典として使用することとした。郵便番号データでは、一つの区画を「町域」と呼んでいるが、市町村の町と区別するため、ここでは「地点」と呼ぶ。

### 2.3.2.3 場所のクレンジング

原典情報に含まれる場所を地名辞典によって高い割合で地点に変換するために、実験で使用する工事場所に対して、次のクレンジングによって欠損の補完や誤記の修正などを実施した。

#### ① 郡名，都道府県名の補完

都道府県名と町村名が存在し、「郡名」を欠くデータについて「郡名」を補完した。次に、市区名または郡と町村名の組が存在し、都道府県名を欠くデータについて「都道府県名」を補完した。ただし、同一名の伊達市（北海道，福島県）と府中市（東京都，広島県）に対しては発注者名等から判断する。

#### ② 路線名，路線番号による補完

路線名に施工場所が登録されていたケース，路線名，路線番号から地点を推定できる場合には路線名，路線番号を利用した。

#### ③ 不要な文字列や空白の除去

「他」や「地先」などの地名以外の文字列や空白を除去した。

### 2.3.2.4 郵便番号データのクレンジング

高い割合で地点を確定させるために地名辞典（郵便番号データ）についても次の処理を施した。なお、一つの場所に対して得られた地名が複数ある場合は、文字列数が最長で一致したものを採用することで、より正確な地点を選ぶこととした。

#### (1) 揺らぎの吸収

- ✓ 市区町村名と町・字名の中の「字」，「大字」の有無が揺らいでいても対応させる。
- ✓ 「ツ」と「ッ」，「つ」と「っ」，「ケ」と「ヶ」，「エ」と「エ」等の表記が揺らいでいても対応させる。

#### (2) 工事場所の登録時期と地名辞典の整備時期の違いの吸収

- ✓ 削除された郵便番号データと追加された郵便番号データの両方に対応させる。

#### (3) 地名辞典のデータ標記の特殊性への対応

- ✓ 複数行にまたがって一つの括弧をなすデータ行も含め、全ての括弧を閉じた形にした上で処理する。
- ✓ 読点（、）や範囲指定（～）を分解し、省略されている部分を補って独立したデータ行として追加する。

#### (4) 照合速度の向上

- ✓ 「以下に記載の無い場合」などの高層ビルの各階毎に割り当てられたデータ行等を削除する.
- ✓ (町村) 一円といった具体的な町域名を表さない要素を削除する

表 2.3-4 に示すように、以上の対策を施した地名辞典により、場所の変換割合は約 3% 上昇した。

表 2.3-4 場所へのクレンジングの効果

| 年度  | 2006   | 2007   | 2008   | 2009   | 2010   |
|---|--------|--------|--------|--------|--------|
| 工事件数(件)                                   | 85,984 | 82,285 | 77,894 | 87,077 | 83,872 |
| 対策前の郵便番号データにて整合<br>(工事件数比 %)              | 71.78  | 74.37  | 75.02  | 81.46  | 84.00  |
| 郡、都道府県、路線名、路線番号による<br>補完後、郵便番号データにて整合 (同) | 74.87  | 77.34  | 78.00  | 84.29  | 87.09  |

### 2.3.2.5 地点の経緯度への変換

後述する経由地点の補間、GIS システムでの表示のため、地点に変換後、その代表点の経緯度をジオコーディング (GeoCoding) にて求める。なお、本研究では、公開されているジオコーディング用 Web-API を用いており、各地点の代表点の選定はそれらの仕様に拠ることとなる<sup>15),16)</sup>。このため、地図上に表示させたときに地図を拡大すると道路上に正しく表示されないという課題がある。

これまでの、原典情報の場所から地点への変換の概要を図 2.3-1 に示す。

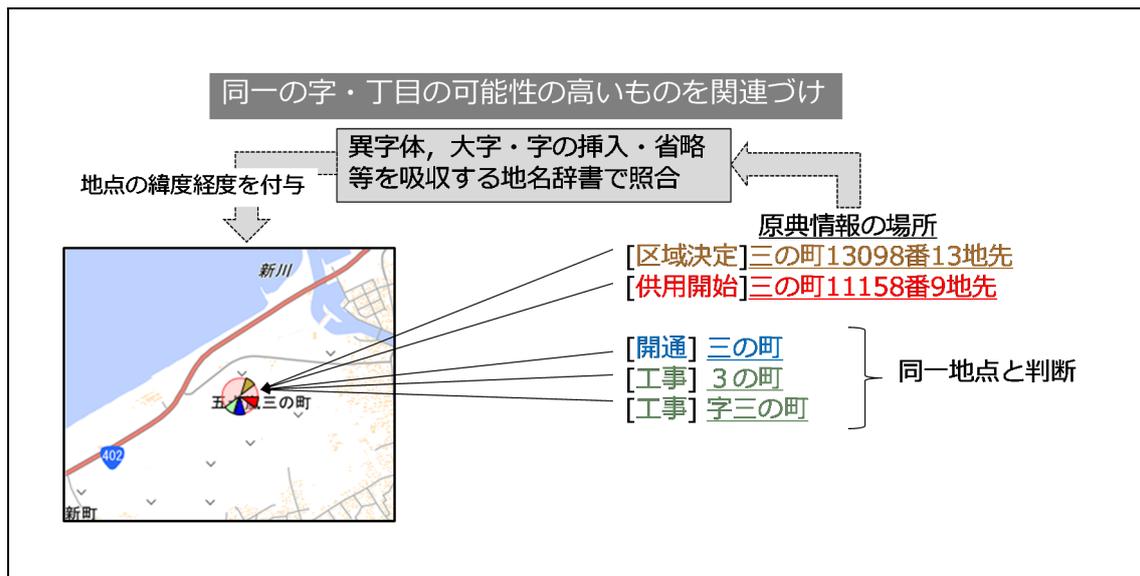


図 2.3-1 原典情報の場所から地点への変換の概要

### 2.3.2.5 起点と終点の経由地点の補間

原典情報が有する場所の情報は一つ乃至は起点と終点の二つのみである。原典情報（1）に起点と終点がある場合にそれぞれを地名辞典と照合して得られる地点は一つで完結しているとは限らず、起点と終点が異なる地点に属する場合がある。そして、起点と終点の地点に原典情報（1）を関連づけただけでは、二点間の中に別の地点（経由地点と呼ぶ）があっても、そこには原典情報は関連づかない。しかし、経由地点に別の原典情報（2）が関連づいている場合があり、このとき、原典情報（1）と原典情報（2）は経由地点を媒介に関連づけられる必要がある。

原典情報を関連づけられるべき地点に漏れなく関係づけるため、起終点が異なる地点であるときには経由地点を補間する。ただし、工事情報では近接しているときに限り、起点終点間の地点が連続して原典情報に関連するものと考えてこととした。近接していると判断する距離は3kmに設定した。これ以上離れている場合には、①登録のミス、②管内全域を工事場所とする工事（維持管理等の工事がこれに当たる）、③管内で点在する工事（道路標識等の工事がこれに当たる）、のいずれかと判断し、経由地点の補間を行わない。経由地点の補間の概念を図2.3-2に示す。

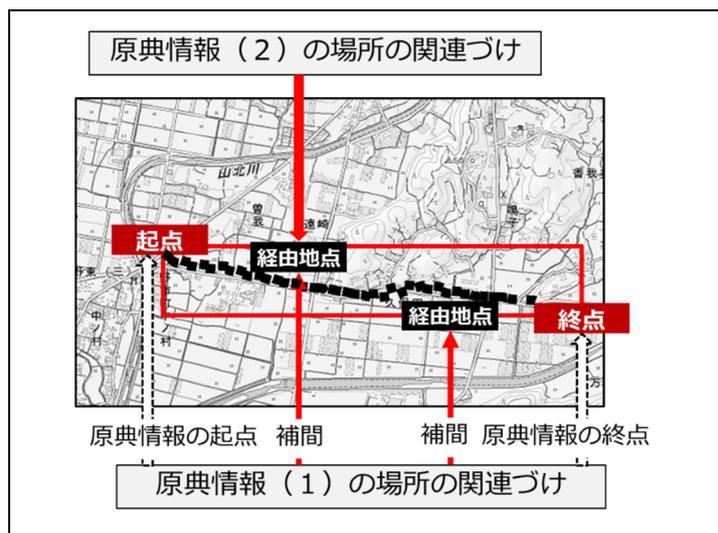


図 2.3-2 経由地点の補間の概念

経由地点の補間を含めた、場所から地点への主な流れは次の通りである。

- ① 起点終点の代表点の緯度経度をジオコーディングによって得る。
- ② 原典情報に起終点があって、異なる地点の場合には2点を対角線とする矩形領域を設定する。ただし、工事情報であるときには2点間の距離が3km以下の場合に限る。
- ③ 矩形領域の中の緯度経度の代表点を有する地点を検索する。
- ④ 検索の結果、該当する地点がある場合には、その地点を原典情報に関連する地点として起点終点に追加する。

矩形領域では無く、路線に沿って経由地点を補間することとすれば地点の関連づけは、より正確なものとなる。そのためには、路線と地点のデータが必要となる。この点については今後の課題とした。

場所が点在する工事、例えば、道路標識工事等のそれぞれの場所が道路変化イベントの把握に重要となることが十分に考えられる。そのためには、点在している工事場所、それぞれを収集することが必要となる。この情報は契約内容の詳細から得られるが、その入手には手間がかかる。このため、課題3の情報の流通に依存する課題となる。

#### 2.3.2.5 路線の併用

公示の情報には路線が含まれている。そして、工事情報においては項目として存在する工事情報（例えば、工事实績情報）や項目ではないが件名等の一部に路線が含まれている場合がある。原典情報から路線が得られる場合には、地点に加えて路線毎に集約したものを一つの地点として取り扱う。路線が不明の場合には、関連づけが失われないようにその地点の各路線に原典情報を関連づける。

号線の番号(数字)を工事件名あるいは工事場所から得る場合には、これらの項目に、「号」（「路線」と記載されていないことがあるため）及びその直前に連続する数字（半角又は全角又は漢数字）を取り出す。ただし、契約番号が○号と記載されている場合、○○号他（あるいは外）○号と記載されている場合があるので、号が出現した順序と前後の文字列でフィルタリングを行う。路線名が地名で記載されている場合には「線」の直前の文字列を発見する。ただし、「電線」、「下水幹線」などの固有の名称を除くフィルタリングが必要である。

場所から地点と路線のセットを生成して原典情報を関連づけるアルゴリズムを付録 1 に示す。

#### 2.3.3 時間による整理

次に、原典情報を時間によって整理することについて述べる。道路の建設は、段階的に基礎的で下層の部分から上部構造へと進み、それらは全て同一時期に完了するというものではない。工事の内容と完了の時期は、建設対象物の規模、自然条件、施工業者、予算の確保の状況などの様々な要素によることとなる。ある交差点の立体交差化の時期の前後に行われたその地点での工事の状況を例として図 2.3-3 に示す。

日時：平成 22 年 3 月 27 日（土）13:30～14:30

場所：国道 407 号日高バイパス 道路内

内容：埼玉県知事出席のもと、次の次第で執り行います。

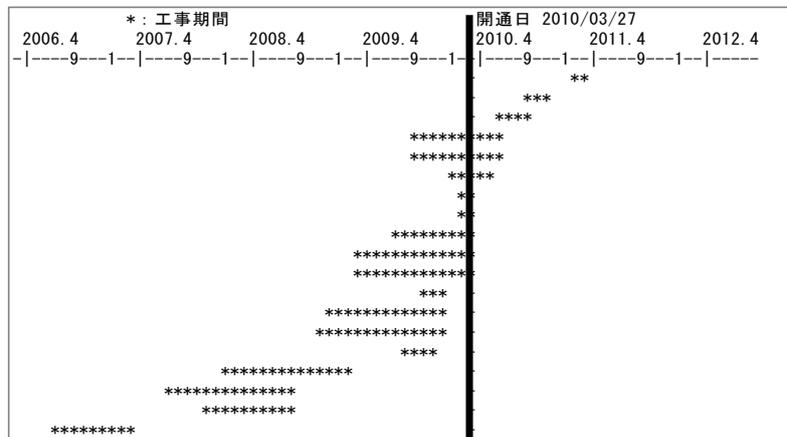
式辞、工事経過報告、感謝状贈呈、祝辞、テープカット、祝賀パレードなど

なお、式典に先立ち、11 時 45 分から地域の皆様による記念行事（踊りなど）を行ないま

す。




(a) 開通の状況



(b) この地点での工事期間と開通日

| 工事件名                     | 工事完了日    |
|--------------------------|----------|
| 舗装指定修繕工事（高萩交差点工区）        | 20101130 |
| 舗装指定修繕工事（きめ細かな交付金）（高萩工区） | 20100930 |
| 総簡加）道路改築工事（舗装工その 2）      | 20100618 |
| 総簡加）道路改築工事（舗装工その 1）      | 20100618 |
| 自転車歩行者道整備工事（高萩工区）        | 20100531 |
| 道路改築工事（舗装工その 3）          | 20100331 |
| 道路改築工事（舗装工その 4）          | 20100330 |
| 道路改築工事（市道橋上部工）           | 20100330 |
| 道路改築工事（その 3）             | 20100325 |
| 道路改築工事（その 4）             | 20100325 |
| 総簡加）舗装指定修繕工事（高萩工区）       | 20091228 |
| 総B除）道路改築工事（その 1）         | 20091225 |
| 総A加）道路改築工事（その 2）         | 20091215 |
| 舗装指定修繕工事（高萩工区）           | 20091110 |
| 簡A総）道路改築工事（その 2）         | 20090220 |
| 道路改築工事                   | 20080829 |
| 地方特定道路（交通安全）整備工事         | 20080801 |
| 道路改築工事（擁壁工）              | 20070331 |

(c) 工事件名と工事完了日（工事完了日は yyyymmdd で表記）

図 2.3-3 同一地点で異なる時期に完了した複数の工事の例<sup>17)</sup>

このような時間差のある複数の工事情報を道路変化イベントの予測に役立てることが考えられる。例えば、次のようなものである。

- (a) 同一地点での工事の発生状況を利用して予測すること（図 2.3-4），
- (b) 予測に使用する工事情報の期間（情報収集期間）の長さを調整すること（図 2.3-5），
- (c) 道路変化イベントの時期と情報収集時期の時間差（時期の差）を利用すること。

ただし、(b)において情報収集期間を長くすることにより多くの情報を取得できる可能性は高まるが、道路変化イベントとの関係が薄れる危険性がある。また、(c)において時期の差を長くとるほど、道路変化イベントが発生する時期より早期に予測が可能となる可能性はあるが、やはり道路変化イベントとの関連が薄れるために予測精度が低下する可能性がある。これらについては後述の実験で確かめる。工事の発生状況の概念を図 2.3-4 に、情報収集期間と時期の差の概念を図 2.3-5 に示す。

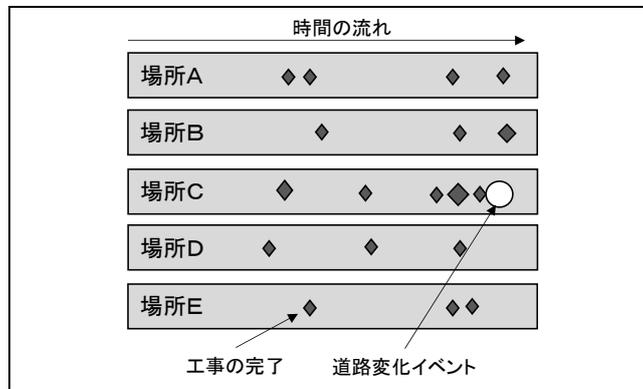


図 2.3-4 各地点の工事の発生状況のイメージ

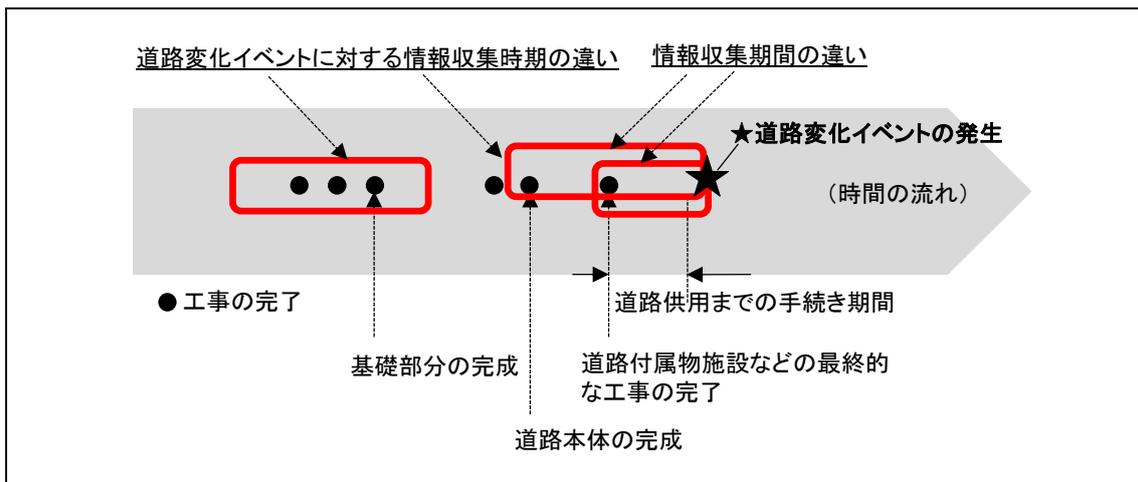


図 2.3-5 情報収集期間と時期の差のイメージ

原典情報に含まれる時間の項目を表 2.3-5 に示す。順序は一般的に左から右へと進行する。複数の時間に関する項目を有する原典情報については、道路変化イベントにより関係するものを表の\*印つき（白抜き）の項目とした。これを原典情報の時間による整理に用いる。工事情報については工事完了日であり、道路の供用については供用日、道路開通情報では開通日である。

表 2.3-5 原典情報に含まれる時間と関連づけに採用した項目

| 原典情報の種類 | 道路区域の決定 | 道路区域の変更 | 工事      | 道路の供用の開始 | 道路開通情報 |
|---------|---------|---------|---------|----------|--------|
| 時間 1    | * 告示日   | * 告示日   | 工事開始日   | * 供用日    | * 開通日  |
| 時間 2    |         |         | * 工事完了日 | 告示日      | 公表日    |

### 2.3.4 場所と時間による関連づけ

これまで述べた場所と時間によって関連づける方法をまとめて図 2.3-6 に示す。これによって場所（地点・路線）毎に点過程で得られる原典情報が時間の順に整理される。

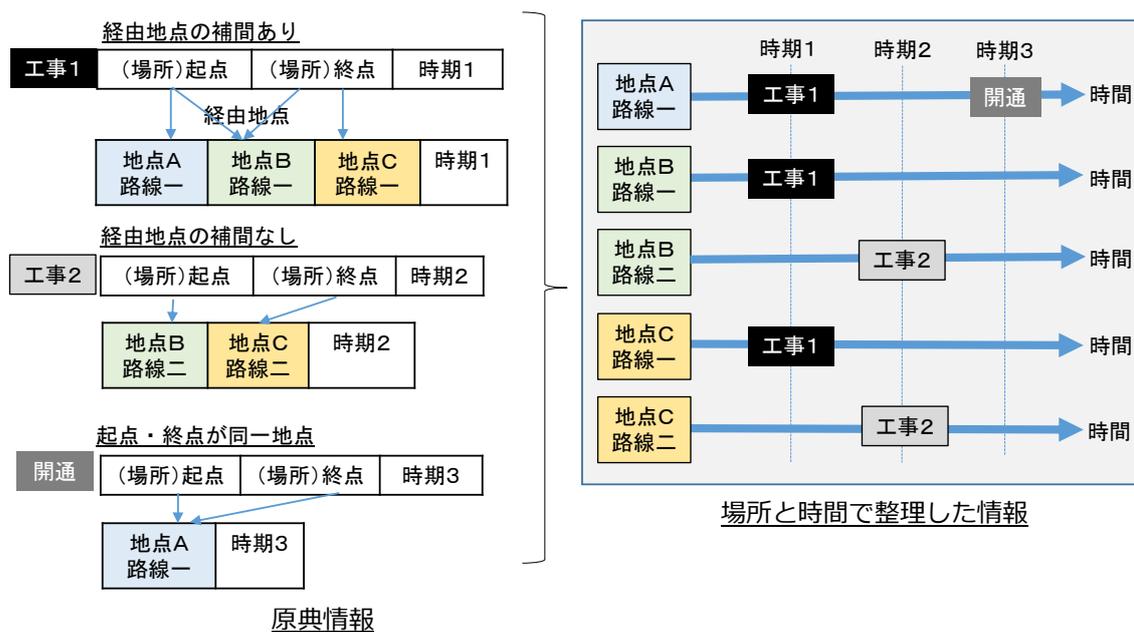


図 2.3-6 場所と時間で関連づける方法

## 2.4 統合的システムのデザイン

本節では、場所と時間で整理された原典情報を直接に確認するための統合的なシステムについて述べる。本研究では下の理由から異なる種類の原典情報のデータセットを統合的に扱うシステムを構築した。

- ① データセットをデータベースに集約することにより、データセットを横断的に利用して工事情報から道路変化イベントの予測について検討することができる。
- ② 原典情報は場所に関係したものであるから、地理情報システム上に可視化することにより、原典情報の発生状況や関連といった状況の把握が直観的に可能となる。
- ③ WEBシステムとすることで、システムへのアクセスが容易となるため、多くの利用者にとって原典情報の関連づけの利点が理解しやすくなる。
- ④ 原典情報の登録システムを備えれば、データの収集から利用までの仕組みが一体化される。

情報の入力を公開するには、誤情報への対策やセキュリティ対策等の処理が必要となり大きな負担が生ずる。このため、現時点では情報の登録は本論文執筆者が行うこととした。このように④は備えていないものの、本研究において①から③を備えた統合的なシステムを構築した。これを「道路開通情報集約システム」と名付けてインターネット上に構築した。なお、道路法には「道路開通」という言葉はないが、一般にはこの言葉が最も浸透していると思われたためシステムの名称に用いた。道路開通情報集約システム（以下、「システム」という）の構成と利用目的を図2.4-1に示す。

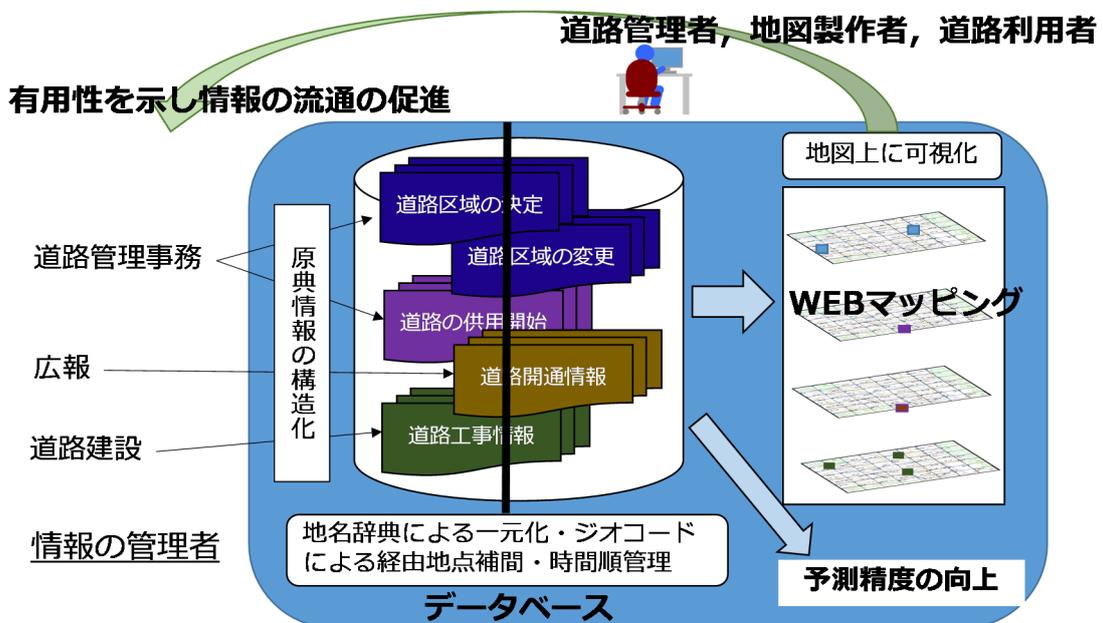


図 2.4-1 道路開通情報集約システムの構成と利用目的

## 2.4.1 原典情報の投入から地図上での表示までの流れ

原典情報の取得から地図上での表示までの流れは以下の通りである。

- ① 公表された原典情報のスクレイピングする
- ② 原典情報をデータテーブルの構造に従って構造化する
- ③ 構造化した原典情報（データセット）をデータベースに投入する。
- ④ 各レコードに含まれる場所を解析して地点に変換する。このとき変換できなかった場所はリストアップされる。
- ⑤ 全てのレコードの対応する地点を求め、地点毎に関係したレコードの基本情報（メタデータ）を集約する。
- ⑥ 各レコードの詳細データは、原典情報の種別ごとに集約する。
- ⑦ 地点毎に集約したデータと原典情報の種別毎に集約したデータは、道路管理者毎にまとめて JSON 形式のファイルとして生成する。
- ⑧ 地点毎に集約したデータは、WEB マッピングシステム（OpenLayers3）により、直観的に把握できるよう地図上に表示する<sup>18)</sup>。背景地図には、地理院地図の標準、写真及びオープンストリートマップが選択できるものとした。

## 2.4.2 ユーザインタフェース

システムのユーザインタフェース画面を図 2.4-2 に示す。

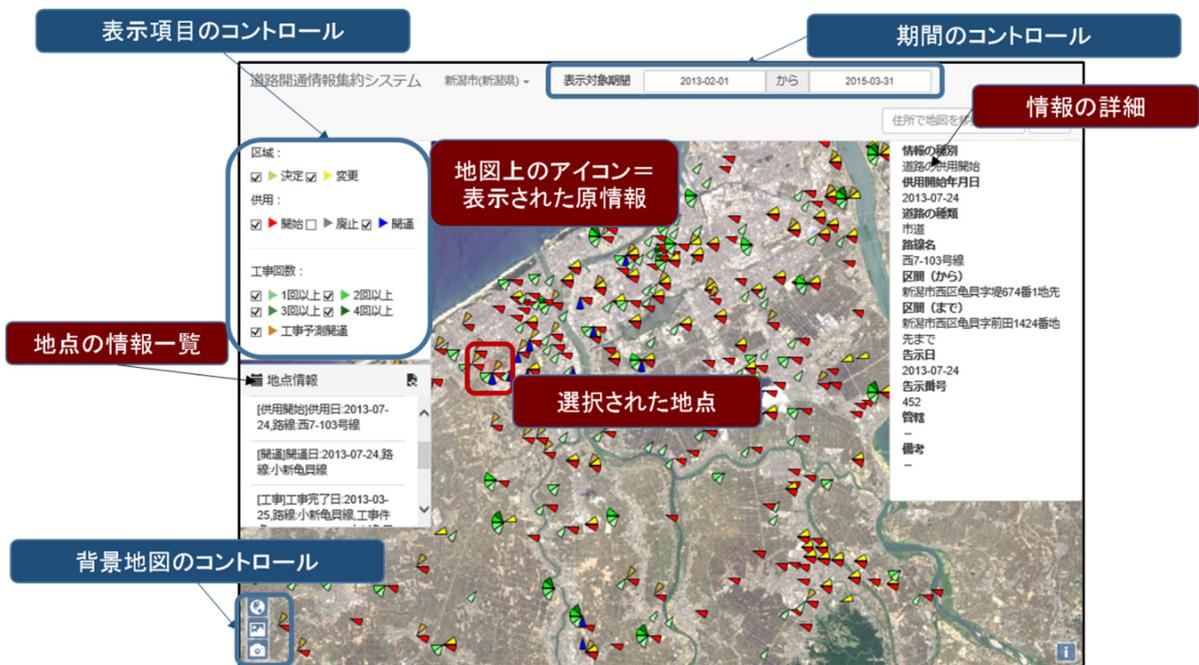


図 2.4-2 ユーザインタフェース

ユーザインタフェースは、次の機能を有している。

(1) 原典情報の地図表示機能

道路管理者毎に選択した期間の単独または複数の原典情報を表示する。同一地点に複数の原典情報を表示する場合には、重ならない配置としている（図 2.3-4 参照）。

(2) 背景地図の切り替え機能

背景地図は WMS（Web Map Service）仕様に沿ったものであれば交換が可能である。地図の更新の状況や実際の状況を確認できるよう複数の背景地図を切り替えられるものとした。一つ目は国土地理院の写真地図、二つ目は同院の標準的な電子国土、三つ目はオープンストリートマップである。図 2.4-3 にその例を示した。

空中写真を背景画面としたものでは、工事（緑）、区域、供用の開始の公示が行われた場所に工事中の道路が見られる。国土地理院の基盤地図とオープンストリートマップでは道路が完成していることがみてとれる。空中写真では撮影時期が地図の更新に追いついていないため工事中のものが掲載されていることが分かる。このように、原典情報を示すアイコンが多数表示されている場所を調査することで直観的に状況を把握することができる。



図 2.4-3 同一地点への複数の原典情報の表示と背景地図の切り替え

### (3) 地点情報の表示, 詳細情報の表示

地図による直観的な状況把握と並行して, 特定の場所の詳細なデータを確認する用途に対応するため, 場所を選択するとその詳細データを2段階で表示する機能を備えた. 多くの原典情報がある場合には, その情報の全てを一度に表示すると見にくいので, 原典情報の主要な事項のみをリスト化し画面左に表示している(図 2.4-2 参照, 図 2.4-4). 詳細情報な原典情報は地点情報の該当するものを選択すると画面右側に表示される(図 2.4-2 参照).



図 2.4-4 地点情報の例

### (4) 工事完了件数, 道路変化イベントの予測の表示

工事については選択した期間の完了件数を制御した表示ができる. また, 選択した期間の工事の集積から決定木によって道路変化イベントを予測し, 地図上に表示することができる(図 2.4-2 参照, 図 2.4-5). これらの機能を備えた理由と詳細については第 5 章に述べる.



図 2.4-5 表示項目のコントロール

## 参考文献

- 1) 道路法令研究会：第5次改訂道路管理の手引き，ぎょうせい，ISBN-13: 978-4-324-09907-0，2014.11
- 2) 川崎市：道路法の手続きに関する事務取扱要領細則，<http://www.city.kawasaki.jp/templates/outline/cmsfiles/contents/0000008/8289/file10746.pdf>，2015/12/1 アクセス
- 3) 国土交通省土地・建設産業局建設業課入札制度企画指導室：入札契約適正化法に基づく実施状況調査の結果について（平成23年調査），[http://www.mlit.go.jp/totikensangyo/cons/t/1\\_6\\_bt\\_000154.html](http://www.mlit.go.jp/totikensangyo/cons/t/1_6_bt_000154.html)，2013.1.1 アクセス
- 4) 一般財団法人日本建設情報総合センター：工事实績情報公開，<http://www.ct.jacic.or.jp/teikyoku/search.html>，2012. 10. 13. アクセス.
- 5) 一般財団法人日本建設情報総合センター：新しいコリンズ・テクリスの概要，JACIC 情報 102，Vol. 26，No. 2，pp.25-26，2011. 7.
- 6) 国土交通省関東地方整備局：土木工事共通仕様書（平成23年度）1-1-5 CORINS への登録，[http://www.ktr.mlit.go.jp/ktr\\_content/content/000011773.pdf](http://www.ktr.mlit.go.jp/ktr_content/content/000011773.pdf)，p.7，2012.10.13.アクセス.
- 7) 国土交通省：調査の概要・用語の定義，[http://www.e-stat.go.jp/SG1/estat/GL08020103.do?\\_pdfDownload\\_&fileId=000005141210&releaseCount=1](http://www.e-stat.go.jp/SG1/estat/GL08020103.do?_pdfDownload_&fileId=000005141210&releaseCount=1)，2012. 3. ，2012. 10.13. アクセス.
- 8) 国土交通省：建設工事施工統計調査完成工事高等時系列表，<http://www.e-stat.go.jp/SG1/estat/Xlsdl.do?sinfid=000012903080>，2012. 3，2012. 10. 13. アクセス.
- 9) 一般財団法人日本建設情報総合センター：工事实績情報公開，<http://www.ct.jacic.or.jp/teikyoku/cgi-bin/DS0001.cgi>，2016. 1. 5. アクセス
- 10) 一般財団法人日本建設情報総合センター：JACIC NET，<http://www.jacicnet.jacic.or.jp/jacic-net/>，2012. 10. 13. アクセス.
- 11) 日本郵便株式会社：郵便番号データダウンロード，<http://www.post.japanpost.jp/zipcode/download.html>，2012.8.16.アクセス.
- 12) 国土交通省国土政策局：位置参照情報ダウンロードサービス，<http://nftp.mlit.go.jp/isj/index.html>，2012.8.16.アクセス.
- 13) 国土地理院：平成25年全国都道府県市区町村別の面積を公表，<http://www.gsi.go.jp/kihonjohochousa/kihonjohochousa60008.html>，2014.9.10.アクセス.
- 14) 国土地理院：平成25年度都道府県別面積，<http://www.gsi.go.jp/KOKUJYOHO/MENCHO/201310/ichiran.pdf>，2014.9.10. アクセス.
- 15) Yahoo! : Yahoo!ジオコーディング API，<http://developer.yahoo.co.jp/webapi/map/openlocalplatform/v1/geocoder.html>，2013.10.1.アクセス
- 16) Google : Google Maps GeoCoding API，<https://developers.google.com/maps/documentation/geocoding/intro>，2013.10.1.アクセス

- 17) 埼玉県：国道 407 号線日高バイパス全線開通，<https://www.pref.saitama.lg.jp/a1005/kaituzyouhou/documents/405889.pdf>，2012.4.7 アクセス。
- 18) Tyler Mitchell 著、大塚 恒平、たくぼ あきお、丹羽 誠、真野 栄一、森 亮訳：入門 Web マッピング，オライリー，2006

### 第3章 原典情報の関係の基礎分析

前章で提案した手法により、各地点で発生した点過程（ポイントプロセス）の原典情報を時間の経過とともに関連づけられることとなった。本章では、この方法で整理した複数の道路管理者の実際の情報を使用して、原典情報の関係を基礎分析した。

#### 3.1 公示情報の間の関係分析

道路が供用に至るには、路線が認定された後、具体的な道路の予定区域を道路区域として決定（または変更）し、それに合わせて区域を告示し、その区域の権原（この場合は道路を建設することが法律上正当とされるための根拠となる原因）を確保し、工事によって道路を建設し、道路が利用できる状態になった後に道路を供用する、という流れが標準的なものとされている<sup>1)</sup>。ただし、権原は道路区域の決定の前に取得することもできる。この流れに従うと、道路区域に関する告示の時期に対して、道路の供用の開始に関する告示の時期は、権原の取得と工事が完了した後となるので、数箇月から数年以上後ということになる。実際に、道路区域情報から一定期間を経て供用開始情報が公表されているであろうか。ここでは、その調査のため、道路法に基づいて公示される①道路区域に関する情報と②道路の供用の開始に関する情報の間の場所と時期の関係を調査した。もし、その間の時期が一定であれば、道路区域情報から供用開始を知ることができる。

調査のもう一つの目的は、情報の相互補完の可能性を知ることである。道路区域情報には道路を構成する敷地の幅や長さが示されるが供用の日付はなく、他方、供用開始情報には道路の構成の詳細は無いが供用日が明示されている。この両者を組み合わせることができれば、供用される日付とその区間の規模を得られることとなる。

これらの情報の概要を表 3.1-1 に示す。

表 3.1-1 道路区域情報と供用開始情報

| 情報種別  | 路線の指定・認定（参考）        | 道路区域の決定                                 | 道路区域の変更 | 供用の開始                                   |
|-------|---------------------|---|---------|---|
| 情報の内容 | 路線名、起点、終点、重要な経過地を示す | 行為制限する区域を示すことが目的であり、路線名、地名、幅員、延長が記載される。 |         | 道路の供用を宣言することが目的であり、供用日、路線名、地名、供用日が記載される |
| 法的根拠  | 道路法第九条（路線の認定の公示）    | 道路法第十八条（道路の区域の決定及び供用の開始等）               |         |   |

