Transport Policy xxx (xxxx) xxx-xxx



Contents lists available at ScienceDirect

Transport Policy



journal homepage: www.elsevier.com/locate/tranpol

Discovering regulatory concerns on bridge management: An author-topic model based approach

Qi Wen, Maoshan Qiang*, Bingqing Xia, Nan An

Tsinghua University, Institute of Project Management and Construction Technology, State Key Laboratory of Hydroscience and Engineering, Beijing 100084, China

A R T I C L E I N F O

Keywords: Bridge management Author-Topic Model (ATM) Text mining Policy trend analysis

ABSTRACT

The management of bridges has major influence on the safety of transport system. Bridge management (BM) practices are carried out in compliance with policies and regulations. This study aims to identify the major aspects of BM from policy documents. We conducted a two-round retrieve to construct a comprehensive dataset of 263 related policies and regulations issued in the 31 provinces/municipalities of mainland China in recent 30 years. The Author-Topic Model (ATM) text mining approach was adopted to identify the key topics in the policies. The revealed 12 topics correspond to 12 major aspects of BM. Some topics, such as *bridge maintenance and safety responsibility system*, are gaining increasing attention in recents, while the prevalence of some topics, such as *bridge maintenance department configuration*, is decreasing. The Granger causality tests suggest that remedial measures attract immediate regulatory attention in response to accidents, while proactive measures contribute to bridge collapse accident reduction, however, in a relatively long term. These findings provide practical implications for policy makers by reminding them that proactive measures should be emphasized alongside remedial measures.

1. Introduction

Bridges are pivotal but also vulnerable components of transport networks (Agbelie et al., 2017). The extent to which bridges are resilient to external perturbations has major impact on the safety of transport system (Zhu et al., 2010; Dong and Dan, 2015). In China, recent ten years have witnessed more than 200 bridge collapse accidents, 80% of which are ascribed to overloading, lack of maintenance and quality defects (Chen, 2014). Thus, bridge management (BM) is attracting growing academic and regulatory attention, especially with the increasing global population, urbanization and climate change trends, which collectively pose significant challenges (Deng and Wang, 2015). There is a pressing need to formulate effective BM policies to ensure that bridges are sufficiently resilient to external shocks from daily operation and natural disasters (Chow et al., 2015).

In China's infrastructure construction industry, professionals tend to pay much more attention to construction than maintenance. Multiple bridge construction projects are usually managed simultaneously by the same project company. However, under the pressure of schedule and budget, project companies may be unable to invest required resources on bridge maintenance during the trial operation phase (Andrić and Lu, 2016). After commissioning, most bridges are transferred to bridge maintenance departments, which are funded by corresponding local governments. According to the National Statistics Bureau of China (2015), there are totally 735 thousand bridges in China, 86 thousand (11.7%) of which are in poor condition. However, the mileage drive on the bridges in poor condition accounts for only 6.5% of the total mileage. This indicates that most of them are small. The maintenance of many bridges, especially small bridges in rural areas, is not supported with sufficient funds (Chen, 2014; Pan and Zhongyong, 2013). Similar problems of poor bridge condition have also been widely reported in many other parts of the world (Orcesi and Cremona, 2010; Andrijcic et al., 2013; Zhu et al., 2010). For example, according to ASCE's survey on the infrastructure system of the US, the average age of the nation's 607,380 bridges is above 40 years, and about one in nine bridges are structurally deficient (Herrmann, 2013); in France, the quality assessment of engineering structures suggests that, among the over 9 thousand bridges, nearly 10% of bridges suffer from structural deterioration (Orcesi and Cremona, 2010). Corresponding policies are desperately needed to enhance BM (Liu and Dan, 2005).

In order to effectively regulate BM practices, China's Ministry of Housing and Urban-Rural Development (MOHURD) and Ministry of Transport (MOT) issued numerous policies and regulations, such as the

* Corresponding author.

E-mail address: qiangms@tsinghua.edu.cn (M. Qiang).

http://dx.doi.org/10.1016/j.tranpol.2017.04.012

Received 24 March 2016; Received in revised form 13 April 2017; Accepted 29 April 2017 0967-070X/ \odot 2017 Elsevier Ltd. All rights reserved.

Notice on Enhancing Urban Bridge Safety Management, Regulations on Bridge Maintenance Management and Technical Standard for Urban Bridge Maintenance. However, in practice, BM subjects to numerous local economic and political conditions. There still exist many managerial problems, especially the unclarified BM responsibilities (Yin et al., 2011). For example, the middle route of the South-North Water Transfer Project includes more than 1200 newly constructed cross-channel bridges, most of which have been commissioned for more than seven years but not transferred to local maintenance departments. The managerial responsibilities of these bridges are not clearly defined and, as a result, many of them are deteriorating without adequate maintenance and regulation on overloading. This, in essence, ascribes to the lack of understanding on the policies and regulations on BM by the stakeholders. In this light, improving bridge operation condition is not only a technical problem, but also a managerial and regulatory problem (Boin and McConnell, 2007; McDaniels et al., 2008; Rattanachot et al., 2015). When formulating BM policies, policy makers need to identify related policies from numerous previous policies and integrate the experience embodied in them. Obtaining knowledge from the massive policy texts and identifying key aspects of BM become the major challenges in this process (Dai et al., 2014).

This study aims to identify the major aspects of BM from policy documents by text mining approach and analyze the trends of topic prevalence. Theoretically, the findings provide a comprehensive review of the major regulatory concerns on BM in recent years. The revealed key aspects of BM can assist policy makers in integrating previous experience during the BM policy making processes. Bridge operators can also utilize the findings to operationalize BM policies into implementable BM procedures, and in turn enhance BM practices.

2. Bridge management: a political perspective

Informed by the importance of BM, many research efforts have been devoted to studying BM from technical point of view. Many studies proposed versatile technical solutions to various aspects of BM. For example, bridge inspection has been a hot topic attracting wide attention from researchers and practitioners. Recent bridge collapse accidents motivated the joint investigation efforts of ASCE and AASHTO to examine the current bridge inspection and rating practices (ASCE/SEI-AASHTO Ad-Hoc Group on Bridge Inspection, Rating, Rehabilitation, and Replacement, 2009). Sweeney and Unsworth (2010) described two effective bridge inspection approaches and emphasized on designing bridge inspection mechanisms according to their conditions. Washer et al. (2016) further developed a risk-based approach to determine bridge inspection intervals and procedures based on the intensity of their operational risks. Other important aspects of BM have also been intensively investigated in previous studies, such as the weigh-in-motion technology for truck weight regulation (Hang et al., 2013), bridge performance rating (Dong and Dan, 2015; Ikpong and Bagchi, 2016) and maintenance strategy optimization (Durango and Madanat, 2002; Orcesi and Cremona, 2010). The recently prevailing concept of infrastructure resilience (Bruneau et al., 2003; Cimellaro et al., 2010; Bocchini et al., 2013) further provides valuable insights for BM (Ikpong and Bagchi, 2016; Dong and Dan, 2015) and implies a socio-technical perspective to enhance BM practices (Mattsson and Jenelius, 2015). However, existing literature is dominated by studies focusing on technical issues, while there is a lack of political and managerial perspectives on technology implementation (Andrijcic et al., 2013).

