Data Article

Complete bibliographic data, cluster assignments and combined citation network of emergency response operations research extant literature

J.P. Minas\textsuperscript{a}, N.C. Simpson\textsuperscript{b,c,*}, Z.Y. Tacheva\textsuperscript{d}

\textsuperscript{a} School of Business, Ithaca College, 953 Danby Rd, Ithaca NY, 14850, USA
\textsuperscript{b} Department of Operations Management and Strategy, School of Management, University at Buffalo (SUNY), 351 Jacobs Management Center, Buffalo, NY 14260-4000, USA
\textsuperscript{c} Stephen Still Institute for Sustainable Transportation and Logistics, University at Buffalo (SUNY) USA
\textsuperscript{d} School of Information Studies, Syracuse University, 900 South Crouse Ave Syracuse, NY 13244 USA

\textbf{A R T I C L E   I N F O}

\textbf{Article history:}
Received 21 May 2020
Revised 14 June 2020
Accepted 17 June 2020
Available online 21 June 2020

\textbf{Keywords:}
Disaster management
Emergency response
Literature sample
Citation analysis
Bibliographic network
Edge file

\textbf{A B S T R A C T}

The Emergency Response Operations Management (EROM) Literature Sample is a collection of 644 papers winnowed from over 5,000 related articles through application of a binary classification tree, collecting the state-of-the-art in decision models of emergencies in progress. References are scraped from each of these 644 publications, to create a dataset describing a total of 14,821 papers linked by 23,175 citation relationships, the analysis of which is presented in, “Modeling Emergency Response Operations: A Theory Building Survey” in \textit{Computers and Operations Research} [1]. Bibliographic research communities are identified within the data set by framing the task of network partitioning as a cluster ensemble problem from machine learning [2]. This data may be used in several ways, including as an extended reference section for [1], as all 644 papers studied could not be individually cited there. Other examples of potential reuse are further study of a particular research cluster, comparative study of the structural characteristics of this bibliographic

\textbf{DOI} of original article: 10.1016/j.cor.2020.104921

* Corresponding author.
\textit{E-mail address:} nsimpson@buffalo.edu (N.C. Simpson).

https://doi.org/10.1016/j.dib.2020.105908

2352-3409/© 2020 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license. (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Specifications table

| Subject | Management Science and Operations Research |
|---------|------------------------------------------|
| Specific subject area | Decision Models for Emergency Response, Citation Analysis |
| Type of data | Tables (Excel) |
| How data were acquired | Queries of Scopus and Web of Knowledge databases (target sample) |
| | Adobe Acrobat and CERMINE (compiling target sample references) |
| | Gephi (network partitioning assignments) |
| | MAXQDA (content analysis for cluster naming) |
| Data format | Raw and analyzed |
| Parameters for data collection | Web of Knowledge: subject area Operations Research and Management Science, collect all published or in-press journal articles marked by topic of crisis, disaster, emergency, or incident. |
| | Scopus: subject area Decision Science, collect all published or in-press journal articles marked by same four topics. |
| Description of data collection | The combined Scopus and Web of Knowledge queries identified 5607 journal articles, each examined in a binary sorting process that extracted 644 models of decision making during emergency response. For each member of the target group, references of older papers were extracted manually, while the references of younger papers were collected using CERMINE’s online Reference Extraction application. A unique integer identifier was assigned to each paper, and its full bibliographic information was compiled into a single text string, for sorting and manual removal of duplicates. |
| Data source location | Global; all sources of operations research published in English, as indexed by the combined reach of Scopus and Web of Knowledge |
| Data accessibility | Repository name: Mendeley Data |
| | Data identification number: ygdrcknxjf.1 |
| | Direct URL to data: http://dx.doi.org/10.17632/ygdrcknxjf.1 |
| Related research article | J.P. Minas, N.C. Simpson, Z.Y. Tacheva, Modeling emergency response operations: A theory building survey, Comput. Oper. Res., 119 (2020). |
| | https://doi.org/10.1016/j.cor.2020.104921. |

Value of the data

• This data is the result of an exhaustive search for published decision models addressing the response phase of an emergency. This literature is not easily detected, as it does not associate with any earmark keywords, and its placement is diffused across 109 different journals.

• Operations researchers interested in the modeling of emergency response can use this data to quickly locate existing literature. Readers of [1] can use this data verify the inclusion and assigned research cluster of any EROM paper known to them, as page space restrictions precluded the citing of all 644 publications in [1].

• Researchers interested in citation analysis may compare network characteristics captured here with citation networks in differing subject areas. Researchers interested in a particular EROM research cluster discussed in [1] can obtain full data on that literature and its shared citations. The combined citation network of the EROM research sample can likewise serve as a testbed for further development of robust cluster detection algorithms.
Table 1
Description of article-level data provided in EROM Literature Sample.xlsx.

| Variable Name (w/ Spreadsheet Location) | Description |
|----------------------------------------|-------------|
| Cluster (Column A) Bibliographic Data (Columns B – E) | Bibliographic cluster assignment [1]. Article author(s), article title, title of journal, and year of publication, in that order. |
| Methodology (Column F) | Operations research methodology used in study; MP= Mathematical Programming, MIL= Mixed Integer Programming; DP= Dynamic Programming |
| Disaster Type (Column G) | Hazard modelled in the article; General=All Hazard/Non-specific Hazard |
| How Many Processes? (Column H) | Count of emergency response processes represented in article’s decision model. |
| What Processes? (Column I) | Inventory of emergency response processes modelled. |
| Objective Function (Column J) | Count of emergency response objectives in decision model. |
| Objective(s) (Column K) | Identification of objective(s) in decision model. Specified object is minimized, unless otherwise stated. |
| Time Horizon (Column L) | Representation of time in decision model. (Single Point= no passage of time; Multiple Stages: Rolling Horizon; Continuous.) |
| Probabilistic Information (Column M) | Type of uncertainty modelled in article, per taxonomy in [2]. (No; Prognostic; Observational; Inferential) |
| Exogenous information (column N) | Modeling of the revelation of additional information over passage of time. (no, adversarial (resulting from hazard), non-adversarial) |
| Transition function (column -O) | Modeling of system state as a function of previous system state, per taxonomy in [2]. (s= stochastic elements, d= deterministic elements, e= exogenous information) |

