

山口大学大学院 学生会員 ○三角 俊介
(有)ミツワ電器 正会員 江本 久雄

(株)リョーセンエンジニアズ 非会員 絹谷 一郎
山口大学大学院 フェロー会員 宮本 文穂

1. はじめに

近年、老朽化した橋梁の増加が危惧されており、橋梁の長寿命化を目的とした効率的な維持管理対策が求められている。その維持管理対策のひとつとして、点検データからの損傷要因の抽出¹⁾が期待されている。著者らはデータマイニング手法のラフ集合を用いて橋梁の損傷要因を抽出しようと試みてきた²⁾。しかし、ラフ集合は計算量が膨大になり、処理が終了しないことがある。そこで本研究では、著者らが作成したラフ集合解析ソフトを用いたデータの前処理方法を提案し、計算量が多くなるデータの場合にも処理が可能であるかの検討を行なった。適用事例として、サンプル数(694件)と条件属性数(28件)のプラントの配管サポート耐震検討データを用いた。また、ハードウェアの環境による改善が行なわれるかについても検討を行なった。

2. データマイニング

データマイニングとは、データから未知で有用な知識(ルール)を導き出すプロセスのことである。原則として次の6段階の手順からなる。①データの選択、②データの洗浄、③データの補強、④データのコード化、⑤データマイニング、⑥報告であり、以上の一連の作業を行うことで知識抽出を行なう。また、各段階においてフィードバックすることが可能であり、それを行なうことでより精度の高いデータマイニングを行なうことが可能となる。

3. ラフ集合

ラフ集合では、同値関係、類似関係などによる集合を知識と考え、与えられた集合を上近似と下近似の2つの近似方法により表現する。上近似とは可能的に近似を考え、下近似とは必然的に近似を考える手法である。この上近似と下近似の対をラフ集合といい、これによってあいまいな情報を近似すべきであるという考えが、ラフ近似と呼ばれる。If-Thenルールで表わすことができ、対象の集合を粗く表現でき、非線形データや非数値データに強いという特徴を持つ。

4. 処理改善方法の提案

4. 1 提案手法

本研究では、ラフ集合における計算量の問題を解決するために、前処理方法を提案した。本手法は影響が強いと考えられる条件属性を選び、データを一定件数ずつのグループに分け、複数回処理を行うことでより多くのデータを処理しようというものである。手順は以下の通りである。

①決定属性決定

決定属性を一つ決定する。

②サンプル数/条件属性数決定

データマイニングに用いるデータのサンプル数と条件属性数を決定する。

③パターン作成

条件属性を絞り込むためのパターンを作成する。

④ルール獲得

絞り込んだ条件を用いて、データマイニングを行ない、ルールを得る。

⑤決定属性値毎分割

得られたルールを決定属性の属性値ごとに分割する。

⑥不必要情報削除

ルールに必要な情報や、使えなくなった情報を削除しておく。

⑦ルール集計/統合

ルールを集計し、重複をカウントしたり、矛盾を削除したりする。それらを決定属性の属性値ごとにルールを分けて統合する。

4. 2 64bit 環境の構築

現在、主流のコンピュータ環境は、32bitアーキテクチャである。この環境で扱えるメモリ使用量は理論上 $2^{32}B(=4GB)$ である。扱えるメモリ使用量に上限があるため、計算量がそれを超えた場合、処理が行えない。そこで、コンピュータを64bit環境にすることにより、扱えるメモリ使用量が理論上 $2^{64}B(=16EB)$ となり、より大きなデータが処理可能となる。

表 1 使用した条件属性・決定属性

条件属性 (10個)	データ件数 部材数	サポート個数 SizeMax	支持条件 ケーブルトレイ	階高範囲 スナバ	サポート条件 支持装置
決定属性	判定(総合)				

表 2 データマイニングの結果(一部)

「判定(総合)」が 「NG」となるルール	データ 件数	サポ ート 個数	支持 条件	階高 範囲	サポ ート 分類	部材数	SizeMax	ケー ブル トレイ	スナ バ	支持装置	条件の 件数
ルール1										アンカー	32
ルール2			床				A				21
ルール3			床			D					18
ルール4			床	D							12
}											

例：ルール2 If [支持条件=床] and [SizeMax = A] then [判定(総合)=NG]

表 3 クラスタリング対応表

階高範囲(mm)		サポート個数・部材数・ ケーブルトレイ・スナバ(個/件)		SizeMax(インチ)	
EL+30,000 以上	A	10以上	A	10以上	A
EL+20,000 以上 EL+30,000 未満	B	6以上10未満	B	3以上10未満	B
EL+10,000 以上 EL+20,000 未満	C	2以上6未満	C	0大なり3未満	C
EL+0 以上 EL+10,000 未満	D	1	D	不明・空白	Z
-EL+1 以下(負)	Y	0	E		
? 不明	Z	不明	Z		

5. プラント配管耐震事例への適用

本研究で提案した前処理方法を実際の事例に対し適用して、有効性の検討を行なった。

使用したデータは(株)リョーセンエンジニアズより提供して頂いた、古いプラントの配管サポート耐震評価を現在の評価基準で判定したものである。判定結果は「OK(損傷無し)」か「NG(損傷有り)」である。元データの条件属性は22個、決定属性候補は4個あり、検討には「判定(総合)」を用いた。

元データに対して提案手法を適用することにより、使用する条件属性・決定属性は表1に示すとおりとなった。データマイニングの結果は表2のとおりである。ここでは『「判定(総合)」が「NG」となるルール』に注目した。また、表3は条件属性のクラスタリング対応表である。

得られた結果の中でも特徴的なのは「支持装置」が「アンカー」となったときである。アンカーボルトは規格がJISではなく比較的新しいJSSであることから、耐震基準が曖昧となり、古いプラントに用いられているアンカーボルトに対しては、評価ができなかった可能性も考えられる。専門家によると、「現在では常識化してきている知識の一つ」とのことであるが、少ない条件の中からも傾向を抽出した点は、データマイニングの知識抽出の観点からも、有効であったといえる。

6. まとめ

以下に、本研究で得られた成果をまとめる。

- ① 本研究で提案した前処理方法を用いて、プラント配管耐震事例へ適用した。計算量の多いデータを用いても一定の評価ができるルールを得ることができ、前処理方法の有効性を確認できた。
- ② 「支持装置=アンカー」のとき、損傷が発生しやすいというルールが得られ、ラフ集合の妥当性も確認できた。
- ③ 64bit 環境を構築したことにより、より多くのデータが処理でき、短時間で処理できることが確認できた。

謝辞

本研究を行うにあたって、多大な協力をして頂いた九州工業大学大学院 冬野聖君に感謝致します。

参考文献

- 1)古田均, 広兼道幸, 田中成典, 三雲是宏: 橋梁の損傷要因診断事例からのラフ集合を用いたルール型知識の獲得方法, 構造工学論文集 Vol.44A, 1998.3
- 2)江本久雄, 絹谷一郎, 三角俊介, 河村圭: 鋼橋点検データからの知識抽出のためのラフ集合論を用いたアプリケーション開発, 第9回土木・建築ソフトコンピューティング応用シンポジウム講演論文集, 2007.2