An Integrated Framework for Data-driven Transportation Infrastructure Planning

Sudipta Chowdhury, Department of Civil and Environmental Engineering, University of Connecticut

Advisor: Jin Zhu, Ph.D., Department of Civil and Environmental Engineering, University of Connecticut

Background

➢ Multitude of factors affect transportation infrastructures
➢ These factors may have a coupling effect
➢ These factors and their inter-relationships need to be identified during transportation infrastructure planning
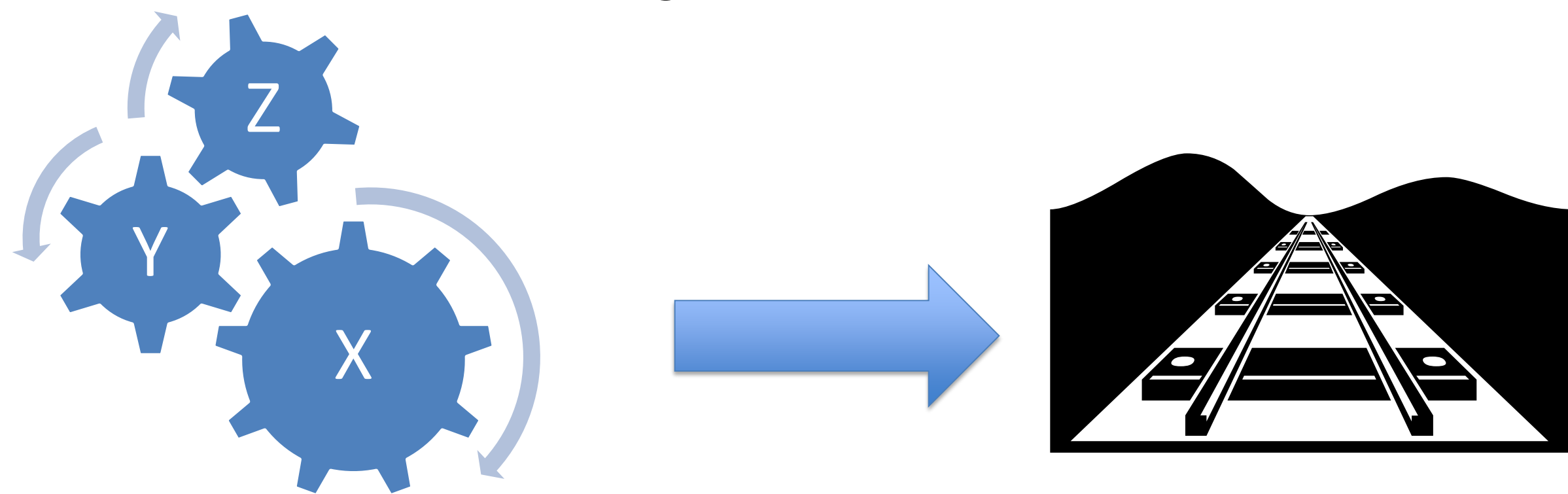

Figure 1. Coupling effect of factors on transportation infrastructures

Challenges

➢ Two primary challenges [1]
  ❖ Identification of factors from various heterogeneous data sources
  ❖ Identifying inter-relationship among the factors

Research Objectives

➢ Initiate the development of an integrated planning framework that can help with future-proofing planning and decision-making of transportation infrastructure

Taxonomy & Association Rule

➢ Taxonomy establishes the hierarchical list of factors that affect transportation infrastructures
➢ Association rule mining captures the inter-relationships among factors
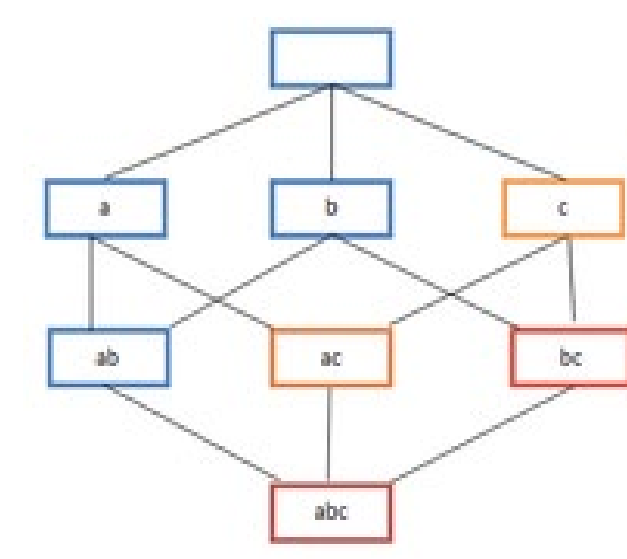

