

Structural Topic Models を用いた橋梁点検所見記述の分析

金沢大学 正会員 小川福嗣
金沢大学 正会員 近田康夫

1. はじめに

現存する橋梁の老朽化対策が急務となっており、2014年には道路法改正や定期点検要領の公表など定期点検実施に向けた体制が整えられるなど老朽化対策の本格実施に向けて取り組みが行われている。また、橋梁の各管理者は、計画的な維持管理を行うために橋梁マネジメントシステムを導入し点検結果等の蓄積が進められてきた。蓄積された点検データを活用して、現状の健全性の詳細な評価や劣化予測に活用する研究は多数行われている。この際、これまでは主に各部材の損傷状態に関するグレード評価値などが利用されてきた¹⁾。しかし、個々の部材評価の積み上げでは橋梁全体を十分に評価されたとは言えないのが現状である。点検結果には定性的・定量的なグレード評価以外に点検を踏まえた専門技術者による総合的な所見に関する記述が記録されている。これらの所見記述には定量的な評価値では表現しきれない専門技術者の知見が含まれ、橋梁全体の健全度評価にも影響すると考えられる。そのため、これらの記述内容を分析することで今後の維持管理に活用できる可能性がある。テキスト分析技術で用いられる基本モデルとしてトピックモデルがある。所見記述にトピックモデルを援用することによって橋梁の損傷の特徴や記述方法により分類が可能となる²⁾。本研究ではトピックモデルを解析対象の文章以外の情報も取り込み可能なよう拡張された Structural Topic Models を用いて分析を行った。

2. 方法

解析対象とする所見記述データは2006年～2016年に富山県で実施された橋梁点検結果の記録を利用する。富山県ではすべての部材で発生した損傷を詳細に把握できるように行われる詳細点検と、比較的小規模な橋梁や簡易な構造の橋梁に対して実施可能な概略点検がある。本研究では専門技術者による所見記述が詳細に記載されている詳細点検の結果760橋分の所見記述データを解析対象とする。

形態素解析にはMeCabを利用し、分割された形態素のうち損傷状態などを表現すると考えられる名詞を解

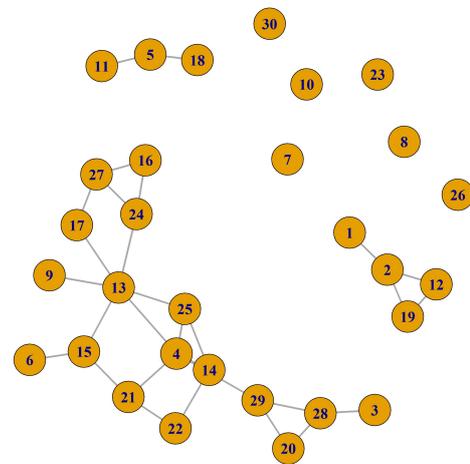


図-1 トピック間の関係

析対象とした。語彙数は1,820語、単語数は37,818語である。橋梁ごとの解析対象となった単語は平均で57単語であった。形態素解析を行ったデータに対する分析に用いられる基本モデルとしてはLatent Dirichlet Allocation (LDA)がある。橋梁の所見記述が記載される際には、前提条件として橋梁の点検結果で得られた健全度情報が考慮されていると考えられる。そのため、本研究ではLDAをトピック間の相関およびトピック分布への各種補助情報の与える影響を考慮できるように拡張された Structural Topic Models³⁾を用いて分析を行った。橋梁の属性情報として上部工と下部工の健全度(0～100の範囲内で評価)及び橋長を取り入れた。トピック数は、各トピックが意味を持つかどうかの指標である Semantic Coherence や一般的な頻出単語が各トピックにどの程度集約されているかを表す exclusivity の指標を参考に30とした。

3. 結果

トピック間の関係を図-1に示す。ノードの番号はトピック番号を表しており、トピック間の相関係数が0.01以上の場合にはノード間にリンクが張られている。図からトピックはひとつの大きなグループとその他のグループに大別されている。最も大きなグループのトピックの上位頻出単語には特段特徴は見られないことから、管理橋梁に全般に確認できる一般的な損傷に関するト