There is a pressing need for harmonizing engineering technology with socio-economic modeling in order to integrate the cross-disciplinary collaborations between policy makers, operators and researchers (Andrijcic et al., 2013; Dong and Dan, 2015). In practice, BM is carried out in compliance with policies and regulations (Dai et al., 2014). Government is the major source of funding on BM. For example, in China, the total investment on infrastructure maintenance and

management in 2015 was 3.83 trillion RMB, 68.3% of which was from Chinese government, with the rest from tolls and private investors (National Bureau of Statistics of China NBSC, 2015). According to the Article Six of Regulations on Bridge Maintenance Management issued by Ministry of Transport, bridges should be managed by bridge maintenance departments under the supervision of local transport departments, except for the toll bridges managed by the operators. Since toll bridges only accounts for 1.8% of the total bridge mileage, the majority of bridges are managed by government departments. Hence, government plays a pivotal role in BM in terms of investment and management responsibility, and government policies act as the standard and enabler of BM. The significance of policy making in BM has also been widely echoed in other countries, e.g. the US (Thompson et al., 2003), Canada (Ikpong and Bagchi, 2016), France (Orcesi and Cremona, 2010), etc. As pointed by Dong and Dan (2015), only when integrated in decision making processes, can engineering technology actually contribute to BM practices.

From a contingent perspective, different BM policies should be formulated according to the attributes and operational conditions of bridges (Sweeney and Unsworth, 2010; Washer et al., 2016). Numerous policies and regulations on BM has been issued by different levels of governments and national ministries. Thus, how to design BM policies for specific contexts and, at the same time, incorporate the knowledge from the mass of previous policies is an important research question (Coglianese, 2004). Moreover, policies at different levels have different influence on BM practices, and how to embed and operationalize superordinate policies into specific regulations at lower level needs investigation (Andrijcic et al., 2013).

As evidenced in previous studies, there is a growing awareness of integrating previous experience in policy making processes (Prosser and Peters, 2010; Marsden and Stead, 2011). However, most previous studies are built on manual content analysis or case analysis methods. which involve considerable subjectivity and are unable to handle massive policy texts (Zirn and Stuckenschmidt, 2014). Moreover, despite the fact that BM practices are carried out within the framework of related policies, there is a lack of systematic review of these policies. With the advent of the big data era, policy making is becoming a more and more information and knowledge intensive process. Coglianese (2004) proposed the grand concept of E-rulemaking aimed at improving the efficiency and effectiveness of policy making processes by extensive information collection and processing. According to the key research questions with respect to E-rulemaking, methods to extract knowledge from previous policies and reveal the trends of regulatory concerns are promising. Heeding this call, this study aims to discover experience and knowledge from massive policy texts on BM by the Author-Topic Model (ATM) text mining approach.

3. Methodology

In order to achieve the research objective, we designed the analysis procedure as illustrated in Fig. 1 and explained in the following subsections.

3.1. Data collection and preprocessing (steps 1-3)

Since various kinds of bridges (mainly road bridges, train bridges and municipal bridges) have similar structural component and are managed in compliance with almost the same set of policies in China, we take a broad perspective to include all the policies relevant to BM. To construct a dataset of BM policies as comprehensive as possible, we conducted two rounds of retrieve with all online databases that we can approach. This study is a part of a research project to formulate BM measures based on previous BM policies, and was conducted in close collaboration with South-North Water Transfer Project Management Bureau. In consultation with experienced BM experts in the Bureau, we used "bridge management", "bridge maintenance", "bridge resilience"

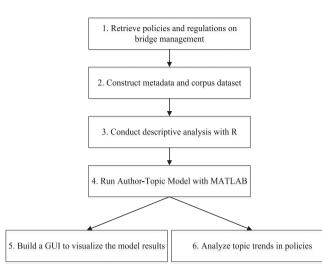


Fig. 1. The analysis procedure of this study.

and "bridge operation" as the keywords for search. In the first round, we searched the 12 prestigious databases frequently used in political studies, including WestLaw, PKULaw, Lawyee, Chinacourt, CNKI, WangFang, etc. Each of the three researchers independently retrieved 4 of the 12 databases, and obtained 22, 29 and 105 policies respectively. A seminar was held to compare and merge the search results. We found 127 unique policies after removing the redundant results and 2 irrelevant policies. The comparison among the three groups of results shows the significant difference among the databases. Generally, WestLaw and WangFang databases (based on which an author obtained 105 relevant policies) are much more comprehensive than others. However, none of the databases is perfectly comprehensive, and the overlap among the three groups of results was little, which indicates little consistency among the results and substantial possibility of omitting relevant policies. Therefore, we carried out the second round of retrieve based on the database of the MOT, MOHURD and the 31 province/municipalities' (in mainland China) Departments of Housing and Urban-Rural Development (DOHURD) and Departments of Transport (DOT). This time, we found 232 relevant policies. After eliminating the 96 policies already existing in the results of the first round, we obtained totally 263 BM policies. The results of the two rounds overlap to a large extent, which indicates that the two different approaches arrived at consistent results. We also compared the results with the policy list independently summarized by the experts from the Bureau based on their management experience. The retrieved policies fully cover the list, so they can be analyzed to represent the BM policies frequently used in practice.

The corpus dataset was constructed by splitting the policies into articles and matching metadata (date, province, file number and the policy to which the article belongs) with the articles. Zirn and Stuckenschmidt (2014) pointed out that using the whole policy as the unit of analysis is problematic for topic models, since it includes an arbitrary mixture of too many topics in a policy and makes it hard to identify meaningful topics. They proposed to split the policy text into thematically coherent parts to better satisfy the assumption of topic model that documents are meaningful co-occurrences of keywords. The policy texts in this study are composed of multiple articles (paragraphs expressing independent meanings and numbered sequentially), which naturally split the whole policy text into thematically coherent parts. On average, the retrieved policy texts include 2759 words and 14 articles. We split them (into totally 3759 articles) and defined their author attribute as the policy to which they belongs. As elaborated in Section 4, this data preparation procedure also enables the meaningful analysis on which policy best represents each aspect of BM.

program. Unlike English texts, Chinese texts are not naturally split by the spaces between words, so we need to conduct word segmentation in the data preparation stage. As suggested by previous studies (Jiang et al., 2016), we executed word segmentation with the R package Rwordseq to identify keywords from texts and, at the same time, eliminate meaningless stopwords, such as "of" ("的"), "and" ("和") and "etc." ("等"). In order to improve the success rate of keyword identification, we attached several dictionaries, which include professional terms on "bridge engineering", "transportation" and "policy and regulation", from Sogou Lab¹ to *Rwordseq*. Thereafter, using the *tm* package (Meyer et al., 2008), we constructed the document-term matrix (DTM) where DTM_{ii} denotes the word frequency of term *i* in document (policy article) *i*. Corresponding to DTM, we also constructed the author-document matrix (ADM) where ADM_{ii} is a dummy variable indicating whether document (article) j belongs to author (policy) i. Based on DTM, we can conduct descriptive analyses, such as word frequency analysis.

