

(60) 工事实績情報を用いた道路開通の予測に関する提案

A Proposal of Prediction for Opening roads from Construction-Records

小林 亘¹

Kobayashi Wataru

抄録：新たに開通した道路では、開通情報に基づき道路地図やカーナビゲーションの地図データの更新が行われることによって、道路が一層有効利用され、移動経路や移動時間が短縮される。しかし、開通情報の収集には大きな労力と時間を要している。これを改善するため、本研究では公共工事の品質確保に利用されている工事实績情報を用いた道路の開通予測を提案する。具体的には工事实績情報の工種等を空間軸による集約、時間軸による分割を行って説明変数とし、教師あり機械学習を適用することによって道路の開通を予測しようとするものである。埼玉県の道路開通にて検証し、提案が有効であることを示した。

キーワード： 工事实績情報、道路地図、機械学習、決定木、ナイーブベイズ分類器

Keywords : Construction-Records, Road update Information, Machine Learning, Decision Tree, Naive Bayes classifier

1. はじめに

近年、大量のデータを活用して価値のある情報を生み出そうとする動きが従来にも増して活発化している。公共工事の分野においてもデータの蓄積は進んでおり、その収集、蓄積に要した費用、時間、労力は少なくなることから、これを従来の用途を超えて、より広い範囲で役立てることができれば、社会により大きく貢献できる可能性がある。本研究は、公共工事に関するデータのうち「工事实績情報」を用いた道路開通の予測について提案を行うものである。

1. 工事实績情報の現況

公共工事の発注機関では、企業や技術者の工事实績や契約状況等を確認するため、公共工事の工事实績に関する情報（「工事实績情報」）を収集、蓄積、提供する「工事实績情報システム」が利用されている。1993年、「中央建設業審議会」において工事实績情報のデータベース整備の必要性が建議され、1994年から工事实績情報システムの運用が開始されて以来、工事实績情報の蓄積は順調に進んでおり、2011年3月における登録件数は約345万件、利用している発注機関は、国土交通省、47都道府県、19政令市、791市区町村、高速道路株式会社など合計1030となっている¹⁾。これらの発注機関では、契約に使用する土木工事共通仕様書²⁾において、「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」を受注者に求めているものが一般的である。国土交通省をはじめとして利用機関が多いこと、対象工事が請負金額500万円以上であり比較的小規模な公共工事を含むことから考えて、工事实績情報システムが扱っている公共工事の割合は高いと考えられる。これを確認するため、統計法に基づく基幹統計である「建設工事施工統計調査」³⁾との比較を行った。平成18年度から平成22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請負金額（建設工事施工統計調査が対象としていない発注機関を除く）の比率は、107.4%～98.7%であった。このことから、公共工事において工事实績情報システムが扱っている割合は極めて高いと言え、工事实績情報を用いた道路開通の予測が可能となれば、その有用性は高い。

2. 道路の開通情報の把握と課題の定式化

道路法第十八条では、道路の開通に代表される道路の供用について、「・・・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、かつ、これを表示した図面を道路管理者の事務所に置いて一般の縦覧に供しなければならない。・・・」と定めている。公示に基づいて開通の情報を収集するには、全国の道路管理者の公示を収集し、そこから道路の開通に係るものを抽出することが必要となる。この方法に対し、本研究で提案する工事实績情報からの開通の予測が可能であれば、公示に先立って、より

1：正会員 工修（一財）日本建設情報総合センター 建設情報研究所

（〒107-8416 東京都港区赤坂 7-10-20, Tel :03-3505-0436, E-mail : kobayashi_w@jacic.or.jp）

計画的に道路地図等の更新の準備を進めることができ、また、公示の収集に比べて容易に全国を網羅する情報を収集できる可能性がある。

この課題を、工事実績情報に含まれるデータを説明変数(x)、道路の開通を目的変数(y)として、適切な関数 $y=f(x)$ を発見することと考え、次にその具体的な方法を検討した。

3. 説明変数

工事実績情報には、(1)件名、(2)請負金額、(3)工期、(4)発注機関、(5)契約形態、(6)請負者、(7)工事の分野、(8)建設業許可業種、(9)入札参加資格区分、(10)工種_工法_型式、(11)施工場所、(12)技術者などが登録されている⁴⁾。