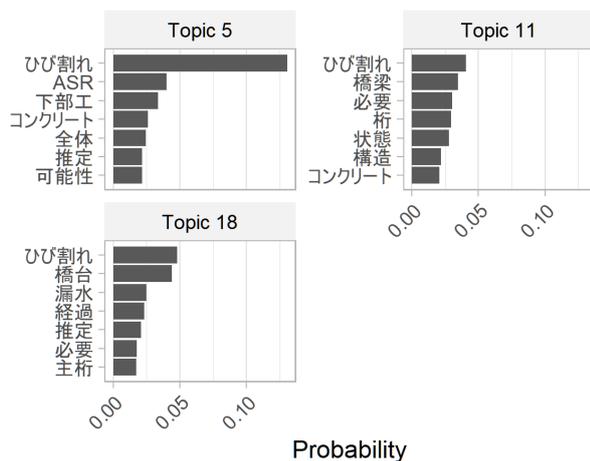


図-2 トピックごとの上位頻出単語

ピックとしてグループが形成されていると考えられる。そのため、そのほかのトピックに着目することにより比較的特異な損傷を生じている橋梁が抽出できると考えられる。本稿ではトピック5, 11, 18のトピックグループに着目する。これらのトピックにおける上位頻出単語とその出現確率を図-2に示す。各トピックの単語には「ひび割れ」、「橋台」や「コンクリート」が確認でき、コンクリート部材の損傷形態を表すトピックであると考えることができる。また、各所見記述文章にもっとも割り当てが高いと推定されたトピックを各橋梁（所見記述）の主要トピックとして割り当てた場合、トピック5が主要トピックとして推定された橋梁は27橋中19橋、トピック11が主要トピックと推定された橋梁は11橋中5橋において、アルカリ骨材反応やアル骨といったASRに関する記述が確認され、ASRに関連したトピックでもあったと考えられる。一方、トピック18主要トピックと推定された橋梁14橋は2橋にしかASRに関する記述は確認できなかったが、ひび割れが広範囲に確認できるなどひび割れに関する記述が多く確認された。このことから、これらのトピックは主に橋梁全体へのひび割れが主要な記述内容となっているトピックとしてグループが形成していると考えられる。そのほかのグループも同様にトピック8は「コンクリートの剥離が生じており鉄筋露出が見られる」といった比較的重大な損傷が生じている記述が多く確認されるなど、他の記述に比較して特徴的な損傷に関するトピックグループが形成されていることが確認できた。

上部工および下部工の健全度とトピックの割り当ての関係について確認する。本稿では特に健全度との関

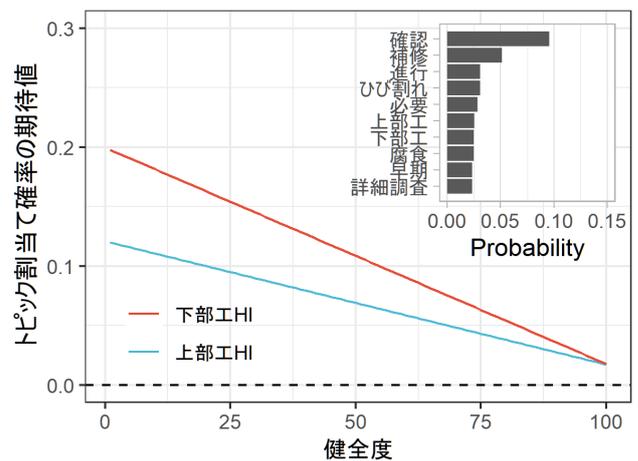


図-3 トピック10と上部工及び下部工健全度の関係

係が明確に確認されたトピック10を抽出し、健全度の影響を図-3に示す。図中右上はトピック10の上位頻出単語を示す。図からトピック10は上部工及び下部工の健全度が低いほどトピックの出現確率が高くなっている。つまり、トピック10の割り当てが高い橋梁は全体的に損傷が見られ健全度が低い傾向にあることが想定される。トピック10の出現確率上位の単語に「早期」、「補修」、「必要」、「進行」があり、損傷が進行しており補修が必要な段階であることが想定される単語となっている。また、トピック10をメイントピック推定された橋梁の健全度を確認すると他のトピックをメインとする橋梁に比べて健全度が低い傾向にあるが確認された。

4. おわりに

本研究では橋梁点検結果のうち専門技術者による所見の記述に対して、所見記述のテキストデータ以外の属性情報を考慮できるStructural Topic Modelsを適用して分析を行った。テキストデータからも橋梁の分類や健全度の傾向を把握できる可能性があることが確認でき、特異な損傷がトピックとして抽出できる可能性があり今後の橋梁維持管理に役立てられる可能性があることを示すことができた。

参考文献

- 1) 大竹雄, 流石堯, 本城勇介, 村上茂之, 小林孝一. 統計的手法を用いた橋梁点検データベースに基づく橋梁健全度評価に関する基礎的研究. 土木学会論文集A2(応用力学), Vol. 67, No. 2, pp. I.813-I.824, 2011.
- 2) 小川福嗣, 近田康夫. トピックモデルを用いた橋梁点検結果所見記述の維持管理への活用に関する基礎的分析. 材料, Vol. 69, No. 3, pp. 197-203, 2020.
- 3) Margaret E. Roberts, Brandon M. Stewart, and Dustin Tingley. stm: An R package for structural topic models. *Journal of Statistical Software*, Vol. 91, No. 2, pp. 1-40, 2019.