##### 3.1.1 使用した情報

インターネットを通じて公報が入手できる 4 つ（2 つの県、1 つの政令市、1 つの市）の

自治体を選び、クローリングシステムを構築して PDF ファイルを取得し、これらを構造化しシステムに投入した。取得した PDF には、道路法に基づく告示のみが記されていてテキスト抽出が機械的に可能なもの、道路法に関連しない情報が混在して段組みがなされていて自動的に構造化が行えないものがあった。システムに投入した公示情報の数量を表 3.1-2 に示す。情報の入手先は付録に示した。1つの告示に複数の道路区域、供用開始の案件が含まれている場合、それぞれの案件を1件（1レコード）としている。表 3.1-2 の件数は投入データのデータベースレコード数（道路区域の変更についてはメタデータのレコード数）から得ている。地点・日という単位は、同じ地点であっても異なる日に発生した原典情報があれば、それぞれを1件としてカウントすることを表す。ただし、同日で同地点に同一種類の複数のレコード（例えば、番地のみが異なるレコード）があるときは1件とカウントしている。地点・日は生成されたデータベースに投入したレコードを地点変換して得られる JSON ファイルから求めている。

1件に異なる地点（地点レベル）が含まれているものが多い場合や経由地点の補間がなされるレコードが多く含まれると件数に対して地点・日が増加し、同一の地点（地点）で同日に多くの告示案件があるものは逆に減少することとなる。

表 3.1-2 公示情報の関係分析に使用した情報数

| 自治体名        | 道路区域の決定 |      |        | 道路区域の変更 |      |        | 供用の開始 |      |        |
|-------------|---------|------|--------|---------|------|--------|-------|------|--------|
|             | 件数      | 地点・日 | 変換不能地点 | 件数      | 地点・日 | 変換不能地点 | 件数    | 地点・日 | 変換不能地点 |
| 島根県 2013 年度 | 1       | 2    | 0      | 178     | 174  | 5      | 145   | 139  | 5      |
| 高知県 2014 年度 | 2       | 3    | 0      | 128     | 120  | 0      | 68    | 72   | 2      |
| 新潟市 2013 年度 | 69      | 70   | 0      | 105     | 120  | 0      | 158   | 174  | 4      |
| 金沢市 2014 年度 | 97      | 51   | 6      | 74      | 77   | 8      | 178   | 266  | 3      |

### 3.1.2 調査結果

#### 3.1.2.1 供用開始に対する道路区域の決定の発生

供用開始の地点で発生した道路区域の決定の数を、供用日との時期の差ごとに集計し、全地点の供用開始日の合計数（M）で除して、割合を求めた結果を図 3.1-1 に示す。表 3.1-2 のように2つの県では道路区域の決定が少ない。このため図 3.1-1 にも県の割合は現れていない。しかし、2つの市では全供用の 40%程度において同じ日に道路区域の決定の告示がなされていることが分かる。

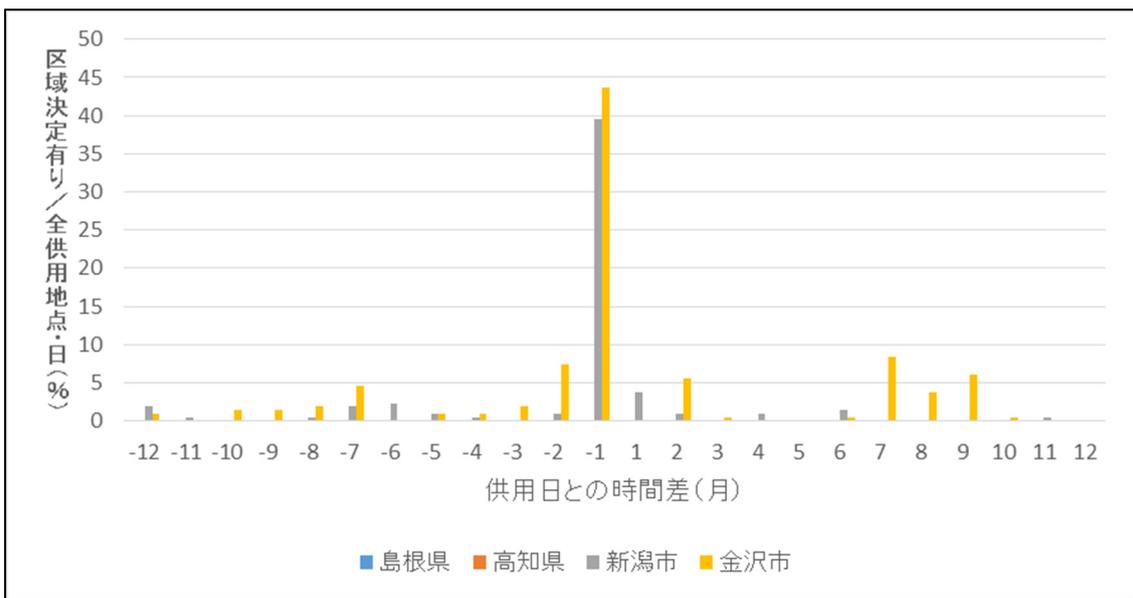
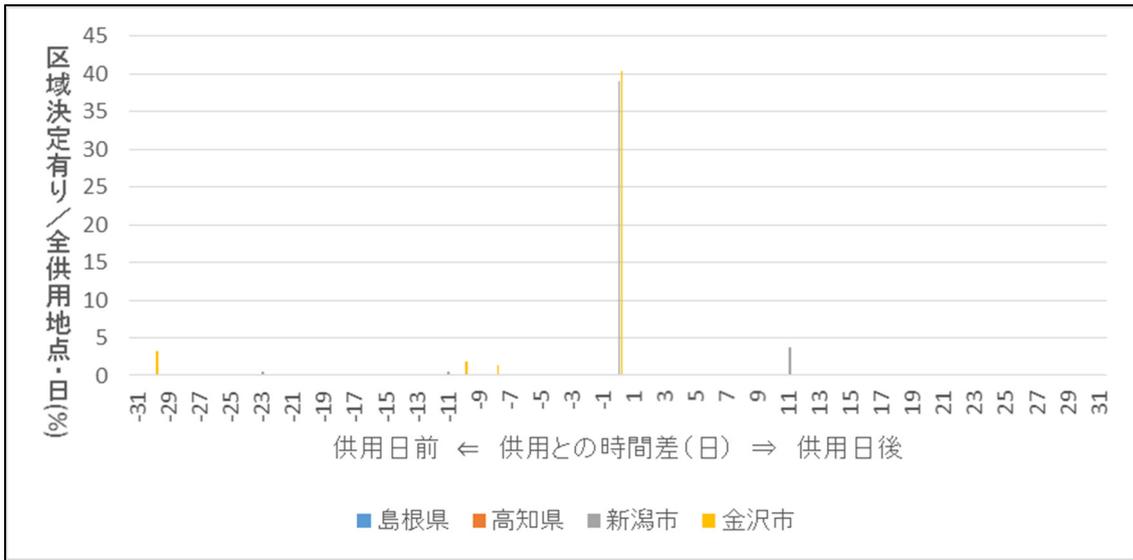


図 3.1-1 供用開始の地点・日に対する同一地点での区域決定の発生時期別の割合

### 3.1.2.2 供用開始に対する道路区域の変更の発生

道路区域の変更に対しても同様の調査を行った。その結果を図 3.1-2 に示す。道路区域の変更は、表 3.1-2 のとおり 2 つの県でも 100 以上の地点・日でなされており、図 3.1-2 において調査した全ての県・市において、全供用の地点・日の 30%以上で同じ日に道路区域の変更の告示がなされていることが分かる。

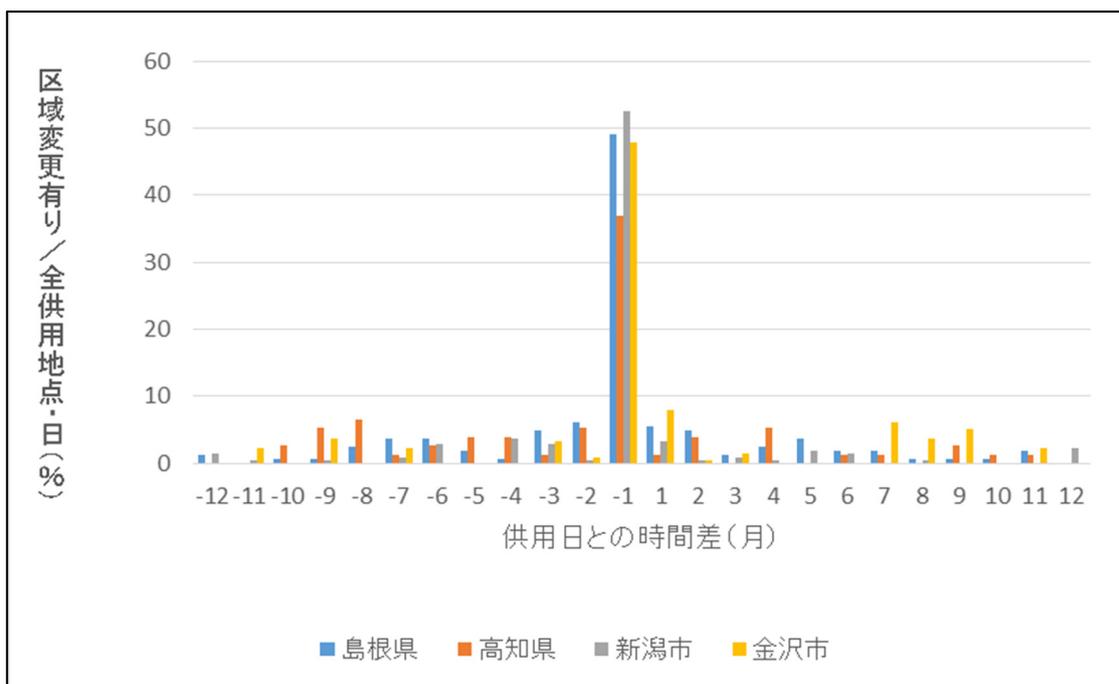
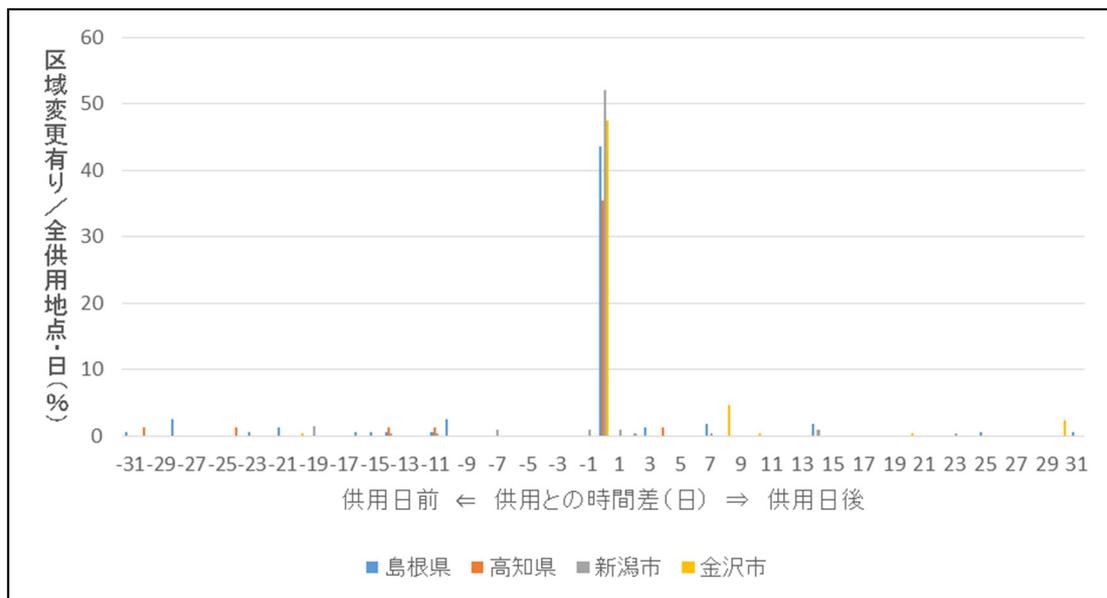
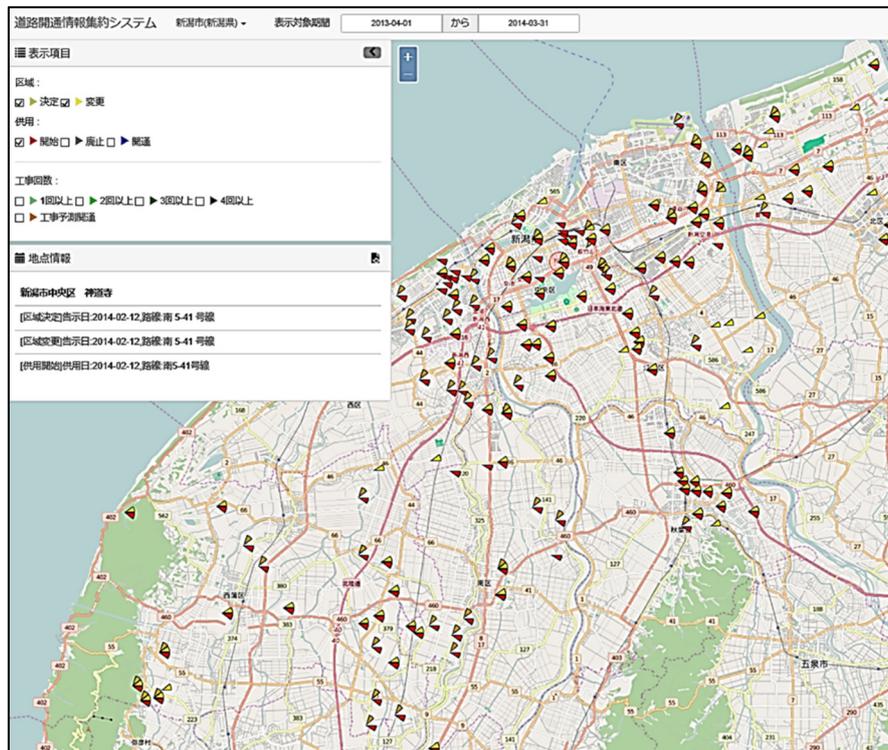


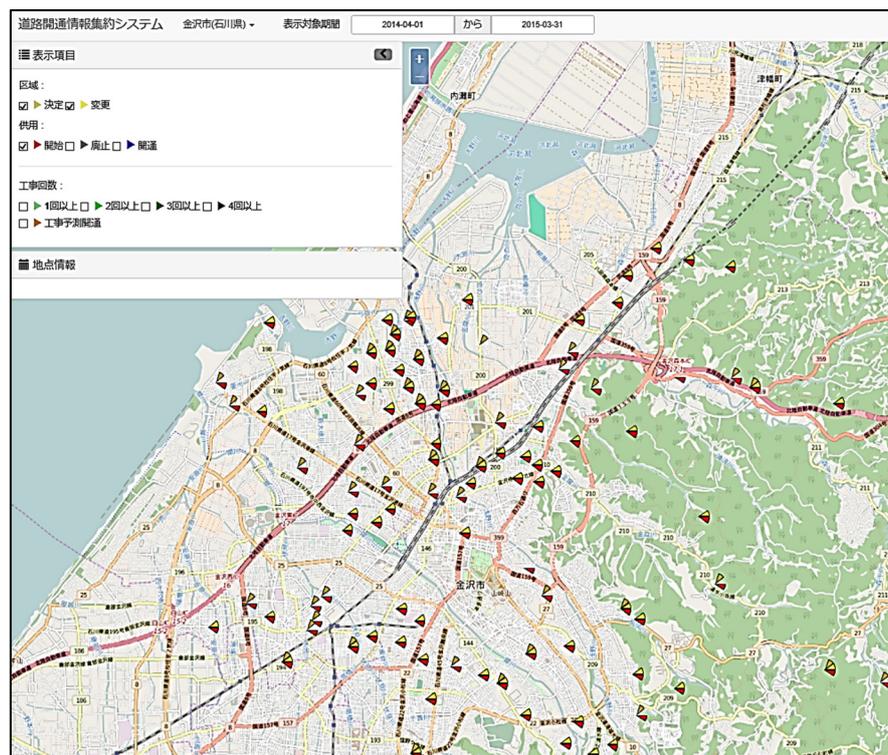
図 3.1-2 供用開始の地点・日に対する同一地点での区域変更の発生割合

### 3.1.2.3 システム画面による確認

図 3.1-3 に、2つの市のシステムでの供用の開始と道路区域情報を重ねて表示させた画面を示す。供用の開始は赤、道路区域の決定は茶色、変更は黄色のアイコンで示される。表示対象期間内の同一年度の多くの地点で供用開始と道路区域に関する情報があることが見られる。



(a) 新潟市



(b) 金沢市

図 3.1-3 システム画面による供用開始と道路区域情報の対比

### 3.1.3 供用開始と道路区域情報の関係のまとめ

#### (1) 道路区域情報から供用開始の予測可能性

道路区域情報と供用開始の手続きの流れについて、例えば、最高裁第一小法廷判決昭和四十四年十二月四日では、「・・・ところで、道路法に定める道路を開設するためには、原則として、まず路線の指定または認定があり、道路管理者において道路の区域を決定し、その敷地等の上に所有権その他の権原を取得し、必要な工事を行って道路としての形体をととのえ、さらに、その供用を開始する手続に及ぶことを必要とするものであつて、・・・」とある。このような原則であるが、実際には、道路区域情報から一定期間を経て供用開始情報が公表されている傾向は確認できず、両者の 30%から 40%は同時に公表されていた。このため、道路区域情報から供用開始を予測できる可能性は低いと考えられる。

ただし、道路区域情報を供用開始と同時にに行えないというものではない。なぜなら、工事区域の決定（または変更）の公示は、区域の決定（または変更）の事実を一般に知らせるための行政上の行為であり、公示は区域の決定（または変更）行為の効力発生要件ではないからである。しかし、供用の開始の公示は供用の開始の行為が有効に成立するための効力発生要件であるため、道路を利用に供するためには必ず行わなければならない。このようなことから、実際には事務作業をまとめて処理がなされていると推察している。

#### (2) 情報の補完の可能性

道路区域情報と供用開始情報はおおよそ 30%から 40%が同時に同じ場所について公表されている。これらの地名で示されている区間の一致が一致すれば、両者は同じ区間に対する情報であると考えることができるため、道路区域情報に含まれる延長と幅員を、供用開始情報の供用日と組み合わせて、供用開始された区間の規模と日を知ることができると考えられる。

## 3.2 道路の供用の開始と工事完了件数の関係の分析

道路の供用の開始の前には、その区域で工事が行われると限らないが、①道路を供用するには一般交通の用に供して差し支えの無いように整備される必要があり、そのために工事が行われること、②工事が完了して供用可能となっている道路を、理由なく供用の開始の告示をしないで長期間放置することはないであろう、という理由から道路の供用の開始の前に当該区域での工事が行われ、逆に工事により通行可能となった区域は可能な限り速やか供用を開始するケースが多く見られるのではないかと考えた。

### 3.2.1 使用した情報

前節の調査で投入した原典情報（道路の供用の開始）と対比するため同じ自治体に対して、公開されている工事情報（2500 万円以上）を収集した。情報の入手先は、告示と同様に付録にまとめて示した。システムに投入した情報の数量を表 3.2-1 に示す。

表 3.2-1 工事情報の関係分析に使用した情報数

| 件数  | 2010 年度 | 2011 年度 | 2012 年度 | 2013 年度 | 2014 年度 | 合計   | 地点数  | 未変換数 |
|-----|---------|---------|---------|---------|---------|------|------|------|
| 島根県 | 462     | 399     | 367     | 358     | 338     | 1924 | 1784 | 324  |
| 高知県 | 267     | 207     | 211     | 291     | 218     | 1194 | 1023 | 236  |
| 新潟市 | 107     | 98      | 93      | 84      | 111     | 493  | 446  | 66   |
| 金沢市 | 26      | 8       | 14      | 16      | 27      | 91   | 85   | 11   |

### 3.2.2 調査結果

全ての供用の地点のうち、同一年度に完了工事が観測された地点の割合を図 3.2-1 に示す。完了工事が観測された地点の全ての供用地点・日に対する割合を供用日との時期別に並べたものを図 3.2-2 に示した。

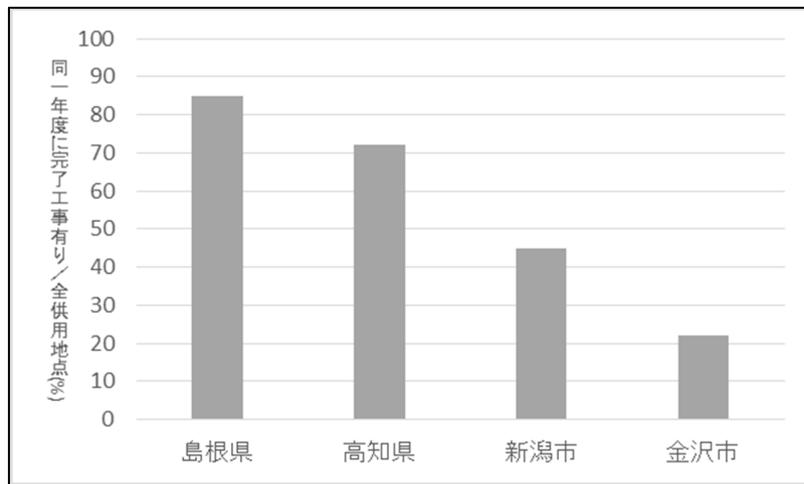


図 3.2-1 同一年度に 2500 万円以上の完了工事があった供用地点の割合

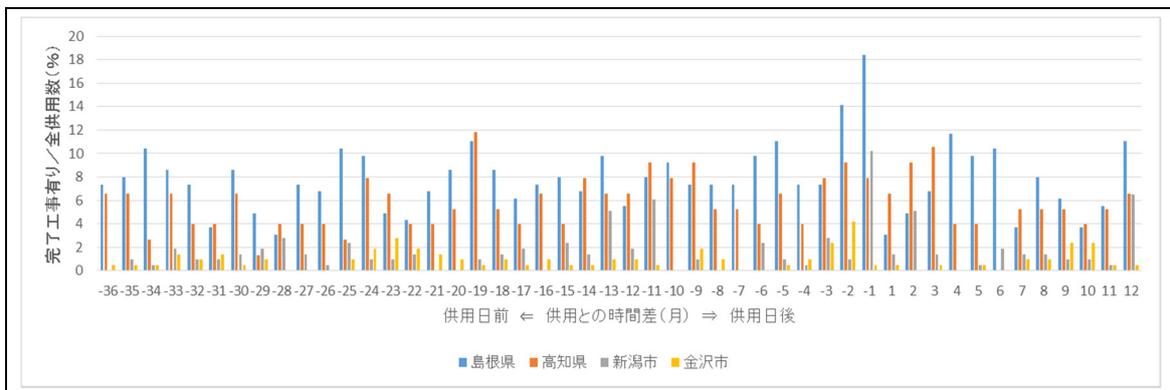


図 3.2-2 供用地点に対する工事の完了状況

供用の開始と工事との関係を表示した画面例を図 3.2-3 に示す。供用開始は赤、工事はその地点で 1 件のものは薄い緑、回数が増えるにつれて濃い緑のアイコンが追加される。



図 3.2-4 供用の開始と工事の対比

### 3.2.3 供用開始と工事完了件数の関係のまとめ

供用の開始の前には公共事業費の大きな自治体ほど多くの割合で工事が行われていた。ただし、ここでは調査対象の工事をその請負額が 2500 万円以上のものとしていることに注意する必要がある。供用開始日と工事完了日との関係では、供用開始地点に多くの完了完了時が観測された島根県において、供用開始と同時期に通常より多くの完了工事が見られた。

### 3.3 道路の開通と工事の関係の分析

#### 3.3.1 使用した情報

前節で調査した自治体のうち、開通情報を発見できなかった金沢市を除く3自治体に対して開通情報を投入し、開通と工事との関係を調査した。島根県については、農道が明示されていた。道路法の適用を受けない道路の開通について工事情報から発見できる可能性を調査するため、農道と一般道とを別に示す。システムに投入した開通情報を表3.3-1に示す。

表 3.3-1 使用した開通情報数

|                   | 件 | 地点 | 未変換地点 |
|-------------------|---|----|-------|
| 島根県 2013 年度 (一般道) | 2 | 2  | 0     |
| 島根県 2013 年度 (農道)  | 2 | 3  | 0     |
| 高知県 2014 年度       | 3 | 6  | 0     |
| 新潟市 2013 年度       | 4 | 6  | 0     |

#### 3.3.2 調査結果

全ての開通の地点のうち、完了工事が観測された割合を供用と合わせて図3.3-1に示す。開通は供用に比べて規模が大きいためその地点で工事が観測される割合が高い。開通日に対してそれぞれの地点において月ごとに完了工事が観測されたものについて、全ての開通地点に対する割合を図3.3-2に自治体別に示す。この図と表3.3-1の開通地点数を考え合わせると、開通日の前に同一の地点で複数の完了工事が発生していることが分かる。

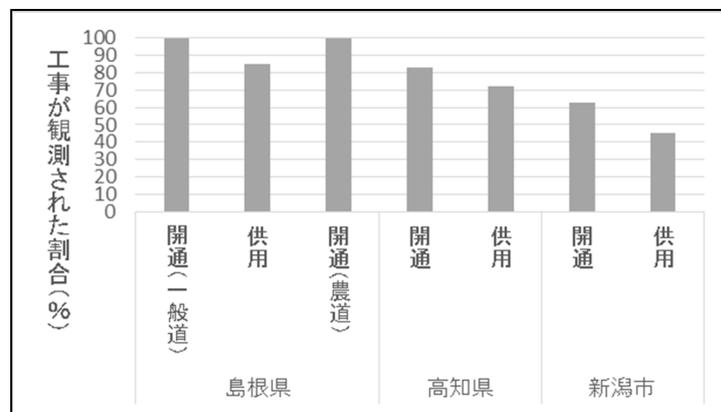


図 3.3-1 同一年度に 2500 万円以上の工事が観測された開通地点、供用地点の割合

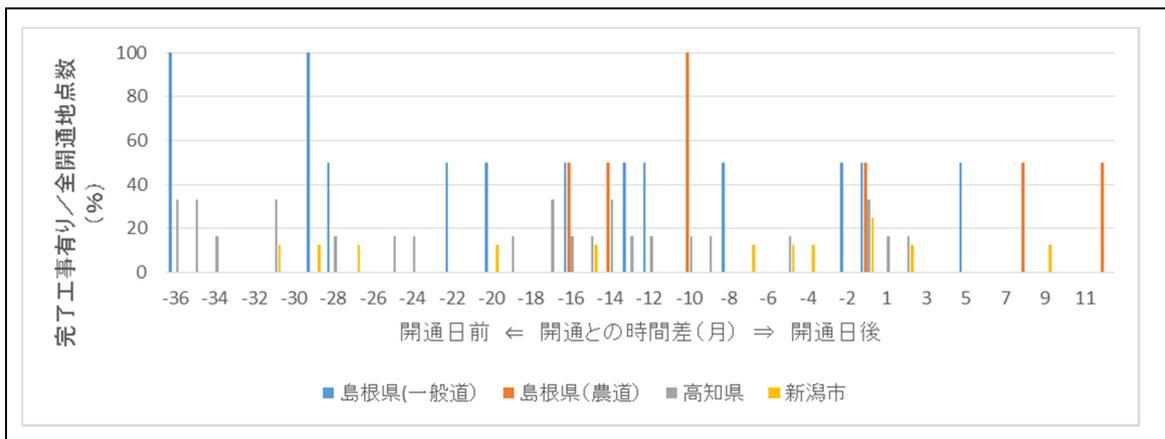


図 3.3-2 開通地点における工事の全開通数に対する割合

工事の「件数」と開通との間の対応関係を把握するため、供用の開始、開通（一般道、農道）があった場所における工事完了の調査年度の総件数を調査し図 3.3-3 に示した。開通地点では地点当たりの工事数が多いことが分かる。開通と工事（件数）を対比させた画面の例を図 3.3-4 に示す。

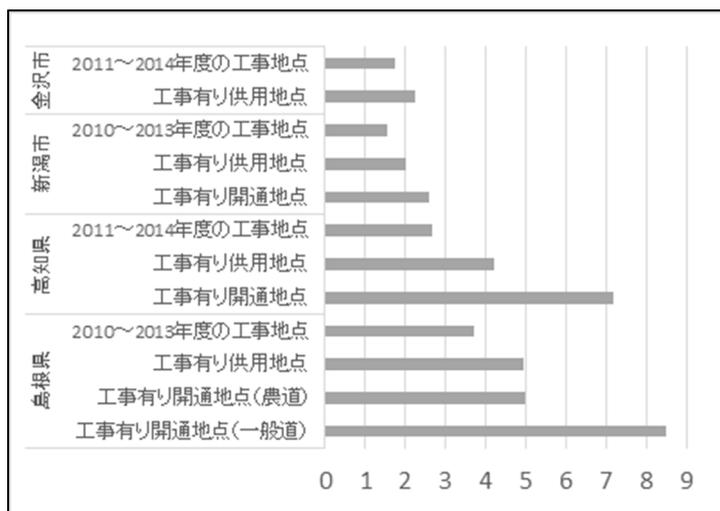


図 3.3-3 供用・開通・全ての地点の地点当たりの調査年度の工事数



図 3.3-4 開通と工事の対比画面例

### 3.3.3 道路開通と工事完了件数の関係のまとめ

供用開始地点、開通地点では地点当たりによくの工事が完了しており、特に開通地点においてはより多くの工事が完了していた。このため、工事完了件数から道路開通の予測の可能性はある。

## 参考文献

- 1) 国土交通省道路局路政課：道路法令 Q & A（道路供用開始の手続きと道路法第 90 条について），[http://www.hido.or.jp/14gyousei\\_backnumber/2014data/1407/1407houreiQ&A.pdf](http://www.hido.or.jp/14gyousei_backnumber/2014data/1407/1407houreiQ&A.pdf), 2013.10.1. アクセス.

各データの入手先はまとめて付録に示した.

## 第4章 工事情報からの道路変化イベントの簡易的な予測

### 4.1 工事情報からの予測により得られる効果

工事情報から道路変化イベントを予測することによって次の効果が得られる。概念を図4.1-1に示した。

#### (1) 道路変化イベントの発生までのリードタイムの確保

工事は道路変化イベントの発生前に先立って行われる。加えて、工事情報は入札契約より前あるいは契約直後に得られるものがある。このため、道路変化イベントの発生まで工事期間以上のリードタイムを確保できる。また、長期的な建設工事（下層・基盤部分の工事）を予測に用いられればさらに長期間のリードタイムを確保することができる。

#### (2) 効率化

多くの情報公開請求が工事図面に対してなされている。また、地図製作者からの依頼で工事図面を抽出する際にも、予測によって請求する工事を絞り込むことができれば、請求者、提供者の双方の作業が効率化される。また、道路変化イベントの発生を直接示す告示情報や開通情報がインターネット上に公開されていない場合があるが、工事情報については異なる方針、ルートで公開されており、入手方法の多様化が図られる。

#### (3) 農道、林道等の把握

農道や林道等は道路法で定められていないが利用に供されている。これらは、道路法に基づく告示がなされない。しかし、これらについても建設工事が行われるため、工事により道路変化イベントを把握できる。

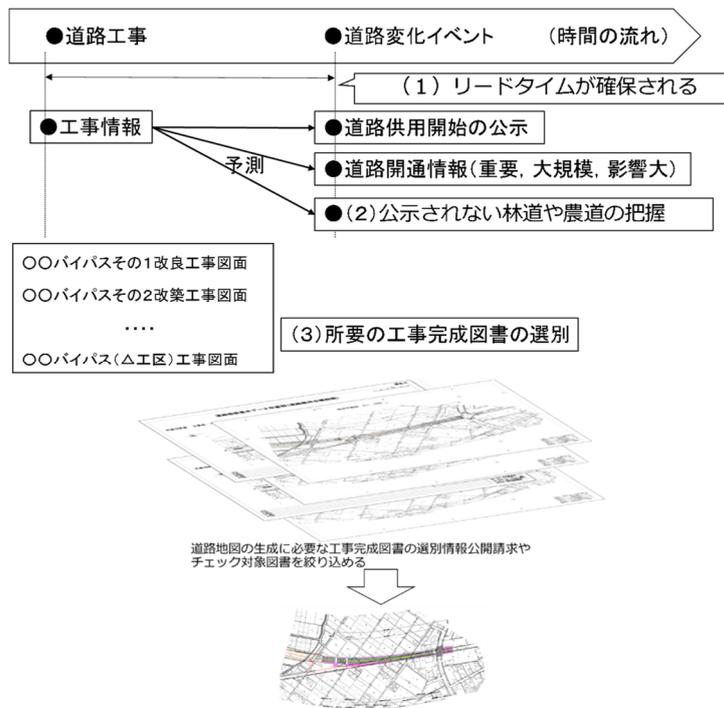


図 4.1-1 必要な図面の絞り込みの概念

## 4.2 予測の評価方法

### 4.2.1 実際と予測結果の組み合わせの分類

ある地点は、実際の道路変化イベントの有無と予測された道路変化イベントの有無によって、次の4種類に分けられる。

- ① 道路変化イベントが有り、有りとして予測されたもの。
- ② 道路変化イベントが有るが、無しとして予測されたもの。
- ③ 道路変化イベントが無いが、有りとして予測されたもの。
- ④ 道路変化イベントが無く、無しとして予測されたもの。

注目するクラス（グループ）を陽性（Positive）とするのが一般的な呼び方である。予測結果に道路変化イベントが有るものを陽性（Positive）、無いものを陰性（Negative）と呼び、実際と予測とが一致するものには真（True）、一致しないものには偽（Negative）を付加して表現する。これにより、①は真陽性（True Positive）、②は偽陰性（False Negative）、③は偽陽性（False Positive）、④は真陰性（True Negative）と呼ばれる。

①から④を一つの表に整理したものが表 4.2-1 混同行列（confusion matrix）である。

表 4.2-1 混同行列

|               | 予測結果が道路変化イベント有り（陽性, Positive） | 予測結果が道路変化イベント無し（陰性, Negative） |
|---------------|-------------------------------|-------------------------------|
| 実際に道路変化イベント有り | 真陽性<br>TP (True Positive)     | 偽陰性<br>FN (False Negative)    |
| 実際に道路変化イベント無し | 偽陽性<br>FP (False Positive)    | 真陰性<br>TN (True Negative)     |

### 4.2.2 予測結果の評価

実際の道路変化イベント有りを正しく予測し捕捉する割合は、真陽性率（True Positive Rate, TPR）あるいは再現率（Recall）または感度（Sensitivity）と呼ばれ、式 (4.1)で表される。これは道路変化イベントを把握できた割合であり1に近づくことが望ましい。これに対して、実際の道路変化イベント無しを正しく捕捉する割合は、真陰性率（True Negative Rate, TNR）あるいは特異度（Specificity）と呼ばれ、式(4.2)で表される。実際に道路変化イベントが無いにも係らずその中から有りとして予測した割合は、偽陽性率（False Negative Rate）であり（式(4.3)）、この値が高いほど無駄な詳細調査（問い合わせや資料要求）が増加するため、偽陽性率はゼロに近づくことが望ましい。なお、偽陽性率 = 1 - 真陰性率である。

前章の表 3.2-1、表 3.3-2 から、道路変化イベントの有る工事地点は全ての工事地点に比べて少ないことがわかっている。このため、道路変化イベントの有る地点と無い地点を合計した値で評価せずに、それぞれを独立して評価することとした。なぜなら、合計して評価すると全てを無しとして予測することで良い評価が得られてしまうからである。

$$\text{真陽性率, } TPR = \frac{TP}{TP + FN} \quad (4.1)$$

$$\text{真陰性率, } TNR = \frac{TN}{FP + TN} \quad (4.2)$$

$$\text{偽陽性率, } FPR = \frac{FP}{FP + TN} \quad (4.3)$$

予測方法の有効性の尺度には受信者操作特性（Receiver Operating Characteristic, ROC）を用いる。ROC は横軸に偽陽性率，縦軸に真陽性率をとり，判定の閾値（カットオフポイント）を媒介変数として変化させた時の偽陽性率，真陽性率の値をプロットして得られる ROC 曲線で表される。予測方法が有効であるほど ROC 曲線は左上にシフトすることになる。定量的には ROC 曲線の下面積（Area Under the Curve, AUC）で示され，完全に予測できたときには AUC は 1 となり，ランダムな予測のときには 0.5 となる<sup>1), 2), 3)</sup>。

