

## 社会基盤の維持管理に関する知識ベースの構築

東京大学 学生会員 ○水野 裕介 東京大学 正会員 阿部 雅人  
 東京大学 フェロー 藤野 陽三 (株)BMC フェロー 阿部 允

### 1. はじめに

社会基盤の維持管理においては、取り扱う情報が多岐にわたり、意思決定を行う際には、構造諸元や履歴等の構造物自身が持つ情報に加えて、地盤状況や周辺環境などを併せて検討する必要がある。

橋梁に関しては、アメリカにおいて各州が包括的な橋梁マネジメントシステム(BMS)の運用を行っており、また日本においても BMS の研究・開発が行われている。BMS には継続的に膨大な情報が蓄積されるが、解析等に用いられる情報はわずかで、データの運用・再利用が十分に行われていない。そこで本研究では蓄積された情報の再利用・有効活用に着目したシステムの構築手法の検討を行う。

### 2. 知識ベースによる情報活用システム

本研究では図1に示すように知識ベースを導入した情報活用システムを提案する。データベースに蓄えられる情報は、維持管理業務を記録するためのものである。知識ベースはデータベースからデータマイニングにより構築され、維持管理に関するデータ間の相関やパターンを知識表現として保持する。知識ベースは、検査員の判断支援や技術者・管理者の意志決定支援のほか、教育用のコンテンツとしての活用が見込まれる。

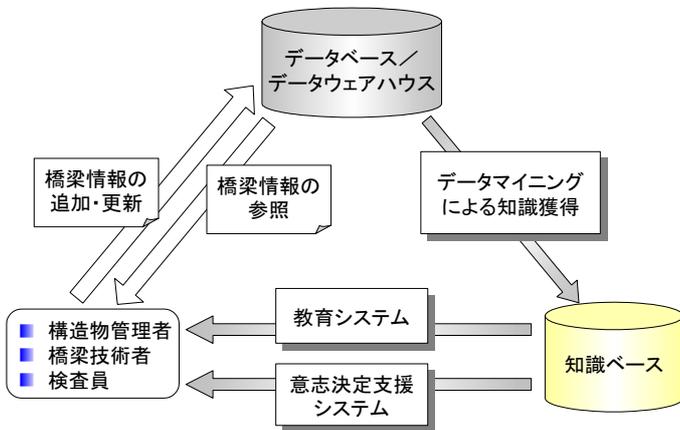


図1 情報の活用に着目したシステム

### 3. データマイニングによる知識獲得

#### 3.1 データマイニング

データマイニングは、離散化や数値化などのデータ変換を行う前処理、マイニングアルゴリズムの適用、ルールの可視化、検証等を行う後処理の3つの部分から構成される。

#### 3.2 CART アルゴリズムによる決定木の構築

決定木は、分類や予測を行う際に用いられている手法であ

り、ある注目する属性(従属変数)に関する重要な知識を、木構造で表現したものである。

CART(Classification and Regression Tree)は、決定木を構築する最も一般的なアルゴリズムの一つで、各ノードにおける分岐が2となる2進決定木である。

分岐の評価を行う尺度として、多様性を表す評価関数を用いる。各ノードではこの評価関数をもっとも減少させるような最良分岐を行う。

CART では、評価関数として多様性を表す指標の一つである Gini index を用いる。従属変数が2値で表現される場合 Gini index は次式のようになる。 $p_1$  はノード内で従属変数がある一方の値である割合を示す。

$$p_1(1-p_1)$$

### 3.3 橋梁の検査記録へのデータマイニングの適用

#### (a) 対象とするデータ

本研究では、橋梁の検査データ(表1)をもとに決定木によるルール抽出を行った。対象とする構造物を鋼鉄道橋とし、橋梁形式を上路プレートガーダー、組み立て構造をリベット形式とし、変状としてき裂の発生の有無を従属変数とし、それ以外の項目を独立変数としてデータマイニングを行った。橋梁連目数は389であり、そのうちき裂が発生している連目数は63であった。対象データの項目は、橋梁情報を管理するデータベースから得られたもので、データの欠損や誤りの少ないものを選択した。

表1 データ構造

項目	データ型	備考
設計荷重	テキスト型	
図面番号	テキスト型	
構造形式	テキスト型	
単複	テキスト型	単線/複線の区別
スパン	浮動小数点型	単位[m]
使用開始年	整数型	西暦
年間通トン	浮動小数点型	交通量, 単位[ton]
斜角の有無	セット型	「あり」または「なし」
曲線	セット型	「あり」または「なし」
溶接補強	セット型	「あり」または「なし」
き裂の発生	セット型	「あり」または「なし」

#### (b) 決定木の導出

導出された決定木を図2に示す。根ノードでは「スパン」による分岐が行われ、次に子ノードでは「図面番号」、「年間通トン」による分岐が行われている。最終的に7つの葉ノードが生成されている。

この決定木を導出に用いた橋梁データに適用すると表2の

キーワード 維持管理, 知識ベース, データマイニング, データベース

連絡先 〒113-8656 東京都文京区本郷7-3-1 東京大学大学院工学系研究科社会基盤工学専攻 TEL: 03-5841-6099

ようになる。決定木による予測と対象データを比較すると、予測の精度は正:361件 (92.80%)、誤:28件 (7.20%)となった。対象データのき裂発生連目は63 (16.20%)であり、決定木による予測は有意なものと考えることができる。