工事実績情報は多数の項目を有するが、道路開通を直接に示す項目はない。このため、説明変数に用いる項目を選定する必要がある。ここでは、項目「工事の分野」を用いて道路関係工事を抽出した後、その中から項目「工種_工法_型式」（以下、「工種等」）を説明変数として開通の予測を試みることにした。項目「工種等」は、登録にあたり選択肢が提示されているうえ、全ての工事実績情報に登録されている。このような安定した品質を持つ技術的な項目であることがその理由である。選択肢の数は、工種において 77、工法_型式において 137 である⁴⁾。ここでは、工種、工法の施工上の目的や体系⁵⁾⁶⁾⁷⁾、実際に各工事実績情報に登録されているデータの共起状況、クラスター分析結果を総合して 18 の説明変数に集約した。図-1 に検証に用いた工事実績情報の集約前の工種等の共起状況を示す (KH-coder を使用)。整理後に説明変数として採用した工種等を表-1 に示す。

4. 空間軸による集約

工事 1 件につき、1 件の工事実績情報が基本的に対応している。このため、工事を単位として予測するならば、各工事実績情報を独立して処理できる。しかしながら、道路の事業計画等には、複数年度で道路建設事業が実施されることが示されているものが少なくなく、さらに、事業費と 1 件の工事金額、1 件の工事の施工内容から考えて、道路の建設は複数の工事によって行われることを否定できない。このため、工事単位では予測に必要な説明変数を収集できない可能性があると考えた。これに対応するため、本研究においては工事が行われる「場所」を単位として工種等を集約することとした。これによって、道路の開通に関する工事が複数である場合にも単独である場合にも説明変数を確保することができる。

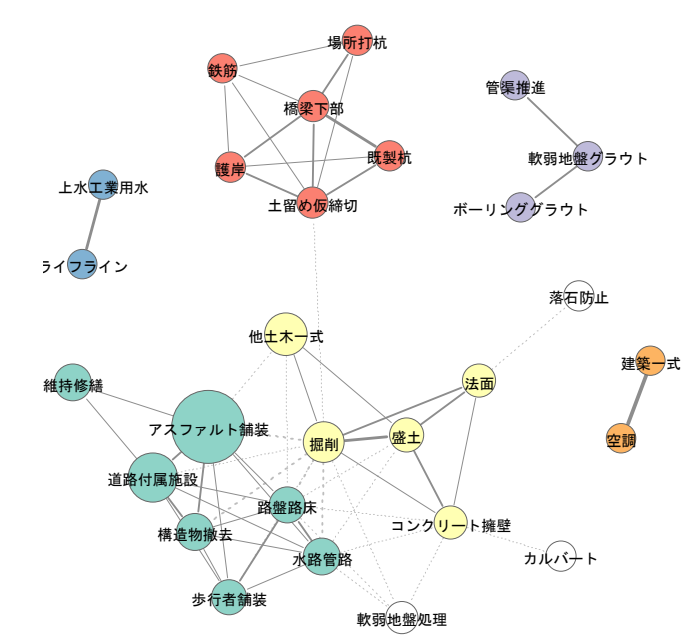


図-1 工種等の共起ネットワーク

表-1 説明変数

工種等	内容
Foundation	フーチング、ケーソン、グラウト等
Earth	掘削、盛土などの土工
PipeLine	水路管路、カルバート
RoadBed	路床路盤
Pavement	アスファルト舗装、コンクリート舗装
Wall	コンクリート擁壁、法面
RoadEquipment	道路付属施設
Works	その他土木一式
Bridge	橋梁の上部、下部、橋脚、床版等
CCBOX	情報ボックス、キャブ、電線共同溝等
PedestrianPV	歩行者舗装
Tunnel	トンネル、管渠推進
Construction	その他とび土工コンクリート
WaterProof	防水
Environment	砂防、地滑り、落石防止、護岸等
Facilities	建築、各種施設、設備
Removal	撤去
Maintenance	維持修繕

説明変数を工事場所毎に集約するため、各工事実績情報に含まれる工事場所を正確に突合することが重要となる。工事実績情報には、起点施工場所、終点施工場所、起点の緯度経度、終点の緯度経度（緯度経度は平成 21 年度から）及び路線が登録できる。これを用いて 7 桁の郵便番号で示される区域（日本郵便では「町域」と表現）レベルで工種等を集約するが、町域レベルで工事場所を登録していないものなどがあり、データクレンジングや路線名と自治体名から工事場所を推定するなどによって工事場所の特定を行うこととした。

5. 時間軸による分割

道路建設のために同一の場所で複数の工事が行われる場合、工事の実施時期や順序が予測に影響を与える可能性があると考えた。このため、道路開通に関係する場所においては道路開通日、道路開通に関係しない場所においては予測したい日（予測基準日）から遡って一定期間毎に工種等を分割することとした。

