

道路維持管理業務支援システムにおけるデータマイニングとその評価

一瀬邦継*1 中西正樹*2 寺田守正*3 藤澤裕樹*4 吉澤憲治*4 金田重郎*4

同志社大学工学部*1 京都府山城広域振興局*2

京都府流域下水道事務所*3 同志社大学大学院工学研究科*4

1. はじめに

近年、情報システムを用いたナレッジの収集とその活用に注目が集まっている。情報システムを利用することで、業務に対する膨大な量のデータにデータマイニングを行える。そして、データマイニングにより生成された知識は、その組織内のナレッジとして共有できると期待されている。そこで著者らは、京都府山城北土木事務所（以下「土木事務所」と呼ぶ）の道路維持管理業務を支援するシステム“京都道守くん”を開発・導入し、実証実験を行った[1]。本稿では、そのデータマイニングから得られた課題を基に、情報システムからのアプローチを行い、新しい組織体制の提案までを述べる。

2. データマイニングの適用

2.1 背景

道路維持管理業務（以下「管理業務」と呼ぶ）は、住民通報や職員巡視で発見された、道路における施設破損や不法投棄など（以下「案件」と呼ぶ）に対応する業務であり、自治体が行う重要な住民サービスのひとつである。管理業務はその特性から、受動的で対症療法的な業務になりやすい。これでは、業務が非計画的なものになり、業務効率が上がらないばかりでなく、案件そのものの削減に繋がらない。また、住民満足の視点からも好ましくない。このような問題を解決できる可能性を見出すために、データマイニングの適用に至った。これにより、案件発生を予測し、予防的で計画的な管理業務を実現できると考えたからである。

2.2 データマイニングの処理

前節を受けて、データマイニングを、土木事務所が管理する宇治エリアの道路で発生した過去 5 年間の案件データ 800 件に対して行った[2]。その結果、複数のクラスターが生成され、特定の案件には一定の周期があると確認できた。

3. データマイニングから得られた課題

前述の通り、特定の場所に似たような案件が集中しているのが明らかになった。しかし、データマイニングの結果からは、新しい業務知識や 2.1

Datamining in Road Management System and its Evaluation

*1 Faculty of Engineering, Doshisha University

*2*3 Kyoto Prefecture

*2 Graduate School of Engineering, Doshisha University

節で述べた問題を解決できるような業務知識を得られなかった。その理由として考えられるものを以下に記述する。

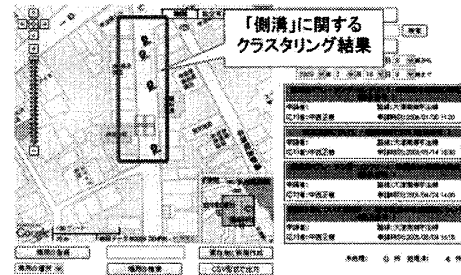


図1 「側溝」に関するクラスタリング結果

3.1 原因情報の欠如

例えば「側溝から水が溢れた」という場合、なぜそのような問題が起こったのかという理由、つまり原因究明は重要である。これまで利用されてきた情報システム“京都道守くん”に蓄積されたデータには、「原因」という文脈付けられたデータ属性が存在しなかった。そのため、原因の入力が必須条件でなく業務担当者の判断に依存し、その案件がなぜ発生したのかわからないものが存在した。第三者が、「その案件が頻発している」という結果だけを見ても、「なぜその案件が発生したのか」という文脈情報が少ないために、案件が頻発する解決策をシステム上では得られなかった。

3.2 「種別」選択的的確性の課題

データマイニングから案件の特徴を抽出する際、「何に関する案件なのか」という情報は、非常に大きな意味を持つ。そこで本システムでは、「案件の種別」という属性を持っていた。しかしその種別の選択も、業務担当者によって業務経験の長さや価値観などに依存し、「種別」として選択するものが各業務担当者で異なっていた。周知の「醜いアヒルの子の定理」によれば、「もの」と「もの」の類似性は一般的には規定できないとある。種別の選択は、無意識の内に何らかの「価値観」に基づいて行われるため、業務担当者は、必ずしも的確に種別選択を行わない。ひとつの価値観の下に種別選択を行うのは、業務担当者が複数人いる以上、本質的に不可能である。