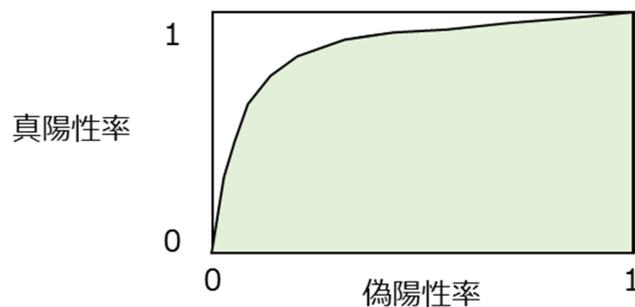


図 4.2-1 ROC 曲線

### 4.3 工事件数による簡易的な開通予測

前章で見たように開通のあった地点では平均的な工事地点に比べて工事件数が多い。この節では，開通との工事件数の関係についてさらに調査し，工事件数による開通の予測を試みる。

開通情報として公表される区間の建設作業（工事）は一般的に通常より大きな規模となると考えられる。このような大規模な建設作業は，次の理由から複数の工事に分離，分割して発注され，工事件数が多くなると考えられる。その根拠の一つは「官公需についての中小企業者の受注の確保に関する法律」，いわゆる官公需法の第 4 条第 2 項中小企業者に関する国等の契約の基本方針の作成等である。その中では中小企業・小規模事業者が受注し易い発注とする工夫の一つとして分離・分割発注の推進が示されている。また，「地方自治体では，

地元の中小企業を保護するために、工事を細かく分割して、それぞれについて入札をするという分割発注を行うことが多い。」ことが大野(2003)により整理されている<sup>4),5),6)</sup>。

また、このほかの理由としては、大規模な建設作業となる場合には、必要となる技術や資材は多くの種類に及ぶと考えられる。そして、工事実施者（工事請負業者）は建設業法に示された28に分類された業種ごとに建設業許可を受けなければその分野の工事を行えないため、専門的な工種ごとに分離発注されることが考えられる<sup>7)</sup>。

このように工事件数は建設作業の量と多様性の両者を総合的に示す指標になると考えられ、開通として公表される区間では工事件数が増加していると考えられる。

本節では、道路上での道路管理者発注工事を抽出し、その工事件数によって開通の予測を行った。

#### 4.3.1 調査データ

ここでは、道路開通情報を公開している埼玉県、愛媛県、新潟市の3つの地方自治体について500万円以上の公共工事を対象に開通と工事との対応関係の実際を調査した。使用したデータは次の通りである。

表 4.3-1 使用した工事情報

|     | 工事件数    | 地点   | 場所未変換工事件数 |
|-----|---------|------|-----------|
| 埼玉県 | 4,405 件 | 1300 | 414 件     |
| 愛媛県 | 3,734 件 | 686  | 450 件     |
| 新潟市 | 1,704 件 | 501  | 62 件      |

工事完了期間：2006年4月から2011年3月

表 4.3-2 使用した開通情報

|     | 開通区間数 | 地点 | 未変換 | 道路開通区間長    | 平均区間長  |
|-----|-------|----|-----|------------|--------|
| 埼玉県 | 16    | 28 | 0   | 0.2～7.0 km | 1.3 km |
| 愛媛県 | 11    | 20 | 0   | 0.6～1.8 km | 1.0 km |
| 新潟市 | 14    | 18 | 0   | 0.2～2.5 km | 1.1 km |

開通情報調査対象期間：2007年度～2010年度

#### 4.3.2 開通の有無と工事件数に関する調査

##### (1) 開通区間の工事件数

個々の開通区間における工事件数を調査した。開通区間は、道路開通情報に示された一つないし複数の開通場所（単位は地点）から構成される。開通までの12箇月間に完了した工事件数の全開通区間数に対する割合を図4.3-1に示す。

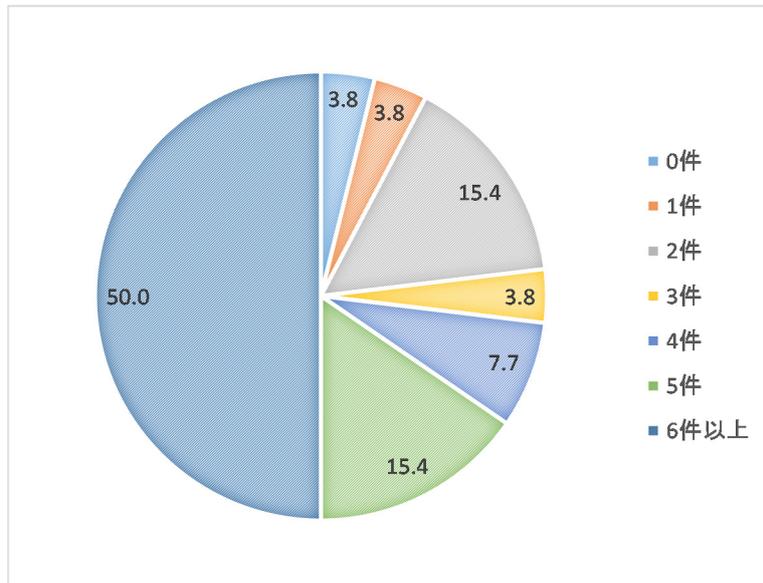


図 4.3-1 工事完了件数（開通地点における開通月～開通 11 箇月前までの 12 箇月間）

96%の開通区間においてこの期間に完了した工事があった。そして、92%で複数の工事が観測された。このように、開通区間において多くの工事が発生していることを定量的に示すことができた。このルールに当てはまらなかった事例は以下の通りである。

① 大規模なイベントで工期と開通時期がずれたケース

新潟市では調査年度に国民体育大会が開催され、これに対応するため開通が前倒しされたとの事例があった。

② 大規模な道路の開通に整合させて開通がずれ、件数が少ないケース

大規模な自動車専用道のインターチェンジ新設に伴ってその直下の 200 メートルの県道の開通が行われたもの（埼玉県川島町平沼 0.2km）では、周辺の道路整備が埼玉県とは別の道路管理者によって整備されたため、埼玉県の工事の件数が少なく、また、大型の道路開通の影響を受けたため工事時期と開通時期が離れたと考えられる。このようなケースの場合には、道路変化イベントの情報は、大型の工事（自動車専用道など）から得て、その周辺の他の道路管理者の情報をあわせて調査することがよいと考える。

③ 災害復旧のための緊急工事であり件数が少ないケース

愛媛県名取トンネルの開通は大きな工事の割に件数の少なく 3 件であった。これは地すべりで旧の名取トンネルの片側出入口を塞ぎ、別の出入口を緊急に建設する工事であったため主要な工事が 1 件で行われた。本検討では工事件数のみを対象としており、工事請負額は考慮していないが、本件の請負額は約 14 億円であり、平均的なものの数十倍の規模であった。このため、極めて工事請負額が高額なものなどについては、複数件数として扱うなどの特別な規則によってデータ化することにより件数のみで検

討する手法を補正することが考えられる。

④ 工事場所が正しく登録されていないケース

寺村バイパス（愛媛県）である。寺村バイパスは内子町寺村に位置し、場所を特定するためには、地点のレベルである「寺村」が必要である。工事実績情報を調査したところ、開通日の前後2年間に工事場所の登録として「内子町」だけのものが5件あり、いずれも路線は寺村バイパスと同じ国道380号線、件名も道路改築工事であった。うち一件については請負業者の施工実績のホームページにあった写真を拡大して得られた銘板にあった水元橋から施工場所が内子町寺村すなわち寺村バイパスであることを確認し工事場所を補完したが、それ以外の工事場所については確認できなかった。このため開通に関連した工事としては利用していない。このように、工事場所の登録が不完全であるために場所による集約が行われず、結果的に件数が少なくなるものがある。

⑤ 不明

工事数が1であったものが1箇所（埼玉県鶴ヶ島市五味ヶ谷 0.2km）、2であったもので上記以外のものが1箇所（新潟市の「都市計画道路河渡線（はなみずき）（暫定開通）」）あり、これら実際の工事件数が少なかったのか、工事データの登録上の問題であったか、工事時期が開通時期と大きく異なるのかは不明である。

(2) 非開通場所における年間工事件数とその出現割合

図4.3-2は、地点当たりの工事完了数を全ての非開通場所について示したものである。これによると年間（年度単位）の工事数が1件だけという場所が46~64%を占めており、場所当たりの工事数が多いものの割合は指数関数的に減少している。各自治体の各年度の値（図4.3-3）を見ても、この傾向は年度によるばらつきが少ないことがわかる。

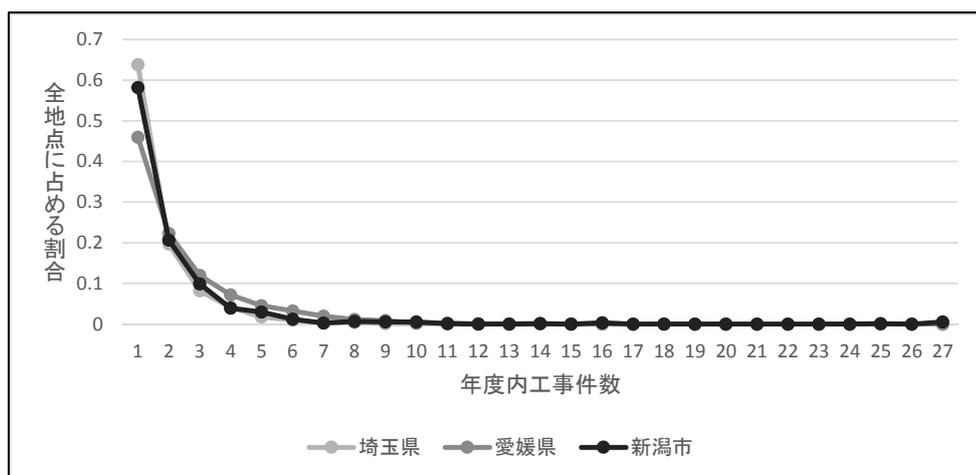
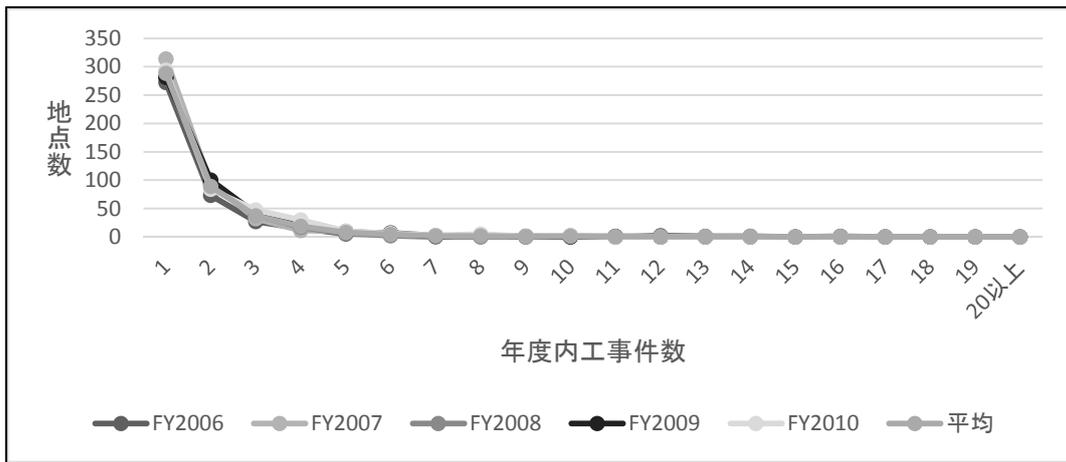
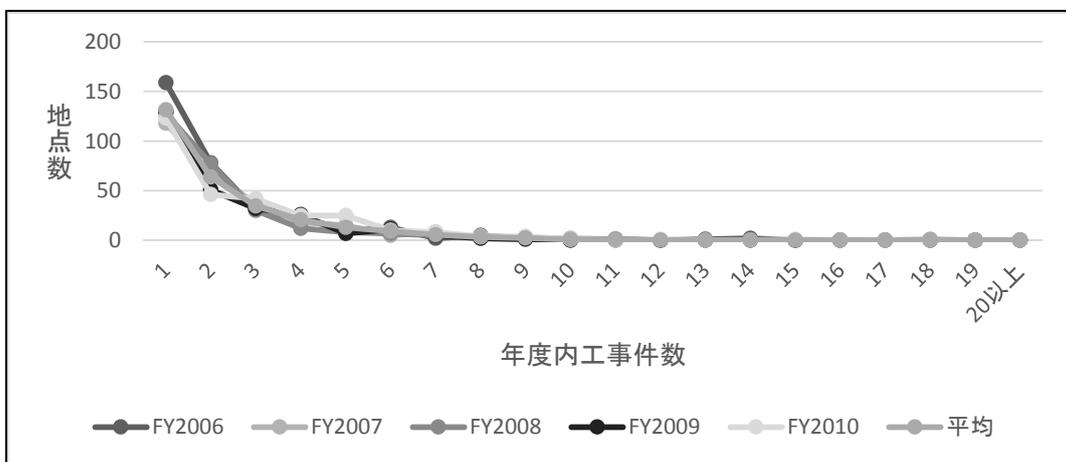


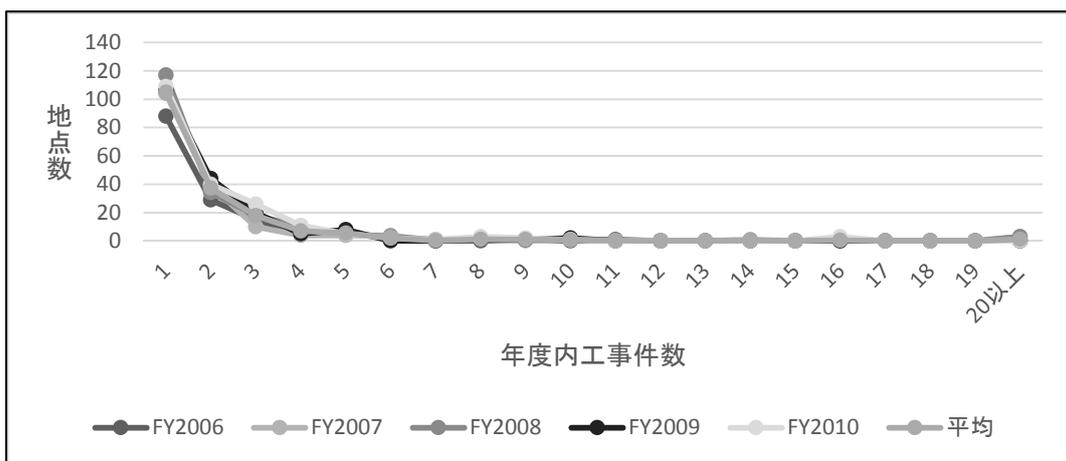
図4.3-2 非開通場所の年度当たり工事完了数（2006年度~2010年度の平均）



(a) 埼玉県



(b) 愛媛県



(c) 新潟市

図 4.3-3 非開通地点の年度内工事件数

### 4.3.3 工事件数からの予測と評価

各地点の12箇月間の工事件数によって開通を予測するという単純なルールを定め、これに基づいて3自治体の開通を予測した時の真陽性率と真陰性率を表4.3-3に示す。

表 4.3-3 年間の工事件数から開通を予測した場合の評価（3自治体の平均）

| 開通を予測するルール |                                   | 評価（契約額 500 万円以上） |      |
|------------|-----------------------------------|------------------|------|
|            |                                   | 真陽性率             | 真陰性率 |
| ルール 1      | 各地点での工事完了件数が 12 箇月以内で 2 件以上を開通とする | 0.92             | 0.57 |
| ルール 2      | 各地点での工事完了件数が 12 箇月以内で 3 件以上を開通とする | 0.77             | 0.78 |
| ルール 3      | 各地点での工事完了件数が 12 箇月以内で 4 件以上を開通とする | 0.73             | 0.88 |

表 4.3-3 のルール 1 は、12 箇月間に 2 件以上の工事があった場所を開通と予測するものである。表 4.3-3 の真陽性率によると開通場所の 92% が把握されている。そして、真陰性率によると、開通では無い地点を正しく予測したもの（この場合は開通では無い地点のうち工事件数 1 件であった地点ということになる）は 57% であったことを示している。1 - 真陰性率、すなわち、 $100\% - 57.2\%$  である  $42.8\%$  は偽陽性率であり、その地点は、開通では無いにも関わらず開通と予測されたことを意味する。ルール 3 は工事完了が 4 件以上の地点を開通と予測するものである。この時、実際の開通地点の  $73.1\%$  を捕捉し、開通でない場所を開通とした割合は  $12.4\%$  ( $1 - 0.876$ ) であった。

契約額が 2500 万円以上であった工事のみを使用した場合には表 4.3-4 となり、真陽性率は低下した。

表 4.3-4 年間の工事件数から開通を予測した場合の評価（3自治体の平均）

| 開通を予測するルール |                                   | 評価（契約額 2500 万円以上） |      |
|------------|-----------------------------------|-------------------|------|
|            |                                   | 真陽性率              | 真陰性率 |
| ルール 1      | 各地点での工事完了件数が 12 箇月以内で 2 件以上を開通とする | 0.69              | 0.65 |
| ルール 2      | 各地点での工事完了件数が 12 箇月以内で 3 件以上を開通とする | 0.54              | 0.84 |
| ルール 3      | 各地点での工事完了件数が 12 箇月以内で 4 件以上を開通とする | 0.45              | 0.91 |

これらを ROC 曲線として図 5.3-4 に示した。契約額 500 万円以上のときには AUC は 0.84 であったが、契約額 2500 万円以上のときには AUC は 0.73 であった。このように低い契約額のものも含めて工事情報を収集することで、より良く分類される。

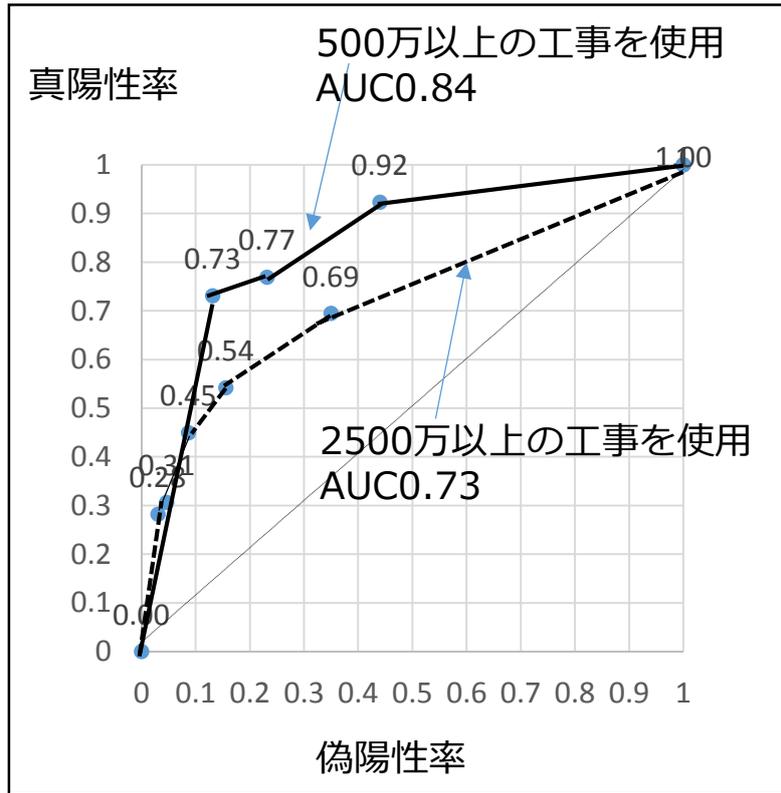


図 5.3-4 件数による道路開通予測の ROC 曲線

工事件数による道路開通の予測は、開通時期は年単位の予測となるが、次の利点がある。

1) 情報の入手のし易さ

予測に必要な項目は、工事の場所と時期だけであり、工事件名、請負会社、請負金額などが不要である。このため情報が入手し易いという利点がある。

2) 計算の容易さ

地点毎に工事件数を集積するだけで予測を行うことができる。

3) 分かり易さ

多数の工事の行われた地点では道路変化イベントの可能性が高いことは、人間の直観に近く、このため、結果が受け入れやすい。

## 参考文献

- 1) 加藤直樹・羽室行信・矢田勝俊著：データマイニングとその応用，朝倉書店，ISBN-13: 978-4254275520，2008， pp.82-90（4.3 モデルの判別性能の評価）
- 2) 元田浩・,山口高平・,津本周作・,沼尾正行著：データマイニングの基礎，オーム社，ISBN-13: 978-4274203480，2006， pp.199-208（5.2 統計学的検定）
- 3) Willi Richert・Luis Pedro Coelho 著，斎藤康毅訳：実践機械学習システム，オライリー，ISBN978-4-87311-698-3，2014， pp.105-109（5章クラス分類）， pp.180-184（9章クラス分類 III）
- 4) 中小企業庁：「中小企業者に関する国等の契約の方針（解説）平成 26 年度」
- 5) 国土交通省：公共工事コスト縮減対策に関する新行動指針（平成 12 年 9 月）
- 6) 大野泰資：公共工事における入札・契約方式の課題，会計検査研究，No.27(2003.3)， p p.159-174
- 7) 国土交通省：建設業の許可とは，[http://www.mlit.go.jp/totikensangyo/const/1\\_6\\_bt\\_000080.html](http://www.mlit.go.jp/totikensangyo/const/1_6_bt_000080.html)，2015.1.1. アクセス

## 第5章 機械学習を用いた工事内容による予測

### 5.1 機械学習を利用した予測の概要

本章では、工事内容を利用して道路変化イベントの予測を行う。工事件数の予測では供用開始の予測、早期の予測を行っていないが、工事内容を利用することでそれらの予測についても行える可能性がある。

工事内容には、建設の段階等に応じて多様性がある。工事内容を示す変数の数は多く次元の統計的な表現では表しにくい。このため、本研究では一般的な統計処理に比べて高次元のデータ空間から未知のルールを発見する用途（いわゆるデータマイニング）に適している「機械学習」(Machine Learning)を利用する。そして、(a)予測する対象が道路変化イベントと明確であること、(b)第2章で述べた方法によって道路変化イベントに関連づいた工事を予め知ることのできるデータセットが得られること、から「教師あり学習(Supervised Learning)」を利用する。

予め分類した（既知の）データセット、本研究においては道路変化イベントの有無の分かっている工事の内容のデータセットは教師データ（Training data）、教師データから生成される分類のルールは（学習）モデル（Learning model）、新たに分類を試みる（未知の）データはテストデータ（Test data）、分類された結果は目的変数（objective variable）あるいは従属変数（dependent variable）など、その道路変化イベント別のグループはクラス（Class）と呼ばれる。

工事内容を利用する道路変化イベントの予測は次のような流れで行う。

- ステップ1 予め道路変化イベントの有無が分かっている場所毎に一定期間内に完了した工事の内容を集めてデータセット（教師データ）を得る、
- ステップ2 これらのデータセットをもとに、工事の内容から道路変化イベントの有無を分類するルール（学習モデル）を生成する、
- ステップ3 このルールによって道路変化イベントの有無の分からない場所と時期について、その一定期間内の工事の内容（テストデータ）から道路変化イベントの有無を予測する、

この予測において、道路変化イベントとして供用（の有無）と主要開通（の有無）をとり、また、工事内容を収集する時期と道路変化イベントの時期の間隔を近いものと遠いものとすることで、道路変化イベントの大きさ別の予測と、工事完了から道路変化イベントまでの時間差別の予測を行える。この種類を表 5.1-1 に示す。予測の方法の概要を図 5.1-1 に示す。

表 5.1-1 クラスと時期による予測の種別

| 予測の対象とする<br>道路変化イベント<br>(クラス) | 道路変化イベントの予測の対象時期                       |   |
|-------------------------------|--|---|
|                               | 短期予測 (St)                              | 長期予測 (Lg)                                   |
| ①道路開通                         | イベント発生日の前後の工事を教師データとし、直近の道路変化イベントを予測する | イベント発生日の2年前の前後の工事を教師データとし、2年後の道路変化イベントを予測する |
| ②道路の供用の開始                     | 近い時期の道路の供用の開始                          | 2年後の道路の供用の開始                                |

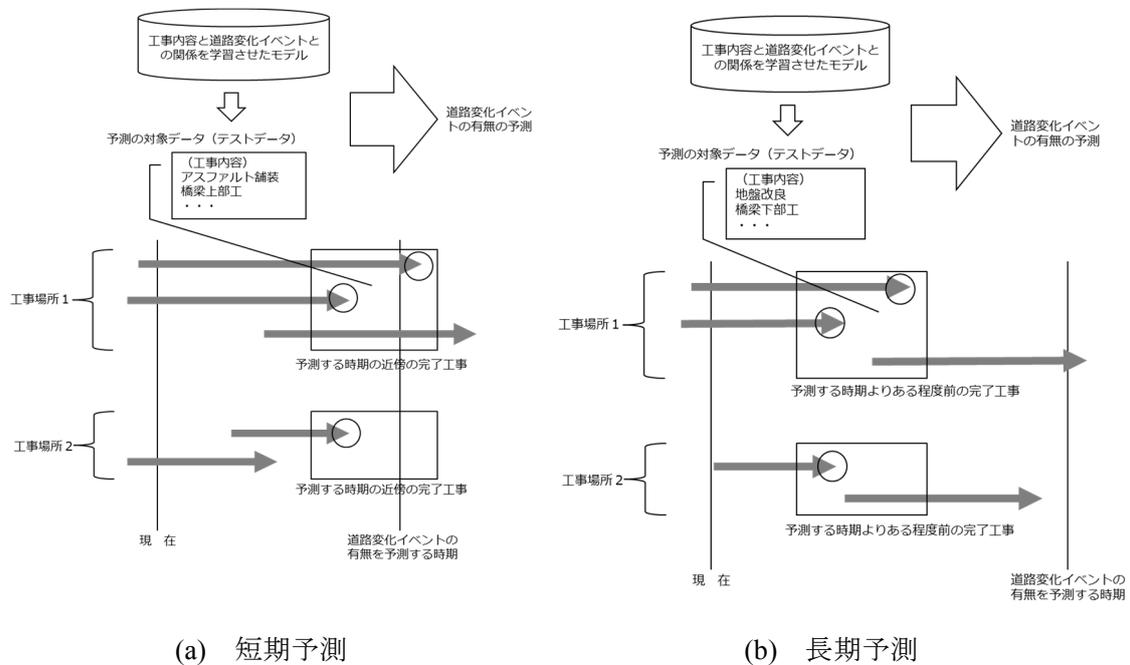


図 5.1-1 予測の方法の概要

## 5.2 予測に用いる変数

機械学習において予測するための変数は、説明変数 (explanatory variable) あるいは独立変数 (independent variable) または素性 (feature) などと呼ばれる。ここでは「説明変数」と呼ぶこととする。工事内容は、入札公告等では概要に示されている。工事实績情報においては「工種」、「工法・型式」がそれに相当する。これらは工事件名とは異なり固有表現や省略表現が排除され語句の揺らぎが少ないという利点がある。

| 1. 工事概要  |                         |  |                       |
|----------|-------------------------|--|-----------------------|
| (1) 工事名  | 吉野沢道路改良工事 (入札保証金納付対象案件) |  |                       |
| (2) 工事場所 | 宮城県本吉郡南三陸町歌津字吉野沢 地内     |  |                       |
| (3) 工事内容 | 掘削工(切土工)                |  | 50,000 m <sup>3</sup> |
|          | 盛土工                     |  | 50,000 m <sup>3</sup> |
|          | 函渠工 B4.7×H4.7           |  | 52.5 m                |
|          | 排水構造物工                  |  | 1 式                   |

図 5.2-1 工事概要の例

### 5.2.1 工事内容から説明変数への変換

#### (1) 次元圧縮

工事内容は、予め整理された土木用語辞典等との一致により項目を抽出する。得られた多様な用語のままに説明変数を設けると、説明変数の種類が多い(高次元の)データセットとなる。

高次元のデータを使うことにより教師データについては、その分類の精度を向上させられる可能性がある。しかし、説明変数の種類を増やすことによって、「次元の呪い(The curse of dimensionality)」と呼ばれる現象によって、未知のテストデータに対する精度が向上しないことが知られている。なぜなら、説明変数の増加にしたがって複雑になるモデルを生成するためには、それに見合っただ教師データの数も増加させなければならないが、教師データの数を増加させられるとは限らず、十分な教師データが無ければ適切なモデルを生成させられないためである。より詳しく説明すると、説明変数の間の関係を示す係数は、説明変数の数の積で増加する。しかも、次元数の増加とともに球面集中現象によって教師データは説明変数からなる次元空間の周辺部に多く位置することとなる<sup>1)</sup>。本研究においては細分化された用語に従って教師データを道路変化イベントの有無で分類すると、分類されたグループ(クラス)に占める一つの用語が関係するデータセットの数が少なくなる。このために多くの教師データが必要となる。具体的に、例えば、軟弱地盤処理工事には、置換工法、プレローディング工法、表層混合処理工法、バーチカルドレーン工法、サンドコンパクション工法、ロッドコンパクション工法などがあり、その中から現場の条件に適したものが採用されるが、これらを別々に扱くと得られる教師データは極めて疎(スパース)なものとなる。それに対して、道路建設において同じ目的を持つこれらを軟弱地盤処理工事としてまとめると、もとの細分化された工法より多くの出現回数となる。

また、説明変数を少なくすることで計算量を減らすことができる。例えば、用語の種類を137とし、そこから5種類の工事内容を選ぶ組み合わせは、 ${}_{137}C_5 = 387,610,812$ (約3.9億通り)となる。用語の種類を減じて17とした場合には ${}_{17}C_5 = 6,188,10$ とした場合には ${}_{10}C_5 = 252$ となる。

さらに、道路工事では通常は利用されることのない工事内容が特別な事情で含まれた教師データをもとに生成された学習モデルでは汎用的な予測の精度が低下してしまう。この

ため、無関係な工事内容を除去することも未知のテストデータに対する精度の向上には必要である。

このような理由から、抽出した用語のままではなく工事内容をまとめた説明変数を作成し予測に使用する。この操作は次元圧縮あるいは次元削減など(Dimension Reduction)と呼ばれる<sup>2)</sup>。また、意味のある部分集合だけを選択する手法を特に特徴選択ということもある。次元圧縮の例を付録2に示し、説明変数リストを表5.2-1に示す。次元圧縮の考え方は次の通りである<sup>3), 4), 5), 6), 7), 8), 9)</sup>。

- ① 道路変化イベントに関係が無いと考えらえる工事内容の削除
- ② 目的や段階が同じ工事内容の併合

表 5.2-1 説明変数

| No. | 説明変数                | 名義尺度の例            |
|-----|---------------------|-------------------|
| 1   | アスファルト舗装            | No                |
| 2   | 路盤路床                | No                |
| 3   | 総合土木工事              | Yes               |
| 4   | 隧道・コンクリート舗装         | No                |
| 5   | 橋梁上部                | Yes               |
| 6   | 橋梁下部                | No                |
| 7   | 基礎工事                | No                |
| 8   | 専門土木工事              | No                |
| 9   | 土工事                 | No                |
| 10  | 道路付属物               | Yes               |
| 11  | 歩道・管路               | No                |
| 12  | 斜面法面                | No                |
| 13  | 防水工事                | No                |
| 14  | 水路                  | No                |
| 15  | 撤去                  | No                |
| 16  | 維持保守                | No                |
| 17  | 緑化工事                | No                |
|     | CLASS (道路変化イベントの有無) | Positive/Negative |

## (2) 尺度水準

データの尺度水準には定性的なものとして①名義尺度、②順序尺度があり、定量的なものとして③間隔尺度、④比例尺度がある。1件の工事から得られる工事内容

を説明変数に変換したとき、同一の説明変数が複数得られる場合があり、また、同一場所の情報収集期間内の異なる工事から同一の説明変数が得られる場合もある。これらの回数を計数して得られるものが比例尺度となる。比例尺度の数値の大小を利用するには、工事の規模（契約金額）、工事の中での工事内容の割合が必要と考えられる。例えば、契約額の多い工事1件で多くの割合を占める説明変数（例えば、アスファルト工事）より、契約額が少ない多数の工事で小さな割合を占める同じ説明変数の方が尺度が高くなるといったケースを処理する必要がある。このため、本研究では、説明変数に重複がある場合には重複を取り除き、それぞれの説明変数の出現（有無）のみを取り扱う名義尺度を用いることとした。例を表 5.4-1 に合わせて示した。

### 5.3 教師データ

道路変化イベントの有無を識別するための教師データには、注目するクラス（道路変化イベント有り）に属する「正例」とそうではないクラス（道路変化イベント無し）に属する負例の二つが必要である。正例には、道路変化イベント有りを特徴づけるデータを選定することが必要であり、負例には逆に道路変化イベントの無いことを特徴づけるデータが必要である。

#### 5.3.1 正例の選定

正例、すなわち、道路変化イベント有りを特徴づける工事データは次のように選定した。

(1) 道路変化イベントの種類

道路変化イベントの種類は、供用開始と開通の2種類とした。ただし、供用開始と開通は別々に予測する。

(2) 場所

道路変化イベントが発生した場所で行われた工事とした。

(3) 時期

① 短期予測（道路変化イベントから数箇月以内）

道路変化イベントを含む時期の工事とした。数箇月は工事のおおよその工期である。

② 長期予測（道路変化イベントより2年+数箇月前）

道路変化イベントの2年+数箇月前の時期に行われた工事とした。

(4) 工事内容を収集する期間

「図 3.2-2 供用地点に対する工事の完了状況」, 「図 3.3-2 開通地点における工事の全開通数に対する割合」から、工事内容を収集する期間の長さに定まったルールが見出せなかった。このため、短期予測については、道路変化イベントの5箇月前から1箇月後の6箇月間、3箇月前から1箇月後の4箇月間、1箇月前から1箇月後の2箇月間の3種類をとった。長期予測では、道路変化イ

イベントの2年前に対して、その5箇月前から1箇月後の6箇月間、3箇月前から1箇月後の4箇月間、1箇月前から1箇月後の2箇月間の3種類をとった。なお、道路変化イベントと工事完了日は、年月日の日のレベルでデータが得られるが、これらが恒常的に日レベルで整合が図られているとは考えにくいため、月を単位として工事内容を収集する。

### 5.3.2 負例の選定

負例についても、道路変化イベントが無い場所の工事内容を選定するという違いがあるあることを除けば基本的な考えは正例の選定と同様である。しかし、正例においては道路変化イベントのあった時期（年月日）が確定しているのに対して、負例には依拠すべき時期が無い。このため、正例のための道路変化イベントを収集した場所と期間を利用して、道路変化イベントの無い可能性の高い場所と時間の工事内容を収集した。

道路変化イベントが全ての日に（全ての時間において）同じ確率で発生すると仮定して、道路変化イベントを調査した結果をもとに、その地点で道路変化イベントが発生していない確率が半分未満である地点と期間の工事内容を負例に採用する。道路変化イベントの事前情報収集期間を $\tau_B$ （日）とし、事後情報収集期間を $\tau_A$ （日）とする図 5.3-1 のとき、ある地点が道路変化イベントの有無を把握している機関に対して  $d1 \sim d4$  の時期のとき、道路変化イベントの有無を把握している日数は表 5.4-2 となる。これが情報収集期間全体の日数( $\tau_A + \tau_B + 1$ )の 0.5 未満となる  $d1 \sim d4$  の期間のデータを使用すれば良いこととなる。詳しくは付録 3 に示した。

