**The Usage of Association Rule Mining towards Future-proofed Transportation Infrastructure Planning**

Sudipta Chowdhury, S.M.ASCE and Jin Zhu, Ph.D., A.M.ASCE

Department of Civil and Environment Engineering, University of Connecticut, USA

**Abstract**

Future-proofed transportation infrastructure planning is a complex process that requires consideration and integration of various external and internal factors. Although existing studies have explored the different factors for transportation infrastructure planning, few studies have investigated the inter-relationships among them. Therefore, this study proposed an association rule mining (ARM) based methodology to study the inter-relationships among key factors for transportation infrastructure planning. First, a list of factors was identified from 48 published documents on future-proofed transportation infrastructure planning via two topic modelling techniques: Latent Dirichlet Allocation and Non-negative Matrix Factorization. ARM was then used for discovering relationships among these factors. Specifically, two quantitative association rule mining metrics: confidence (frequency of association) and lift (strength of association) were used. Results showed that stronger associations existed between certain factors, e.g., a significant association was found between societal trends with environmental performance. It implies that in order to achieve better environmental performance of transportation infrastructure, capturing and taking advantage of societal trends could be useful, since societal trends such as less dependency on personal vehicles can significantly reduce the environmental impact of transportation infrastructures (e.g., less emission). Such results demonstrate the potential of using ARM to discover inter-relationships among a list of factors based on large text data. It could help transportation planners to better understand the inter-relationships among different factors and use them towards a more integrated transportation infrastructure planning.

 

Figure 1. Sample factors and their complex inter-relationship that can influence the planning process

**Acknowledgements:** This project was financially sponsored by U.S. DOT under its Transportation Infrastructure Durability Center (TIDC).

**References**

[1] Chen, Y., Bordes, J. B., & Filliat, D. (2017). Comparison studies on active cross-situational object-word learning using Non-Negative Matrix Factorization and Latent Dirichlet Allocation. *IEEE Transactions on Cognitive and Developmental Systems*, *10*(4), 1023-1034.

[2] Abdel-Basset, M., Mohamed, M., Smarandache, F., & Chang, V. (2018). Neutrosophic association rule mining algorithm for big data analysis. *Symmetry*, *10*(4), 106.