Figure 2. Taxonomy

Figure 3. Association rule

References
[1] Skorobogatova, O., & Kuzmina-Merlino, I. (2017). Transport infrastructure development performance. *Procedia Engineering*, *178*, 319-329.   [2] Khanzadeh, M., et al. (2018). Porosity Prediction: Supervised-learning of Thermal History for Direct Laser Deposition. Journal of manufacturing systems, 47, 69-82.
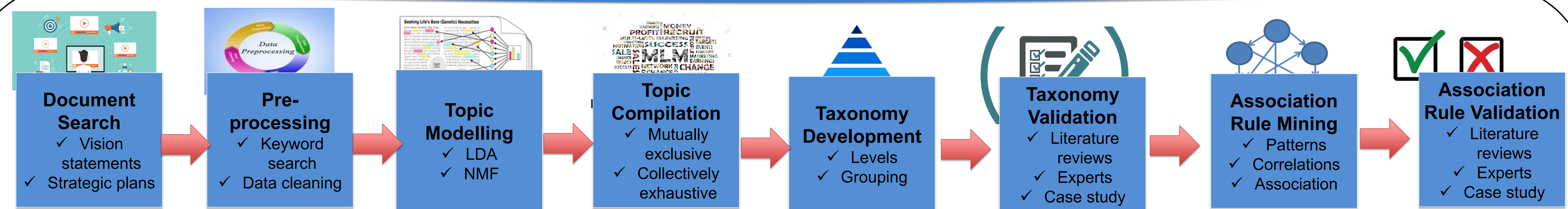
Methodology

Document Search ✓ Vision statements ✓ Strategic plans → **Pre-processing** ✓ Keyword search ✓ Data cleaning → **Topic Modelling** ✓ LDA ✓ NMF → **Topic Compilation** ✓ Mutually exclusive ✓ Collectively exhaustive → **Taxonomy Development** ✓ Levels ✓ Grouping → **Taxonomy Validation** ✓ Literature reviews ✓ Experts ✓ Case study → **Association Rule Mining** ✓ Patterns ✓ Correlations ✓ Association → **Association Rule Validation** ✓ Literature reviews ✓ Experts ✓ Case study

Figure 3. Development steps

Document Search and Pre-processing
➢ 50 transportation strategic planning documents published by transportation agencies were collected

Latent Dirichlet Allocation (LDA) [2]
➢ P(word | topics): Prevalence of word across topics
➢ P(topics |doc): Prevalence of topics in the document

Non-negative Matrix Factorization (NMF) [2]
➢ $X(:,j) \approx \sum_{k \in r} W(:,k) H(k,j)$ with $W \geq 0$ and $H \geq 0$
➢ The goal is to minimize $\sum_{i,j}(X - WH)_{ij}^2$

Topic Compilation
➢ No topics were repeatedly counted
➢ All important topics were included

Taxonomy Development
➢ "Bottom-up" approach was used
➢ Four levels in total were identified

Taxonomy Validation
➢ Expert validation
➢ Compare with literature

Association Rule Mining
➢ Association strength among topics were identified
➢ Lift and confidence measure were calculated