3.2. Author-Topic Model (step 4)

Based on the constructed dataset, we compiled an R program to transform DTM and ADM into the format ready for ATM analysis.

Author-Topic Model (ATM) is a probabilistic model for discovering the intellectual structure of a corpus (a collection of documents) from multiple authors (Steyvers et al., 2004). It takes the raw texts and the authors (categorical attributes) of documents as the input, and outputs the major topics (represented as sets of keywords) of each document and the major topics concerned by each author. In achieving this, it takes the words in the documents as the units of analysis, and models the corpus by a four-level structure.

First, the corpus contains multiple documents (d) that express independent meanings. Second, the documents are written by their authors (A_d) . The "author", in essence, can be defined as any document-level categorical attribute (x). For example, Morchid et al. (2015) adopted ATM to analyze Twitter documents and defined the "author" attribute as the country where the Twitter document was from. In this study, the corpus is composed of BM policies, which is further composed of multiple articles. Corresponding to the configuration of ATM, each article is treated as an individual document. The policy to which an article belongs is defined as the categorical "author" attribute of the article (document). For example, the Regulations on Bridge Maintenance Management issued by Ministry of Transport in 2007 contains 50 articles, each of which prescribes an independent regulatory measure on BM. They were analyzed as 50 documents in the corpus with the same "author" attribute (i.e. belonging to the same policy). This model configuration not only coincides with the fact that the 50 articles were written by the same group of policy makers, but also makes it convenient to find the major concerns of each policy based on the model results (see Section 4 for more details).

Third, each author (policy) is concerned with a collection of topics (z) and represented by a probability distribution over topics (θ). This coincides with the fact that each policy elaborates on several key aspects of BM. Fourth, each topic is represented by a probability distribution over keywords (φ) describing the underlying meaning of the topic. By this four-level model configuration, ATM imitates the generative processes of real-word policy articles in a simplified manner described as follows (see Fig. 2 for a graphical illustration):

- 1. For each article d in the corpus, find the policy (x) to which the article (d) belongs, so that the topic distribution of the article (d) is determined by the topic distribution of the policy (x).
- Sample a topic from the multinomial distribution of topics in policy
 (x) z|x~Multinomial (θ_x). The topic distribution parameter θ_x of policy

Descriptive analyses on the dataset were conducted with an R

¹ http://www.sogou.com/labs/dl/w.html/

Transport Policy xxx (xxxx) xxx-xxx

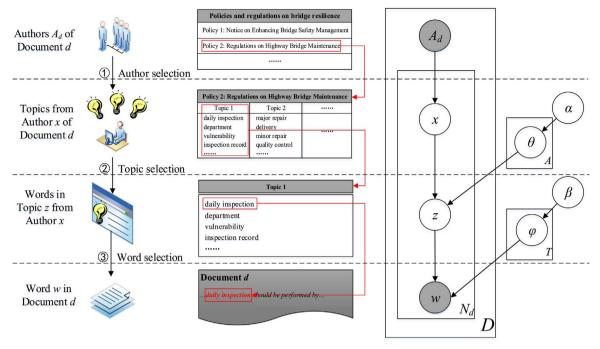


Fig. 2. The generative process of ATM.

(*x*) is sampled from the Dirichlet prior distribution $\theta_x \sim Dir(\alpha)$.

Sample a word from the multinomial distribution of words in topic (z) w|z~Multinomial (φ_z), The word distribution parameter φ_z of topic (z) is sampled from the Dirichlet prior distribution φ_z~Dir(β).

In this way, a word w in article d of policy x is generated. By repeating the steps 1–3, article d can be generated word by word, and, in the same vein, the whole corpus can be generated. Based on the probabilistic procedure illustrated in Fig. 2, the probability of generating word w in article $d_m(m=1 \sim D)$, where D denotes the total number of articles in the corpus) from topic z can be calculated as:

$p(w, z, \theta_{x_m}, \varphi_{z}\alpha, \beta, A_{d_m}) = p(w\varphi_{z})p(\varphi_{z} \mid \beta)p(z \mid \theta_{x_m})p(\theta_{x_m} \mid \alpha)p(x_m \mid A_{d_m})$

Since each document (article) belongs to a unique author (policy), the probability $p(x_mA_{d_m}) = 1$. Meanwhile, the word w can be generated by more than one topic, so the total probability of generating word w in the *n*th position of article d_m is obtained by marginalizing the latent topics z_i and omitting the hyperparameters α and β :

$$p(w_{m,n} = w\theta, \varphi, x_m) = \sum_{j=1}^{T} p\left(w\varphi_{z_j}\right) p(z_j \mid \theta, x_m)$$

The likelihood of article d_m is given by (where N_d denotes the length of d_m):

$$p(d_m\theta, \varphi, x_m) = \prod_{n=1}^{N_d} p(w_{m,n}\theta, \varphi, x_m)$$

Based on this, the complete-data likelihood of the whole corpus D is given by:

$$p(D\Theta, \Phi, A_D) = \prod_{m=1}^{D} p(d_m\theta, \varphi, x_m) = \prod_{m=1}^{D} \prod_{n=1}^{N_d} p(w_{m,n}\theta, \varphi, x_m)$$

where matrix Θ and Φ specify the parameters of the probability distribution of topics and keywords; A_D specifies the corresponding relationships between the articles and the policies.

By maximizing the complete-data likelihood, we can obtain the estimation of Θ and Φ . The distribution of $z \mid x_m$ given by Θ indicates the proportion of topics in policy x_m , and reflects the key aspects of BM in policies. The distribution of $w \mid z$ given by Φ indicates the keywords of

topic *z*, and reflects the underlying meaning of the topics. Using Bayes formula, the distributions of $x_m \mid z$ and $d_m \mid z$ can also be calculated to identify the representative policies and articles focusing on topic *z*. With these results, we can achieve the research objective of identifying the key aspects of BM in policies.