6. 予測手法

予測のための手法として、本研究では、入手できた道路開通情報を訓練データとする「教師あり機械学習」を採用し、各々の場所に対して道路開通の有無を分類する問題と考えることとした。教師あり機械学習の手法には、決定木学習、ナイーブベイズ分類器、サポートベクターマシン、アンサンブル学習など様々なものがある。後述する検証においては、決定木、ナイーブベイズ分類器を使用した。

決定木学習は、root(根)を含む各内部ノードが「説明変数」を表し、leaf(葉)が説明変数に対する「予測値、分類値」を表す木構造を構築することによって問題を解くものである。後述する検証で用いたC4.5アルゴリズムは、決定木のノード選定に「情報利得比」を用いている。情報利得比は、分割によるエントロピーの平均的な減少量である「情報利得」を、分割すること自体を伝える情報量である「分割情報量」で除したものである。

ナイーブベイズ分類器は、説明変数が独立しているとの仮定のもと、ベイズの定理を適用して確率的に分類するものである。

7. 実データによる検証

以上の提案方法を用いて道路の開通の予測を試みた。概要は次の通りである。

(1) 予測対象

埼玉県が建設した埼玉県内の道路の開通

なお、埼玉県では土木工事共通仕様書において工事実績情報の登録を求めているが、工事金額 500 万円以上の道路工事は工事実績情報に登録されていると考えられる

(2) 目的変数（クラス）の訓練（教師）データ

埼玉県がインターネットで公表している平成 20 年度～平成 22 年度の道路開通 19 件

(3) 工事実績情報

工事完了日が平成 18 年度(平成 18 年 4 月 1 日)から平成 22 年度(平成 23 年 3 月 31 日)の間の埼玉県内における埼玉県発注の工事実績情報（技術

表－2 データセット（説明変数）の例

町域	開通又は予測基準日以前に完了した工種等		
	0-1 年間	1-2 年間	2-3 年間
久喜市 菖蒲町 菖蒲	Works//Pavement//RoadEquipment//PedestrianPV//PipeLine//Foundation//Earth//RoadBed//Facilities	Pavement//Earth//Works//Foundation//RoadBed//PipeLine	Foundation//Earth//Works//RoadEquipment//RoadBed//Pavement

者情報は削除済み) 4363 件

(4) 空間軸の集約方法

郵便番号を利用して町域単位に集約

(5) 時間軸の分割方法

道路開通区間：道路開通日を基準日として工事実績情報の工事完了日を年単位に分割

非道路開通区間：予測基準日を基準日として同様に分割

(6) 説明変数

工事実績情報の項目「工種_工法_型式」を 18 種類に集約し年毎に 3 年分を使用

(7) 採用した機械学習手法（WEKA を使用）

・決定木学習（C4.5 アルゴリズムを用いた weka.classifiers.trees.J48）

・ナイーブベイズ分類器

（weka.classifiers.bayes.NaiveBayesSimple）

(8) 検証方法

10-分割交差検証（10-fold cross-validation）：全データをランダムに 10 分割し、1 つをテスト事例、残る 9 個を訓練事例とし、10 回検証する。

得られたデータセットは、589 の場所（町域）に対応しており、内訳は開通を含むデータセットが 27、開通を含まないデータセットが 562 であった。整理されたデータセット（説明変数）の一例を表－2 に示す。

8. 検証結果

(1) 決定木学習

決定木学習(C4.5)を用いた結果は次の通りである。

- ・正しく分類された数とその割合 564 (95.8 %)
- ・誤って分類された数とその割合 25 (4.2 %)

開通場所とその他の場所について、適合率(Precision＝適合数／予測数)、再現率(Recall＝適合数／正解数)、F 値(F-measure＝適合率と再現率の調和平均)を表－3 に示した。

(2) ナイーブベイズ分類器

ナイーブベイズ分類器による結果を表－4 に示す。

表－3 場所別の予測性能（決定木 C4.5）

目的変数(クラス)	適合率	再現率	F 値
開通場所	0.563	0.333	0.419
開通以外の場所	0.969	0.988	0.978

表－4 場所別の予測性能（ナイーブベイズ分類器）

目的変数(クラス)	適合率	再現率	F 値
開通場所	0.388	0.704	0.5
開通以外の場所	0.985	0.947	0.966

- ・ 正しく分類された数とその割合 551 (93.6 %)
- ・ 誤って分類された数とその割合 33 (6.4 %)

9. 考察

（1）提案手法の有効性について

ナイーブベイズ分類器では、49 の町域が予測され、うち 19 について道路の開通が公表されている。埼玉県内の郵便番号は 2929 あり⁸⁾、これが町域数に相当するとなると、調査対象区域が 1.67% まで減じられたこととなる。開通を漏れなく把握することが道路地図の更新には重要であるため、開通場所に対し適合率が低下しても再現率を向上させることが今後の課題である。