Association Rule Validation
➢ Expert validation
➢ Compare with literature

Results

Figure 4. Proposed taxonomy

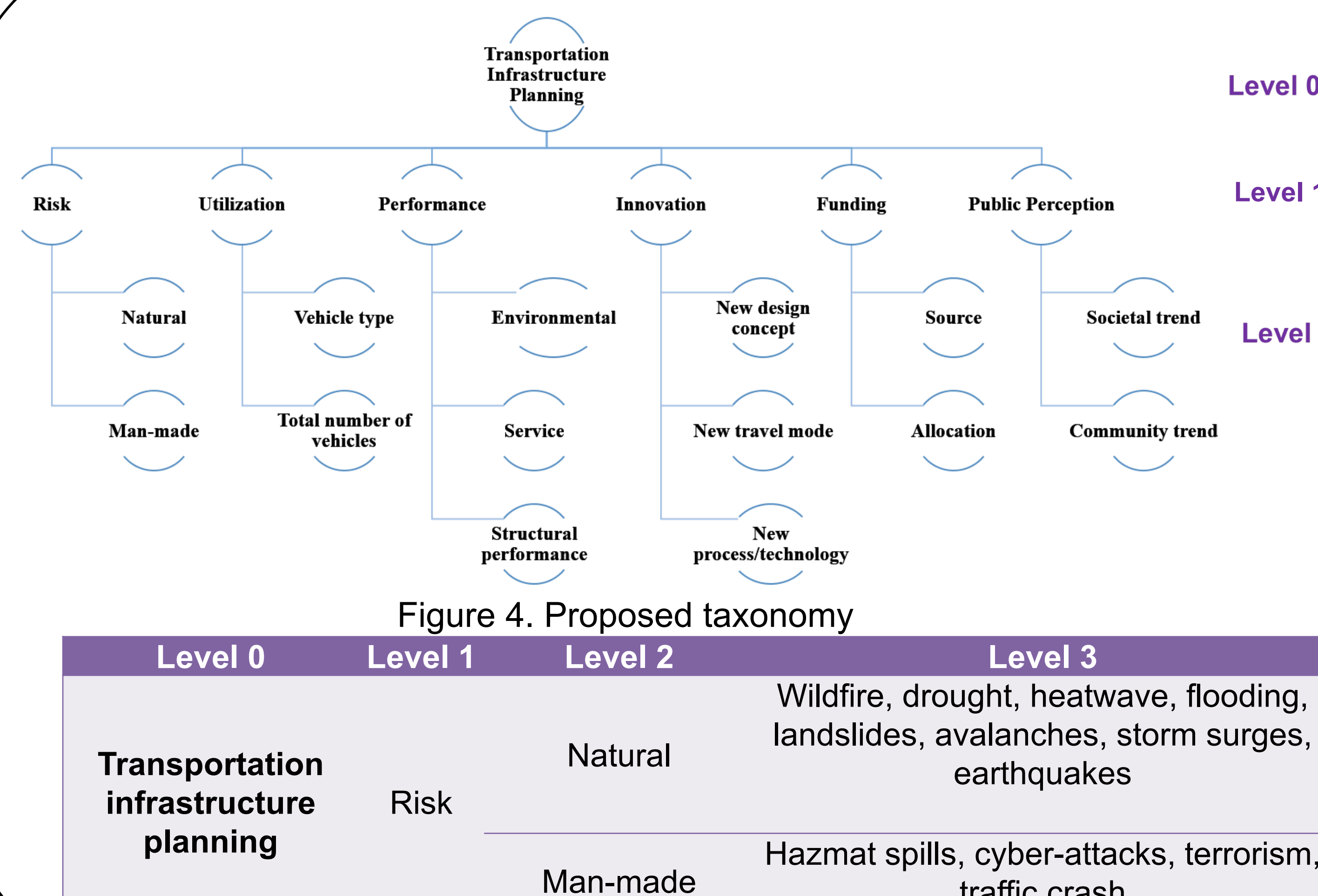| Factors	Strongest Association
Funding allocation	Service performance, new design concept
Environmental performance	Societal trend, new travel mode
Vehicle type	Structural condition performance, new process technology
Community trend	Man-made disruption, new design concept
Total number of vehicles	Societal trend, structural condition performance
Man-made disruption	Community trend, new process/technology

Figure 6. Sample association strengths

Level 0	Level 1	Level 2	Level 3
Transportation infrastructure planning	Risk	Natural	Wildfire, drought, heatwave, flooding, landslides, avalanches, storm surges, earthquakes
		Man-made	Hazmat spills, cyber-attacks, terrorism, traffic crash

Figure 5. Sample taxonomy construction

Man-made disruption

Community trend	New process/technology	New design concept
Lift: 1.80	Lift: 1.39	Lift: 1.05
Conf: 28.5%	Conf: 22.0%	Conf: 16.6%

Figure 7. Sample lift & confidence values with man-made disruption

Contributions

➢ A formalized structure to collect, organize, analyse, and utilize information from different sources
➢ It could help lay the foundation for data-driven infrastructure planning and management initiatives

Acknowledgement