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

Sudipta Chowdhury, S.M.ASCE[1], and Jin Zhu, Ph.D., A.M.ASCE[2]

[1] PhD student, Department of Civil and Environment Engineering, University of Connecticut, Storrs, CT, USA; e-mail: sudipta.chowdhury@uconn.edu
[2] Assistant Professor, Department of Civil and Environment Engineering, University of Connecticut, Storrs, CT, USA; e-mail: jzhu@uconn.edu

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 impacts 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.