（2）機械学習手法について

全体の正解率では決定木学習 (C4.5) がナイーブベイズ分類器を上回ったものの、重要である「開通場所の再現率」ではナイーブベイズ分類器が決定木学習を上回った。誤予測の原因となる説明変数の欠落に対してナイーブベイズ分類器が決定木より寛容であるためと考えている。

（3）析出された予測ルールについて

決定木において情報利得比の高い説明変数は、予測対象場所で 2～3 年前の間に完了した橋梁関係工事、0～1 年前に完了した道路附属施設、2～3 年前に完了した基礎工事などであった。道路開通の前にはアスファルト舗装が行われているが、アスファルト舗装は維持補修においても広く行われているため、決定木のノードに採用されていなかった。路盤路床がノードとなるのではないかと筆者は考えていたが、決定木のノードに採用されていなかった。検証に用いたデータセットは橋梁に関係した開通が多かったこと、開通ではない道路の拡幅などは目的変数に含まれていないことが原因と考えている。

（4）誤予測の原因について

a) 開通の予測漏れ

開通が予測できなかったデータセットには、下記の原因で道路開通に必要な説明変数が不足していたと考えている。

（工事に起因するもの）

- ・ 工事現場の条件により、通常の道路建設で行われる工種等が無かった
- ・ 開通時期より相当以前に工事が行われていた（登録に起因するもの）
- ・ 施工場所あるいは工種等が正しく登録されていなかった

b) 開通場所以外を開通と予測

予測された場所が公表された道路開通区間ではない原因を次のように考えている。

- ・ 道路の新設に匹敵する大規模な工事（大規模な拡幅や災害復旧など）が行われた
 - ・ 開通予定であるが未公表の区間を判定した
 - ・ 開通された区間であるが公表リストになかった
- 開通以外であっても大規模な道路改良等は、道路地図の更新に重要であるため、不要な予測結果とは言えず、拡幅等の訓練データが入手できれば、詳細な調査が可能であると考えている。

10. おわりに

本研究では、空間軸と時間軸という客観的な指標を用いてデータを処理し、これに機械学習を適用することによって、工事实績情報から道路開通の予測が可能であることを示した。今後、他の道路管理者の道路開通に対して本提案手法を適用し、予測の汎用化を更に進めること、精度の向上のために開通地点データの少なさへの補正、誤判定の事例について調査を進めたい。

謝辞：本研究にあたり、東京大学空間情報科学研究センター柴崎亮介教授、関本義秀特任准教授に貴重なご助言をいただきました。ここに感謝の意を表します。

参考文献

- 1) (財団法人) 日本建設情報総合センター: JACIC 情報 (特集コリンズ・テクリス新時代), No.102, pp.25-26, 2011 年 7 月
- 2) 国土交通省, 埼玉県等: 土木工事共通仕様書 (平成 23 年 4 月), 埼玉県版 pp.95, < <http://www.pref.saitama.lg.jp/page/doboku-kouji-jitumuyouran.html> >, (入手 2012.6.25)
- 3) 国土交通省: 建設工事施工統計調査 (平成 20, 21, 22 年度), 第3表 発注者別, 工事種類別一元請完成工事高 < http://www.mlit.go.jp/statistics/details/kkoji_list.html >, (入手 2012.3.9)
- 4) (財団法人) 日本建設情報総合センター: コリンズ・テクリスに関する項目定義書, < <http://ct.jacic.or.jp/corporation/known/xml/index.html> >, (入手 2012.3.9)
- 5) 国土技術政策総合研究所建設システム課: 工事工種体系ツリー (平成 23 年度改訂版), < <http://www.nilim.go.jp/lab/pbg/theme/theme2/sekisan/tree/t1-4.pdf> >, (入手 2011.11.1)
- 6) 愛知県建設部道路建設課: 道路構造の手引き (平成 23 年 4 月), < <http://www.pref.aichi.jp/dourokensetsu/tebiki/index.html> >, (入手 2011.11.1)
- 7) 岐阜県県土整備部道路建設課: 岐阜県道路設計要領 (平成 19 年 1 月改正), < http://www.pref.gifu.lg.jp/kendo/michi-kawa-sabo/doro/gbridge/d_spc.html >, (入手 2011.11.1)
- 8) 日本郵便: 郵便番号データダウンロード, < <http://www.post.japanpost.jp/zipcode/download.html> >, (2012.7.1 入手)