1. Data description

The spreadsheet file EROM Literature Sample.xlsx contains bibliographic data for the 644 papers featured in [1], identified as decision models related to emergency response. The first column provides each paper’s research cluster assignment, and columns to the right of bibliographic information provide categorizations including each paper’s dominant methodology and an inventory of modeling axioms. Table 1 provides more detail on each data variable.

Complete bibliographic data on 14,821 papers (the EROM Literature Sample and the pooled references of its membership) is provided on the ‘Node Identities’ sheet of the spreadsheet booklet EROM Citation Network- Full Bibliographic Data.xlsx. Each of these papers is assigned a unique integer as the identifier, used in the booklet sheet ‘Citation Network Edges’. ‘Citation Network Edges’ provides the links between the original EROM Literature Sample and this group’s combined set of references. The format is that of an edge file, in which the left column is the source node (EROM Literature Sample paper) and the right column is the target node (reference), describing 23,175 directed relationships between citers and cited. Fig. 1 provides a network visualization of ‘Citation Network Edges’, shaded according to observed research clusters identified through robust cluster analysis. Fig. 2 provides a subgraph of Fig. 1, illustrating these clusters in terms of the 644 EROM Literature Sample papers only, omitting all non-EROM citations. Fig. 3 displays two further subgraphs of Fig. 1.

2. Experimental design, materials, and methods

In February 2019, the subject area of ‘Decision Science’ in Scopus and the subject area of ‘Operations Research and Management Science’ in Web of Knowledge were queried for all articles identified by at least one of four keywords: emergency, disaster, incident and crisis. The 6935 results were manually reviewed to confirm 5607 unique journal articles. These 5607 articles were then split through a binary sorting tree to ultimately identify 644 instances of decision
Fig. 1. The combined citation network of the EROM sample, consisting of 14,821 papers linked by 23,175 citation relationships. Variations in shading among the nodes indicate the observed research clusters.

Fig. 2. The EROM sample network, consisting of the 644 papers of the EROM Literature Sample, linked by 2478 citation relationships within that sample. Variations in shading among the nodes indicate the observed research clusters.
models related to emergency response, defined as decision making addressing an emergency in progress. Fig. 4 displays the binary sorting tree used to identify the sample, as well as the disposition of the balance of 5607 papers initially retrieved through keyword search. These 644 target articles were designated the EROM Literature Sample, and collected as pdf files for the bibliographic network building phase.

Bibliographic network data was obtained for the EROM Literature sample by scraping the reference section of each of the 644 pdfs, and compiling these observations in Microsoft Excel. For many of the older EROM papers, observations were collected through the Export function of Adobe Acrobat, converting the manuscript to spreadsheet format and extracting the reference sections. For the younger half of the sample, we uploaded each manuscript's pdf file on the Content Extractor and Miner platform CERMINE (http://cermine.ceon.pl/index.html) and used its online Reference Extraction application to export each paper's bibliographic references to an Excel spreadsheet. The results of a random subsample were manually spot-checked to verify the application's accuracy. Each unique instance of a publication, be it an EROM sample paper or a
paper cited by the EROM sample, was assigned a unique integer as a short form identity, for use in input files read by network analysis software Gephi (gephi.org).

Robust cluster analysis was conducted through ensemble clustering [3], featuring the partitioning heuristic of Blondel et al. [4], available in Gephi. To detect instances of bibliographic clustering, the network composed of the EROM sample and all papers cited by that set was partitioned ten times from ten randomized starting positions. Bibliographic clusters are defined by a consensus function: papers x and y belong to a single cluster IFF x and y are both contained at an intersection of all ten partition sets. The ten largest clusters identified were agnostically named after the two word phrase that appeared most often within the content of that cluster, provided that rate was three times greater than the EROM sample as a whole. MAXQDA was the content analysis software used to tally the frequency of these phrases.

3. Ethics statement

This research does not involve human subjects or animal experimentation. The authors confirm their commitment to upholding the principles of ethical publishing as described by Elsevier and Data in Brief.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2020.105908.

References

[1] J.P. Minas, N.C. Simpson, Z.Y. Tacheva, Modeling emergency response operations: a theory building survey, Comput. Oper. Res. 119 (2020) https://doi.org/10.1016/j.cor.2020.104921.
[2] W.B. Powell, A unified framework for optimization under uncertainty, in: Aparna Gupta, Agostino Capponi, J. Cole Smith, Harvey J. Greenberg (Eds.), Optimization Challenges in Complex, Networked and Risky Systems, INFORMS, 2016, pp. 45–83. https://doi.org/10.1287/educ.2016.0149.
[3] A. Strehl, J. Ghosh, Cluster ensembles - a knowledge reuse framework for combining multiple partitions, J. Mach. Learn. Res. 3 (2002) 583–617.
[4] V.D. Blondel, J.L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks, J. Stat. Mech. 10 (2008) P10008.