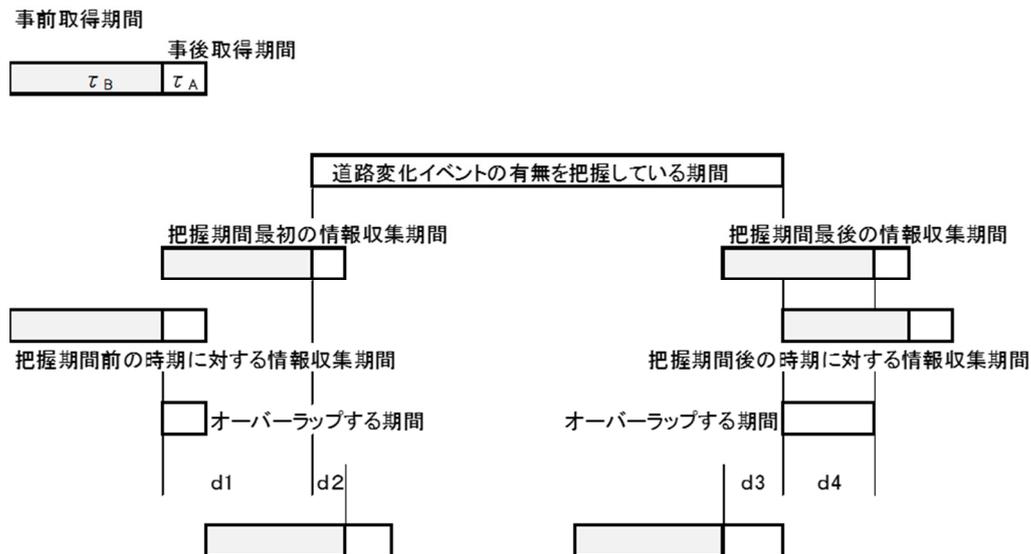


図 5.3-2 道路変化イベントの把握期間と収集期間

表 5.3-2 情報収集期間のうち道路変化イベントを把握している日数

| 場合  | 把握日数の算出式   |
|---|--|
| (1) 調査日が道路変化イベント把握期間より d 1 日前のとき          | $(\tau_A + \tau_B - d 1) \times (\tau_A + \tau_B - d 1 + 1) / 2$                                     |
| (2) 調査日が道路変化イベント把握期間の中 (期間開始の d 2 日後) のとき | $(\tau_A + \tau_B + d 2 + 2) * (\tau_A + \tau_B + 1 - d 2) / 2 + (\tau_A + \tau_B + 1) \times (d 2)$ |
| (3) 調査日が道路変化イベント把握期間の中 (期間完了の d 3 日前) のとき | $(\tau_A + \tau_B + 1) \times (d 3) + (\tau_A + \tau_B + d 3) * (\tau_A + \tau_B + 1 - d 3) / 2$     |
| (4) 調査日が道路変化イベント把握期間より d 4 日後のとき          | $(\tau_A + \tau_B - d 4 + 1) \times (\tau_A + \tau_B - d 4 + 2) / 2$                                 |

道路変化イベントの把握期間に対する時期と把握データ個数

5.3.3 正例と負例の均衡化

道路変化イベントに関係するデータは関係しないものに比べて少ない。このようなクラス間の数の不均衡によって生ずる問題 (class imbalanced problem) を解消するために、オーバサンプリングと呼ばれる手法、すなわち、ここでは数の少ない正例を繰り返し用いて数の多い負例と同数の正例を生成し、正例と負例の均衡化を図った<sup>10), 11), 12), 13), 14)</sup>。

5.4 使用するアルゴリズム

2006 年の IEEE International Conference on Data Mining で確認されたデータマイニングのトップ 10 アルゴリズムのうち教師あり学習のアルゴリズムは、①C4.5 とその発展形、②サポートベクターマシン、③AdaBoost、④ナイーブベイズ、⑤CART であった<sup>15)</sup>。①と⑤は決定木のアルゴリズムであり、③はアンサンブル学習と呼ばれる複数の分類器を用いるものである。線形分類器とは、説明変数の線形結合の値に基づいて分類を行う確率的分類器であり、これらのうち単独の分類器を線形分類器であるか否かによって分類すると、④ナイーブベイズ分類器は線形分類器であり、決定木である①、⑤は非線形分類器、②は線形、非線形のどちらのものもある。

分類器は、テストデータを分類した時に生ずる誤差 (汎化誤差) を小さくすることを目標とする。変数  $X$  に対する予測値が  $Y = f(X) + \varepsilon$ 、ここで  $\varepsilon$  は正規分布  $N(0, \sigma^2)$  に従う真のエラー項、という関係があるとき、学習モデル  $\hat{f}(X)$  で点  $x$  を推定したときの二乗誤差は式 (5.1) で表され、これは 3 つの項からなる式 (5.3) となる。式 (5.3) で、第一項は真の平均  $f(X)$  と予測値の差であるバイアスの二乗を示し、第二項は学習モデルを生成した教師データの違いに由来する分散を表すバリエーションであり、第三項は削減不能な誤差である。

$$Err(x) = E \left[ (Y - \hat{f}(x))^2 \right] \tag{5.1}$$

$$Err(x) = (E[\hat{f}(x)] - f(x))^2 + E[\hat{f}(x) - E[\hat{f}(x)]]^2 + \sigma_{\epsilon}^2 \quad (5.2)$$

$$Err(x) = Bias^2 + Variance + Irreducible\ Error(Noise) \quad (5.3)$$

バイアスとバリエンスはトレードオフの関係にあるため、バイアスとバリエンスの和が小さくなるアルゴリズムを選ぶ必要がある。バイアスはモデルの表現力に由来する誤差であり、バリエンスはデータセットの選び方に由来する誤差であることから、単純な線形分類器ではバイアスは大きくなり、バリエンスは小さくなる。他方、複雑な非線形分類器ではバイアスは小さくなるが、バリエンスは大きくなる<sup>16)</sup>。

本実験においては、それらの代表例として、①ナイーブベイズ分類器、②決定木及びナイーブベイズ分類器を改良した③AODE (Averaged One-Dependence Estimators)、決定木を弱分類器とするアンサンブル学習である④ランダムフォレストを実験に使用した。その理由は、これらが線形、非線形の代表的な分類器であることに加え、ナイーブベイズ分類器においては、各説明変数の有無に対する道路変化イベントの有無の確率が示され、決定木においては効果的な分類を可能とする説明変数が示されることによる。すなわち、どちらも分類結果と教師データの関係を明示できる。これらのアルゴリズムの種類と特徴を表 5.4-1 に示す。

表 5.4-1 使用する分類器の種類と特徴

|        | 分類器 (アルゴリズム) | <i>Bias</i> 誤差                   | <i>Variance</i> 誤差 |
|--------|--------------|----------------------------------|--------------------|
| 線形分類器  | ナイーブベイズ分類器   | 大                                | 小                  |
|        | AODE         | ナイーブベイズ分類器の <i>Bias</i> 誤差を下げたもの |                    |
| 非線形分類器 | 決定木          | 小                                | 大                  |
|        | ランダムフォレスト    | 決定木の <i>Variance</i> 誤差を下げたもの    |                    |

#### 5.4.1 ナイーブベイズ分類器 (Naïve Bayes classification)

##### (1) ナイーブベイズ分類器<sup>17),18)</sup>

説明変数が相互に独立しているとの仮定のもとに確率を用いて分類を行うものがナイーブベイズ分類器である。ナイーブベイズ分類器を用いて工事の内容から道路変化イベントの発生を識別する方法を説明する。

ある場所におけるある期間中の道路変化イベントの発生が、次の2種類、事象  $A_1$  は道路変化イベント有り、事象  $A_2$  は道路変化イベント無し、に分類されるとする。ちなみに、両方の事象は同時に発生することはない、かつ、両方を合わせると全ての事象である。この場所で、ある期間中の完了工事の内容、具体的には説明変数の組み合わせとなるが、これを  $B$  として、 $B$  から確率的に  $A_1$  または  $A_2$  が分類できるものとする。このとき、ベイズの定理を利用して、 $B$  から確率の高いものが  $A_1$  であるか  $A_2$  であるかを求め、この両者の中でより高い確率の事象  $A_i$  を予測結果とする。この推定方法は最大事後確率推定 (MAP 推定, Maximum a posteriori 推定) と呼ばれる。これがナイーブベイズ分類器を用いて工事の内容から道路変化イベントの発生を予測する仕組みの概略である。

計算方法を次に述べる。ベイズの定理 (Bayes theorem) を式(5.4)に示した。 $P(x)$  は  $x$  の確率、 $A$  は事象、 $B$  は条件である。本研究においては事象  $A$  が道路変化イベントの発生であり、条件  $B$  が工事内容となる。

$$P(A|B) = \frac{P(A) \cdot P(B|A)}{P(B)} \quad (5.4)$$

式(5.4)において、 $P(A)$  は事象  $A$  (道路変化イベントの発生有りまたは無し) の起こる確率 (事前確率, prior) を、 $P(B)$  は条件  $B$  (ある工事内容の組み合わせ) の生ずる確率を、そして、 $P(B|A)$  は、事象 (道路変化イベントの発生有りまたは無し)  $A$  のもとで条件 (工事内容)  $B$  が発生する条件付き確率 (尤度, likelihood) を表している。そして、これらにより、 $P(A|B)$ 、すなわち、条件 (工事内容)  $B$  のもとで事象 (道路変化イベントの有無)  $A$  の確率 (事後確率, posterior) が計算できることを示している。

道路変化イベントの事象は、有りか無しの2つであり、両方の事象は同時に発生することはない ( $A_1 \cap A_2 = \emptyset$  (空集合))、かつ、両方の事象を合わせると全てとなる ( $A_1 \cup A_2 = \Omega$  (全体事象))。この場合、ある事象  $A_i$  が条件  $B$  のもとで起こる条件付き確率は、確率  $P(B)$  が  $A_i$  のそれぞれの条件のもとの確率の和であることから、式(5.5)と表すことができる。

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2)} \quad , i = 1, 2 \quad (5.5)$$

説明変数  $B$  の各要素を  $b_j$  (総数は  $n$  個) とすると、 $P(B|A_i)$  を求めるためには  $b_j$  の全てのとり得る組み合わせについての計算が必要となる。これは極めて困難なものとなる。この問題を回避するため、全ての説明変数の組み合わせを計算するのではなく、異なる説明変数の間に独立性があるものと仮定することによって  $P(B|A_i)$  の算出を容易にしたものがナイーブ (=単純) ベイズ分類器である。独立性の仮定により式(5.6)が成り立つとすると、 $P(B|A_i)$  は式(5.7)で表すことができる。

$$P(b_j|A_i, b_k) = P(b_j|A_i) \quad , (j \neq k) \quad (5.6)$$

$$P(B|A_i) = \prod_{j=1}^n P(b_j|A_i) \quad (5.7)$$

式(5.5)に式(5.7)を代入すると式(5.8)が得られる. 式(5.8)において,  $P(A_i|B)$ の大きさを比較して最大事後確率を求める場合には, 全て計算に共通する右辺分母は無視できるため式(5.9)の関係を利用すればよいこととなる.

$$P(A_i|B) = \frac{P(A_i) \prod_{j=1}^n P(b_j|A_i)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2)}, \quad i = 1, 2 \quad (5.8)$$

$$P(A_i|B) \propto P(A_i) \prod_{j=1}^n P(b_j|A_i), \quad i = 1, 2 \quad (5.9)$$

## (2) スムージング<sup>19)</sup>

事象 $A_i$ の総数を $C_i$ ,  $C_i$ における説明変数 $b_j$ の存在数を $D_j$  とすると, その尤度は式(5.10)となる.

$$P(b_j|A_i) = \frac{D_j}{C_i} \quad (5.10)$$

このとき,  $D_j$ がゼロとなる説明変数が一つでもあった場合には, 他の説明変数の確率によらず式 (5.9) は常にゼロとなる. これはゼロ頻度問題 (sparseness problem) と呼ばれる. これを避けるための方法の一つがスムージング (smoothing, 平滑化) である. スムージングでは, 出現回数にバイアス $\alpha$ を与えることによって尤度がゼロとなることを防ぐ. 式(5.11)はスムージングを導入した尤度の計算式であり, 分母には説明変数 $b_j$ における条件付き確率の合計が1となるよう調整した値を加える.  $\alpha$ を1としたものはラプラススムージングと呼ばれる.

$$P(b_j|A_i) = \frac{D_j + \alpha}{C_i + (b_j \text{の値の種類}) \times \alpha} \quad (5.11)$$

## (3) ナイーブベイズ分類器による道路変化イベントの予測計算例 例として, 表 5.4-3 の場所工事データを考える.

表 5.4-3 計算のための教師データの例

| 教師データ | 説明変数 |       |       |          | クラス |
|-------|------|-------|-------|----------|-----|
|       | 基礎工事 | 盛土・切土 | 路盤・路床 | アスファルト舗装 |     |
| 1     | 有り   | 有り    | 有り    | 有り       | 開通  |
| 2     | 無し   | 有り    | 有り    | 有り       | 開通  |
| 3     | 無し   | 無し    | 有り    | 有り       | 開通  |
| 4     | 無し   | 有り    | 無し    | 有り       | 開通  |
| 5     | 有り   | 無し    | 無し    | 無し       | 非開通 |
| 6     | 有り   | 有り    | 無し    | 無し       | 非開通 |
| 7     | 無し   | 有り    | 有り    | 無し       | 非開通 |
| 8     | 無し   | 無し    | 無し    | 有り       | 非開通 |

開通、非開通の事前確率  $P(A_1)$  と  $P(A_2)$  は、それぞれ教師データ 8 例において 4 例ずつであることから、 $P(A_1) = 0.5$ 、 $P(A_2) = 0.5$  である。尤度は開通、非開通における説明変数の発生数からラプラススムージングを用いて表 5.4-4 のように算出される。

表 5.4-4 ナイーブベイズ分類のための計算サンプル

(条件)

| 条件付き確率(尤度)               |  | 説明変数 |       |       |          |
|--------------------------|--|------|-------|-------|----------|
|                          |  | 基礎工事 | 盛土・切土 | 路盤・路床 | アスファルト舗装 |
| 開通<br>の<br>確率<br>条件<br>下 | (説明変数を含む開通の数+1)<br>/(開通数+2)<br>$P(\text{説明変数有}   \text{開通})$       | 2/6  | 4/6   | 4/6   | 5/6      |
|                          | (説明変数を含まない開通数プ<br>ラス1)/(開通数+2)<br>$P(\text{説明変数無}   \text{開通})$    | 4/6  | 2/6   | 2/6   | 1/6      |
| 非開通<br>確率<br>条件<br>下の    | (説明変数を含む非開通の数+1)<br>/(非開通数+2)<br>$P(\text{説明変数有}   \text{非開通})$    | 3/6  | 3/6   | 2/6   | 2/6      |
|                          | (説明変数を含まない非開通の<br>数+1)/(非開通の数+2)<br>$P(\text{説明変数無}   \text{非開通})$ | 3/6  | 3/6   | 4/6   | 4/6      |

|             |       |       |        |       |
|-------------|-------|-------|--------|-------|
| 開通・非開通の事前確率 | 開通の確率 | 0.500 | 非開通の確率 | 0.500 |
|-------------|-------|-------|--------|-------|

(計算結果)

| テストデータ  | 説明変数 |       |       |          | 事後確率  |       |
|---------|------|-------|-------|----------|-------|-------|
|         | 基礎工事 | 盛土・切土 | 路盤・路床 | アスファルト舗装 | 開通    | 非開通   |
| テストデータ1 | 無し   | 無し    | 有り    | 無し       | 0.308 | 0.692 |
| テストデータ2 | 有り   | 無し    | 有り    | 無し       | 0.182 | 0.818 |
| テストデータ3 | 無し   | 無し    | 有り    | 有り       | 0.816 | 0.184 |

### 5.4.2 AODE (Averaged One-Dependence Estimators)

ナイーブベイズ分類器は全ての変数が独立であると仮定していた。これに対して、独立の仮定を緩和して、各変数の一つを順に選び、それとクラスの組み合わせを条件として、他は独立とするという考えを採用したものが One-Dependence Estimators, ODE である<sup>20)</sup>。ナイーブベイズ分類器では分類器を構成する各確率の条件となる変数の個数が 1 であるのに対して、ODE では 2 となる。ODE はナイーブベイズ分類器に比べてバイアスは小さくなるがバリエーションが大きくなる傾向がある。そこで AODE では全ての ODE の推定結果を平均することでバリエーションを削減している。データセット  $X$  を構成する説明変数の値を  $x_i$  ( $i$  は 1 から  $n$ ) とするとき、ベイズの定理から式(5.12)が成り立つ。式(5.12)を全ての説明変数について平均したものが式(5.13)である。ここで  $F(x_i)$  は、 $x_i$  が教師データにおいて出現したデータセット数であり、その最低値を  $m$  としている。式(5.13)を用いて式(5.14)によってクラス分類をするものが AODE である。

$$P(y, X) = P(y, x_i)P(X|y, x_i) \quad (5.12)$$

$$P(y, X) = \frac{\sum_{i: 1 \leq i \leq n \cap F(x_i) \geq m} P(y, x_i)P(X|y, x_i)}{|\{i: 1 \leq i \leq n \cap F(x_i) \geq m\}|} P(y, x_i)P(X|y, x_i) \quad (5.13)$$

$$\arg \max_y \left( \sum_{i: 1 \leq i \leq n \cap F(x_i) \geq m} \hat{P}(y, x_i) \prod_{j=1}^n \hat{P}(x_j|y, x_i) \right) \quad (5.14)$$

### 5.4.3 決定木

#### (1) 決定木

決定木は、説明変数に応じて枝分かれする分割ルールによって対象データを分類する、木構造の分類器である。最小の木構造を求めることは計算量的に困難 (NP 困難) であることが知られている。このため、分割する時点で最も効果的に分類できると判断した分割を行い、さらに分割の必要があるならば、分割されたデータを再帰的にさらに分割していくが、いったん分割したデータをもとに戻すことは無い貪欲アルゴリズム (Greedy Algorithm) によって木構造を完成させるというトップダウン方式の分割統治法 (Divide and Conquer) を採用している。決定木には、分類のルールや分類後のクラスに属するデータ数、道路変化イベントの分類では、分類に用いた工事内容とその結果を明示できるという特徴がある。

分類の流れは、最初に木構造の「根 (root)」となるノード (節点) に分類の対象とするテストデータを置き、分類が行き止まりとなるノード (葉, leaf) が同じクラスとなることを

目標とする。図 5.4-5 に決定木の例を示す。

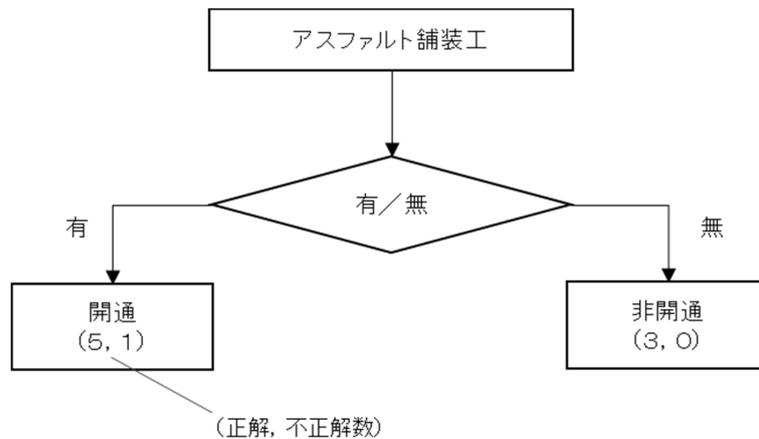


図 5.4-5 決定木の例

このように根から葉へ至る分岐においてクラスの多様性を減少させていく分割を効果的に行うことを目指して、分割に使用する説明変数を選択するための多様性を表す指標が幾つか提案されている。多様性の指標として、Gini 係数が CART アルゴリズム (Breiman, 1984) に、ID3 (Quinlan, 1986) には分割前後の平均情報量 (エントロピー) の利得が、そして、ID3 の改良型として発表された C4.5 (Quinlan, 1993) には情報量利得比が用いられている。本論文では、この中で最も新しい C4.5 アルゴリズムを使用した。以下に、C4.5 アルゴリズムのもととなっている ID3 と実験で使用した C4.5 アルゴリズムについて説明する<sup>21), 22)</sup>。

## (2) ID3 アルゴリズム

平均情報量 (エントロピー) は無秩序さや不確実性を表す指標であり、不確実な事象が確実になると減少する。ID3 は各説明変数で分割する前後の平均情報量の差 (利得) を算出し、最も高いもの、つまり、最も不確実さを減少させたものを分割に使用する。

確率的な事象の情報を符号化するために必要なビット数は、 $P(i)$  をその事象の確率とするとき式 (5.15) で表され、これは選択情報量と呼ばれる。平均情報量は、選択情報量に発生確率を乗じたものの総和として式 (5.16) で求められ、情報の不確実さを表す。ただし、生起確率がゼロである  $P(i)$  に対しては、 $H$  はゼロとみなす<sup>23)</sup>。

$$I(i) = -\log_2 P(i) \quad (5.15)$$

$$H = -\sum_{i \in \Omega} P(i) \log_2 P(i) \quad (5.16)$$

分割によって多様性が減少すると、分割された後の集合の不確かさは減少するため平均情報量はゼロに近づく。最も効果的な分割は、分割前の平均情報量から分割後の平均情報量を減じたもの（これを情報利得、information gain という）を最大とするものと考え、各説明変数の分割前後の平均情報量を式(5.16)で計算し、式(5.17)によって最も高い情報利得を示した説明変数によって分割を繰り返すアルゴリズムが ID3 である<sup>24)</sup>。式(5.17)において $H_0$ は分割前の平均情報量、 $H(i)$ は説明変数 $i$ で分割した後の平均情報量であり、 $IG(i)$ はその時の情報利得を表す。

$$IG(i) = H_0 - H(i) \quad (5.17)$$

### (3) C4.5 アルゴリズム<sup>25)</sup>

ID3 では、細かな分割によって分割後の不確かさを減らして高い情報利得を得ることが可能である。例えば、個別のデータに付与された識別コードによる分割がそれに当たり、分割後には一つのデータしかないため不確かさはなく、その平均情報量はゼロとなる。このような価値のない分割を防ぐため、分岐数の多い説明変数の選択を抑制できるよう分割情報量 (split information) を ID3 に導入したものが C4.5 である。分割情報量は式(5.18)で表され、データが分割される分岐の情報を示すために必要となる情報量であり、より複雑な分岐では分割情報量が多くなる。ここで、 $T$  は分岐する枝の集合、 $t$  はその枝に割り振られるデータの割合を示す。C4.5 アルゴリズムが評価に用いる情報利得比 (information gain ratio) は、情報利得を分割情報量で除したものとして式(5.19)で与えられる。

$$SI = - \sum_{t \in T} P(t) \log_2 P(t) \quad (5.18)$$

$$IGR \text{ (情報利得比)} = \frac{IG \text{ (情報利得)}}{SI \text{ (分割情報量)}} \quad (5.19)$$

### (4) 枝刈り

教師データに適合しすぎた状態を過学習 (overfitting) といい、過学習になると未知のデータに対して安定した性能が発揮されないこととなる。決定木では「枝刈り」(pruning) という手法によって決定木の大きさを抑制し、モデルが特定の教師データに過剰に適合することを防ぐ。C4.5 では未知のデータを含むデータ集合全体を母数 (parameter) と考えて、母数に対して、葉におけるエラー率を考える。エラー率は成功と失敗の二事象で構成されるため二項分布に従い、これに対して、信頼度 (confidence factor) を用いて母数におけるエラー率の区間推定を行う。C4.5 では、母数の推定エラー率の区間の中で最悪なケースを採用し、枝刈り前より枝刈り後のエラー個数が低下する場合にはその枝を刈る<sup>26)</sup>。信頼度は C4.5 では 0.25 が一般的に用いられる。

(5) 工事内容に適用した例

ナイーブベイズ分類器で用いた例，表 5.4-3 を再び使用して決定木の例を示す．  
根の平均情報量は式(5.20)で算出される．

$$-\frac{4}{8} \times \log_2 \left( \frac{4}{8} \right) - \frac{4}{8} \times \log_2 \left( \frac{4}{8} \right) = 1 \quad (5.20)$$

これに対して，4つの説明変数のそれぞれによって分割した場合の平均情報量，情報利得，情報利得比を表 5.4-4 に示す．この例ではアスファルト舗装によって分割する場合に最も情報利得比が高いため，アスファルト舗装によって根を分割し，他の説明変数に対して再帰的に計算を行い，決定木を構築する．これにより得られた決定木が本項の冒頭の図 5.4-5 である．

表 5.4-4 決定木のための情報利得比の計算例（最初の分岐の決定）

| 計算項目  |  | 説明変数   |        |        |              |
|---|--|--------|--------|--------|--------------|
|   |  | 基礎工事   | 盛土・切土  | 路盤・路床  | アスファルト舗装     |
| 事例数 $\geq$  | それぞれの説明変数を含む部分集合の開通の割合 $P_1$   | 1/3    | 3/5    | 3/4    | 4/5          |
|   | それぞれの説明変数を含む部分集合の非開通の割合 $P_2$  | 2/3    | 2/5    | 1/4    | 1/5          |
|   | それぞれの説明変数を含む部分集合の平均情報量 ( $H_a$ )<br>$-P_1 \times \log_2 P_1 - P_2 \times \log_2 P_2$   | 0.918  | 0.971  | 0.811  | 0.722        |
| 事例数 $\leq$  | それぞれの説明変数を含まない部分集合の開通の割合 $P_1$   | 3/5    | 1/3    | 1/4    | 0/3          |
|   | それぞれの説明変数を含まない部分集合の非開通の割合 $P_2$  | 2/5    | 2/3    | 3/4    | 3/3          |
|   | それぞれの説明変数を含まない部分集合の平均情報量 ( $H_b$ )<br>$-P_1 \times \log_2 P_1 - P_2 \times \log_2 P_2$ | 0.971  | 0.918  | 0.811  | 0.000        |
| それぞれの説明変数を使用した場合の平均情報量<br>$N/(N+M) \times H_a + M/(N+M) \times H_b$ |  | 0.951  | 0.951  | 0.811  | 0.451        |
| 根の平均情報量(1.000)に対する情報利得  |  | 0.049  | 0.049  | 0.189  | 0.549        |
| 分岐後のデータ数  |  | (3, 5) | (5, 3) | (4, 4) | (5, 3)       |
| 分割情報量   |  | 0.954  | 0.954  | 1.000  | 0.954        |
| 情報利得比   |  | 0.051  | 0.051  | 0.189  | <b>0.575</b> |

#### 5.4.4 ランダムフォレスト

複数の分類器（弱識別器）を総合して高い性能を得ようとする手法はアンサンブル学習あるいは集団学習と呼ばれる。ランダムフォレストは、複数の決定木を弱識別器とする森を構成し、その多数決により識別を行うアルゴリズムである。決定木は低バイアスを実現できる反面、高バリエーション（過学習）となるアルゴリズムであり、多数決によりバリエーションを下げ、未知のテストデータに対する予測精度を向上させることを意図している。

ランダムフォレストは、ランダムに抽出した教師データとランダムに選択した説明変数を用いて相関の低い決定木群を作成することであり、これによって過学習を防いでいる。具体的な手順は次の通りである<sup>27),28)</sup>。

- ① ランダムフォレストを構成する決定木（ここでは $n$ 本とする）は独立して学習を行う。個々の決定木が使用する教師データ（ $B_1, \dots, B_n$ ）は、重複を許してランダムに抽出する（ブートストラップ）。
- ② 個々の教師データの説明変数は、分岐の度にランダムに選択したものを使用する。
- ③ 生成した $n$ 個の決定木の多数決によって識別を行う。

ランダムフォレストの主要なパラメータは次の2つである。

- ① 作成する決定木の数  
作成する決定木の数を決定する方法は、予測に用いる木の数を増やしていき結果が安定する数を利用すればよいとされている。
- ② 使用する説明変数の数  
使用する説明変数は説明変数の数が $N$ の時 $\sqrt{N}$ が推奨値となっている。

### 5.5 機械学習を用いた工事内容による予測の実験

#### 5.5.1 実験の方針

実験では次の点を明らかとすることを目的として条件を制御する。

##### (1) 汎化誤差の大きさ

教師データに対する予測と、テストデータに対する予測の結果を比較する。十分な数のデータセットが得られない時は、教師データの一部をテストデータとする交差検証（cross validation）による評価がしばしば行われる。本実験においては正例をオーバサンプリングして教師データを生成しており、教師データの一部を抽出しても残りのデータとの重複を生ずる可能性が懸念されるため、テストデータと教師データは重複の無いものを準備した。

##### (2) 道路開通と供用開始の予測可能性

道路開通と供用開始の予測の可能性を実験によって評価する。そのため、教師データは道路開通の有無で正例と負例を構成したものと、供用開始の有無で正例と負例を構成したものを準備した。

(3) 長期予測の可能性

道路変化イベントの時期に近い工事データで予測する短期予測と、道路変化イベントより 2 年前の時期の工事データで予測する長期予測を実施し、長期予測の可能性を実験によって評価する。

(4) 情報収集期間の長さの評価

情報収集期間の長さを 2 箇月、4 箇月、6 箇月としてその長さによる違いを評価する。

(5) アルゴリズムとパラメータ

機械学習のアルゴリズムとパラメータを変えて、この用途に適したアルゴリズムとパラメータについて評価を行う。

(6) 教師データを生成するためのデータ量

教師データを生成するためのデータ量、特に数の少ない正例において良い予測を行うことについて評価を行う。

### 5.5.2 実験方法

説明変数、教師データ、テストデータ、アルゴリズムの基本的な事柄については前述したので、ここでは実験の詳細な設定のみを示す。

(1) 短期予測と長期予測、情報収集期間

短期予測は実験結果では次の表記としている：St

長期予測（2 年前）は実験結果では次の表記としている：Lg

情報収集期間は短期予測、長期予測のそれぞれで次の 3 種類である。

- ① 指定した時期の事後の 1 箇月と事前の 1 箇月の組み合わせ(実験結果表記:11)
- ② 指定した時期の事後の 1 箇月と事前の 3 箇月の組み合わせ(実験結果表記:31)
- ③ 指定した時期の事後の 1 箇月と事前の 5 箇月の組み合わせ(実験結果表記:51)

(2) 教師データ

データは、道路工事の内容を説明変数に尺度化したものである。

正例は道路変化イベントの地点と年月日に、時期の差を加えた情報収集期間に完了した工事内容である。

負例は、道路変化イベントの無い可能性が 50%以上の情報収集期間の工事内容である。

教師データは正例と負例をオーバサンプリングによって平衡化したものである。正例の数は負例（道路開通時期を含まない工事情報）に比べて少ないので、正例をオーバサンプリングして正例と負例の均衡化を図っている。このため、教師データの多様性は正例に依存することとなる。

### (3) テストデータとクラスの付与

実際の運用においてはいつどこで道路変化イベントが発生するかは不明である。これと同じ状況下で予測を行うことを想定して、一日ずつ予測のためのデータを全ての場所について生成する。すなわち、予測を行いたい日に時期の差（道路変化イベント当日と2年前）に対して情報収集期間（事前5, 3, 1箇月と事後1箇月）の間に完了した工事内容を説明変数に変換したものを生成する。予測日に供用のあった場所のデータについてはクラスを陽性とし、供用はあるが当日では無く情報収集期間中であるものはクラスを定めず、それ以外、つまり供用の無い場所及び供用の有った場所でも情報収集期間外はクラスを陰性とした。この様子を図 5.5-1 に示す。

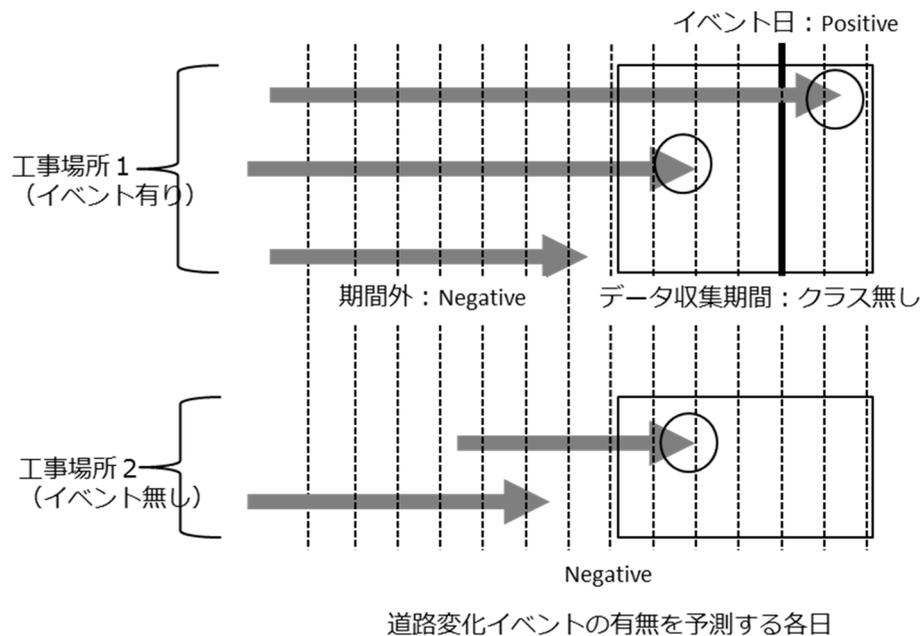


図 5.5-1 テストデータのクラスの付与

### (4) 予測に用いたアルゴリズムとプログラム

予測に用いたアルゴリズムは次の通りである。

- ① ナイーブベイズ分類器      `weka.classifiers.bayes.NaiveBayes`  
ラプラススムージング実施
- ② AODE      `weka.AODE.classifiers`  
ラプラススムージング実施  
説明変数の最低発生数 (m) = 1
- ③ 決定木(C4.5)      `weka.classifiers.trees.J48`
- ④ ランダムフォレスト      `weka.classifiers.trees.RandomForest`  
使用する変数 = 5, 使用する木 = 10

## (5) 実験結果の凡例

実験結果は次の表記で表現する.

### ① アルゴリズム

NB: ナイーブベイズ分類器

AODE: AODE

DT: 決定木

RF: ランダムフォレスト

### ② 短期・長期の別

短期: St

長期: Lg

### ③ 情報収集期間の長さ

11: 情報収集基準日の事前1箇月・事後1箇月,

31: // 事前3箇月・事後1箇月,

51: // 事前5箇月・事後1箇月

## 5.5.3 道路開通に対する予測実験

開通を公表している自治体として島根県, 高知県, 愛媛県, 新潟市を選び, それらの工事情報から道路開通の予測の実験を行った.

### 5.5.3.1 道路開通予測に用いたデータ

道路開通予測に用いたデータを表 5.5-1, 表 5.5-2 に示す. 本実験では説明変数の数を 17 としている. これに対して道路開通の件数は 1 自治体で年間数件程度である. このため, 自治体単独で教師データを構成した場合, 説明変数とクラスとの関係を示す十分な数が収集できないと考えた. そのため本実験では, 複数の自治体のデータを合算して教師データを生成することとした.

これら 4 つの自治体の 2011 年から 2013 年 (自治体により異なる) までのデータにより教師データを作成し, 収集した教師データとは重ならない年度の工事データに対して予測計算を行ったものが, 表のケース A である. これと比較するためにより多くの正例からなる教師データ (ケース B) を生成した. それは予測対象のテストデータを除く 7 種類のデータを合算したものである.