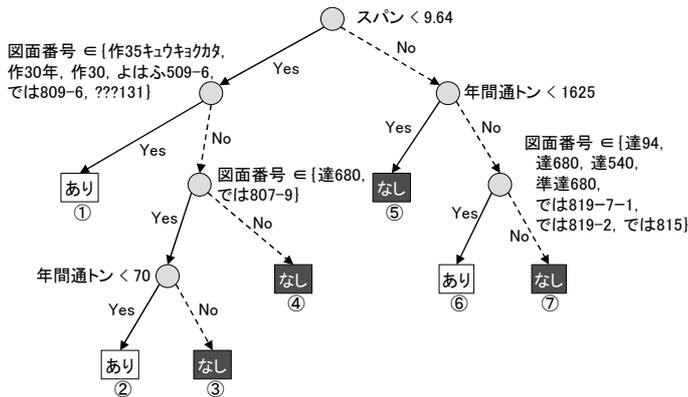


図2 導出された決定木

表2 決定木による予測

		決定木による予測	
		き裂あり	き裂なし
対象データ	き裂あり	39	24
	き裂なし	4	322

(c) プロダクションルールへの変換

決定木の根ノードから葉ノードへのルートを以下のようにIF-THEN形式のプロダクションルールに変換した。なお、ルールに示された番号は、図2の葉ノードに対応する。

- ルール1: IF (スパン < 9.64) AND (図面番号 ∈ {作35キョウカガ, 作30年, 作30, よはふ509-6, よはふ506-13, では809-6, ???131} ) THEN き裂あり
- ルール2: IF (スパン < 9.64) AND (図面番号 ∈ {達680, では807-9} ) AND (年間通トン < 70)) THEN き裂あり
- ルール3: IF (スパン < 9.64) AND (図面番号 ∈ {達680, では807-9} ) AND (年間通トン ≥ 70)) THEN き裂なし
- ルール4: IF (スパン < 9.64) AND (図面番号 ∈ {トケツ, 達970, 達540, 達1084, では508, 達1084, 達4.5', では816, では808, 806-7, Value} ) THEN き裂なし
- ルール5: IF (スパン ≥ 9.64) AND (年間通トン < 1625)) THEN き裂なし
- ルール6: IF (スパン ≥ 9.64) AND (年間通トン ≥ 1625) AND (図面番号 ∈ {達94, 達680, 達540, 準達680, では819-7-1, では819-2, では815})) THEN き裂あり
- ルール7: IF (スパン ≥ 9.64) AND (年間通トン ≥ 1625) AND (図面番号 ∈ {曲では864(8)-1, では828-2, では819-6, では816-50, では816, DG831-1, DG825-2, Value} )) THEN き裂なし

上記のルールの評価指標として、確信度とサポートを表3に示す。表中のルール該当データは、対象データ中でそれぞれ

れのルールの前提部(IF節)に該当するデータ数を表す。整合データはルール該当データのうち、ルールによる予測と実際のき裂発生が整合したデータ数で、矛盾データは予測と実際のき裂発生が一致しなかったデータ数である。確信度は、整合データをルール該当データ中の割合で示され、個々のルールの精度を表す。サポートは、整合データを対象とする全データ中の割合で示され、対象データ中でのルールの有効性を表す。

表3 ルールの評価

ルール	ルール該当データ	整合データ	矛盾データ	確信度(%)	サポート(%)
1	14	14	0	100.00	3.60
2	4	4	0	100.00	1.03
3	8	6	2	75.00	1.54
4	21	20	1	95.24	5.14
5	294	273	21	92.86	70.18
6	25	21	4	84.00	5.40
7	23	23	0	100.00	5.91
計	389	361	28		

4. 抽出されたルールの検証

抽出されたルールの中で、ルール5はサポートがもっとも大きく、確信度も高いため有効なルールと考えることができる。また、ルールの内容を検討すると、スパンがある程度以上で、交通荷重(年間通トン)がある値以下の橋梁にき裂が発生しないことは、橋梁維持管理に従事する技術者の経験と照らし合わせても妥当と判断できる。

ルール2, 3を見ると、ルール5とは異なり、交通荷重の少ない橋梁にき裂が発生する傾向が示されている。この傾向は、橋梁技術者の経験・知識に一見矛盾するように考えられるが、ルールの解釈方法により、この矛盾を解消できる。その解釈とは、交通荷重のきわめて少ない線区は閑散線区であり、適切な維持管理が欠落していることである。

ルール1では、ある特定の図面番号を持つ橋梁にき裂が発生しやすい傾向が読みとれる。これらの橋梁は補強桁や明治期に製作されたものであり、経年劣化が進んでいるものと考えられる。

5. まとめ

データベースに蓄えられた情報を有効活用するシステムのフレームワークを提案し、また、知識を抽出するデータマイニングを行った。その結果、いくつかのルールが生成され、その有効性を検討した。鋼鉄道橋を対象とし、き裂発生の有無について決定木を構築した結果、いくつかの有効なプロダクションルールが得られ、これらのルールは橋梁技術者の検証を行うことで、膨大なデータから効率的に知識を抽出する有効性を示した。

参考文献

- 1) M.J.A.Berry, G.S.Linoff: Mastering Data Mining, John Wiley & Sons, Inc., 2000.
- 2) 水野ら: データベース技術を用いた社会基盤構造物に関する維持管理データ管理手法の提案, コンクリートの耐久性データベースフォーマットに関するシンポジウム論文集, 土木学会, pp.77-82, 2002年12月