The above likelihood maximization is a high dimensional optimization problem. Traditional variation expectation maximization (VEM) method is slow to converge with high computation complexity. Griffiths and Steyvers (2004) proposed a Gibbs Sampling algorithm to estimate topic models, which substantially improved the computation efficiency and addressed the overfitting problem. Steyver et al. (2004) further applied the Gibbs Sampling method to ATM and built a MATLAB package to implement the algorithm. This method has been widely adopted in subsequent studies on bibliometric analyses (Steyvers and Smyth, 2004), computer software codes (Linstead et al., 2007) and Twitter documents (Morchid et al., 2015). These studies suggest that, by considering the author attribute of the documents, ATM significantly outperformed traditional topic models.

In this light, we utilized ATM and the MATLAB package developed by Steyvers et al. (2004) to analyze the dataset in this study. As described above, the outputs of the ATM include the distribution of words in topics ($w \mid z$), the distribution of topics in articles ($z \mid d_m$), the distribution of topics in policies ($z \mid x_m$), and the representative articles ($d_m \mid z$) and policies for topics ($x_m \mid z$). We can infer the underlying meaning of topics to reflect key aspects of BM, based on these results. This is a relatively subjective process involving the judgments of researchers. However, with the information on the representative policies and articles, ATM makes the process more objective and precise than traditional topic models (Steyvers and Smyth, 2004; Morchid et al., 2015).

3.3. Time-series analysis on topic prevalence (step 5 and 6)

The prevalence of each topic in each year can be calculated as the proportion of words, corresponding to that topic, in the policies issued in that year (Steyvers and Smyth, 2004). Based on the ATM results, we can derive the topic prevalence data in each year and analyze the topic trend by time-series analysis.

Granger causality analysis is a time-series analysis technique to

Q. Wen et al.

reveal the Granger causality relationship between variables based on vector auto-regression (VAR) model (Granger, 1988). Granger causality model regresses a vector of variables $(X_t, Y_t, Z_t,...)$ on their lagged terms { { $\{X_{t-i}\}_{i=1}^{k}, \{Y_{t-i}\}_{i=1}^{k}, \{Z_{t-i}\}_{i=1}^{k}, ... \}$ (the length of lag k is determined by information criteria). According to the axiom that the cause precedes the effect, the Granger causal relationship between any two variables in the vector (e.g. X and Y) is determined by the following procedure:

- 1. If $\{X_{t-i}\}_{i=1}^{k}$ have significant effects on Y_t , while in turn $\{Y_{t-i}\}_{i=1}^{k}$ do not significantly influence X_t , then X is the Granger cause of Y and vice versa.
- 2. If $\{X_{t-i}\}_{i=1}^k$ have significant effects on Y_t and $\{Y_{t-i}\}_{i=1}^k$ also significantly influence X_t , then X and Y are the Granger cause of each other.
- 3. If both the effects of $\{X_{t-i}\}_{i=1}^{k}$ on Y_t and the effects of $\{Y_{t-i}\}_{i=1}^{k}$ on X_t are insignificant, then there is no causal relationship between X and Y.

In this study, we performed Granger causality analyses on the time series data of topic prevalence and the number of bridge collapse accidents to explore how the regulatory concerns on specific aspects influence and are influenced by bridge collapse accidents.

4. Results and discussions

4.1. Descriptive analyses on BM related policies in China

237 out of the 263 retrieved policies are issued by local governments (and others are issued by the central government). Fig. 3A illustrates how these policies are distributed across the 31 provinces/ municipalities in mainland China. Fig. 3B illustrates the number of bridges in the provinces/municipalities, according to the data in Yearbook of China Transportation & Communications (2014). As shown in the figures, there are more bridges in southeast provinces where there are also more policies on BM. The Pearson correlation (0.501 with significant level p < 0.05) between the number of policies and the number of bridges also suggests that provinces with more bridges also tend to issue more policies on BM. Fig. 3C shows the total number of policies on BM in mainland China over the recent 30 years. There has been a rapid growth in the number of policies, especially in



Fig. 4. Wordcloud of key words in BM policies.

the recent 10 years, which reflects the growing awareness on the importance of BM.

By transforming the policy texts into document-term matrix, we identified the words with the highest frequency and visualize them by a wordcloud. Fig. 4 illustrates the 100 most frequent words in the corpus. As shown in Fig. 4, besides the keywords for retrieve (e.g. "bridge", "management", "maintenance"), "road", "department", "work", "urban", "engineering", "facility" and "accident" are the most frequently used words in policies on BM.

4.2. Key aspects of BM

The number of topics should be determined before running ATM,

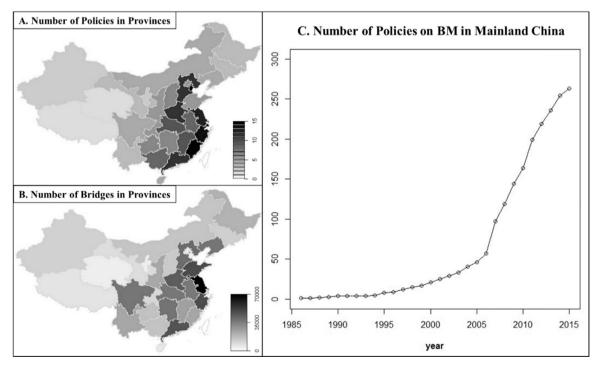


Fig. 3. Bridges and policies on BM in mainland China. 5

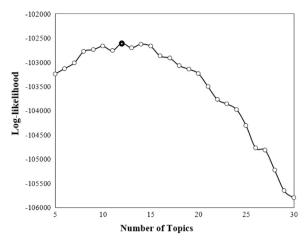


Fig. 5. Log-likelihood of models with different number of topics.

and a common approach is to choose the number of topics with the highest likelihood. We ran ATMs with 5–30 topics, and found that the maximum likelihood is achieved by the model with 12 topics. Based on the keywords of the 12 topics (Fig. 6), we made preliminary interpretation of their underlying meaning. However, as shown in Fig. 5, the likelihood of the models with 10–15 topics are similar, indicating that these models are also good candidates for modeling the intellectual structure of the corpus. As suggested by Jiang et al. (2016), manual interpretation and inspection should be performed to check if the model with the highest likelihood in probability space is also meaningful in semantic space. Thus, we carried out two semantic tests on the 12-topic model (Jiang et al., 2016). First, we checked if the topic names coincide with the representative articles and policies. Second, we

checked if the topic distribution in articles is consistent with the content of the articles.

Fig. 7 is an illustrative example of semantic test on a representative article of topic 1 bridge maintenance and safety responsibility system, which is from the Regulations on Jiangsu Jiangyin Highway Bridge. By reading it manually, we found that the article is mainly about bridge administration and safety responsibility system and largely coincide with topic 1. The proportion of topics also coincide with the content, which mainly focuses on the responsibility system (topic 1), and relates to bridge department configuration (topic 6) and highway bridge management (topic 7). Similar approach was also applied to the five most representative articles of all the 12 topics. The results of the tests indicate strong consistency between the article contents and the topic names, albeit some minor inconsistency, based on which we modified the names of the topics. The semantic test suggests that the 12-topic model has sufficient explanatory power to model the intellectual structure of the corpus. After this test and the modification process, the topic names were finally determined as shown in Fig. 6.