表 5.5-1 道路開通情報のデータ数

| 自治体(年度)   | 件数 | ケース A (正例少)               | ケース B (正例多)                               |
|-----------|----|---------------------------|---|
| 新潟市(2013) | 4  | まとめて一つの教師データとする           | テストデータ以外の7種をまとめたものを教師データとする(正例がより多く含まれる). |
| 島根県(2011) | 7  |                           |   |
| 愛媛県(2013) | 2  |                           |   |
| 高知県(2012) | 2  |                           |   |
| 新潟市(2014) | 3  | それぞれをテストデータとして結果を平均して表示する | テストデータに対する予測結果は平均化して表示する                  |
| 島根県(2013) | 1  |                           |   |
| 愛媛県(2014) | 2  |                           |   |
| 高知県(2014) | 3  |                           |   |

表 5.5-2 工事データ件数 (2010 年度~2015 年度 8 月) (契約額 2500 万円以上)

| 自治体 | データベース登録件数 | 地点数  | 変換件数 | 変換割合  |
|-----|------------|------|------|-------|
| 島根県 | 2186       | 2042 | 1927 | 88.2% |
| 高知県 | 1307       | 1196 | 1115 | 85.3% |
| 新潟市 | 582        | 555  | 525  | 90.2% |
| 金沢市 | 114        | 114  | 105  | 92.1% |
| 愛媛県 | 1116       | 997  | 917  | 82.2% |

### 5.5.3.2 道路開通予測の実験結果

工事情報から未知の道路開通を予測する実験の結果は次の通りである。

#### (1) 線形分類器による短期予測結果

正例数、情報収集期間を変化させて、線形分類器であるナイーブベイズ分類器、AODE によって道路開通の短期予測を行った結果を図 5.5-2 に示す。実験結果から次のことが得られた。

##### 1) 情報収集期間

図 5.5-2 は情報収集期間の長さによって大きく3つに分けており、11 (2 箇月)、31 (4 箇月)、51 (6 箇月) に大きな違いは見られなかった。

##### 2) 分類アルゴリズム

情報収集期間それぞれの中は大きく2つに分けており、左がナイーブベイズ分類器 (NB)、右が AODE を表している。ナイーブベイズ分類器と AODE は同じような傾向を示している。

##### 3) 正例数

さらにナイーブベイズ分類器、AODE の実験結果のそれぞれについて、正例の数が 10 前後のもの (ケース A) と 20 前後のもの (ケース B) を左右に示している。

これを見比べると正例の数を 10 前後から 20 前後に増やすことで真陽性率，真陰性率とも向上していた．このときの数値を表 5.5-3 に示す．正例数をおおよそ 20 にしたときの真陽性率は 0.58～0.77，真陰性率は 0.81～0.86 であった．

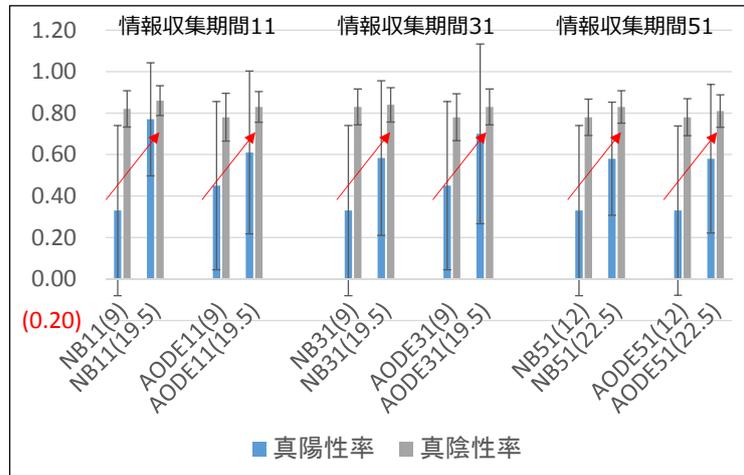


図 5.5-2 線形分類器による短期予測の真陽性率・真陰性率・標準偏差

アルゴリズム名と情報収集期間に続く括弧内の数値は教師データの正例数

表 5.5-3 正例数と予測の評価

| 教師データの正例数  | ケース A 少ない: 9～12               | ケース B 多い: 19.5～22.5                |
|------------|-------------------------------|------------------------------------|
| ナイーブベイズ分類器 | 真陽性率: 0.33<br>真陰性率: 0.78～0.82 | 真陽性率: 0.58～0.77<br>真陰性率: 0.83～0.86 |
| AODE       | 真陽性率: 0.33～0.45<br>真陰性率: 0.78 | 真陽性率: 0.58～0.70<br>真陰性率: 0.81～0.83 |

## (2) 非線形分類器による短期予測結果

非線形分類器では正例データの数，情報収集期間のほかに，教師データに適合させる程度を変えることができる．決定木分析，ランダムフォレストでは木を深くすればするほど教師データに適合した決定木となり，教師データに対しては良い評価が得られる．しかし過学習となると，未知のデータに対する分類性能が低下する可能性がある．図 5.5-3 は決定木について，木のサイズ（木のノードとリーフの合計数），正例の数，情報収集期間の長さの 3 つのパラメータを変化させた実験結果を 1 つの図に示したものである．図 5.5-4 は決定木を弱分類器とするランダムフォレストについて，木の深さ，正例の数，情報収集期間の長さの 3 つのパラメータを変化させた実験結果を 1 つの図に示したものである．これらの図から次のことが得られた．

### 1) 情報収集期間

図 5.5-3 は情報収集期間の長さによって大きく 3 つに分けており、11 (2 箇月)、31 (4 箇月)、51 (6 箇月) に大きな違いは見られなかった。

2) 分類アルゴリズム

決定木とランダムフォレストは同じような傾向を示している。

3) 正例数

正例数を 10 前後 (ケース A) から 20 前後 (ケース B) に増やしても必ずしも真陽性率、真陰性率は向上しなかった。

4) 木のサイズ・木の深さ

決定木を深さまたはサイズを大きくすると過学習により汎化性能は低下した。

St51 正例数 9 の教師データに対する数値を表 5.5-4 に示す。学習を制限した場合の真陽性率は 0.85~0.92、真陰性率は 0.45~0.60 であった。

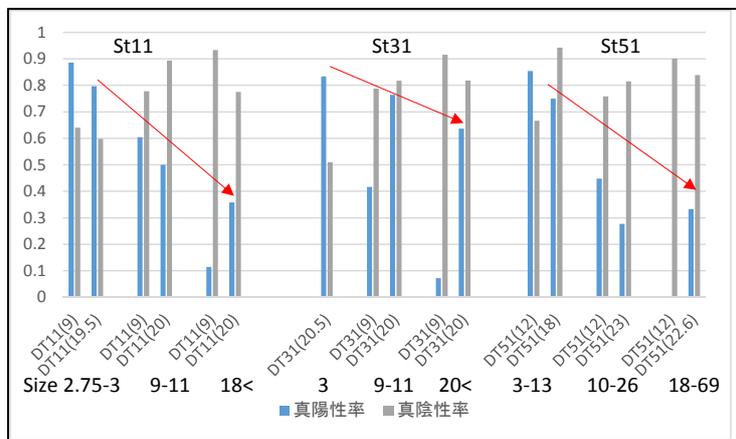


図 5.5-3 決定木による短期予測の真陽性率・真陰性率

アルゴリズム名と情報収集期間に続く括弧内の数値は教師データの正例数

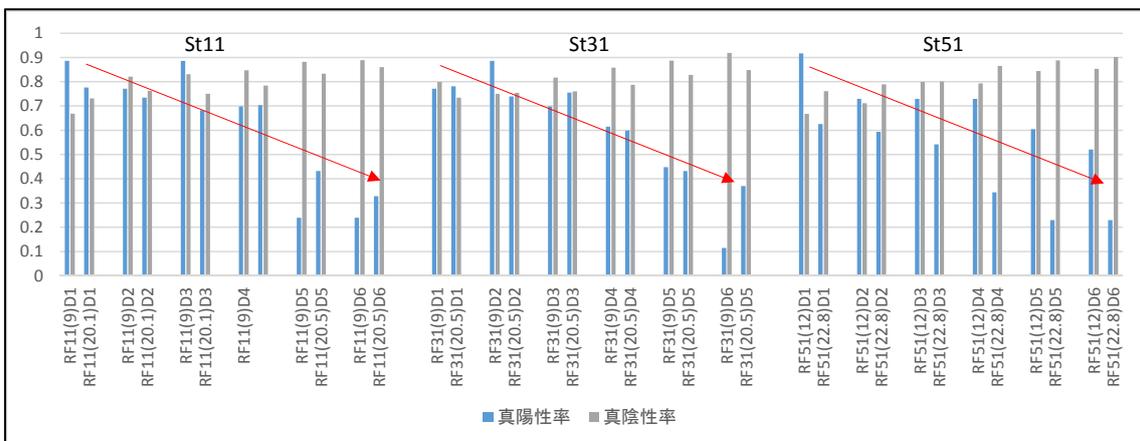


図 5.5-4 ランダムフォレストの短期予測の真陽性率・真陰性率

アルゴリズム名と情報収集期間に続く括弧内の数値は教師データの正例数, Dn : n は決定木の深さ

表 5.5-4 木のサイズ又は深さと予測の評価 (St51 正例数 9 のとき)

| 学習の抑制状況   | 制限                                     | 制限を緩和                                  |
|-----------|--|--|
| 決定木       | (木のサイズ平均 13)<br>真陽性率:0.85<br>真陰性率:0.67 | (木のサイズ平均 26)<br>真陽性率:0.45<br>真陰性率:0.76 |
| ランダムフォレスト | (木の深さ 1)<br>真陽性率:0.92<br>真陰性率:0.67     | (木の深さ 5)<br>真陽性率:0.60<br>真陰性率:0.84     |

St51 正例数 9 で生成された決定木の例を図 5.5-5 に示す。図 5.5-5 は、道路開通の短期予測において、第一段階で基礎工事が含まれているものを開通では無いと判断している。基礎工事が無いとき、第二段階で水路・管路があり、第三段階で道路付属物が含まれている時を開通、道路付属物が無いときは第四段階でアスファルト舗装が無ければ開通と判断している。第二段階で水路・管路が無いとき、第三段階で土工事が含まれているものを開通では無い、土工事が無く、アスファルト舗装工事が含まれているものを開通と判断していることを示している。これは決定木の一つの例である。

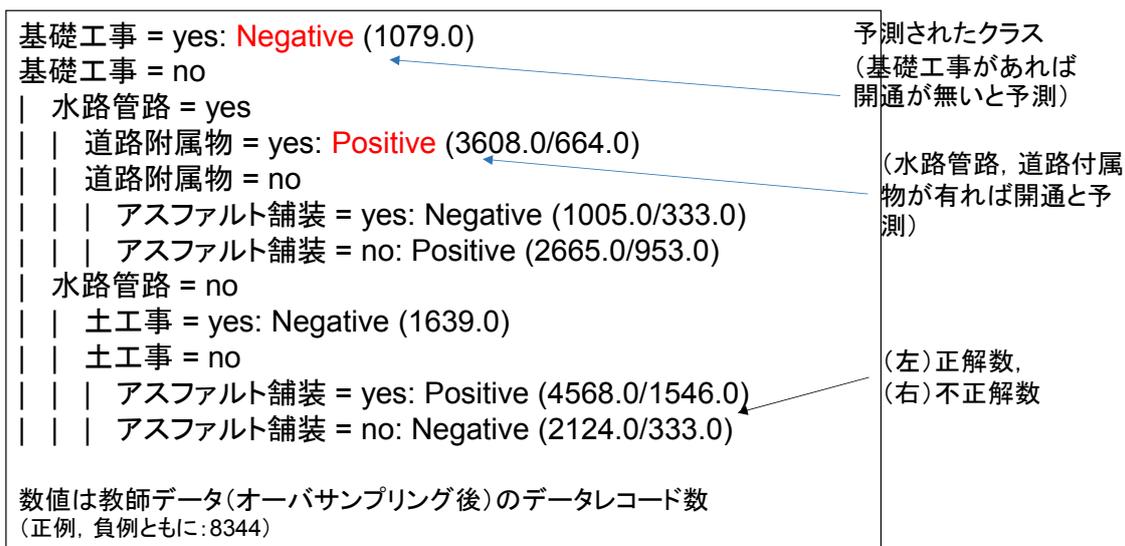


図 5.5-5 St51 で生成された決定木の例 (木の大きさ 13, 葉の数 7)

### (3) 長期予測結果を加えた全体の実験結果

長期予測は開通より 2 年前の工事を予測に利用する。情報収集期間を 2 箇月 (Lg11), 4 箇月 (Lg31) としたものでは、十分なデータが得られず、長期予測については Lg51 (6 箇月) のものだけとなった。これまで示した短期予測の実験結果に Lg51 を加えた、道路開通の予測全体の実験結果を図 5.5-7 に示す。長期予測では短期予測に比べて精度は低下している。長期予測では開通からの時期が離れた工事によるため、短期予測に比べて正例の特徴が分散化しているためと考えられる。

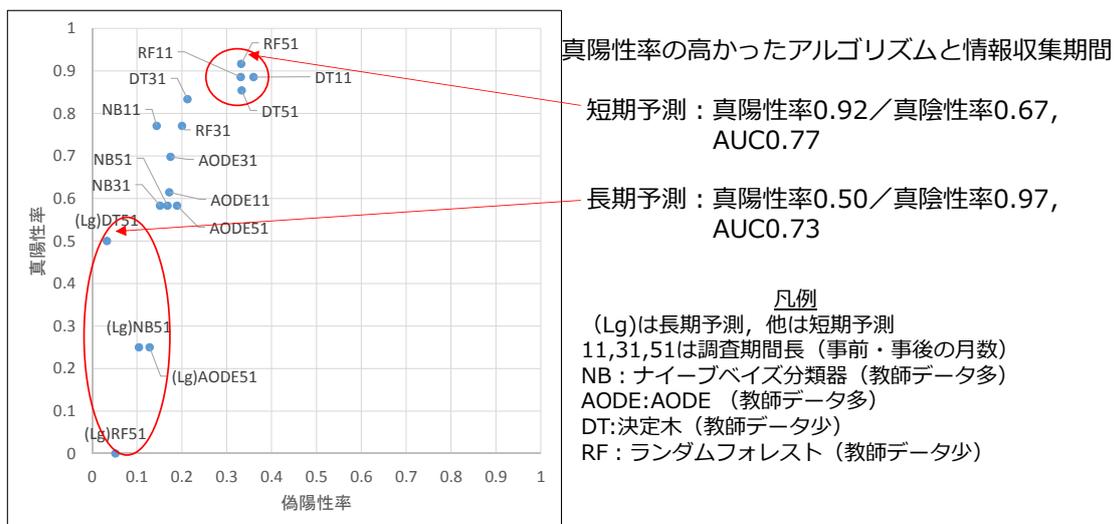


図 5.5-7 道路開通に対する予測結果

#### (4) 道路開通予測のまとめ

これまでの実験結果をまとめたものは次の通りである。

- ✓ 短期予測の精度は長期予測を上回る。
- ✓ 情報収集期間により結果に差があるものの、必ず優位になるという期間はなかった。
- ✓ 線形分類器では一定数以上の教師データ (この場合は正例) が必要である。
- ✓ 非線形分類器では過学習を避ける必要がある。
- ✓ 短期予測ではランダムフォレスト (情報収集期間 51, 木の深さ 1) において真陽性率 0.92, 真陰性率 0.67, AUC0.77 が得られた。
- ✓ 長期予測では決定木 (情報収集期間 51) において真陽性率 0.50, 真陰性率 0.97, AUC0.73 が得られた。

### 5.5.4 供用開始に対する予測実験

道路開通と同様に、供用開始について情報収集期間・教師データ・予測アルゴリズム等を変化させて短期・長期の予測を行った。

#### 5.5.4.1 供用開始予測に用いたデータ

供用開始の件数は今回採用した自治体においては年間数十回以上であった。使用した供用開始データは表 5.5-6 の通りである。同一自治体で異なる年度のデータを教師データ、テストデータとして評価検証するために、2013 年度、2014 年度のデータを準備した。ただし、既に述べたようにテストデータと教師データは混在させていない。

表 5.5-6 供用開始データ数

| 年度   | 道路管理者 | 道路の供用の開始 |      |     | ケースA   | ケースB                            |
|------|-------|----------|------|-----|--|---------------------------------|
|      |       | 登録       | 変換件数 | 地点  | 正例少  | 正例多                             |
| 2013 | 島根県   | 145      | 143  | 139 | 同じ自治体<br>で年度の異<br>なるものを<br>教師データ<br>とテストデ<br>ータとする | まとめて教<br>師データと<br>する            |
|      | 高知県   | 71       | 70   | 75  |  |                                 |
|      | 新潟市   | 158      | 155  | 174 |  |                                 |
|      | 金沢市   | 315      | 314  | 266 |  |                                 |
| 2014 | 島根県   | 157      | 153  | 152 |  | 夫々をテス<br>トデータと<br>して結果を<br>平均する |
|      | 高知県   | 68       | 68   | 72  |  |                                 |
|      | 新潟市   | 249      | 246  | 200 |  |                                 |
|      | 金沢市   | 178      | 172  | 131 |  |                                 |

使用した工事データは、表 5.5-2 工事データ件数（2010 年度～2015 年度 8 月）（契約額 2500 万円以上）と同じである。

#### 5.4.4.2 供用開始予測の実験結果

道路開通の予測と同じく Lg は長期予測（2 年前），St は短期予測，11 は事前 1 箇月事後 1 箇月，31 は事前 3 箇月事後 1 箇月，51 は事前 5 箇月事後 1 箇月を表す。

##### (1) 線形分類器による予測結果

図 5.5-8 は、教師データをそれぞれの自治体毎に生成して、教師データと同じ自治体であるが年度の異なる供用開始を予測したもの（ケース A）と、教師データを複数の自治体をまとめて生成し、教師データとは異なる年度の各自治体の供用開始を予測したもの（ケース B）を対比したものである。図は真陽性率、偽陽性率とそれぞれの標準偏差を示している。

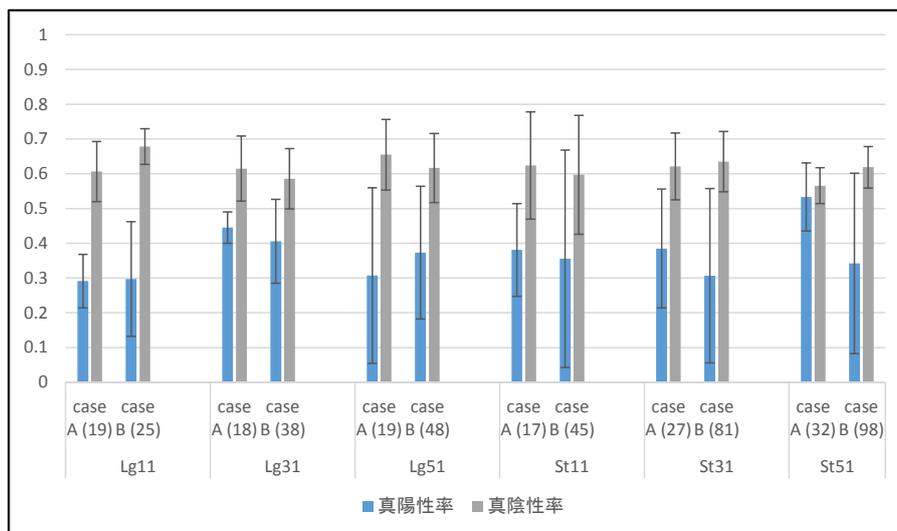


図 5.5-8 ナイーブベイズ分類器による教師データの異なる供用開始予測

1) 教師データ（正例数と構成）について

ケース A（同じ情報収集期間の右側）とケース B（左側）では、ケース B の正例がより多い。道路開通の予測では正例の数の多いものの評価が勝っていたが、供用開始の予測では、すべての情報収集期間においてケース B が勝っているわけではない。その原因は、供用開始では、(a)ケース A において説明変数の数と同程度の正例は確保できていること、(b)ケース B では国道、県道、市道が混在したデータを教師データとしていることが考えられる。原因(b)から、予測精度の向上のためには、教師データの生成において、規模や自然条件などを制御した正例・負例の収集が必要である可能性がある。

次に、ケース A（同一自治体で異なる年度の教師データ・テストデータの組み合わせ）について、ナイーブベイズ分類器と AODE による予測の結果を真陽性率、真陰性率と AUC（ROC 曲線下の面積）によって評価したものを図 5.5-9 と図 5.5-10 に示した。図は長期・短期と情報収集期間で分類している。

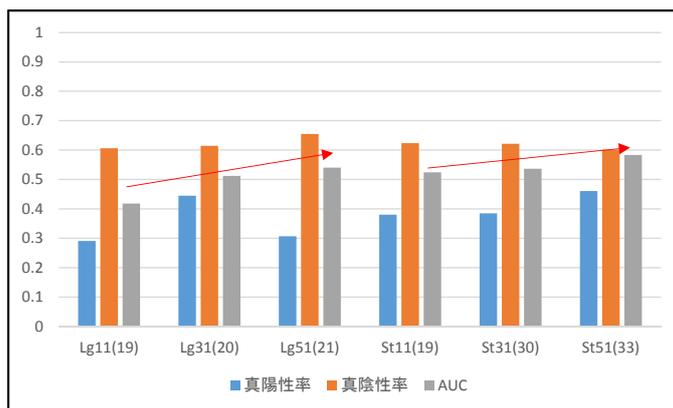


図 5.5-9 ナイーブベイズ分類器による供用開始予測（ケース A）

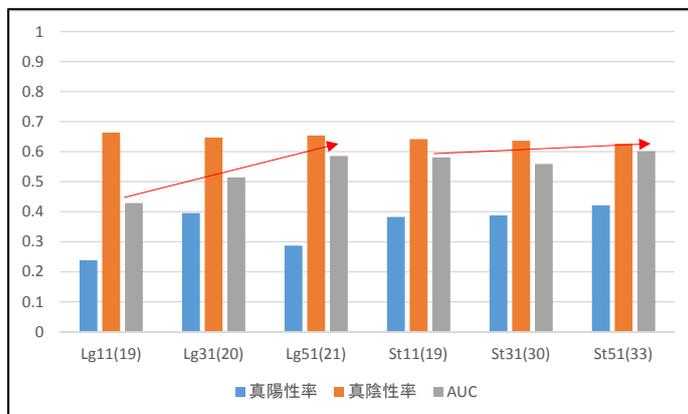


図 5.5-10 AODE による供用開始予測（ケース A）

この結果からさらに次のことが得られた。

## 2) 長期予測と短期予測

図 5.5-9, 図 5.5-10 は長期予測と短期予測によって左右に大きく 2 つに分けている。真陰性率の変化は少ないが, 真陽性率では概ね短期予測が長期予測より良い評価を示した。

## 3) 情報収集期間

情報収集期間の 3 種類の AUC は右上がり (図中の赤線) を示している。長期予測の Lg31 を除くと, 真陽性率も情報収集期間の長いものが良い結果を示した。

## 4) アルゴリズム

ナイーブベイズ分類器による予測と AODE による予測には大きな違いは見られなかった。

## (2) 非線形分類器による予測結果

長期予測・短期予測, 情報収集期間, 決定木分析では木のサイズ, ランダムフォレストでは木の深さを変化させて, 供用開始の予測を行った結果を図 5.5-10, 図 5.5-12 に示す (教師データはケース A)。図中の横軸はそれぞれ木のサイズ, 木の深さである。

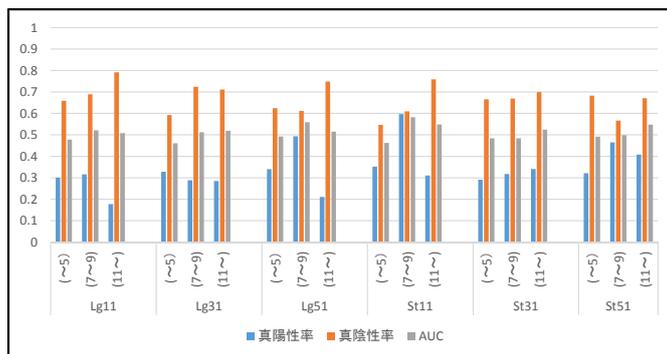


図 5.5-10 決定木分析による木のサイズと供用開始予測 (ケース A)

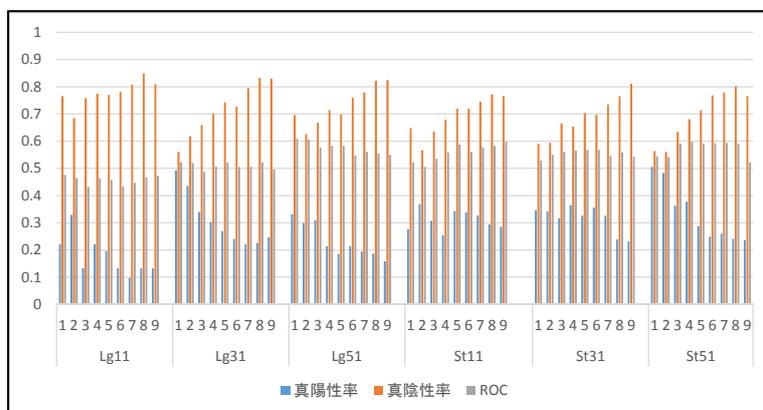


図 5.5-11 ランダムフォレストによる木の深さと供用開始予測 (ケース A)

実験結果から次のことが得られた。

1) 長期予測と短期予測, 情報収集期間

決定木分析 (図 5.5-10) では, 長期予測, 短期予測と情報収集期間により大きな差は見られなかった。ランダムフォレスト (図 5.5-11) では, 短期予測の AUC は全て 0.5 を上回っているが, 長期予測の AUC は 0.5 を下回るものがある。

2) 木のサイズまたは木の深さ, アルゴリズム

道路開通の予測では, 木のサイズまたは木の深さを大きくすると過学習によって評価は低下した。決定木分析による供用開始の予測 (図 5.5-10) では, 木のサイズを 7~9 にしたときに良い評価が得られたものが多い。道路開通では工事件数と規模が供用開始に比べて多いまたは大きいため, 細かな識別をしなくても良い評価が得られたが, 供用開始では道路開通に比べてより細やかな分類が求められている。ランダムフォレストでは, 木の深さを大きくするにつれて真陽性率は低下し真陰性率は上昇している。真陽性率と真陰性率のトレードオフの結果, AUC にはあまり変化が無かった。ランダムフォレストによる既知データと未知データの決定木の深さによって変化する真陽性率を図 5.5-12 に示す。決定木の深さを高めて教師データに適合させるにつれて未知のデータへの真陽性率が低下している様子が見られる。

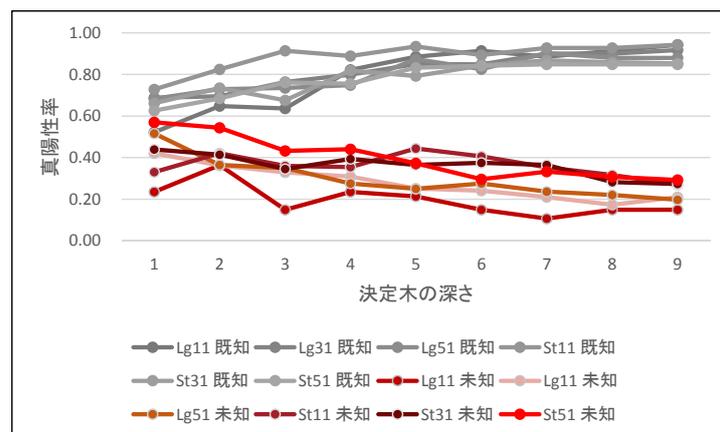


図 5.5-12 ランダムフォレストによる既知データと未知データの真陽性率 (ケース A)

St11 で生成された決定木の例を図 5.5-13 に示す。この決定木は, 斜面法面の工事が有った時には, アスファルト舗装が有った場合, 土工事が無い場合に供用開始有りとしている。斜面法面の工事が無くてもアスファルト舗装がある場合も供用開始有りとしている。斜面法面と土工事があってアスファルト舗装が無い場合, 斜面法面とアスファルト舗装のいずれも無い場合には供用開始は無いと予測している。この決定木では真陽性率 89%であり, 89%の供用開始を予測したが真陰性率は 50%であり, 供用開始ではない地点の半数を供用開始と予測した。

```

斜面法面 = yes
| アスファルト舗装 = yes: : Positive (1704.0/594.0)
| アスファルト舗装 = no:
| | 土工事 = yes: Negative(1730.0/740.0)
| | 土工事 = no : Positive (1682.0/572.0)
斜面法面 = no
| アスファルト舗装= yes: Positive (1466.0/726.0)
| アスファルト舗装 = no: Negative (1558.0/370.0)

```

数値は教師データ(オーバサンプリング後)のデータレコード数  
(左)正解数, (右)不正解数

図 5.5-13 St11 で生成された決定木の例 (木の大きさ 9, 葉の数 5)

(3) 供用開始予測の総合的な結果

供用開始の長期・短期の各情報収集期間の各アルゴリズムによる予測結果を平均したものを図 5.5-14 に示す。道路開通に比べて評価が低いものとなった。

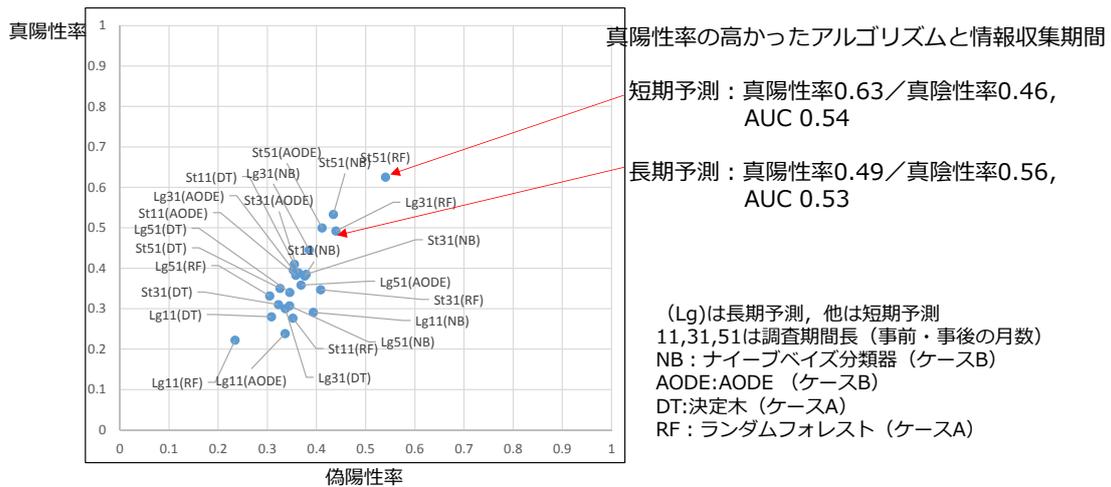


図 5.5-14 供用開始の予測結果

(4) 供用開始予測のまとめ

これまでの実験結果をまとめたものは次の通りである。

- ✓ 短期予測の精度は長期予測を上回ることが多い。
- ✓ 情報収集期間により結果に差があるものの、必ず優位になるという期間はなかった。
- ✓ 線形分類器では真陰性率はあまり変化が無い。
- ✓ 非線形分類器のうちランダムフォレストでは過学習により真陽性率が低下する。決定木分析では木のサイズ 5 から 7 程度のときに真陽性率が良かった。

- ✓ 短期予測ではランダムフォレスト（情報収集期間 51，木の深さ 1）において真陽性率 0.63，真陰性率 0.46，AUC0.54 が得られた。
- ✓ 長期予測ではランダムフォレスト（情報収集期間 31，木の深さ 1）において真陽性率 0.49，真陰性率 0.56，AUC0.53 が得られた。

## 5.6 予測の正確さの向上

機械学習による道路変化イベントの予測において，真陰性率を低下させずに真陽性率を向上させて，予測の正確さを向上させることが課題である。真陽性率が高いことは把握できる道路変化イベントの割合が多いことであり，真陰性率が高いことは道路変化イベントにつながらない予測の割合が少ないことを意味する。本節では，予測の正確さを向上させる方策について述べる。

### 5.6.1 データ：量と質の向上

#### （1）データの量

一般的に機械学習においてはデータ量が多いほど予測の正確さは向上する。また，機械学習以外の手法である「4.3.3 工事件数からの予測と評価」においても 500 万円以上の工事データを使用することにより，2500 万円以上の工事データによる予測に比べて AUC は 0.73 から 0.84 へと上昇した。予測の正確さの向上のためには道路変化イベントと工事情報を数多く収集することが必要となる。今のところ，データの収集については，インターネットに公開されている情報を自らスクレイピングする方法に頼っている。これらの情報を有料で提供する業者もあるが，将来的には，道路管理者から直接的に流通することが社会全体の生産性の向上に有益と考えている。そのための検討については第 6 章に示した。

#### （2）データの質

道路変化イベントに関する情報は道路管理者から提供されるものを使用している。これらには路線と正確な場所が示されている。これに対して，例えば，工事实績情報の工事場所は工事業者が登録している。工事实績情報の使用目的から，工事場所がどの発注者の管内であるかは実績の精査に必要な情報であるため注意が払われるが，町丁目レベルで記載すること重要ではないと考えることができる。このため，工事实績情報には正確な場所が登録されていないことがある（「4.3.2 開通の有無と工事件数に関する調査 ④ 工事場所が正しく登録されていないケース」）。また，点在する工事場所がある場合には道路管理者が場所を示す場合であっても管内全域などと登録されることがある。このように路線と正確な場所が示されていない場合には，情報を路線と地点へ集約することができない。場所から地点への変換ができた割合は「表 2.3-3 工事实績情報の工事場所の登録状況と地名辞典との適合性」，「表 5.5-2 工事データ件数（2010 年度～2015 年度 8 月）（契約額 2500 万円以上）」に示した通りである。工事場所を正確に地点に変換するためには，工事情報の一部からでも場所を推定する技術，実世界の地理識別子の揺らぎ（字体，町名変更，通称など）に対応する

地名辞典，路線と地名の対応辞典，路線に対応したジオコード技術の開発が考えられる。

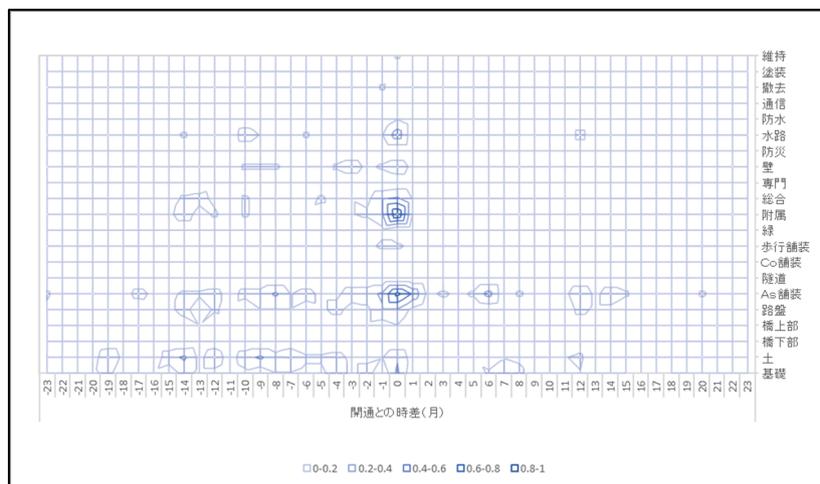
### 5.6.2 説明変数の改良

本研究では工事内容を次元圧縮した名義尺度の説明変数を用いている。その理由は「5.2.1 工事内容から説明変数への変換（2）尺度水準」に示した通りである。工事内容から適切な比例尺度に変換した説明変数を取り出すことができれば，数量的を加味した予測を行うことができる。これにより，工区が分割されて複数の工事が行われている地点と一過性の維持工事などを分離できる可能性が考えられる。

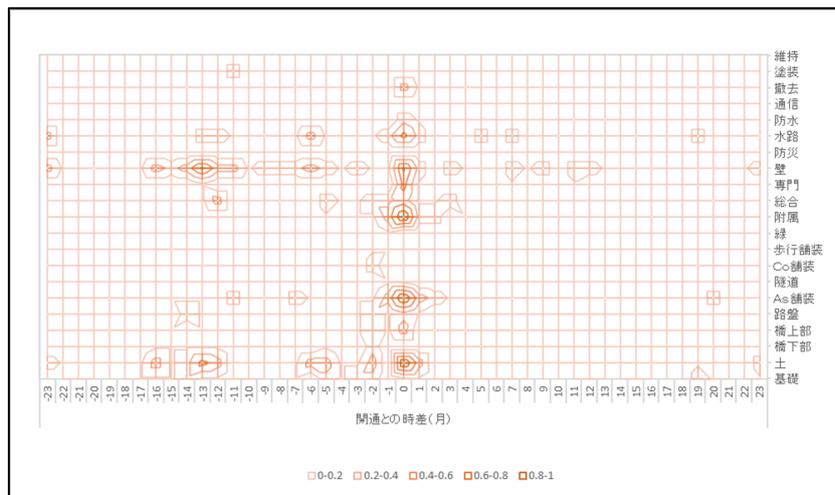
また，工事内容の他，工事契約額を地点数で除した地点当たりの工事額を用いて，建設工事の規模や維持工事との識別を行うこと，工事件名から維持，改築などの単語（形態素）を取り出して併用する手法が考えられる。工事件名は先行研究において道路工事の分類に利用されている<sup>29), 30)</sup>。