4. アプローチ

以上の議論から、著者らは“京都道守くん”に以下のような改良を施した。

4.1 原因項目の追加

3.1 節を考慮して、「原因」をデータ属性として、入力できるようにした。但し、「原因」の入力は、案件を処理するという視点では、必ずしも必要ではない。グループウェアの研究において周知の現象であるが、利用者は入力に必然性を感じない情報は、入力インセンティブが沸かない。そこで、上司への報告書にも「原因」記入を必須とし、心理的に入力を促すよう配慮した。

4.2 近隣案件画面の設置

データマイニング結果では、複数のクラスが抽出され、マイニング自体は成功したと言える。そこで、通報を受けた案件の近隣で、過去に同じような案件が発生しているならば、それを同じ地図上に表示できるような画面を設置した。これにより、業務担当者に「同種の案件が近隣で頻発している」の気づかせられる。そして、そこで発見した過去の案件の、原因や処置方法をフィードバックし、業務担当者間に原因究明の重要性に対する認識が高まると考えた。また、同じ画面に他の業務担当者が、どのような案件内容に対して、同じ種別を選択したのかを確認できる。これによって、ある程度の種別選択の正確性を保てる。



図2 近隣案件画面

5. 考察

今回の社会実験では、既存システムに「原因」というデータ属性を追加すべきであるという結論を得た。データマイニングにより結果そのままでは、ナレッジには繋がらなかったためである。

ここで、データマイニングの代表例である「ショッピングバスケット分析」について考えてみる。「ショッピングバスケット分析」は、「ある物を買う顧客は同時にある別の商品を買う確率が高い」といった結果を抽出できる。そして、その結果を業務の中に直接取り入れ、業務の改善(売上げ)に繋げている。また、この過程の間に「なぜこの結果になるのか」という考慮はするが、推論までに留まっている。

しかし、今回の管理業務におけるデータマイニングでは、「同様の案件が頻発している」という結果を得た。「頻発している」というマイニング結果は、「案件の削減」という業務目標とは方向が逆である。そのため、マイニング結果を直接業務に取り入れられない。「ショッピングバスケット分析」では推論で終わっていた原因究明が、管理業務

では「なぜ頻発しているのか」という視点で、より深く原因究明を行う必要がある。

トヨタカンバン方式の生みの親・大野耐一は、「5つの why」を提唱している[3]。何かトラブルが生じた際、それが何故起きたのかを5段階繰り返して、本質に迫れとの意味と思われる。したがって、今回の事例でも、この「5つの why」と本質は同じだと考えられる。つまり、「業務改善に直接利用できないマイニング結果」を有効に利用するためには、「なぜ」を深く問いかける必要がある。

また、「種別」が正しく選ばれていなかった問題に関しては、近隣案件画面である程度の正確性は保てるが、完全ではないと考えられる。そこで、組織体制として、ナレッジコンシェルジュを置く必要があると考えられる。当該担当者が、入力された「原因」は本質を突いた原因究明を行っているか、「種別」は正しく選択されているかというチェックを行うのである。このようにひとつの価値観の下でデータを整理しなければ、蓄積されたデータも意味を成さず、根本的な解決には繋がらないのではないかと考える。

6. まとめ

本稿では、道路維持管理業務に対して行ったデータマイニング結果を元に、「京都道守くん」の改良を行い、改良に対する考察とナレッジコンシェルジュの必要性の提案までを行った。

業務データを蓄積してデータマイニングを行えば、そのままナレッジとして利用できるとは言い難い。データマイニングから得られる結果が対象業務の目標と逆行する場合には、「なぜ」を問う必要がある。またデータマイニングの正確性を得るためには、ナレッジコンシェルジュの設置が必要である。

本研究に関しては、引き続き社会実験を実施してデータを蓄積し、今回行った改良と、追加した「原因」のデータ属性の有効性も合わせたデータマイニングの評価をしていきたいと考えている。

謝辞

最後に、今回の研究に関してご協力いただいた自治体職員各位に、深く感謝いたします。なお、本稿で示した内容は、当該自治体の公的見解を表すものではありません。

参考文献

- [1] 吉澤憲治他、「Web-GISを用いた道路管理業務支援システム“京都道守くん”の開発」、情報処理学会、SIG-IS、情報処理学会研究報告・2007-IS-99、pp.39-44
- [2] 藤澤裕樹他、「道路管理情報にもとづくデータマイニングの可能性について」、情報処理学会・第70回全国大会6ZF-2、pp.4-637~4-638
- [3] 大野耐一、「トヨタ生産方式—脱規模の経営をめざして」、ダイヤモンド社