The 12 topics reflect 12 key aspects of BM. According to Fig. 6, the proportion of topics are similar in the corpus, indicating that these key aspects gain similar attention in BM policies. Some aspects aim to guide various bridge management practices and directly influence bridge conditions. Others are *indirect measures* that do not have material effects on bridge conditions but are indispensable underpinnings for the direct measures. Among the direct measures, there are *proactive measures* aiming at preventing hazards ex-ante and *reme-dial measures* aiming at eliminating hazards ex-post by accident response, repair and rehabilitation. According to their keywords and representative policies, the 12 key aspects of BM are discussed in detail as follows:

TOPIC 1 Bridge maintenance	0.08	TOPIC 2 Bridge accident	0.07	TOPIC 3 Bridge maintenance	0.09	TOPIC 4 Bridge maintenance	0.09
and safety responsibility system	0.08	response plan		fund management	0.09	technical standard	0.09
Administration	0.39			Road	0.67	Bridge	0.37
Work	0.11	Department	0.07	Agency	0.06	Management	0.16
Institution	0.06	Bridge	0.06	Toll	0.04	Technical condition	0.06
Safety	0.06	Information	0.03 Investment 0.04 Plan		Plan	0.04	
Organization	0.05	Group	0.03	Regulation	0.03	Standard	0.03
Laws and regulations	0.04	Response plan	0.03	Development	0.02	Agency	0.03
MOT on bridge maintenance	0.05	Guilin urban bridge accident	0.06	MOT on bridge maintenance	0.07	Jiangxi bridge maintenance	0.04
Zhejiang DOT on improving	0.03	Hebei major bridge accident	0.05	Zhejiang DOT on bridge	0.02	Guangxi bridge maintenance	0.03
Hubei DOT on bridge operation	0.02	Hebei urban bridge accident	0.05	Guangxi bridge maintenance	0.02	Zhejiang DOT on bridge	0.03
Taizhou bridge maintenance	0.02	Tibet bridge accident response	0.05	Jiangsu rural road operation	0.02	Hubei DOT on bridge	0.03
Guangxi bridge maintenance	0.02	Gansu DOHURD on bridge	0.05	Shanghai rural road	0.02	Shanghai bridge maintenance	0.03
TOPIC 5 Rural bridge	0.10	TOPIC 6 Bridge maintenance	0.08	TOPIC 7 Highway bridge		TOPIC 8 Bridge safety hazard	
management and maintenance	0.10	department configuration		management	0.06	examination and control	0.08
Engineering	0.16	Department	0.30	Tunnel	0.17	Safety hazard	0.13
Construction	0.11	Administration	0.22	Highway bridge	0.15	Work	0.11
Rural	0.11	Regulation	0.07	Operating system	0.09	Event	0.06
Delivery	0.08	Road	0.04	Area	0.07	Critical project	0.06
Project	0.07	Authority	0.03	Agency	0.06	Production	0.05
Quality	0.07	Transport management	0.03	Sign	0.06	Development	0.05
MOT on rural road management.	0.04	Jiangmen municipal facility	0.03	Nanjing Yangzi river bridge	0.04	Hubei DOT on bridge safety	0.04
Jiangsu Department of water	0.04	Guilin urban bridge regulation	0.03	Xiamen highway bridge	0.02	Beijing bridge safety accident	0.04
Jiangsu rural road operation	0.02	Shanghai urban bridge	0.03	Hubei DOT on bridge safety	0.02	Anhui bridge safety	0.03
Hubei DOT on bridge safety	0.02	Jiangsu rural road operation	0.02	Guizhou highway management	0.02	MOT on bridge maintenance	0.02
Shanghai rural road management	0.02	Nanjing Yangzi river bridge	0.02	Hangzhou highway bridge	0.02	Inner Mongolia DOT on bridge.	0.02
TOPIC 9 Urban bridge	0.08	TOPIC 10 Bridge commissioning	0.09	TOPIC 11 Bridge engineer	0.08	TOPIC 12 Bridge archive and	0.10
inspection and repair	0.00	and transfer	0.09	responsibility system	0.00	information management	0.10
Bridge	0.38	Facility	0.20	Work	0.17	Bridge	0.48
Urban	0.36	Authority	0.11	Bridge	0.14	Department	0.16
Repair	0.05	Municipal bridge	0.08	Technical condition	0.14	Supervision	0.08
Evaluation	0.03	Transfer	0.08	Engineer	0.14	Structure	0.04
Inspection	0.02	Department	0.07	Staff	0.06	Archive	0.04
Management	0.02	Construction	0.05	Responsibility	0.05	Profession	0.02
Hubei urban bridge inspection	0.04	Jiangmen municipal facility	0.08	Zhejiang DOT on bridge	0.03	Taizhou bridge maintenance	0.04
Taiyuan urban bridge regulation	0.04	Regulation on urban bridge	0.04	Guilin urban bridge regulation	0.03	Zhejiang DOT on bridge	0.04
Zibo bridge inspection and	0.03	Dalian urban road and bridge	0.03	Hubei DOT on bridge safety	0.03	Zhejiang DOT on improving	0.04
Jiangmen municipal facility	0.03	Chengdu urban facility	0.03	Jiangxi bridge maintenance	0.03	Shanghai urban bridge	0.03
Changsha urban bridge safety	0.02	Nanning urban bridge	0.03	Shanghai bridge maintenance	0.02	Jiangxi bridge maintenance	0.03

Fig. 6. Results of the 12-topic model. Note: Only the most relevant six key words (in bold) and five representative policies (in italic) are presented.

Q. Wen et al.

Transport Policy xxx (xxxx) xxx-xxx

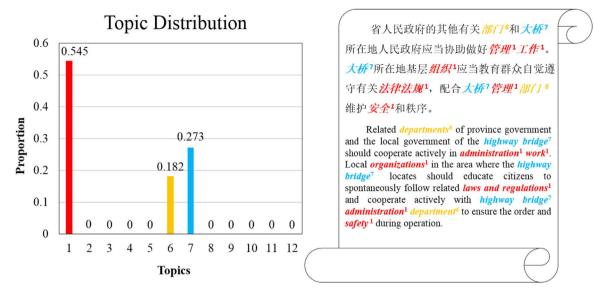


Fig. 7. An illustrative example of semantic test.