### 5.6.3 スパースデータに適したデータマイニングアルゴリズムの採用

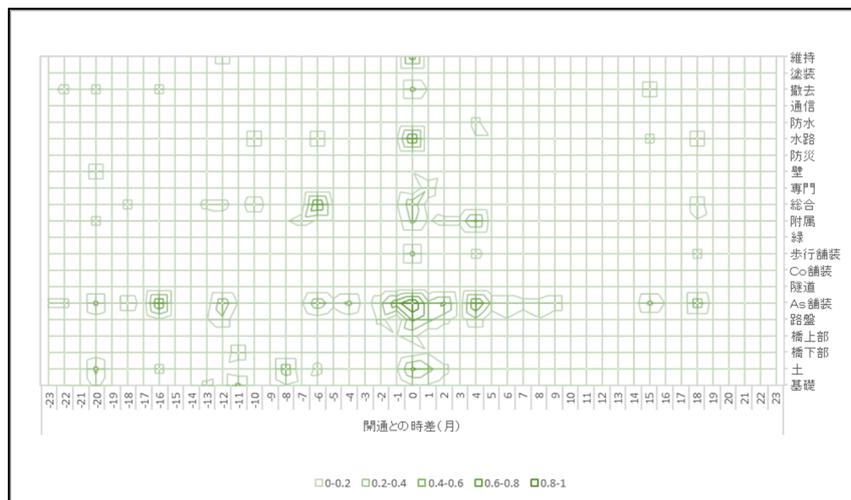
一つの地点で情報収集期間に観測される工事は数件以下であることが多い（「図 3.3-3 供用・開通・全ての地点の地点当たりの調査年度の工事数」）。加えて，1 件の工事情報から得られる説明変数も数個以下である（例「図 5.2-1 工事概要の例」）。この結果，機械学習に投入されるデータは，教師データもテストデータもスパース（sparse，疎）なものとなる。開通時期の前後に観測された説明変数の例を参考に図 5.6-1 に示す。自治体によって工事数・地点数・説明変数の発生数が異なるため数値は各説明変数の最大数で正規化している。



(a) 埼玉県



(b) 愛媛県



(c) 新潟市

図 5.6-1 説明変数の出現状況

(使用データは「4.3.1 調査データ」による)

データが大規模であってそのうち意味のある情報が一部であるものは、大量にパラメータを準備して重要でないものをゼロと推定する「スパース推定」が適用される問題である。スパース推定は例えば正則化 (regularization) によってモデルの複雑さに罰則を科すことで高速計算を実現する。正則化の手法としては L1 正則化の Lasso などが代表的である。

本研究での問題は、これとは異なり、データが疎であるために必要なデータ量が十分に獲得できないことである。これはデータスパースネス問題 (data sparseness problem) と呼ばれ、自然言語処理やレコメンデーションの分野で取り組まれている<sup>31), 32)</sup>。例えば、自然言語処理では、文書分類やトピックモデルの生成において、数十万を超える語 (例えば、広辞苑の第六版で約 24 万語) に対して、通常の文書はごく一部の語のみを使用していること

から、語ごとに発生回数を計数すると、そのほとんどはゼロである大きな行列が得られることとなる。レコメンデーションでは、ユーザにアイテム（商品やコンテンツ）を薦めるにあたり、ユーザ毎に分類してもアイテム毎に分類しても購買履歴のある組み合わせはごく一部に留まるため、値のほとんどがゼロである大きな行列が得られることとなる。これにより新たな値に対して類似のクラスが見つけれられない（cold start problem）やクラス分類に関係するデータ数が少ないためうまく分類できない（reduced coverage problem）といった問題が発生している。この問題に対処する方法が、例えば次元圧縮、スムージングである。本研究においても既に述べたように次元圧縮（「5.2 予測に用いる変数」）、スムージング（「5.4.1 ナイーブベイズ分類器」）を採用している。

### 5.6.3.1 階層型クラスタ分析

情報収集期間の工事内容を集めた工事データ1件を1つの文書と見做して、道路変化イベントを文書のテーマあるいはトピックと見做し、工事内容を文中の語句と見做すことで、本研究の予測は文書分類（ある語句＝工事内容があったときに、道路変化イベントの文書であるか否か）と考えることができる。文書分類には多くの取り組みがなされているが、ここでは階層型クラスタ分析の適用について検討を行う。

ここまで本研究では予測に道路変化イベントの有無をクラス（目的変数）とする、二値分類の教師あり学習を適用してきた。そのために道路変化イベントの有無を外的規準とし、それに従って工事内容を分類したデータを教師データとしてきた。しかし、道路開通や供用開始の有無といったクラスで分類することは適当ではなく、例えば、橋梁や隧道といった構造物毎の分類がより相応しい可能性が考えられる。このような立場に立ち、工事内容をグループ（クラスタ）に分類したいが、どのような規準で分類すべきか不明であるときには、教師なし学習による分類がより相応しい。

階層型クラスタ分析は教師なし学習の代表的な手法の一つである。階層的クラスタ分析は、要素間の距離（非類似度）あるいは類似度にしたがって近いものから要素をまとめていく手法であり、要素がまとめられた理由が明示でき、同一のデータからは同一の結果が得られる、すなわち再現性があるという特徴を持つ。教師なし学習には他にk-means法、自己組織化マップがあるが、これらにはランダム性があるため、計算結果が必ずしも同じにはならない。ここでは、分かり易さから階層的クラスタ分析を用いて検討を行う。用いた階層型クラスタ分析の要素間の類似度の算出にはコサイン類似度を用い、クラスタ間の類似度には群平均化法を採用した。以下にこの手法について説明する。

#### （1）コサイン類似度

文書分類では、文書中の単語をベクトルとして表現し、文書同士をそのベクトル空間上で比較することによって文書間の類似度を計算する手法が広く利用されている。その際にコサイン類似度が用いられることが多い。本研究では情報収集期間中の1地点で得られた工事内容の各説明変数（例えば、アスファルト舗装、道路付属物など）が、それぞれ独立したベクトル空間で扱われる単語と同じ意味を持つことになる。コサイン類似度は、ベク

トル同士の成す角度  $\theta$  の余弦  $\cos \theta$  であり、ベクトルの向きの近さを類似性の指標としたものである。したがって、1 に近ければ類似しており、0 に近ければ似ていないことになる。二つの説明変数からなるベクトル  $X$ 、 $Y$  のコサイン類似度は式(5.21)によって得られる。

$$\text{コサイン類似度}(X, Y) = \frac{\vec{X} \times \vec{Y}}{|\vec{X}| |\vec{Y}|} \quad (5.21)$$

なお、文書分類では、単語による出現のし易さや特徴となる度合いを、TF (Term Frequency : 語の出現頻度 (文書内での出現数 / 文書内の全単語数)) と IDF (Inverse Document Frequency : 逆文書頻度 ( $\log$  (全文書数 / 単語を含む文書数))) を用いて表現して、TFIDF (TF×IDF) により語の重みづけを行うことがある。

#### (2) 群平均化法

群平均化法とは、クラスタ内の各要素について、他のクラスタ内の各要素との類似度を算出し、その平均をクラスタ間の類似度とする計算方法である。クラスタ内の遠いものや近いものに引きずられることがないという特徴がある。

#### (3) 階層的クラスタ分析

階層型クラスタ分析は、全ての場所工事データ (要素) について、要素相互のコサイン類似度を算出し、類似度の最も高い要素をまとめてクラスタを生成する。引き続き、要素間、クラスタと要素、クラスタ間のコサイン類似度を群平均法を用いて算出し、類似度の最も高いものをクラスタ化するという手順を繰り返す<sup>34)</sup>。

#### (4) 階層型クラスタ分析の利用法

本分析の結果の利用法として、道路変化イベントに直接関係するクラスタが見いだせれば最善である。あるいは、関係しないクラスタを見出すことでデータの前処理フィルタを作成できる可能性がある。また、道路変化イベントに関係のある割合の高いクラスタをクラスとする教師あり学習を行って、コサイン類似度以外の指標で改めて分類することが考えられる。

#### (5) 階層型クラスタ分析の適用例

図 5.6-2 は、埼玉県の工事データ (2008-2010 年度) を上記の方法にしたがって階層型クラスタ分析し、類似度 0.8 以上を一つのクラスタとしたときの分析結果である。図では中心的な説明変数を明示し、数の多いものを大きな円で示した。この図によると、例えば、アスファルト舗装 (As 舗装) を中心とするクラスタは、その数は多いが、種類 (バリエーション) は少ないこと、また、橋梁上部 (橋上部) と橋梁下部 (橋下部) はそれぞれを中心とする別のクラスタとなっているがこれらには類似性があることが見て取れる。

このようにクラスタ化された工事内容を各地点の時間経過に当てはめたものを図 5.6-3 に示す。

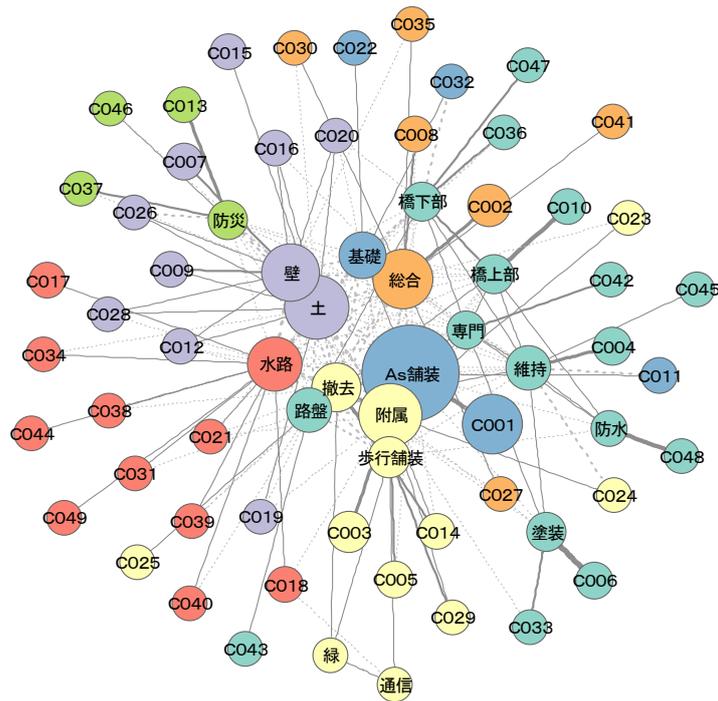


図 5.6-2 工事内容の階層型クラスタ分析の例



図 5.6-3 工事内容クラスタの発生状況の例

### 5.6.3.2 アソシエーション分析

橋梁の建設、隧道の建設といった複数のルールを工事内容から見出すためには、アソシエーション分析の適用が考えられる。アソシエーション分析は、マーケット・バスケット分析（一つの買い物かごに入れられる＝同時に購入される商品の分析）に利用されるもので、ルールの検出のアルゴリズムとしては Apriori が有名である。

情報収集期間の工事内容の説明変数をそれぞれ 1 つの商品と見做し、同時に道路変化イベントも 1 つの商品と見做すことにより、どの商品（＝説明変数）が買われた（＝出現し

た)ときに道路変化イベントが買われた(=発生した)かのルールをアソシエーション分析により見出すことができる。Aprioriで見出されたルールは、 $X \Rightarrow Y$ の形式で表され、左辺(X)の条件部(antecedent: left-hand-side or LHS)、右辺(Y)の結論部(consequent: right-hand-side or RHS)、支持度(support)、確信度(confidence)、リフト(lift)が示される<sup>35)</sup>。支持度( $X \Rightarrow Y$ )はXとYを含む関係(トランザクション)が全体の中に占める割合で定義される。確信度( $X \Rightarrow Y$ )はXとYを含むトランザクションの数を条件部Xを含むトランザクションの数で割った値である。リフト( $X \Rightarrow Y$ )は確信度( $X \Rightarrow Y$ )を指示度(Y)で割った値である。埼玉県工事データと道路開通(いずれも2008-2010年度)にアソシエーション分析を行い、リフト値の順番に並べた一部を図5.6-4に示す。最初のルールでは、道路付属物の新設、路盤路床、道路改良は、工事データの0.136%の出現(support)であるが、これがあつたときには100%開通が現れていることを示している(confidence)。そして、条件部があつたときに開通のあつた確率は一般的な開通の確率に比べて9.25倍高いことが示されている(lift)。

|    | lhs  | rhs     | support     | confidence | lift     |
|----|--|---------|-------------|------------|----------|
| 1  | {道路付属新設,<br>路盤路床,<br>道路改良}                 | => {開通} | 0.001356852 | 1.0000000  | 9.251046 |
| 2  | {アスファルト舗装,<br>盛土,<br>軟弱地盤}                 | => {開通} | 0.002261420 | 1.0000000  | 9.251046 |
| 3  | {掘削,<br>盛土,<br>土留め仮締切,<br>橋梁下部}            | => {開通} | 0.001809136 | 1.0000000  | 9.251046 |
| 4  | {アスファルト舗装,<br>路盤路床,<br>盛土,<br>軟弱地盤}        | => {開通} | 0.001809136 | 1.0000000  | 9.251046 |
| 5  | {アスファルト舗装,<br>掘削,<br>盛土,<br>軟弱地盤}          | => {開通} | 0.001809136 | 1.0000000  | 9.251046 |
| 6  | {アスファルト舗装,<br>路盤路床,<br>掘削,<br>盛土,<br>軟弱地盤} | => {開通} | 0.001356852 | 1.0000000  | 9.251046 |
| 7  | {掘削,<br>土留め仮締切,<br>橋梁下部}                   | => {開通} | 0.003618272 | 0.8888889  | 8.223152 |
| 8  | {掘削,<br>盛土,<br>土留め仮締切}                     | => {開通} | 0.002261420 | 0.8333333  | 7.709205 |
| 9  | {路盤路床,<br>盛土,<br>軟弱地盤}                     | => {開通} | 0.002261420 | 0.8333333  | 7.709205 |
| 10 | {アスファルト舗装,<br>道路付属新設,<br>道路改良}             | => {開通} | 0.001809136 | 0.8000000  | 7.400837 |

図 5.6-4 道路開通に対するアソシエーション分析の例

アソシエーション分析の成果について、どの範囲のルールを道路変化イベントに組み込

むことで、どの程度の評価値（真陽性率など）が得られるかを実験し、工事内容と建設物の関係の知見を増やし、予測の正確さを向上することが考えられる。

#### 5.6.4 工事の順序関係（マルコフ連鎖）を利用する方法

本章で述べた工事内容による予測方法は、情報収集期間内の工事の内容を集めたデータが道路変化イベントの有無を決定づけるとの考えに基づいている。この考えに基づいて確率的分類器であるナイーブベイズ分類器による予測では、工事内容の条件下での道路変化イベントの有無の条件付き確率を求め、MAP 推定を行い、道路変化イベントの有無を予測している。これは工事の内容を集めたデータと道路変化イベントの有無を状態空間とした単純マルコフ連鎖を想定し、その中で、工事内容と道路変化イベントの有無の関係のみを取り扱っていると考えることができる。

工事内容と道路変化イベントの関係を、ある工事内容が観測されたときに次の工事の内容を推定するように拡張することが考えられる。すなわち、ある工事（群）から次の工事（群）、そして道路変化イベントの発生に至る遷移を、マルコフ性を持つ確率過程、すなわちマルコフ過程であると考えられる。具体的な例で示すと、とりえる値（状態空間）が、橋梁の下部工事、橋梁の上部工事、道路変化イベント（開通）、あるいは、基礎工事、路盤路床工事、アスファルト舗装工事、道路付属物工事、道路変化イベント（供用開始）といった集合を考える。このとき、ある地点の次の期間の値（状態）は、現在の値（状態）から決まる、言い換えると、マルコフ性を有する確率過程であると考えられると、状態空間は離散集合で表わされていることから、このマルコフ過程は離散マルコフ過程、すなわちマルコフ連鎖となる。

最初に、これが単純マルコフ連鎖、つまり、1つの時点の状態だけで次の状態が得られるものであると想定する。形式的には式(5.22)で示される。

$$P(X_{n+1} = x | X_n = x_n) \quad (5.22)$$

橋梁下部工 (No.1)、橋梁上部工 (No.2)、工事無し (No.3)、道路変化イベント (No.4) の4つの状態からなるマルコフ連鎖を例として考えることにする。これらの状態が次の時点に遷移していく状態とその確率を図 5.6-5 に示す。これは状態遷移図と呼ばれるものである。ただし、確率はあくまで例である。図 5.6-5(a) の状態遷移図は、橋梁下部工が観測されたときに、次の期間に同じように橋梁下部工が観測される確率を0、橋梁の上部工事が観測される確率を0.5、何も工事がなされない確率を0.3、道路変化イベントが発生する確率を0.2であることを示している。他の状態も確率は異なるが考え方は同様である(図 5.6-5(b)~(d))。

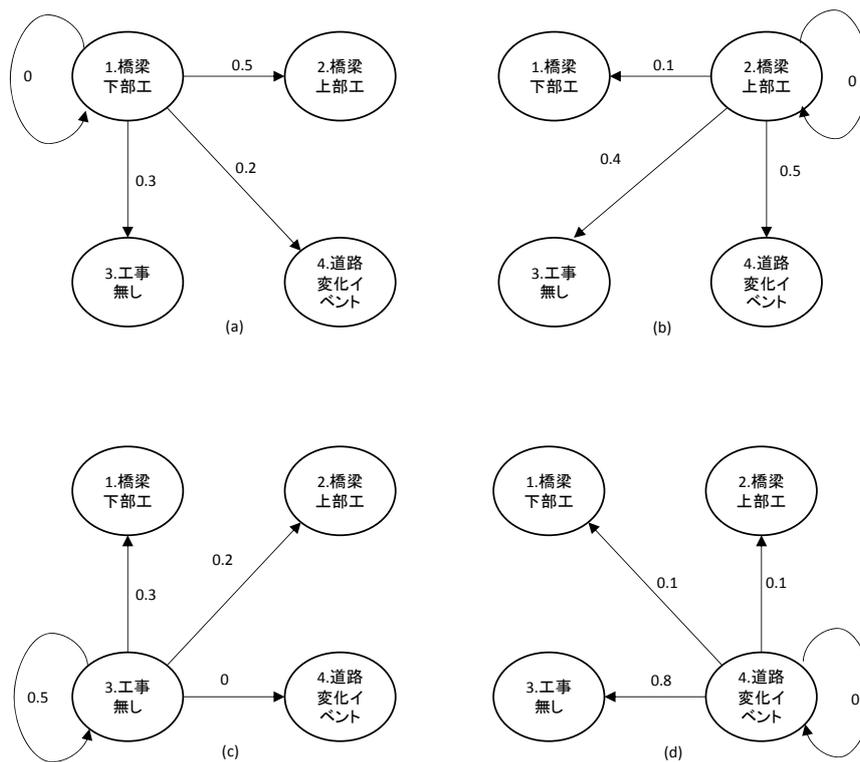


図 5.6-5 状態遷移図の例

$ij$  成分に No.  $i$  から No.  $j$  に遷移する確率を記したものを遷移確率行列という. 図 5.6-5 の例では遷移確率行列は次のように表される.

$$P = \begin{bmatrix} 0 & 0.5 & 0.3 & 0.2 \\ 0.1 & 0 & 0.4 & 0.5 \\ 0.3 & 0.1 & 0 & 0.6 \\ 0.1 & 0.1 & 0.8 & 0 \end{bmatrix}$$

$n$  期間経過に対する遷移確率行列を  $P^{(n)}$  とする. すなわち,  $P^{(n)}$  の  $ij$  成分は式 (5.23) となる.

$$P(X_{n+1} = j | X_n = i) \tag{5.23}$$

このときチャップマン・コルモゴロフ方程式により式 (5.24) となることが分かっている.

$$P^{(n)} = P^n \tag{5.24}$$

これにより、例えば、橋梁下部工が観測されたときに2期あとに道路変化イベントが生ずる確率は $P^2$ の計算により0.25と求めることができる。

$$P^2 = \begin{bmatrix} 0.16 & 0.05 & 0.54 & 0.25 \\ 0.17 & 0.14 & 0.67 & 0.02 \\ 0.19 & 0.21 & 0.49 & 0.11 \\ 0.25 & 0.13 & 0.55 & 0.07 \end{bmatrix}$$

この単純マルコフ連鎖の例では、工事が無いときに次の時期に道路変化イベントが発生する確率は0である。しかし、工事が観測されなかった期間の直前に大規模な工事がなされている地点では、長期間にわたって工事が観測されていない地点に比べて、道路変化イベントが発生する確率は高くなることが考えられる。このように、確率過程は、次の状態が現在を含めた過去N個の状態履歴に依存して決まると考えることもできる。これはN階マルコフ連鎖と呼ばれるものであり、形式的には式(5.25)で表される<sup>36)</sup>。

$$P(X_{n+1} = x | X_n = x_n, \dots, X_0 = x_0) \quad (5.25)$$

複数の工事(群)から道路変化イベントに至る確率過程をN階マルコフ連鎖と考えると道路変化イベントを予測するには、まず、(a)状態空間、(b)情報収集期間の長さ、(c)N階マルコフ連鎖のNの値について、それぞれ検討し設定を行う必要がある。その後、道路変化イベントの発生した地点、発生しなかった地点の工事(群)に関する情報をN+1回の情報収集期間について収集し、これらの情報から状態空間の値の条件付き確率を求め各階の遷移確率行列を生成する。この遷移確率行列を用いて未知の地点の工事情報から道路変化イベントを予測することとなる。この方法によると道路変化イベントだけでなく、ある工事群から次に生ずる工事群の条件付き確率も求めることができる。

## 5.7 件数による予測、機械学習による予測の利用の位置づけ

工事情報から道路開通と供用開始のそれぞれに対して短期予測(道路変化イベントに近い時期)と長期予測(道路変化イベントの2年前)を行った実験について述べてきた。これまでの結果をまとめたものを表5.7-1に示す。この結果に対して、予測へのリクワイメントから評価し、予測方法をどのように利用すべきかについて述べる。

表 5.7-1 予測まとめ

|                         |   |   |   |                        |                       |                        |
|-------------------------|---|---|---|------------------------|-----------------------|------------------------|
| 数字は真陽性率の平均が高かったもの       | <u>短期予測 (St)</u><br>工事の開始前後にデータが入手できるため、工事期間（数箇月程度）のリードタイムが得られる |   | <u>長期予測 (Lg)</u><br>工事期間（数箇月）に加えて2年のリードタイムが得られる |                        |                       |                        |
| 予測方法                    | 工事件数による予測   |   | 工事内容による予測                                       |                        |                       |                        |
| 評価項目                    | 真陽性率  | 真陰性率  | 真陽性率  | 真陰性率                   | 真陽性率                  | 真陰性率                   |
| 道路開通（大規模な道路変化イベント）      | 0.92(500万円以上, 2件以上)<br>0.69(2500万円以上, 2件以上)                     | 0.57(500万円以上, 2件以上)<br>0.65(2500万円以上, 2件以上) | 0.92 (St51, RF, 木の深さ1)                          | 0.67 (St51, RF, 木の深さ1) | 0.5 (Lg51, 決定木)       | 0.97 (Lg51, 決定木)       |
| 供用開始（小規模を含む全ての道路変化イベント） |   |   | 0.63 (St51, RF, 木の深さ1)                          | 0.46 (St51, RF, 木の深さ)  | 0.49(Lg31, RF, 木の深さ1) | 0.56 (Lg31, RF, 木の深さ1) |

(1) 情報の空白地帯への利用

「1.2.4 道路変化イベントの情報の収集の現状」の「表 1.2-6 道路変化イベントの把握方法の種類と特徴」に示したように、これまで使用されてきた道路変化イベントの把握方法のうち、リードタイムが得られる方法は道路管理者への問い合わせだけである。問い合わせへの回答には法的な義務はないため、その対応は道路管理者の考えに依存し、加えて得られた回答を確かめる術がない。

これに対して、工事情報は主として公共工事の入札及び契約の適正化の促進に関する法律によって公表が義務づけられている。そして、工事情報から道路変化イベントを得ることによって、リードタイムが確保できる別の情報収集ルートが得られることとなる。これは、問い合わせに回答が得られない道路管理者の管内では情報が皆無であったことと比べれば、大きなメリットとなる。このため、情報が得られない道路や、情報が不正確であった道路に対して重点的に利用することが考えられる。

(2) 簡易な予測方法の適用範囲

本研究において工事件数と工事内容の2種類の予測手法を提案した。工事件数による予測は、取り扱う情報の項目が少なく、処理ルールが全国的に一律であり、適用が簡単であるという利点がある。このため、問い合わせや詳細調査の重点箇所を一斉に洗い出す用途に向いている。これに対して工事内容による予測は、長期、小規模なものにも適用できる利点がある反面、予測結果にばらつきがあるため、当面は適用範囲を絞って運用して、運用と並行して正確さと利便性を向上させていくことが良いと考える。

### (3) 重要な道路変化イベントの把握

道路開通の短期予測は工事件数による予測でも、工事内容による予測でも良い成績であった。道路開通は一般的な供用開始よりも規模が大きな工事が多く、道路開通を特徴づける工事が発生していたためと推察する。見逃せない重要な道路変化イベントである道路開通については、両者の予測方法を併用することが考えられる。

### (4) 無駄な問い合わせの削減

短期予測は長期予測に比べて道路変化イベントに近い時期の工事を予測に用いている。短期予測に用いた工事は、道路建設の最終段階の工事であることが多いと考えられる。これに対して、長期予測に用いた2年前の工事では、建設対象、自然条件、予算によって内容がばらついていると考えられる。その結果、長期予測の真陽性率は短期予測に比べて低い。しかし、真陰性率は高く、例えば、ランダムフォレストで木の深さを大きくして学習を進めることで真陰性率をさらに高めることができていた。真陰性率が高いということは、予測の結果に基づいて道路変化イベントが示唆された地点にはずれが少ないことを意味する。このため、道路管理者へ誤った問い合わせが少ないことから、回答が得られやすいというメリットがある。

これらを整理した予測手法の適用のフローを図 5.7-2 に示す。

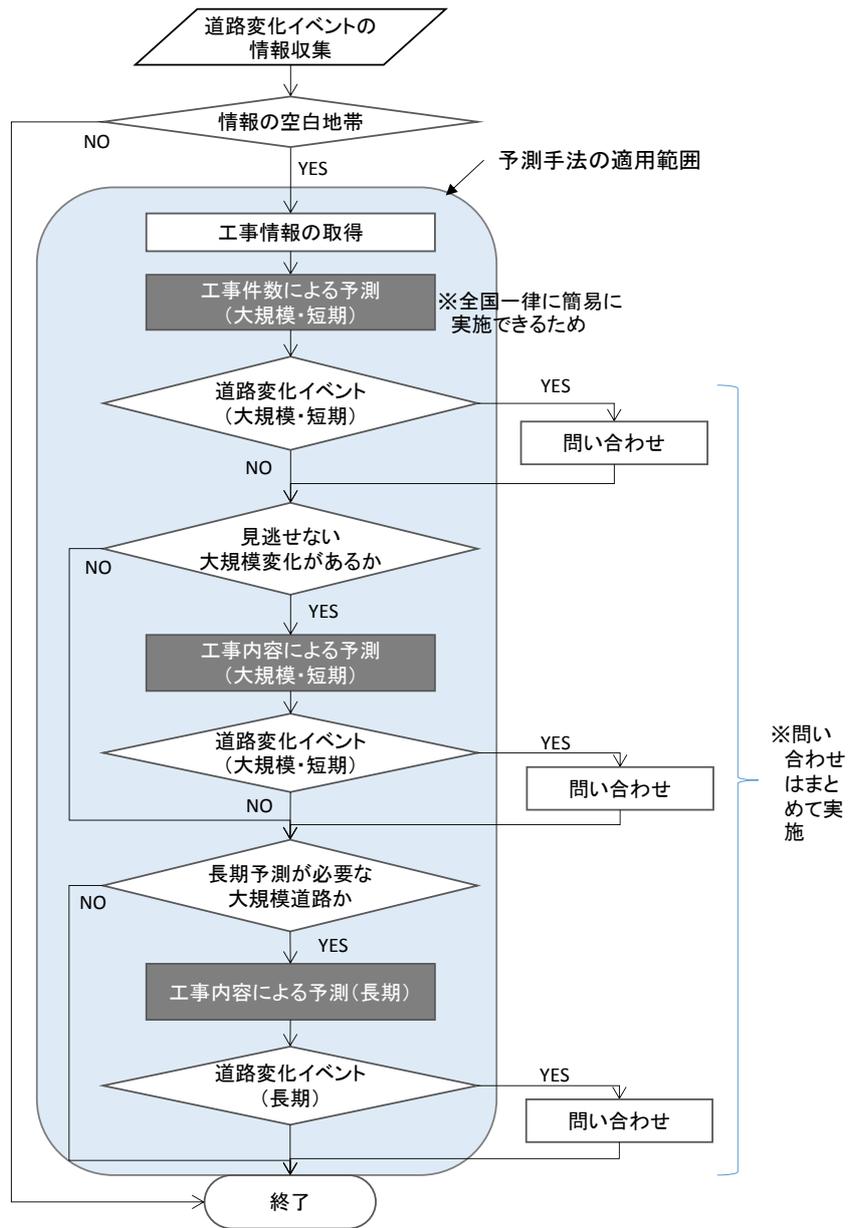


表 5.7-1 予測の適用フロー案

## 参考文献

- 1) C.M.ビショップ,元田浩,栗田多喜夫,樋口知之,松本裕治,村田昇 : パターン認識と機械学習, pp.1-3,pp.33-37, 丸善出版, ISBN978-4621061220, 2012.
- 2) Willi Richert・Luis Pedro Coelho 著, 斎藤康毅訳 : 実践機械学習システム, オライリー, pp.214-233 (11章次元削減), ISBN978-4-87311-698-3, 2014.
- 3) 国土交通省国土技術政策総合研究所 : 1.1 工事工種体系における分類と整備状況 (平成19年度改訂版), <http://www.nilim.go.jp/lab/pbg/theme/theme2/sekisan/tree/t1-1.pdf>, 2014.6.5.アクセス.
- 4) 国土交通省国土技術政策総合研究所 : 1. 2 体系階層 (レベル) の定義, <http://www.nilim.go.jp/lab/pbg/theme/theme2/sekisan/tree/t1-2.pdf>, 2014.6.5.アクセス.
- 5) 国土交通省国土技術政策総合研究所 : 工事工種体系ツリー (平成25年度10月版), <http://www.nilim.go.jp/lab/pbg/theme/theme2/sekisan/tree/t1-5.pdf>, 2014.6.5.アクセス.
- 6) 国土交通省中部地方整備局 : 道路設計要領 (設計編), <http://www.cbr.mlit.go.jp/road/sekk-eiyouyou/index.htm>, 2012. 7. 23. アクセス.
- 7) 石井一郎・元田良孝 : 道路工学, 第8章道路土工~第15章維持補修, 鹿島出版会, 1993.1.
- 8) 一般財団法人日本建設情報総合センター : コリンズ項目定義書 XML 定義書, <http://ct.jacic.or.jp/corporation/known/xml/file/c02x01.xls>, 2012. 10. 13. アクセス.
- 9) 一般財団法人日本建設情報総合センター : コリンズ選択項目詳細-工種・工法, 型式/担当工事内容, [http://ct.jacic.or.jp/corporation/known/xml/file/c02x01\\_16.xls](http://ct.jacic.or.jp/corporation/known/xml/file/c02x01_16.xls), 2012. 7. 5. アクセス.
- 10) Ian H. Witten and Eibe Frank : DATA MINING Practical Machine Learning Tools and Techniques, pp.103, pp.289, ISBN978-0120884070, 2nd ed., ELSEVIER, 2005.
- 11) Rushi Longadge, Snehlata S. Dongre, Latesh Malik : Class Imbalance Problem in Data Mining: Review, International Journal of Computer Science and Network (IJCSN) Volume 2, Issue 1, ISSN 2277-5420, 2013.2
- 12) Nathalie Japkowicz : Learning from Imbalanced Data Sets: A Comparison of Various Strategies, AAAI workshop on learning from imbalanced data sets,\* AAAI Technical Report WS-00-05, 2000.
- 13) Nitesh V. Chawla : Data Mining for Imbalanced Datasets:An Overview, Data Mining and Knowledge Discovery Handbook 2<sup>nd</sup> ed., Part VI Advanced Methods, Chapter45, pp.875-886, ISBN 978-0-387-09823-4, Springer, 2010.4
- 14) Haibo He and Edwardo A. Garcia : Learning from Imbalanced Data, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 21, NO. 9, 2009.9
- 15) Xindong Wu, Vipin Kumar : The Top Ten Algorithms in Data Mining, ISBN 9781420089646, CRC Press, 2009

- 16) Trevor Hastie, Robert Tibshirani, and Jerome Friedman : The Elements of Statistical Learning: Data Mining, Inference, and Prediction.(Second Edition), pp.219-233 (Model Assessment and Selection), Springer, eBook ISBN:978-0-387-84858-7, 2009
- 17) 加藤直樹・羽室行信・矢田勝俊著 : データマイニングとその応用, pp111-115, 朝倉書店, ISBN-13: 978-4254275520, 2008.
- 18) George h. John, Pat Langley: Estimating Continuous Distributions in Bayesian classifiers, in Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann Publishers, 1995
- 19) Willi Richert・Luis Pedro Coelho 著, 斎藤康毅訳 : 実践機械学習システム, pp.119-120 (6章クラス分類II), オライリー, ISBN978-4-87311-698-3, 2014
- 20) Webb, G. I., J. Boughton, and Z. Wang : Not So Naive Bayes: Aggregating One-Dependence Estimators. Machine Learning 58(1), pp.5-24, Netherlands: Springer, 2005
- 21) 加藤直樹・羽室行信・矢田勝俊著 : データマイニングとその応用, pp95-110 (5.決定木), 朝倉書店, ISBN-13: 978-4254275520, 2008
- 22) 元田浩・,山口高平・,津本周作・,沼尾正行著 : データマイニングの基礎, pp.21-29 (2.1 決定木), オーム社, ISBN-13: 978-4274203480, 2006
- 23) C. E. SHANNON : A Mathematical Theory of Communication , The Bell System Technical Journal, Vol. 27, pp. 379–423, 623–656, 1948.
- 24) J.R. QUINLAN : Induction of Decision Trees, Machine Learning 1: 81-106, 1986
- 25) J. Ross Quinlan : C4.5: programs for machine learning, Morgan Kaufmann Publishers Inc., ISBN:1-55860-238-0, 1993
- 26) 中田豊久 : 基礎から学ぶデータマイニング, pp.46-47, コロナ社, ISBN : 978-4-339-02470-8, 2013
- 27) Breiman, Leo (2001). “Random Forests”. Machine Learning 45 (1): 5–32. <http://www.stat.berkeley.edu/~breiman/randomforest2001.pdf>
- 28) 元田浩・,山口高平・,津本周作・,沼尾正行著 : データマイニングの基礎, pp.51-55 (3.1 アンサンブル学習), オーム社, ISBN-13: 978-4274203480, 2006.
- 29) 関本義秀, 中條覚, 南佳孝, 山口章平, 山田晴利, 布施孝志 : 工事発注見通し情報を用いた全国における道路更新情報の自動抽出に向けた試み, 土木学会論文集 D3 (土木計画学), Vol. 68, No. 3, pp.117-128, 2012.
- 30) 中條覚 : 工事入札公告を用いた道路更新情報の自動収集・推定に関する研究, 東京大学博士論文, pp.10, 2012.
- 31) 高村大也, 奥村学 (監修) : 言語処理のための機械学習入門, pp.71 (2.4.3 データスパースネス問題), コロナ社, 2010.
- 32) Xiaoyuan Su and Taghi M. Khoshgoftaar : A Survey of Collaborative Filtering Techniques, Advances in Artificial Intelligence, Volume 2009 (2009), Article ID 421425, 19 pages

- 33) ローネン・フェルドマン, 辻井潤一ほか(翻訳): テキストマイニングハンドブック, p p179 (7.1.5 学習データスパースネスの扱い), ISBN:978-4501548100, 東京電機大学出版局, 2010.
- 34) 金明哲: R によるデータサイエンス-データ解析の基礎から最新手法まで, ISBN 978-4627096011, 森北出版, 2007.
- 35) 加藤直樹・羽室行信・矢田勝俊著: データマイニングとその応用, pp17-42, 朝倉書店, ISBN-13: 978-4254275520, 2008.
- 36) R.B. シナジ著, 今野紀雄, 林俊一訳: マルコフ連鎖から格子確率モデルへ (現代確率論の基礎と応用), ISBN 4431708650, シュプリンガー, 2001.

