

# Crowd Management for the FIFA WC 2022 in Doha: survey-based traffic analysis and forecasting using machine learning methods

Simon Rienks<sup>1</sup>, Fiona Sauerbier<sup>1</sup>, Knut Haase<sup>1</sup>, and Martin Spindler<sup>2</sup>

<sup>1</sup>University of Hamburg, Department of Logisitcs, Transport and Production, Moorweidenstraße 18, 20148 Hamburg, Germany

<sup>2</sup>University of Hamburg, Department of Statistics with application in business administration, Moorweidenstraße 18, 20148 Hamburg, Germany

We conducted a survey on the use of different transportation modes during the FIFA Football World Cup in Qatar. The aim of this research is to identify the most important factors influencing decision behaviour and to estimate a utility function that is as close to reality as possible without using subjective but widely used expert knowledge for this purpose. Derived from this, a no show rate will be determined, which will be used for a queuing model in front of the metro stations.

To analyze the dataset, a combination of several methods is used to provide a data-driven approach. The collected dataset already contains many variables and is extended with interaction terms and binary variables (cuts). These cuts are generated by a deep decision tree to approximate non-linear effects. Overall, the utility function of the multinomial logit model to be estimated is high-dimensional and is therefore not estimated with the maximum likelihood method as in classical decision analysis, but with the maximum likelihood method in combination with a Lasso penalty term. This performs data-driven variable selection and estimates a sparse utility function for this high-dimensional problem. For the given dataset, our approach is superior to the classical method and provides higher accuracy in predicting decision behaviour on unknown data.