- 1. Bridge maintenance and safety responsibility system clarifies the responsibility allocation structure, and is a supportive indirect measure. For example, *Zhejiang Department of Transport (DOT)* on improving bridge safety responsibility system is a representative policy of this topic. It proposes to establish bridge maintenance and safety management responsibility system with the principles of "authority-obligation conformity" and "unified leadership upon multiple levels of administration".
- 2. *Bridge accident response plan* specifies what measures the group of government departments should take to integrate various information and organize accident response activities. The representative policies of this topics are all detailed accident response plans aimed at eliminating the effects of accidents as soon as possible. Thus, this aspect belongs to remedial measures.
- 3. Bridge maintenance fund management involves obtaining funds for bridge maintenance by government investment or bridge toll. Since it provides bridge maintenance activities with sufficient resources, it is an important supportive indirect measure. The representative policy *MOT* on bridge maintenance management points out the problem of "maintaining road without maintaining bridge". It prescribes that the maintenance of both roads and the bridges, on which the roads locate, should be supported with sufficient funds.
- 4. *Bridge maintenance technical standard* specifies bridge maintenance plan, bridge maintenance project management and bridge technical condition re-evaluation process after maintenance. It is an important proactive measure. For example, *Jiangxi bridge maintenance management standard* stipulates that the bridge administration department should check if the technical condition of bridge reaches grade I or II after major repair, minor repair and retrofit projects.
- 5. *Rural bridge management and maintenance* emphasizes on carefully managing the quality of rural bridges during both the construction and the operation phase. In China, small bridges in rural areas account for a large proportion of bridges in bad condition, and the policies on this topic aim to encourage proactive BM measures. For example, *Jiangsu rural road operation management measures* prescribe that local bridge administration departments should take preventative measures and cooperate actively with rural residents to ensure the good operation condition of rural bridges.
- 6. Bridge maintenance department configuration specifies the department in charge of bridge maintenance and management, and

belongs to supportive indirect measures. For example, *Jiangmen municipal facility management regulation* clarifies that bridges in urban areas should be maintained by the administration department of the city, while the maintenance of bridges in other areas should be appointed to local administration agencies.

- 7. Highway bridge management involves special measures to ensure the safety of highway bridges and belongs to proactive measures. For example, Xiamen highway bridge and tunnel management measures prescribe that administration agencies should delimit protected areas, place warning signs and establish special operation management systems for highway bridges.
- 8. Bridge safety hazard examination and control is a proactive measure to prevent bridge safety hazards from various sources and enhance bridges' ability to maintain their functions. For example, *Hubei DOT on bridge safety hazard examination and elimination plan* focuses on examining potential safety hazards in construction technique, construction quality and construction materials.
- 9. Urban bridge inspection and repair aims to identify problems during operation and bring bridges back to their original conditions by timely repair. It is a remedial measure. For example, *Taiyuan urban bridge management regulation* specifies that bridge inspection and repair should be planed and conducted regularly.
- 10. Bridge commissioning and transfer includes the evaluation, commissioning and transfer of bridges, and belongs to supportive indirect measures. For example, *Dalian urban road and bridge management measures* formulate detailed specifications on bridge technical condition evaluation, commissioning, transfer and repair warranty.
- 11. Bridge engineer responsibility system requires bridge administration departments to appoint an experienced bridge engineer to host bridge inspection, repair and archive management processes. It is a supportive indirect measure. For example, *Zhejiang DOT on bridge maintenance and management* specifies the qualification requirements and the responsibilities of bridge engineers.
- 12. Bridge archive and information management system includes comprehensive records on bridges from construction to operation. It is a supportive indirect measure to enable better management of bridge information. For example, *Taizhou bridge maintenance and management measures* prescribe that an individual archive, including information on construction, commissioning, repair and operation condition, should be established for each bridge.

Transport Policy xxx (xxxx) xxx-xxx

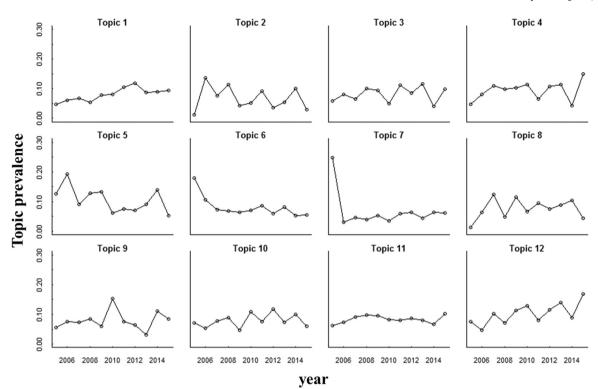


Fig. 8. Topic prevalence trends during 2005-2015.

4.3. Topic prevalence trend analysis

The proportion of words accounted by each topic in policies issued in a specific year was calculated to reflect the prevalence of the topic in that year. The prevalence of the 12 topics during 2005–2015 was calculated, and the topic prevalence trends are illustrated in Fig. 8.

According to Fig. 8, the prevalence of some topics has been increasing (e.g. *bridge maintenance and safety responsibility system* and *bridge archive and information management*). This reflects the growing emphasis on clarifying safety responsibility (Fang et al., 2004; Ismail et al., 2012) and achieving the life-cycle management of bridges based on comprehensive information (Ma et al., 2009; Xie et al., 2014). The prevalence of some topics has been diminishing (e.g. *bridge maintenance department configuration*). This may be ascribed to the fact that bridge maintenance department configuration has been more and more well-defined (Ge and Xiang, 2011).

We also explored the relationship between topic prevalence and the number of bridge collapse accident events by Granger causality analyses, based on the data during 1999–2013 shown in Table 1 (data source: Liu, 2013). The number of bridges in poor operation condition is also included in the vector auto-regression (VAR) model to control the bridge condition factor, which may also influence the number of bridge collapse accidents (data source: National Bureau of Statistics of China). According to the information criteria, all the time lags (*k* in $\{X_{i-i}\}_{i=1}^{k}, \{Y_{i-i}\}_{i=1}^{k}\}$ of the 12 causality test models were determined (as shown in Table 2). The *t*-values of the lagged terms' regression coefficients are listed in Table 2, where the arrows in the first column denote the direction of causality.

Table 1 The number of bridge collapse accident events during 1999–2013.

Year	1999	2000	2001	2002	2003	2004	2005	2006
Number of accidents Year Number of accidents	2007	2008	11 2009 34	9 2010 39	,	17 2012 25	21 2013 15	28

The emphases on some aspects of BM (including 2. Bridge accident response plan, 9. Urban bridge inspection and repair and 11. Bridge engineer responsibility system) grows immediately (within 1 period of time lag) with the increase in bridge accidents. The evolution of the emphasis on these three aspects reflects the event-driven tendency in policy making (Birkland, 1998). It is also interesting to find that two of the three aspects are remedial measures, as discussed in Section 4.2. For example, after the bridge collapse accident in Fenghuang county of Hunan province in Aug 13 2007, 11 provinces issued policies to enhance bridge inspection and conduct timely repair (corresponding to topic 9). However, to our surprise, the emphases on these BM measures do not in turn significantly reduce the number of accidents. These results suggest that the post-hoc introduction of new remedial measures is not associated with a significant reduction in accidents for some reasons, for example, these measures may be not easily enforced consistently in a relatively long time span (Liu, 2013).