## 第6章 情報の流通とシステムの有用性に関する評価

### 6.1 情報の流通への取り組み方針

道路変化イベントを直接的に把握するには供用開始情報、道路開通情報の入手が必要である。そして、道路変化イベントを早期に把握するには工事情報からの予測という方法が考えられ、そのためには工事情報の入手が必要である。本研究で取り扱った原典情報は、法令によって公示や公表が求められているものを主としており、紙媒体による掲示も含めると何らかの形では公開されているものが大部分である。しかし、日本では公開に電子的な手段を用いることは求められておらず、その実施は各行政機関に任されている。この結果、「1.3 先行研究」で示したように、これらの情報は必ずしもインターネットなどの電子的な手段で入手が可能ではなく、そして、電子データで入手したものであってもシステムで利用できるフォーマットへ構造化することは容易ではない状況となっている。このようなことから、研究の実用化のためには、情報の利用を容易にするための環境の実現に取り組む必要がある。

情報ビジネスを生業とする業者が提供する情報が一部にあるものの、原典情報の提供者は主として道路管理者である。このため、情報流通環境の整備には情報発生源である道路管理者の協力を得ることが効果的であると考えた。

この章では、道路管理者に対して道路変化イベントの把握のための情報の流通への可能性について調査した。そして、情報の流通のために本研究で構築したシステムの活用を目指して、システムの有用性について意見を収集した。

### 6.2 流通の対象とする情報について

流通の対象となる情報について、本研究でこれまで対象としてきた原典情報以外に有効な情報が存在する可能性を調査した。

#### (1) 道路変化イベントの発生を把握する手段について

埼玉県下、愛媛県下の82市町村に、道路変化イベントの発生を把握するとしてどのようなものがあるかについて2015年に調査を行った。得られた回答数は41であった。得られた回答の全てにおいて、(a)供用の開始の告示、(b)道路台帳のいずれかまたは両方という内容であった。それ以外の原典情報として、都市計画図を挙げた自治体が1あった。なお、(a)供用の開始の告示については窓口で閲覧のみと付記した自治体が1あった。これにより電子的な手段で入手できないものがあることが自治体の回答からも確認された。(b)道路台帳については、その更新には例えば半年以上の時間がかかるという意見が4件付記された。このように、道路台帳は詳細情報を得る手段として適当であるが、道路変化イベントを速やかに把握する手段として利用することには適当ではないことが自治体の回答からも確認された。

以上のように、道路変化イベントの発生を把握する手段として「表 1.3-1 原典情報の時期と目的」に示したものの以外に新たなものは見つけられなかった。

(2) 予定されている道路変化イベントの情報

道路変化イベントの発生を前もって知るために有効な方法についても、上記と合わせて調査した。回答数は同じく 41 であった。得られた回答を集計した結果、(a)道路区域の告示としたものが 54%と第一位、(b)計画書が 31%で二位、(c)都市計画情報が 7%、(d)無いという回答が 7%であった。(a)道路区域の告示は供用の開始と同時あるいはそれ以降と回答した自治体が 2 あった。第三章に示した調査からも、道路区域の告示は必ずしも供用開始に先立って行われているとは限らないことが確認されている。(b)計画書については、予算により変更される、用地交渉に支障がある、内部資料であるなどの理由から、10 自治体では外部では利用できないとの意見が付記されていた。その他の資料として、事業説明会を挙げた自治体が 1 あった。しかし、事業説明会に参加して情報を得ることは、事業説明会の開催の情報を得ることも含めて困難であると考えられる。

このように、道路変化イベントを前もって知るために網羅性があり、かつ、確実な方法は見つけられなかった。

以上の回答の集計結果を表 6.2-1 に示す。

表 6.2-1 道路管理者のアンケートで挙げられた情報 (N=41, 2015 年度)

| 時期    |           | 早期              |            |         | 道路変化イベント時  | 事後        |
|-------|-----------|-----------------|------------|---------|------------|-----------|
| 情報の種別 |           | 道路区域(告示)        | 事業計画書      | 地元説明会資料 | 供用の開始(告示)  | 道路台帳      |
| 評価項目  | リードタイム    | ○               | ○          | ○       | ×          | ×         |
|       | 入手可能性(公開) | ○               | ×          | ×       | ○          | ○         |
|       | 入手容易性(電子) | △               | ×          | ×       | △          | △         |
|       | 網羅性(義務づけ) | ○               | ×          | ×       | ○          | ○         |
|       | 詳細情報      | ×               | ○          | ○       | ×          | ○         |
|       | 備考        | 供用の開始と同時との回答もあり | 内部資料, 変更有り | 不定期, 現地 | 電子手段とは限らない | 更新に時間がかかる |

6.3 道路変化イベントの公表について

道路変化イベントの公表の実態と公表についての考えを調査した。調査の対象と期間は前節と同じである。回答数は 37 であった。36%の自治体ではホームページで主要な開通について公表しているとの回答があり、26%で公表の必要なしとの回答があった。最も多かったのは公表の意思はあるが余裕が無いとの回答で 38%であった。道路変化イベントの公表に前向きである、公表を実施している、公表の意思はあるとした回答の合計は 73%であった。このような意思に応える手段として、本研究で構築したシステムの利用が考えられる。調査結果を図 6.3-1 に示す。

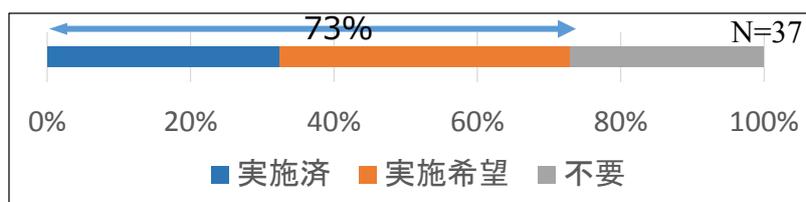


図 6.3-1 道路変化イベントの公表に対する道路管理者の意見

#### 6.4 工事情報の流通について

情報の流通には、情報が利用しやすい形で提供され、その利用が認められる必要がある。そのため、統一的なフォーマット、提供する意思、二次利用について提案並びに調査を行った。

##### 6.4.1 統一的なフォーマットの提案と提供可能性

電子的に提供されている情報であっても、手作業でなければデータ化できないものが少なくないことは先行研究で明らかにされ、本研究においても少なからずのデータを人手で登録している。提供者に対して情報のフォーマットを提示することにより、利用者の利便性が向上するのはもちろんであるが、提供者にとっても、自らがフォーマットを考える必要がなく、他の自治体で利用されているのであれば安心であるなどの理由からメリットがある。このため、本研究で用いたデータフォーマット（第3章で示したデータテーブル）をシステムとともに公開することとした。これらのデータテーブルのうち、項目について行政機関から意見が寄せられる可能性のある工事データテーブルについて、道路管理者から意見を収集した。データフォーマットのうち、公示情報、開通情報については揺らぎが少ないうえ、流通にあたり、項目の問題は無いと判断した。提示した工事データフォーマットを図 6.4-1 に、寄せられた意見を図 6.4-2 に示す。アンケート結果の異議有りの中には、予測精度の懸念なども含まれており、これらを意義とてカウントしても、約 85%の自治体から異議がなかった。異議や条件の中には契約金額については不要ではないかという 4 件あった。

フォーマットには無いが、道路の新設の有無、不明について記入する項目を追加することによって、より良い整理ができると考えられる。

| (データの使用用途)         | (項目)          | (例)              |
|--------------------|---------------|------------------|
| (開通との関連の有無)        | 工事件名          | 市道第〇〇号線道路改良工事    |
| (時期)               | 完了日           | 2020/3/29        |
| (場所)               | 発注機関          | 〇〇県〇〇市           |
| (規模の参考)            | 請負金額(円)       | 12,345,678       |
| (場所)               | 路線・水系名等       | 〇市道第〇〇号線         |
| (開通との関連の有無)        | 工種、工法・型式      | その他土木一式工         |
| (場所)               | 施工場所          | 〇〇県〇〇市〇〇         |
|                    | 緯度経度          | 35, 12, 135. 12  |
| (現道の拡幅の可能性を判断するため) | 夜間工事の有無       | 夜間工事無し           |
|                    | 交通規制の有無       | 交通規制有り           |
|                    | 規制道路の種別       | 市町村道             |
|                    | 規制道路の交通量(台/日) | 5,000以上~10,000未満 |
|                    | 規制車線数         | 規制車線数1/全車線数1     |
|                    | 交通の確保手段       | 切り回し             |

黒地白抜きは本研究で予測に用いた項目

図 6.4-1 道路管理者に提示した工事データフォーマット

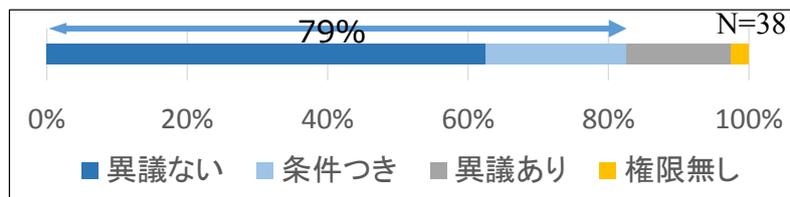


図 6.4-2 工事データフォーマットの実利用の可否について (N=38, 2015年)

## 6.5 システムの有用性の評価

このシステムが地図製作者等の道路管理者以外の情報の利用者だけではなく、情報を提供する側の道路管理者にとって有用であるか道路管理者から意見を聞いた。その結果を図 6.5-1 に示す。既に庁内システムを構築している自治体以外からは、システムによって異なる部署間の情報共有が進むとの意見が寄せられ、その割合は過半数となり、道路管理者にとっても有用なものとなることが期待できる。表示が分かりにくいとされたものは、現在のシステムのアイコンの表示が、工事データの場所の精度に合わせて字丁目の中心となっていることに対して、道路上へアイコンを表示すべきという意見によるものである。これに対し

では、南佳孝らによる「路線情報を加味した道路関連情報の位置特定に関する研究」(土木学会論文集 F3 (土木情報), Vol. 67, No. 1, pp.7-17, 2011)などを参考に改良を検討している。

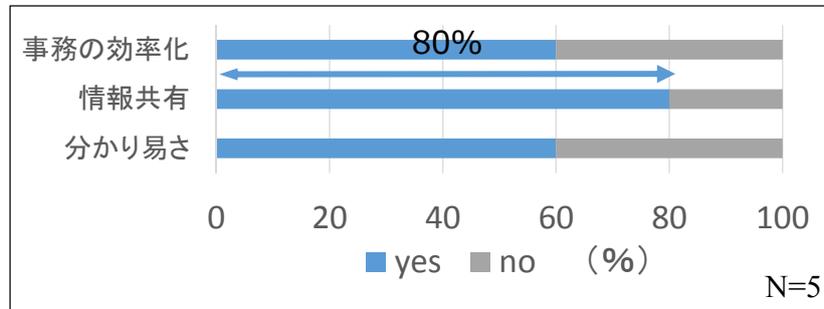


図 6.3-1 道路管理者からのシステムへの有用性の意見 (N=5,2015年)

## 第7章 総括

### 7.1 本研究の成果のまとめ

本研究は、道路変化イベントの早期かつ網羅的な把握の実用化に向けて、道路変化イベントと工事情報を組み合わせるための統合的システムを構築し、機械学習を利用した道路変化イベントの予測と情報の流通の促進に取り組み、以下の内容を明らかにした。

第1章では、まず、道路変化イベントを取り巻く現状と問題を研究の背景として概説した。道路管理者による情報の提供、そして、地図調製者による情報の収集が実施されているにも関わらず、これらがかみあっておらず、道路利用者がカーナビゲーションの地図の更新に不満を示しているという状況を示した。

次に、この問題の改善を図るための先行研究の調査から、主に3点の課題を抽出した。その一つ目は、情報利用の準備のためのリードタイムを得るには道路変化イベントに先立つ工事情報が有効であり、工事情報を活用するには道路変化イベントと対照する方法が必要であること。二つ目は、道路変化イベントの予測には、道路変化の種類、予測時期(2年後)というリクワイアメントに応える予測方法、簡便な予測方法が必要であること、三つめは、実用化には情報の利用を容易にする流通環境が必要であることである。これらを本研究の課題として設定し、研究の取り組みの方向性を示した。

第2章では、最初に、一つ目の課題である、工事情報と道路変化イベント(これらをまとめて原典情報と呼んだ)を対照させる方法を提案した。原典情報は場所に固定したものであり、そして、時間的な順序を有する事柄を示す情報であることから、本研究では、(a)場所と路線による集約、(b)時期による整理、を用いる方法を提案した。場所と路線による集約においては、地名辞典として郵便番号データを選定し、これに改良を加えて2010年度の道路工事データにおいて87%程度の場所を大字レベルの地点に変換できた。また、原典情報の起点と終点の経由地点を補間する方法を合わせて提案した。時期による整理においては、公示日、供用日、開通日、工事完了日を用いて事業の流れにあわせて原典情報を点過程として整理する方法を提案した。

次に、この提案した方法が実現可能であることを検証するため、加えて、情報を可視化して確認できるよう、(a)データベースシステムと(b)WEBマッピングシステムを中核とする統合的システムをデザインし構築した。

第3章では、統合的システムに実際の原典情報を投入し、供用開始情報、道路開通情報、道路区域情報、工事情報の実態を調査分析した。これによって道路変化イベントの把握の方向性を整理した。

調査分析の結果、道路区域に関する情報の告示が供用開始の告示に先立って実施されている傾向は大きくは見られなかった。このため、道路区域情報を道路変化イベントの予測に

利用することは現時点では考えにくいことが分かった。道路区域情報には対象区間の幅員と延長が示されている。他方の供用開始情報には供用開始日は必ずあるが、対象区間の規模を明示する項目は（道路法施行規則では示されていないので）有るとは限らない。調査分析から、道路区域情報の約 4 割は供用開始情報と同日同地点で公表されている傾向が見られた。このため、道路区域情報と供用開始情報とを組み合わせると、道路変化イベントの規模と時期が得られる可能性があることが分かった。

次に、工事情報と道路変化イベントの関係について調査分析を行った。ただし、工事の対象は契約額 2500 万円以上である。道路変化イベントのうち、道路開通情報の該当地点では工事が行われた割合が高く、しかも工事件数が他の地点に比べて多いことが分かった。供用開始の該当地点では、道路開通情報の地点には及ばないものの、工事が行われている地点が、公共事業費の多い自治体ほど高い割合で観測された。

第 4 章では、工事情報から道路変化イベントを予測することの利点、評価方法といった予測に共通的なことを最初に示した。利点として、(a)リードタイムの確保、(b)道路法の対象外の道路の把握、(c)道路変化イベントに関係する工事の絞り込みが得られる、ことを示した。評価方法として、道路変化イベントを把握できた割合である真陽性率（1 に近いほど良い）、道路変化イベントの無い場所を誤って有ると予測した偽陽性率（0 に近いほど良い）、これらを総合的に示す ROC 曲線と AUC を採用することを述べた。

次に、工事完了件数から道路開通を予測する方法を提案し、実際のデータにより検証した。この予測は、公共工事から道路工事を、主として発注者、工事件名、工事場所によって抽出し、その工事場所と工事完了日を第 2 章で提案した関連づけ手法で整理し、一定期間内の工事件数から開通を予測するものである。この予測方法は、工事の内容を入手する必要が無く、簡便に行えるという利点がある。その予測結果は、3 つの自治体の契約額 2500 万円以上の工事を用い、1 年度に 2 件以上の工事がある地点を開通と予測するルールによって真陽性率 0.69、偽陽性率 0.43 が得られた。同じ条件で 500 万円以上の工事を用いたときには真陽性率 0.92、偽陽性率 0.35 であった。このことから、予測の適用の範囲は開通の短期予測（数箇月）と制限はあるものの、簡便な方法によっても予測が可能であることを示した。

第 5 章では、道路変化イベントに対する多様な情報収集のリクワイヤメントに応えるために、工事内容を利用した予測の方法について述べた。工事内容とはアスファルト舗装や橋梁上部工などの工事概要あるいは工法・型式などに示されているものである。これらの種類は多岐にわたるため、予測に用いる説明変数（素性、特徴量）も多変数となる。したがって、高次のデータ空間を対象として予測を行える機械学習（教師あり学習）を適用した。

本研究では、機械学習によって予測するクラスを、規模別に道路開通、供用開始の 2 種類とし、それぞれについて時期別に短期（工事期間の数箇月程度）と長期（道路変化イベ

ントの2年前)の2種類の組み合わせの合計4種類の予測を行うこととした。

機械学習の準備として、工事内容を説明変数に次元圧縮する必要性と方法を整理した。次に、規模と時期の異なる予測クラスごとに教師データとテストデータを生成する方法を提案した。予測は、教師データから予測アルゴリズムによって学習モデルを生成し、テストデータにこのモデルを適用することで行う。機械学習に用いるアルゴリズムには、未知の(教師データと異なる)工事データに対する予測誤差(汎化誤差)を評価するため、誤差に対して特性が異なる線形分類器(ナイーブベイズ分類器, AODE), 非線形分類器(決定木, ランダムフォレスト)を使用した。工事情報を集積する期間(情報収集期間)は3種類(6箇月, 4箇月, 2箇月)とした。

実験の結果、道路開通は供用開始に比べて成績が良いこと、短期予測は長期予測より成績が良いこと、情報収集期間による成績の差には決まったルールが発見できないこと、非線形分類器では過学習により真陽性率が低下することなどの知見を得た。これらをもとに予測の利用フローを作成した。

予測結果の真陽性率は供用開始、長期予測ではおおよそ0.6以下であり、予測の正確さの向上が課題である。データの質・量の充実のほか、説明変数の改良、スパースデータに適したデータマイニングアルゴリズムの適用、工事の順序関係(マルコフ連鎖)を利用する方法を課題への対応案として示した。

第6章では、研究の実用化に必須な事項である、情報の流通に関する取り組みを述べた。情報の流通には、情報の提供者である道路管理者の理解が必要であると考えた。そのため、道路管理者に対してアンケートを実施した。調査の意図は、研究の実用化において、(a)対象とする原典情報の拡張の必要性、(b)(c)道路管理者における情報流通の障害の有無、(d)システムの活用可能性を調査することである。(a)道路変化イベントにおける有用な情報に関する調査では、本研究で扱った原典情報以外に効果的なものは見つからなかった。(b)道路変化イベントの公表への意思に関する調査から、73%の道路管理者が公表の意思を持っていることがわかった。(c)情報の流通に対する問題の有無では、79%の道路管理者から工事データの利用について概ね受け入れられた。(d)本研究で構築したシステムの有用性の評価については、既に庁内システムを除く道路管理者から有用と評価された。

なお、道路管理者への調査は、道路管理者へ道路変化イベントを取り巻く問題点に理解を促すこと、研究の意図を浸透させ、システムの認知を高めることも意図している。

## 7.2 本研究の特徴

### 7.2.1 本研究の独自性

道路変化イベントを把握するための研究は、「1.3 先行研究」に示したようにこれまでも行われてきた。本研究は、その成果を踏まえつつ、以下の点について独自の視点から取り組んだ。

### ① 情報に含まれる時間（時期）の利用

道路変化イベントの予測において次の理由から時間（時期）を利用する手法を開発した。

- (a) 道路建設が複数の工事の積み重ねで行われることを利用するため。
- (b) 道路変化イベントと工事（例えば、地盤改良や基礎工、橋梁下部工などの下層部分の工事、アスファルト舗装や道路付属物、橋梁上部工などの上層部分の工事）との時間差を利用するため。

これによって、工事件数による簡易な道路変化イベントの予測、早期の段階での道路変化イベントの予測（長期予測）を行うことができた。

### ② 統合的で共通して利用できるシステムの構築

道路変化イベントを取り巻く各種の情報は、道路管理、工事発注、広報と自治体内で異なる部署から発出される。これらを統合し、また、自治体が異なっても共通して使えるシステムを構築した。システムの構築によって、研究の成果を分かり易く提示でき、加えて、情報の利用と流通に向けての関係者の理解を助けることができた。

### ③ 情報の流通への取り組み

情報の流通は、社会的・制度的な要素と、技術的な要素の両方に関係する問題である。情報提供の依頼だけでは道路管理者側に解決策と利点が提示されず、対応が進まない。これに対して、本研究では次の方策によって課題の解決を図った。これにより情報の流通を具体的に進めるための基盤を構築することができた。

- (a) 共通的に利用できるシステムを構築して道路管理者の負担を軽減すること。
- (b) 電子的フォーマットを提示し、道路管理者が取り組みやすい方法を提示すること。
- (c) 道路管理者に対しても有用性があることを示すこと、
- (d) 使われ方を示して提供者の意図しない利用法ではないことを道路管理者が確認できることで情報の提供への不安を払しょくすること

## 7.2.2 本研究の有用性

本研究の有用性あるいは価値は次のようにまとめられる。

### ① 原典情報の関連づけ手法の有用性

道路変化イベントの発生までは、道路区域の決定・変更、道路工事、道路の供用へと至るという事業の流れがある。本研究において、道路区域情報、工事情報、供用開始情報、道路開通情報を統合的に扱う手法を考案したことにより、これらの事業の流れの実態を実際の情報によって確認できることとなった。さらに、道路変化イベントの予測に役立つ、時間的な要素が含まれたデータセットが得られるようになった。このデータセットを充実させることで、さらに多くの知見を得ることが期待できる。

## ② 工事情報を用いた予測の有用性

法令に基づく道路変化イベントの公表は実施と同時になっているため、道路変化イベントの発生より前に情報を入手するには理解ある道路管理者からの情報提供に頼らざるを得ない。これに対して、本研究では、公表されている工事情報から道路変化イベントを予測することに取り組んだ。本研究を利用することで、事前に全く道路変化イベントを把握できない地域であっても、問い合わせの対象となる自治体を得ることや具体的な地点についての問い合わせが可能となる。さらに、人によって収集されてきた道路変化イベントを、工事データという異なる方法により検証することができる。

## ③ 統合的システムとフォーマットの有用性

これまでクローリングやスクレイピングと手動による調整を併用してデータを得るために大きな労力が割かれてきた。本研究では、情報の流通を促進するために、システムの可視化機能を用いて道路管理者に対して情報の流通の利点を示し、また、共通して利用できるデータテーブル構造とシステムを提案した。情報の流通を進めるには、具体的なメリットやイメージを情報利用者・情報提供者（道路管理者）に提示することは有用である。

## 7.3 課題

本研究の取り組みにおける課題を次に示す。

### ① データの充実と精度の向上

関連づけた原典情報から事業の実態の調査解析、工事からの道路変化イベントの予測ができることを本研究で示した。その正確さは、データの量と質に依存する。データの内容のうち、調査解析と予測の全てに関わるものは、場所から地点・路線への変換である。この変換ができない場合には、収集した情報が利用されないものとなる。道路工事の場所は、人の居住地ではないため、上り車線に面した場所の地名と下り車線に面した場所の地名が異なる場合がある。また、地名の登録が通称で行われている場合などがある。現在は、携帯端末やブラウザから緯度経度を用いて場所を入力するツールが利用できる環境になっており、このような入力方法の改良により正確な場所の取得が望まれる。また、正確な地点と路線が確定できれば、統合的システムの GIS 上の道路の該当する位置にアイコンを表示させることができる。これはシステムに対する道路管理者からの要望の一つである。

データの内容のうち、工事内容を用いた予測に関係するものには工事内容がある。原典情報に含まれる工事概要や工法・型式に加えて、仕様書の詳細から正確な工事内容の種類を把握すること、さらには工事量が利用できるようになれば、工事内容を量的尺度で正しく反映させることができる。これによって工事内容を用いる予測の精度の向上が考えられる。

### ② 予測の正確さの向上

予測の正確さの向上には、データの質・量の充実のほか、説明変数の改良（工事件名、工事請負金額の利用）、スパースデータに適したデータマイニングアルゴリズムの適用（階層型クラスタ分析、アソシエーション分析）、工事の順序関係（マルコフ連鎖）を利用する方法が考えられる。これらの適用による評価、改善が課題である。

### ③ 情報の流通

道路変化イベントの把握を行うためには、原典情報を網羅的に入手することが必要である。取り扱った原典情報の内容は法令により公開が求められている公示の情報、工事情報と、地方自治体が広く知って欲しいと考えている道路開通情報であり、法令の上での障壁は無い。しかしながら、利用しやすい電子的な形態での情報の流通という点において、カナダ、フランス、韓国など幾つかの国では公開情報の電子化について、法律による強制がなされているのに対して、日本ではオープンデータ、オープンガバメントの動きはあるものの努力義務的なものとなっている。法律の改正などによって電子的な情報が流通することになれば、システムに容易にデータを投入することができる。しかし、既に多くの自治体では作成された原典情報は電子化されており、法律の改正等を待つことなく、業務の改善で対応できるのではないかと考えている。そのための方策として、情報流通の有用性を可視化するためのシステムの構築と、情報流通をし易くするためのデータフォーマットの作成を行ったが、さらに、情報を流通するために関係機関との調整が課題である。

## 7.4 今後の展望

### 7.4.1 詳細調査との連動

本研究では道路変化イベントを把握することを課題とした。その目的は、把握した道路変化イベントによって調査と対象とすべき場所を絞り込み、さらに詳細を調査するという利用を想定している。このため、道路変化イベントの把握から詳細調査まで連動させられれば、より利便性は高まる。詳細な情報としては、道路変化イベントの発生前後であれば、関係する工事の電子納品された道路平面図があり、やや遅れて更新されることの多い道路台帳附図がある。

工事の完成図面が電子納品されていれば図 7.4-1 のように、道路変化イベントの予測時に、予測のもととなった工事の件名やID等を出力し、その情報から完成図面を入手するシステムを考えることができる。現在は電子納品の公開がなされていないが、仮に公開されたとしても、地図の更新に必要な工事を選び出すことは簡単では無く、本研究によってそれを助けることができる。

また、システムは標準的なWEB マッピングプロトコルである WMS (Web Map Service) に対応しており、このプロトコルで道路台帳附図やあるいは、照合したい地図製作者の地図が提供されれば、その上へ原典情報を表示できる。本研究のシステムをこれらと連携させることにより、情報公開請求等により発生している情報の提供者と利用者の負担をさらに減少できる。

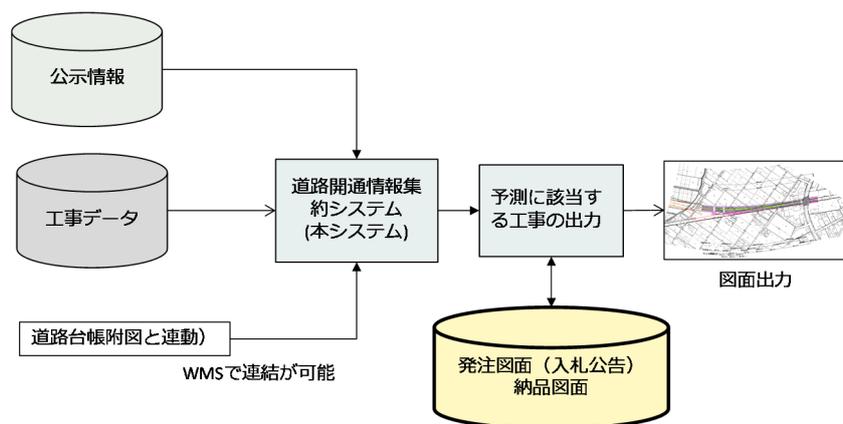


図 7.4-1 電子納品、道路台帳附図との連携システム

#### 7.4.2 道路変化イベント以外の予測への拡張

工事情報をマルコフ連鎖の状態空間と考えることで、道路変化イベントだけでなく、ある工事群から次に生ずる工事群の条件付き確率も求めることができる。工事の予測によって人員資材の効率的な運用や確保に役立てられる可能性がある。また、道路以外の社会インフラへの拡張が考えられる。

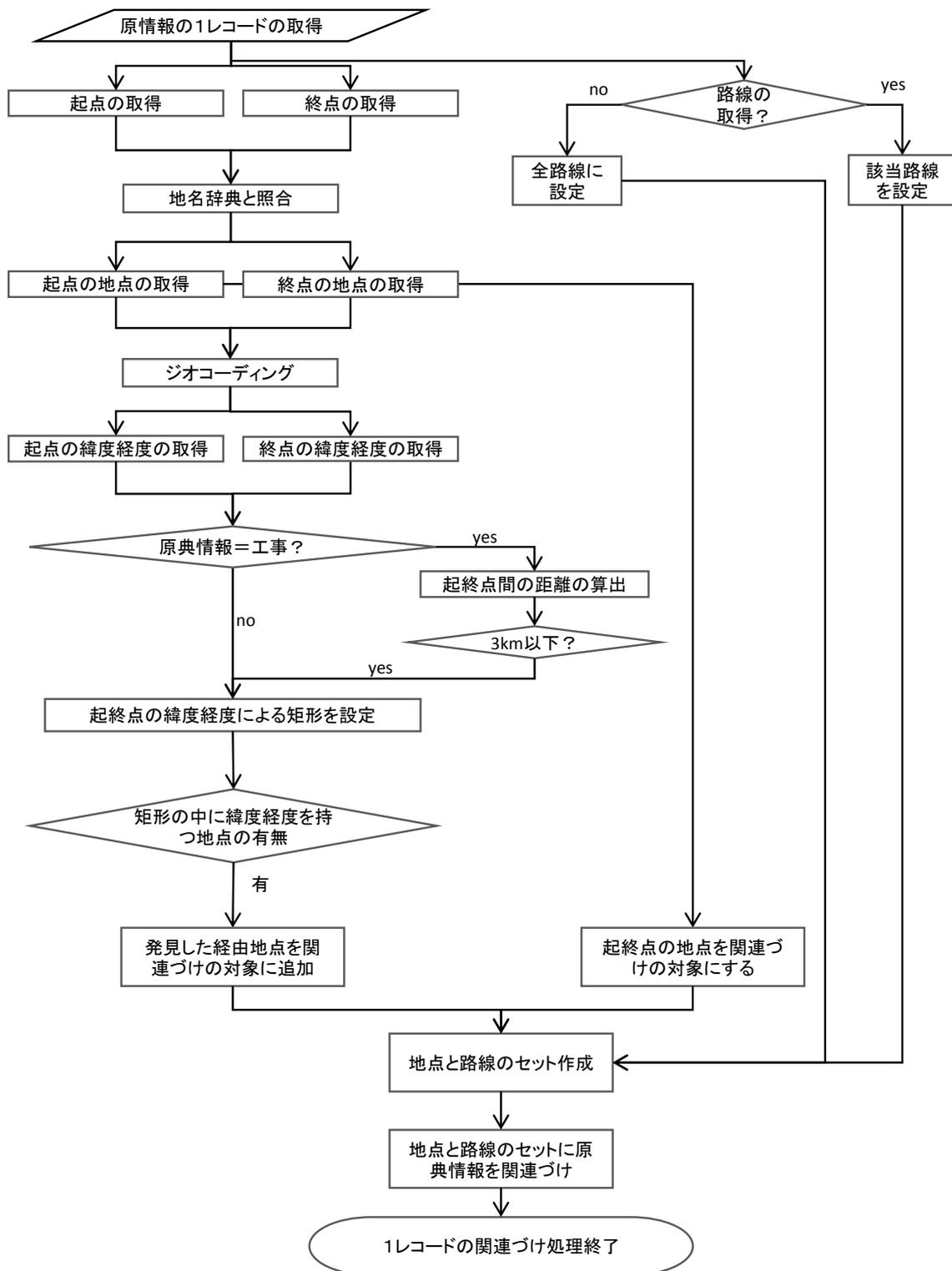
#### 7.5 おわりに

工事データから道路変化イベントが把握を実用化できれば、道路管理者、道路利用者、地図データ等の制作者のそれぞれにとって次のメリットが得られる。道路管理者にとっては、供用した道路について、道路利用者・地図データ等の制作者等への情報提供を効率的に行うことができるほか、情報公開請求への対応が軽減できる。道路利用者にとっては、利用できる道路の情報が、管理者のサイト、地図やカーナビゲーションを通じてこれまでより早く入手できるため、効率的な道路の利用が可能となる。地図データ等の制作者にとっては、全国の道路管理者への問い合わせを効率化できる。

以上に述べてきた道路変化イベントの把握によって、道路管理者が生成する情報が活用され、地図調製者らの情報収集が効率化され、道路利用者による最適な道路利用が促進されて、社会的な便益が一層向上することを願って本論文の結びとする。

# 付録

## 付録1 場所から地点・路線への関連づけ

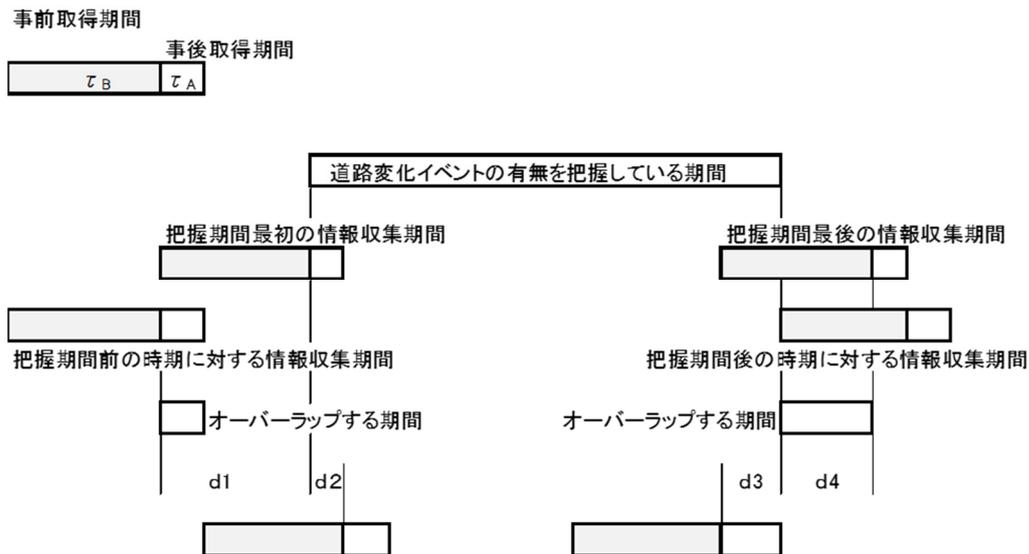


付録2 次元圧縮の例

| 次元圧縮前の工事内容   | 圧縮後の説明変数    |
|--|-------------|
| ニューマチックケーソン工<br>オープンケーソン工<br>カルバート工<br>フーチング工<br>鋼矢板・鋼杭工<br>鋼製セル工<br>敷砂工<br>浮体工<br>基礎工<br>置換工法<br>プレローディング工法<br>表層混合処理工法<br>バーチカルドレーン工法<br>サンドコンパクション工法<br>ロッドコンパクション工法<br>バイブロフロテーション工法<br>石灰パイル工法<br>深層混合処理工法<br>DC工（動圧密工法）<br>サーチャージ工法<br>軽量土工法<br>薬液注入工法<br>ボーリンググラウト工<br>土留め・仮締切工<br>鋼管矢板基礎工<br>既製杭工<br>場所打杭工<br>深礎杭工<br>地中連続壁工 | 基礎工事        |
| NATM, 矢板トンネル, TBM<br>開放型シールドトンネル<br>密閉型シールドトンネル<br>開削トンネル・半地下<br>沈埋トンネル<br>非開削特殊工<br>セメントコンクリート舗装工事  | トンネル・セメント舗装 |
| 歩行者舗装<br>CAB<br>情報BOX<br>電線共同溝   | 歩道・通信管路     |

### 付録3 負例の整理方法

負例は、道路変化イベントが発生していない期間と地点の工事内容を収集し作成する。道路変化イベントの有無を把握している期間に情報収集期間が一部かかるケースは4つある。一部かかっても道路変化イベントに関係する可能性が0.5未満であれば、その期間と地点の情報を負例に用いる。以下は、4ケースについて詳細を示すものである。



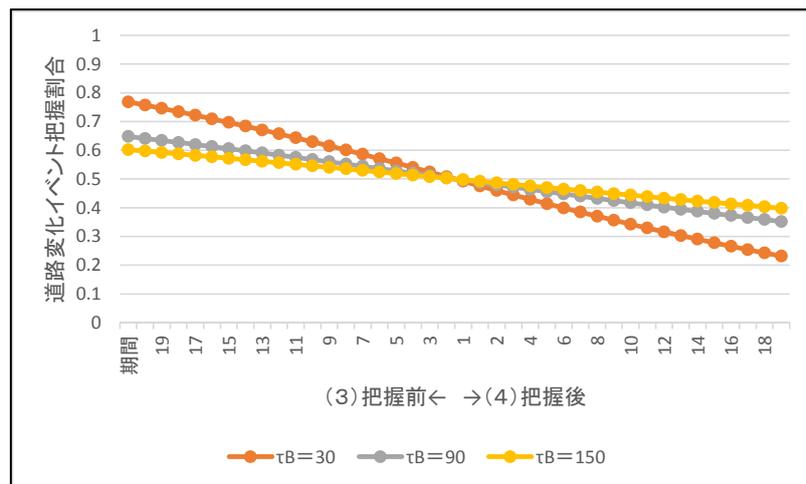
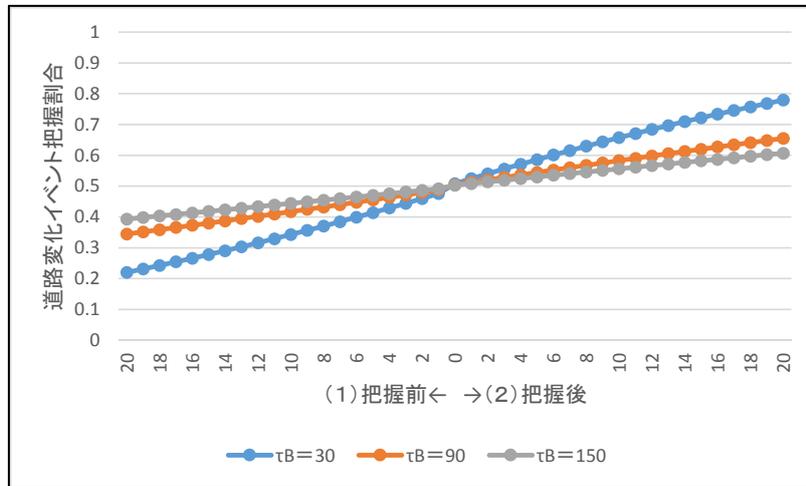
道路変化イベントを把握している期間に対して情報収集期間のとり得る4ケース

上図で $\tau A$ 、 $\tau B$ 、 $d1 \sim d4$ が与えられたとき、情報収集期間のうち道路変化イベントの有無を把握している期間にかかる日数は下記で求められる。

#### 情報収集期間のうち道路変化イベントを把握している日数

| 場合                                       | 把握日数の算出式  |
|--|---|
| (1) 調査日が道路変化イベント把握期間より $d1$ 日前のとき        | $(\tau A + \tau B - d1) \times (\tau A + \tau B - d1 + 1) / 2$                                    |
| (2) 調査日が道路変化イベント把握期間の中(期間開始の $d2$ 日後)のとき | $(\tau A + \tau B + d2 + 2) * (\tau A + \tau B + 1 - d2) / 2 + (\tau A + \tau B + 1) \times (d2)$ |
| (3) 調査日が道路変化イベント把握期間の中(期間完了の $d3$ 日前)のとき | $(\tau A + \tau B + 1) \times (d3) + (\tau A + \tau B + d3) * (\tau A + \tau B + 1 - d3) / 2$     |
| (4) 調査日が道路変化イベント把握期間より $d4$ 日後のとき        | $(\tau A + \tau B - d4 + 1) \times (\tau A + \tau B - d4 + 2) / 2$                                |

把握している日数の全ての調査対象期間の日数に対する割合を  $\tau B$  別にグラフ化したものを下図に示す。グラフにおいて把握割合が 0.5 以上の期間のものを負例に採用する。



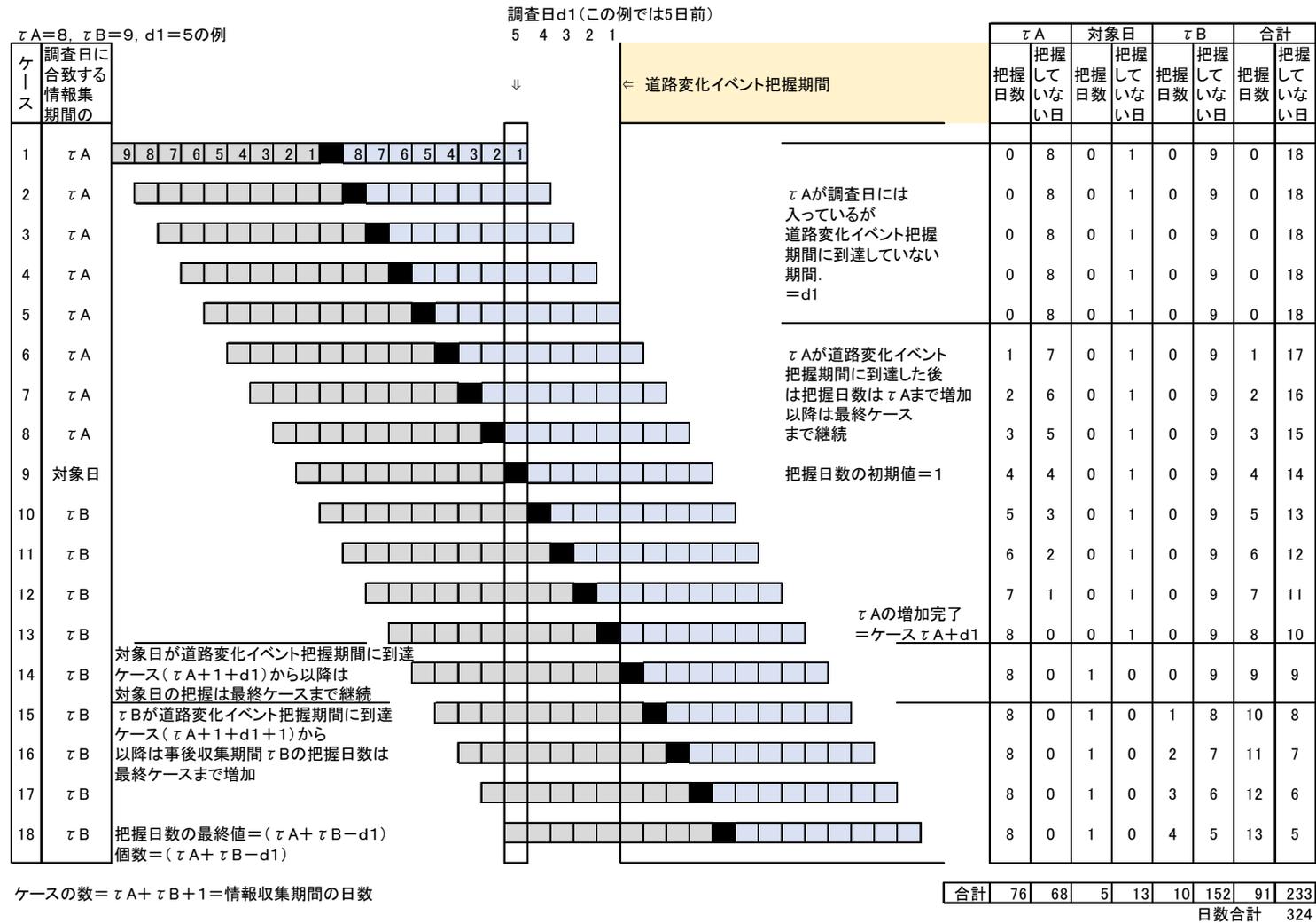
データに占める道路変化イベント把握の割合

次の 4 ケースは、下の情報収集期間に対する具体的な例である。

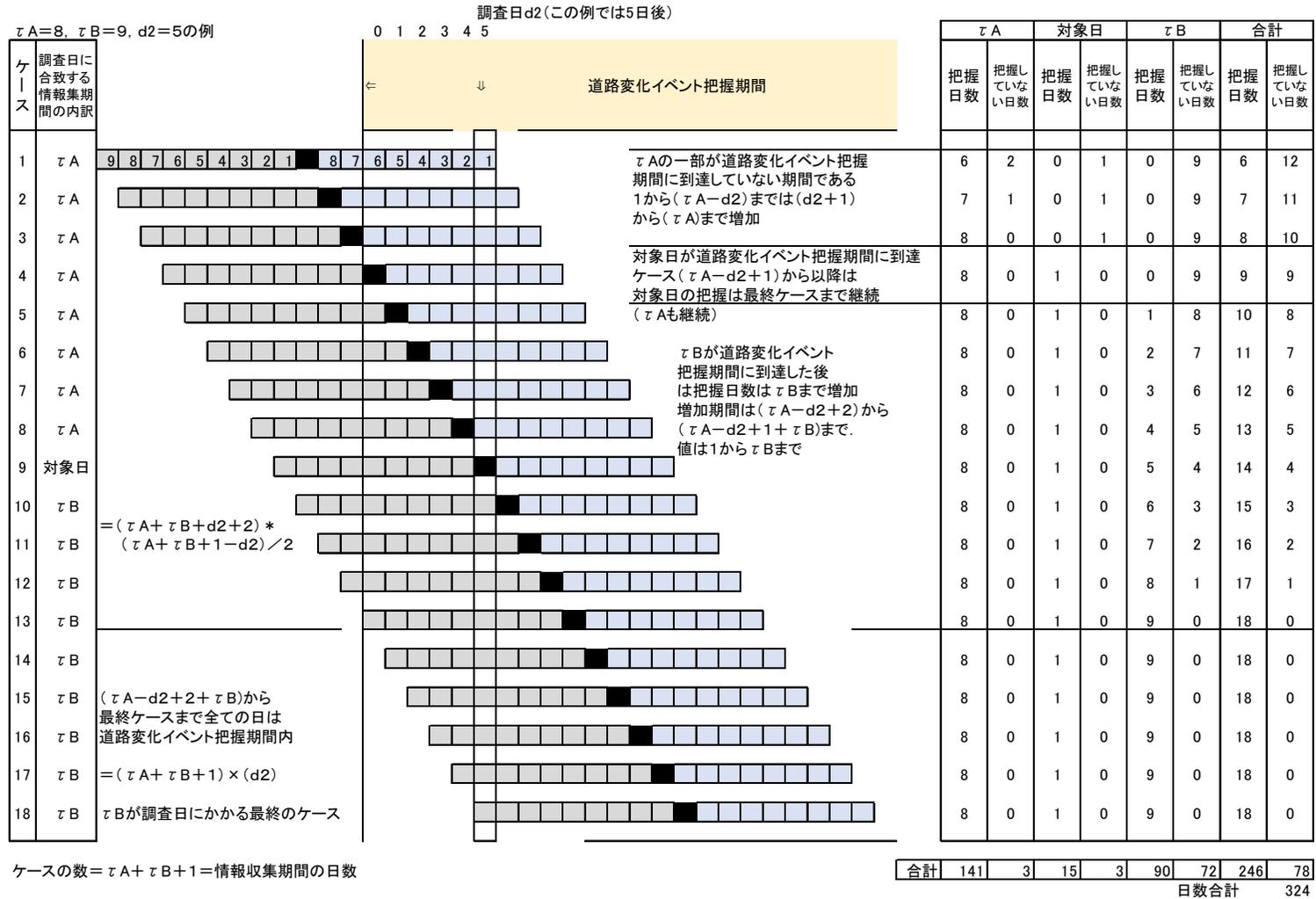


具体例のための情報収集期間のモデル

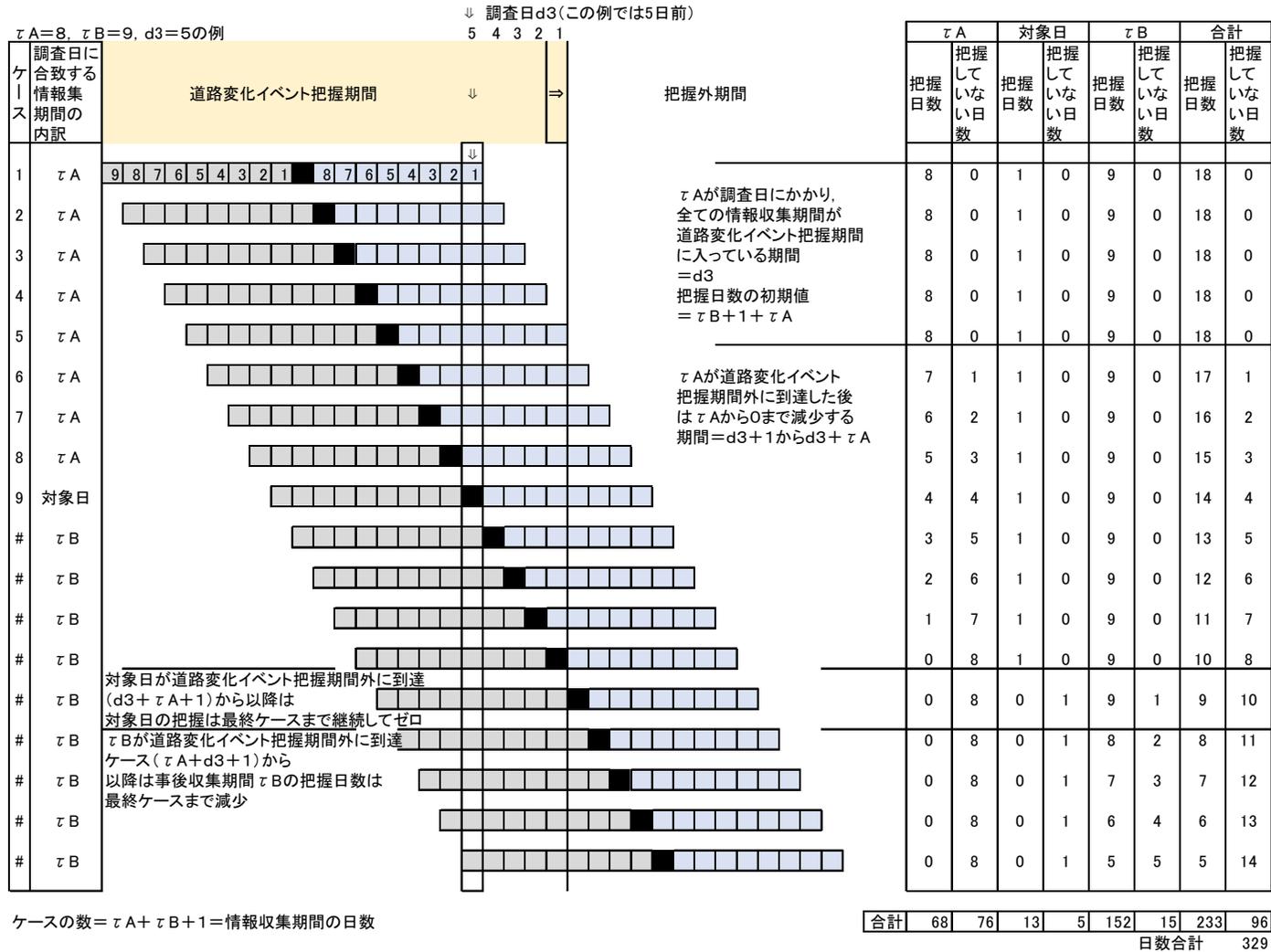
負例を取得する範囲を決めるためのデータ数 (1) 調査日が道路変化イベント把握期間より前のとき



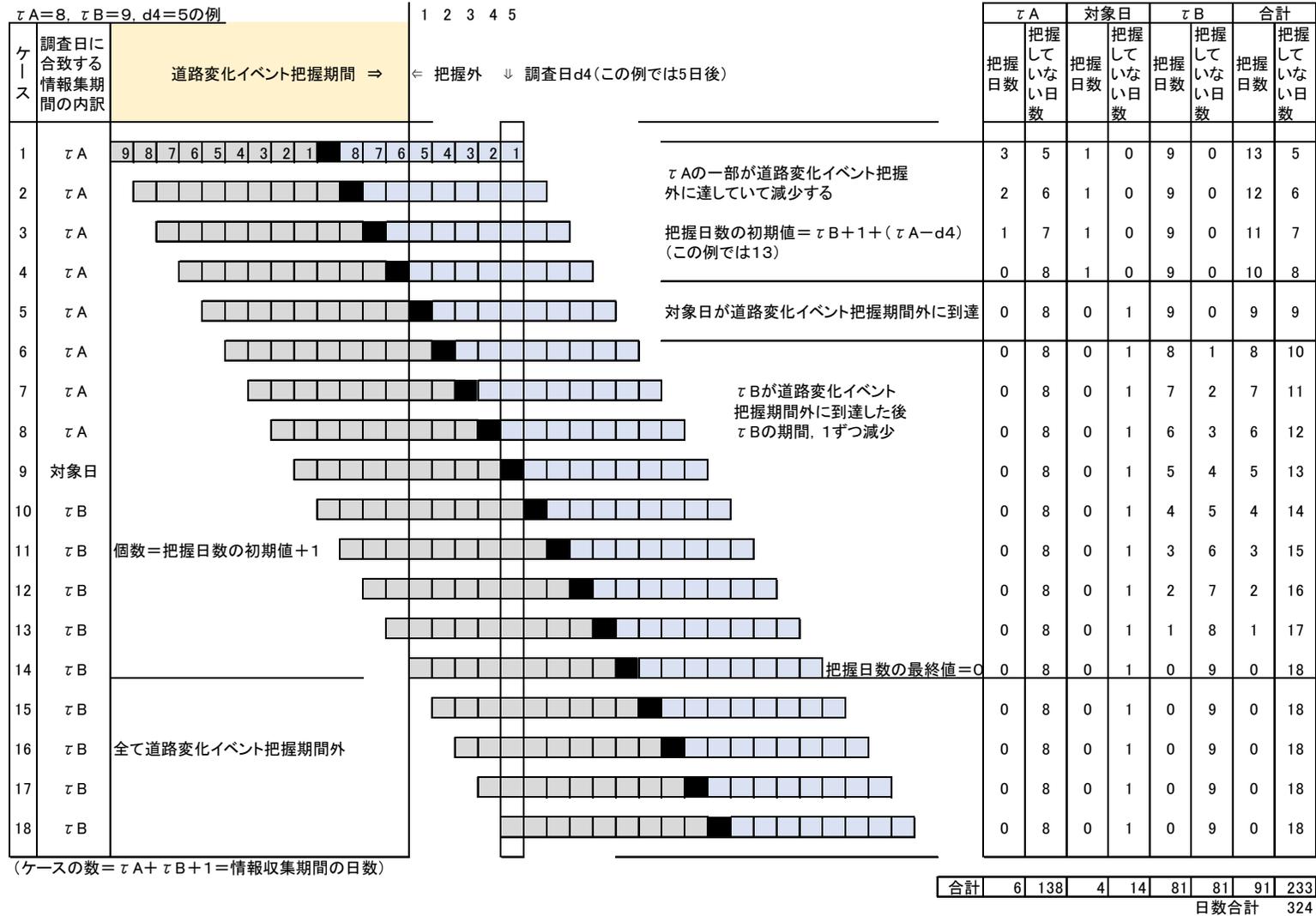
負例を取得する範囲を決めるためのデータ数 (2) 調査日が道路変化イベント把握期間の中 (期間開始直後)



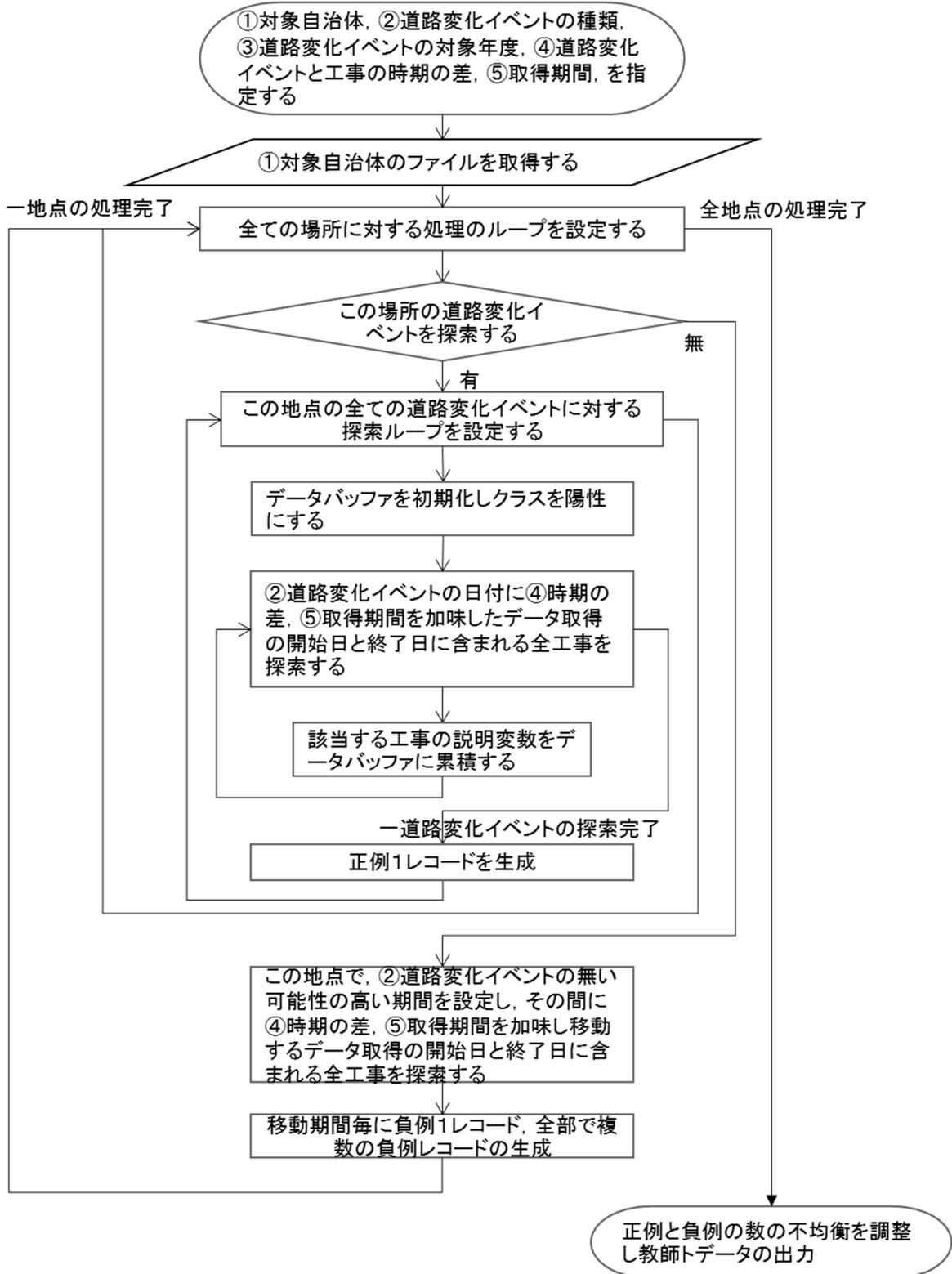
負例を取得する範囲を決めるためのデータ数 (3) 調査日が道路変化イベント把握期間の中 (期間完了直前)



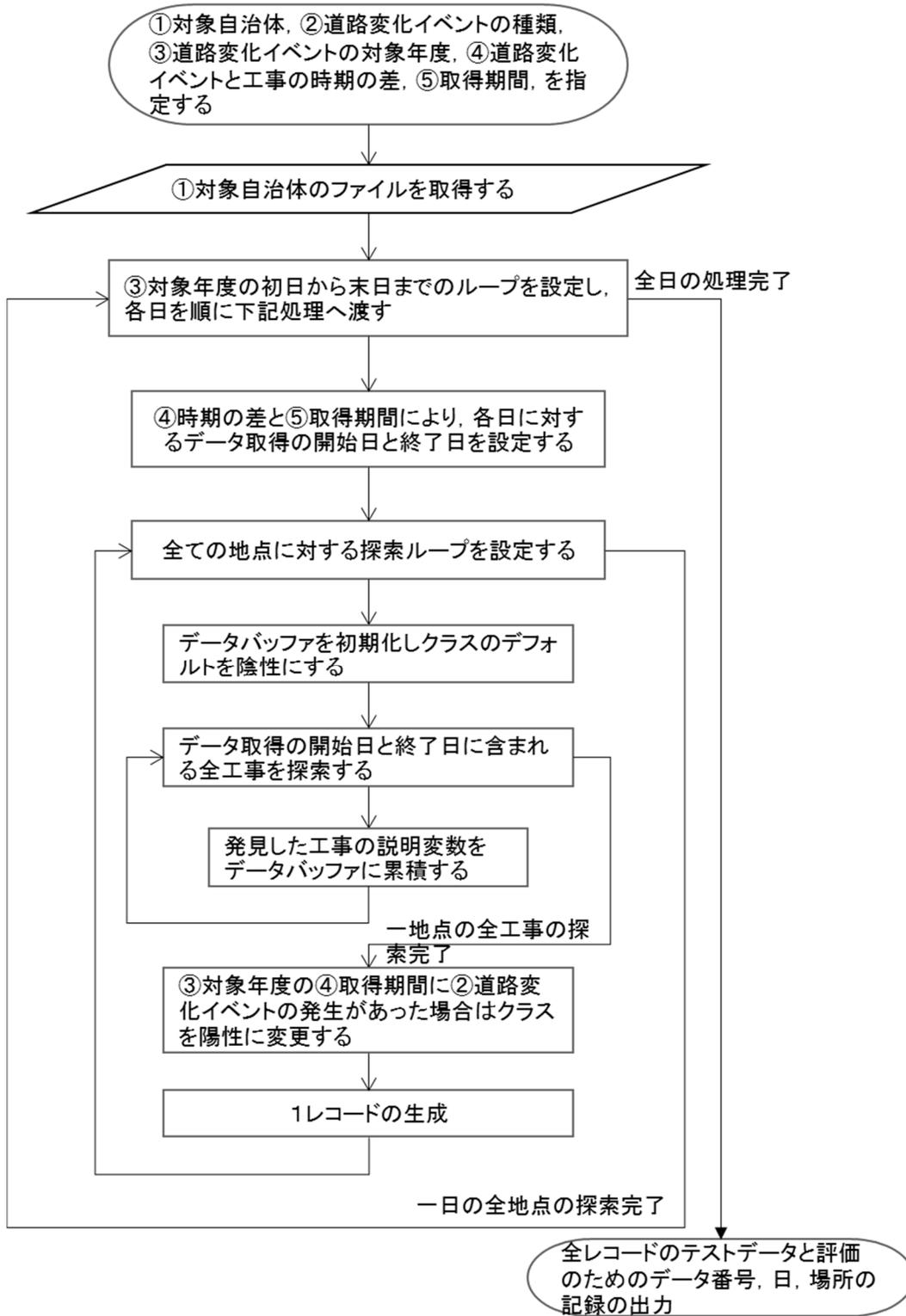
負例を取得する範囲を決めるためのデータ数 (4) 調査日が道路変化イベント把握期間より後のとき



付録4 教師データの生成



付録5 テストデータの生成



## 付録6 データ関係 URL

### 1. 道路法に基づく告示

- 1) 島根県総務部総務課：県報，<http://www.pref.shimane.lg.jp/admin/pref/info/kenpou/>，2015.8.1.アクセス。
- 2) 高知県道路課：平成 27 年度道路法関連告示一覧，<http://www.pref.kochi.lg.jp/soshiki/170701/2015041000092.html>，公開日 2015 年 12 月 22 日，平成 26 年度道路法関連告示一覧，<http://www.pref.kochi.lg.jp/soshiki/170701/kokuji-itiran26.html>，公開日 2015 年 3 月 31 日，平成 25 年度道路法関連告示一覧，<http://www.pref.kochi.lg.jp/soshiki/170701/kokuji-itiran25.html>，公開日 2014 年 4 月 9 日，平成 24 年度道路法関連告示一覧，<http://www.pref.kochi.lg.jp/soshiki/170701/kokuji-itiran24.html>，公開日 2013 年 4 月 29 日，平成 23 年度道路法関連告示一覧，<http://www.pref.kochi.lg.jp/soshiki/170701/2015040100014.html>，公開日 2015 年 5 月 19 日，2015.8.1.アクセス。
- 3) 高知県法務課：高知県公報発行月別閲覧，<http://www.reikisyuutou.pref.kochi.lg.jp/kakokouhou/tsukisentaku.htm>，2015.8.1.アクセス
- 4) 新潟市総務部：行政文書情報の提供・公開窓口（公示文書検索），<https://info-navi.city.niigata.lg.jp/iDiscPub/noticeSearch.do>，2015.8.15.アクセス。
- 5) 金沢市総務局文書法制課：金沢市公報，<http://www4.city.kanazawa.lg.jp/13004/kouhou/>，2015.8.16.アクセス。

### 2. 道路開通情報

- 1) 新潟市：道路開通情報一覧表，<http://www.city.niigata.lg.jp/kurashi/doro/road/dorokaituinfo/index.html>，2010.3.19.更新，2012.8.15.アクセス。（第 4 章）
- 2) 島根県道路建設課：道路開通情報（地区別にリンク），[http://www.pref.shimane.lg.jp/infra/infos/douro\\_kaitu/kaitu/](http://www.pref.shimane.lg.jp/infra/infos/douro_kaitu/kaitu/)，2015.11.3.アクセス（第 3 章，第 5 章）
- 3) 高知県道路課：高知県管理道路の開通情報，<http://www.pref.kochi.lg.jp/soshiki/170701/kaituujyohou.html>，公開日 2014 年 1 月 14 日，（国道 439 号）<http://www.pref.kochi.lg.jp/soshiki/170107/kokudou-439kaitu.html>，公開日 2013 年 2 月 8 日，川登バイパス，<http://www.pref.kochi.lg.jp/soshiki/170701/121229kawanoboribaipas.html>（第 3 章，第 5 章）
- 4) 新潟市道路計画課：道路開通情報一覧表，<https://www.city.niigata.lg.jp/kurashi/doro/road/doroseibi/dorokaituinfo/index.html>，2015.11.26.アクセス。（第 3 章，第 5 章）
- 5) 愛媛県土木部道路建設課：主要道路開通情報，<http://www.pref.ehime.jp/h40400/5744/kaituu/kaituu.html>，2010.3.19.更新，2013.9.29.アクセス。（第 4 章），2015.7.14.更新，2015.11.25.アクセス。（第 3 章，第 5 章）
- 6) 埼玉県道路街路課：道路開通情報，<http://www.pref.saitama.lg.jp/site/kaituszyouhou/>，2012.

#### 10.7.アクセス。(第4章)

- 7) 埼玉県道路政策課：道路開通情報（国道），<http://www.pref.saitama.lg.jp/site/dousei1006/905-20091224-51.html>，2010.3.19.更新，2010.8.15.アクセス。(第4章)
- 8) 埼玉県道路政策課：道路開通情報（県道），<http://www.pref.saitama.lg.jp/site/dousei1006/dousei035.html>，2012.2.3.更新，2012.8.15.アクセス。(第4章)
- 9) 埼玉県道路政策課：道路開通情報（その他），<http://www.pref.saitama.lg.jp/site/dousei1006/dousei059.html>，2012.1.17.更新，2012.8.15.アクセス。(第4章)

### 3. 工事情報

- 1) 島根県：島根県電子調達共同利用システム入札情報サービス，[https://choutatsuweb.pref.shimane.lg.jp/SMN/PPI\\_P/](https://choutatsuweb.pref.shimane.lg.jp/SMN/PPI_P/)，2015.8.13.アクセス。
- 2) 高知県土木部建設管理課：高知県入札情報システム，<http://www.efftis.jp/39000/ebia/contents/>，2015.8.17.アクセス。
- 3) 新潟市財務部契約課：入札・契約結果検索（建設工事），<http://www.n-jouhou.city.niigata.lg.jp/epco/servlet/p?shubetsu=1&job=NyusatsuKeiyakuKojiSearch>，2015.8.14.アクセス。
- 4) 金沢市総務局監理課：入札情報システム（PPI），<https://www.ep-bis.supercals.jp/ebidPPIPublish/EjPPIj>，2015.8.16.アクセス。
- 5) えひめ電子入札共同システム：入札情報公開システム，<https://www.ebid-ppi.pref.ehime.jp/ebidPPIPublish/EjPPIj>，2012.12.1 アクセス。
- 6) 埼玉県：入札情報公開システム，<https://ebidjk2.ebid2.pref.saitama.lg.jp/koukai/do/KF000ShoAction>，2013.12.1.アクセス。
- 7) 一般財団法人日本建設情報総合センター：JACICNET コリンズ一般検索，<http://www2.jacicnet.jacic.or.jp/cgi-bin/CIKojiSrchFrame.cgi>，2015.8.18.アクセス。
- 8) 一般財団法人日本建設情報総合センター：工事实績情報公開，<http://www.jacic.or.jp/jacic-hp/index.php>，2015.8.18.アクセス。

## 発表論文リスト

- 小林亘,柴崎亮介,関本義秀：工事实績情報を用いた道路供用の予測に関する研究, 土木学会  
論文集 F3(土木情報学), 68(2),150-161 (2012)
- Kobayashi,W., Shibasaki,R. & Sekimoto, Y. (2013). PREDICTION FOR LOCATIONS OF CHANGES  
IN ROADS FROM CONSTRUCTION-RECORDS, International Conference on Civil and  
Building Engineering Informatics (ICCBEI 2013), 133-140.,2013
- Kobayashi,W., Shibasaki,R. & Sekimoto, Y. (2015). GENERALIZATION ABILITY IN PREDICTING  
ROAD OPENINGS FROM CONSTRUCTION DATA, International Conference on Civil and  
Building Engineering Informatics (ICCBEI 2015), 71(1-8),.2015
- 小林亘,柴崎亮介,関本義秀：教師あり機械学習による工事データからの道路開通予測, 人工知能学  
会第 29 回全国大会論文集 1H3-1(1-4), 2015

## 謝辞

この研究に取り組むきっかけは、国土交通省から2011年に財団法人日本建設情報総合センターに出向し、東京大学空間情報科学研究センター次世代社会基盤情報寄付研究部門に係ることとなったことに始まる。この寄付研究部門は、(1) 地理空間情報の流通・高次利用に関する技術基盤の開発、(2) 産官学が円滑に地理空間情報をやり取りできるコミュニティの醸成を目標に掲げており、シンポジウムや勉強会の活動の中で、道路変化イベントの把握が十分に、また、効率的に行われていないことを強く認識した。出向先の日本建設情報総合センターでは、入札契約が適切に行われるよう、工事实績情報を収集、保管、管理していたことから、関係部署の協力を得て、これを道路変化イベントの把握に利用することを考えて試行錯誤を始めた。財団法人日本建設情報総合センターの坪香理事を始め、関係部署の方々には研究についてご理解をいただき深く感謝している。

東京大学空間情報科学研究センター寄付研究部門を率いておられた柴崎亮介教授、関本義秀准教授には、研究をまとめるために重要なお助言を多く頂いた。柴崎教授には、研究に取り組んだ初期の段階で機械学習に関するご示唆をいただいた。2011年当時は今ほど機械学習や人工知能が広く知られていなかったが、2015年の現在では一般向けの雑誌等においてもこれらが取り上げられており、柴崎教授のご慧眼に感服するとともに、ご示唆に深く感謝申し上げる。

関本准教授には、研究成果を整理していく過程で、節目節目に的確なお助言をいただいた。入手したデータをクレンジングして統計的な情報や予測情報を取り出すためには、何段階もの工程が必要であったので、いろいろなプログラムを作成していろいろな処理結果を得ることによって隘路にはまり込みがちになる中、関本准教授のお助言が明るい光となり、出口へ導いていただいた。特に処理系のシステム化は、処理作業が大幅に効率化されると同時に、多くの関係者に研究をご理解いただくうえで極めて効果的であり、そのために貴重なご指導をいただき深甚な感謝を申し上げる。

中條覚特任准教授にはこの研究テーマに関連して先行して研究された学位論文とともに適切なお助言、また、学位取得にあたり多くのことをご教示いただいた。

大口敬教授、堀田昌英教授には、研究論文をまとめるための筋道について大変有難いご意見をいただいた。布施孝志准教授にはデータの特性に適したアルゴリズムの改善方策について多大なお示唆をいただいた。

財団法人日本デジタル道路地図協会からは本研究に関連して研究助成とともに、泉理事長、矢口前理事から励ましのお言葉を、また、研究開発部長土居原氏、企画調査部市川氏、情報管理部水上氏をはじめ皆様から実際の道路変化イベントの情報収集における貴重なご意見や参考となるデータをいただいた。多くの地方自治体の方には意見聴取にご協力

をいただき、特に、本研究に利用させていただいたデータを公開されている自治体にはデータの利用を快くお認め頂いた。皆様に感謝申し上げます。

2006年から2008年まで在籍した国土技術政策総合研究所高度情報化研究センターでは、2007年に成立した地理空間情報活用推進基本法に則り、地理空間情報を相互に利用できる仕組みとして地理空間情報プラットフォームを確立するため国土地理院、東京大学空間情報科学研究センターとの共同研究体制を確立し、空間情報連携仕様の作成、地理空間情報プラットフォームの試作版の構築を行ったが、検討の過程において民間企業から行政機関等に対して様々な情報の提供への要望があるものの、用途や形式が具体化されていないことが情報の流通の障壁になっていると感じたこともこの研究を取りまとめるにあたって幾度となく思い返されたことである。また、2011年に設立された社会基盤情報流通推進協議会に設立当初から関わらせていただき、情報を流通させるには法制度、技術、リテラシー、関係者の理解など様々な課題をクリアする必要があることを学ばせていただいた。これらの関係者の皆様に感謝申し上げます。