The emphases on four topics (1. Bridge maintenance and safety responsibility system, 4. Bridge maintenance technical standard, 5. Rural bridge management and maintenance and 8. Bridge safety hazard examination and control) are associated with significant reduction in bridge accidents. Three of the four aspects are proactive BM measures as discussed in Section 4.2. The only exception is topic 1 (a supportive indirect measure), which operationalizes the BM into responsibility systems. According to the results, these four aspects turn out to be conducive to addressing the problems in practice, and hence, to some extent, contributes to accident prevention. However, there are relatively long time lags (above 2 periods) before the effects of the four aspects take place. These time lags may make it harder for policy-makers to evaluate and recognize their long term effects, and this partly explains why the four aspects are not immediately emphasized in response to the increase in accidents.

Taken together, there is a mismatch between the BM aspects that draw immediate regulatory attention and the aspects that contribute to accident prevention in a relatively long term. The findings suggest that emphasizing on remedial measures immediately after accidents is intuitive and important, while proactive accident prevention measures should also be taken to fix the root cause of accidents.

Table 2

Granger causality analysis results.

Topics	1	2	3	4	5	6	7	8	9	10	11	12
Event→Topic _i	-0.83	2.24 [*]	-1.03	0.44	0.93	1.36	1.66	0.28	2.68	-0.14	2.60	-1.58
Event←Topic _i	- 2.5 7	-0.81	-0.34	- 2.58 *	- 2.52	-1.30	0.01	- 2.84 [*]	-0.46	-0.47	-0.22	-1.51
Time lag	2	1	3	3	2	2	2	3	1	2	1	3

* denotes *p* < 0.05.

5. Conclusions and limitations

This study provides a systematic review of BM policies in mainland China. It is of both theoretical and practical implications. The revealed 12 topics in policies reflect 12 aspects of BM, and correspond to indirect, proactive and remedial BM measures. The topic prevalence trend analysis results indicate that there is a growing emphasis on bridge safety responsibility and management information system, while the regulatory attention on bridge maintenance department configuration has been diminishing. These results reveal the general evolution trends of regulatory concerns on BM in China. The Granger causality analysis results further indicate that regulatory emphases on proactive measures contribute to accident reduction, however, in relatively longer terms. These findings remind policy makers that although it is straightforward and essential to emphasize on remedial measures in response to accidents, they should also pay long-term regulatory attention on proactive BM measures, which are the fundamental solutions to accident reduction.

The findings of this study should be viewed with respect to its limitations, which leave the door open for future studies. First, since the dataset were retrieved manually, there may be some policies omitted by the collection of 263 policies. In fact, even with more efforts on data collection, there inevitably may be some omitted policies, which are within the interest of specific users. We plan to address this problem in future studies by developing an ATM policy mining toolkit with the function to import policy documents from users. In this way, policy makers can analyze the policies they are interested in and fully incorporate their own expertise in the analysis process. Second, the findings are based on the Chinese context, and the method can only be applied to Chinese policy documents at current stage. Future efforts will be devoted to generalize the method to other countries and enable the analyses on policy documents written in other languages. To overcome the limitations of Author-topic Model in analyzing multiple languages, some recently developed deep learning based natural language processing methods (e.g. distributed representations of words) can be utilized. Third, the Granger causality analysis model is relatively simple. Only the topic prevalence value is used to reflect the political emphasis on a specific aspect of BM, while the type of policy is not considered. For example, emphasizing on a BM measure in superordinate policies may be more influential than emphasizing on it in specific regulations. The model simplicity herein serves the exploratory research purpose of this study, but future studies are needed to more rigorously and systematically evaluate the effectiveness of different BM policies.

Acknowledgements

The present study was funded by the National Natural Science Foundation of China (General Programs 51179086, 51479100, 51379104) and the grants (2013-KY-5, 2015-KY-5) from State Key Laboratory of Hydroscience and Engineering.

References

Agbelie, K., Labi, S., Sinha, K.C., 2017. Estimating the marginal costs of bridge damage due to overweight vehicles using a modified equivalent-vehicle methodology and inservice data on life-cycle costs and usage. Transp. Res. Part A: Policy Pract. 95, 275-288.

- Andrić, Jelena M., Lu, Da-Gang, 2016. Risk assessment of bridges under multiple hazards in operation period. Saf. Sci. 83, 80–92.
- Andrijcic, E., Haimes, Y.Y., Beatley, T., 2013. Public policy implications of harmonizing engineering technology with socio-economic modeling: application to transportation infrastructure management. Transp. Res. Part A: Policy Pract. 50, 62–73.
- ASCE/SEI-AASHTO Ad-Hoc Group On Bridge Inspection, Rating, Rehabilitation, and Replacement, 2009. White paper on bridge inspection and rating. J. Bridge Eng. 14 (1), 1–5.
- Birkland, T.A., 1998. Focusing events, mobilization, and agenda setting. J. Public Policy 18 (01), 53–74.
- Bocchini, P., Frangopol, D.M., Ummenhofer, T., Zinke, T., 2013. Resilience and sustainability of civil infrastructure: toward a unified approach. J. Infrastruct. Syst. 20 (2), 04014004.
- Boin, Arjen, McConnell, Allan, 2007. Preparing for critical infrastructure breakdowns: the limits of crisis management and the need for resilience. J. Contingencies Crisis Manag. 15 (1), 50–59.
- Bruneau, M., Chang, S.E., Eguchi, R.T., et al., 2003. A framework to quantitatively assess and enhance the seismic resilience of communities. Earthq. Spectra 19 (4), 733–752.
- Chen, Ju, 2014. Analyses on the cause of bridge accident and preventative strategy: a systematic perspective based on the statistics of bridge accidents in 1999–2013. Acad. Forum 37, 40–45.
- Chow, H.F., Szeto, W.Y., David, Z.W., Travis, S., 2015. Quantitative approaches to resilience in transport networks. Transp. A: Transp. Sci. 11 (9), 751–753.
- Cimellaro, G.P., Reinhorn, A.M., Bruneau, M., 2010. Framework for analytical quantification of disaster resilience. Eng. Struct. 32 (11), 3639–3649.
- Coglianese, Cary, 2004. Information technology and regulatory policy new directions for digital government research. Soc. Sci. Comput. Rev. 22 (1), 85–91.
- Dai, Kaoshan, Benjamin, H. Smith, Chen, Shen-En, Sun, Limin, 2014. Comparative study of bridge management programmes and practices in the USA and China. Struct. Infrastruct. Eng. 10 (5), 577–588.
- Deng, Xiaomei, Wang, Yuhong, 2015. Special issue on transport infrastructures to climate change. Transp. Policy.
- Dong, You, Dan, M. Frangopol, 2015. Risk and resilience assessment of bridges under mainshock and aftershocks incorporating uncertainties. Eng. Struct. 83, 198–208.
- Durango, P.L., Madanat, S.M., 2002. Optimal maintenance and repair policies in infrastructure management under uncertain facility deterioration rates: an adaptive control approach. Transp. Res. Part A: Policy Pract. 36 (9), 763–778.
- Fang, D.P., Huang, X.Y., Hinze, Jimmie, 2004. Benchmarking studies on construction safety management in China. J. Constr. Eng. Manag. 130 (3), 424–432.
- Ge, Yaojun, Xiang, Haifan, 2011. Concept and requirements of sustainable development in bridge engineering. Front. Archit. Civil. Eng. China 5 (4), 432–450.
- Granger, C.W.J., 1988. Some recent development in a concept of causality. J. Econometrics 39 (1-2), 199-211.
- Griffiths, T.L., Steyvers, M., 2004. Finding scientific topics. Proc. Natl. Acad. Sci. USA 101 (Suppl. 1), 5228–5235.
- Herrmann, A.W., 2013. "ASCE 2013 Report Card for America's Infrastructure." IABSE Symposium Report, 99(33): 9-10.
- Hang, W., Xie, Y., He, J., 2013. Practices of using weigh-in-motion technology for truck weight regulation in China. Transp. Policy 30, 143–152.
- Ikpong, A., Bagchi, A., 2016. Managing highway bridges against climate-triggered extreme loading in cold regions. J. Infrastruct. Syst. 7, 04016026.
- Ismail, Zubaidah, Doostdar, Samad, Harun, Zakaria, 2012. Factors influencing the implementation of a safety management system for construction sites. Saf. Sci. 50 (3), 418–423.
- Jiang, Hanchen, Qiang, Maoshan, Lin, Peng, 2016. Finding academic concerns of the Three Gorges Project based on a topic modeling approach. Ecol. Indic. 60, 693–701.
- Linstead, Erik, Paul Rigor, Bajracharya, Sushil, Lopes, Cristina, Baldi, Pierre, 2007. Mining eclipse developer contributions via author-topic models. Paper presented at the Mining Software Repositories, 2007. ICSE Workshops MSR'07.
- Liu, Min, Dan, M. Frangopol, 2005. Bridge annual maintenance prioritization under uncertainty by multiobjective combinatorial optimization. Comput.-Aided Civil. Infrastruct. Eng. 20 (5), 343–353.
- Liu, Meiming, 2013. "Analysis of Bridge Accidents." MS thesis. Southwest Jiaotong University.
- Ma, Junhai, Chen, Airong, He, Jun, 2009. General framework for bridge life cycle design. Front. Archit. Civil. Eng. China 3 (1), 50–56.

Marsden, Greg, Stead, Dominic, 2011. Policy transfer and learning in the field of transport: a review of concepts and evidence. Transp. Policy 18 (3), 492–500.

- Mattsson, Lars-Göran, Jenelius, Erik, 2015 Vulnerability and resilience of transport systems: A discussion of recent research.
- McDaniels, Timothy, Chang, Stephanie, Cole, Darren, Mikawoz, Joseph, Longstaff, Holly, 2008. Fostering resilience to extreme events within infrastructure systems: characterizing decision contexts for mitigation and adaptation. Glob. Environ.

Q. Wen et al.

Change 18 (2), 310-318.

- Meyer, David, Hornik, Kurt, Feinerer, Ingo, 2008. Text mining infrastructure in R. J. Stat. Softw. 25 (5), 1–54.
- Morchid, Mohamed, Portilla, Yonathan, Josselin, Didier, Dufour, Richard, Altman, Eitan, El-Beze, Marc, Cossu, Jean-Valère, Linarès, Georges, Reiffers-Masson, Alexandre, 2015. An author-topic based approach to cluster tweets and mine their location. Proc. Environ. Sci. 27, 26–29.
- National Bureau of Statistics of China (NBSC), 2015 (http://www.stats.gov.cn/tjsj/ndsj/) (Jan. 29, 2016).
- Orcesi, A.D., Cremona, C.F., 2010. Optimization of maintenance strategies for the management of the national bridge stock in France. J. Bridge Eng. 16 (1), 44–52.
- Prosser, Brenton, Peters, Colin, 2010. Directions in disaster resilience policy. Aust. J. Emerg. Manag. 25 (3), 8. Rattanachot, Wit, Wang, Yuhong, Chong, Dan, Suwansawas, Suchatvee, 2015.
- Adaptation strategies of transport infrastructures to global climate change. Transp. Policy.
- Steyvers, Mark, Smyth, Padhraic, Rosen-Zvi, Michal, Griffiths, Thomas, 2004. "Probabilistic author-topic models for information discovery". Paper presented at In: Proceedings of the tenth ACM SIGKDD international conference on Knowledge

discovery and data mining.

- Sweeney, P., Unsworth, J.F., 2010. Bridge inspection practice: two different North American railways. J. Bridge Eng. 15 (4), 439–444.
- Thompson, P.D., Sobanjo, J.O., Kerr, R., 2003. Florida DOT project-level bridge management models. J. Bridge Eng. 8 (6), 345–352.
- Washer, G., Connor, R., Nasrollahi, M., Provines, J., 2016. New framework for risk-based inspection of highway bridges. J. Bridge Eng. 21 (4), 04015077.
- Xie, Hui-Bing, Wang, Yuan-Feng, Wu, Han-Liang, Li, Zheng, 2014. Condition assessment of existing RC highway bridges in China based on SIE2011. J. Bridge Eng, 19 (12), 04014053.
- Pan, Xuewu, Zhongyong, Hua, 2013. Reinforcement, reconstruction and maintenance of old highway bridge in rural area. Urban Roads Bridges Flood Control 5, 051.
- Yin, Zi-hong, Li, Yuan-fu, Guo, Jian, Li, Yan, 2011. Integration research and design of the bridge maintenance management system. Proc. Eng. 15, 5429–5434.
- Zhu, S., Levinson, D., Liu, H.X., Harder, K., 2010. The traffic and behavioral effects of the I-35W Mississippi River bridge collapse. Transp. Res. Part A: Policy Pract. 44 (10), 771–784.
- Zirn, Cäcilia, Stuckenschmidt, Heiner, 2014. Multidimensional topic analysis in political texts. Data Knowl. Eng. 90, 38–53.