An Analysis Pathway for the Quantitative Evaluation of Public Transport Systems

Stephen Gilmore\textsuperscript{2}, Mirco Tribastone\textsuperscript{1}, and Andrea Vandin\textsuperscript{1}

\textsuperscript{1} Electronics and Computer Science, University of Southampton, Southampton, UK
\textsuperscript{2} Laboratory for Foundations of Computer Science, University of Edinburgh, Edinburgh, UK

Abstract. We consider the problem of evaluating quantitative service-level agreements in public services such as transportation systems. We describe the integration of quantitative analysis tools for data fitting, model generation, simulation, and statistical model-checking, creating an analysis pathway leading from system measurement data to verification results. We apply our pathway to the problem of determining whether public bus systems are delivering an appropriate quality of service as required by regulators. We exercise the pathway on service data obtained from Lothian Buses about the arrival and departure times of their buses on key bus routes through the city of Edinburgh. Although we include only that example in the present paper, our methods are sufficiently general to apply to other transport systems and other cities.

1 Introduction

Modern public transport systems are richly instrumented. The vehicles in a modern bus fleet are equipped with accurate GPS receivers, Wi-Fi, and on-board communications, allowing them to report their location for purposes such as fleet management and arrival-time prediction. High-frequency, high-resolution location data streams back from the vehicles in the fleet to be consumed by the predictive models used in real-time bus tracking systems.

We live in a data-hungry world. Users of public transport systems now expect to be able to access live data about arrival times, transit connections, service disruptions, and many other types of status updates and reports at almost every stage of their journey. Studies suggest that providing real-time information on bus journeys and arrival times in this way encourages greater use of buses \cite{1} with beneficial effects for the bus service. In contrast, when use of buses decreases, transport experts suggest that this aggravates existing problems such as outdated routes, bunching of vehicles, and insufficient provision of greenways or bus priority lanes. Each of these problems makes operating the bus service more difficult. Bus timetables become less dependable, new passengers are discouraged from using the bus service due to bad publicity, which leads inevitably to budget cuts that further accelerate the decline of the service.

Service regulators are no less data-hungry than passengers, requiring transport operators to report service-level statistics and key performance indicators which are used to assess the service delivered in practice against regulatory requirements on the quality of service expected. Many of these regulatory requirements relate to punctuality of buses.
defined in terms of the percentage of buses which depart within the window of tolerance around the timetabled departure time; and reliability of buses, defined in terms of the number of miles planned and the number of miles operated. The terms schedule adherence or on-time performance are also used to refer to the degree of success of a transportation service running to the published timetable.

With the aim of helping service providers to be able to work with models which can be used to analyse and predict on-time performance, we have connected a set of modelling and analysis tools into an analysis pathway, starting from system measurement data, going through data fitting, model generation, simulation and statistical model-checking to compute verification results which are of significance both to service providers and to regulatory authorities.

The steps of the analysis pathway, depicted in Figure 1, are as follows:

1. Data is harvested from a bus tracking system to compile an empirical cumulative distribution function data set of recorded journey times for each stage of the bus journey. In this paper, we generate inputs to the system using the BusTracker automatic vehicle tracking system developed by the City of Edinburgh council and Lothian Buses [2].
2. The software tool HyperStar [3] is used to fit phase-type distributions to the data sets.
3. A phase-type distribution enables a Markovian representation of journey times which can be encoded in high-level formalisms such as stochastic process algebras. In particular, we use the Bus Kernel model generator (BusKer), a Java application which consumes the phase-type distribution parameters computed by HyperStar and generates a formal model of the bus journey expressed in the Bio-PEPA stochastic process algebra [4]. In addition, the BusKer tool generates an expression in MultiQuaTEx, the query language supported by the MultiVeStA statistical model-checker [5]. This is used to formally express queries on service-level agreements about the bus route under study.
5. MultiVeStA is hooked to the simulation engine of the Bio-PEPA Eclipse Plugin, consuming individual simulation events to evaluate the automatically generated MultiQuaTEx expressions. It produces as its results plots of the related quantitative properties.

We are devoting more than the usual amount of effort to ensuring that our tools are user-friendly and easy-to-use. This is because we want our software tools to be used “in-house” by service providers because only then can service providers retain control over access to their own proprietary data about their service provision. With respect to ease-of-use in particular, making model parameterisation simpler is a crucial step in making models re-usable. Because vehicle occupancy fluctuates according to the seasons, with the consequence that buses spend more or less time at bus stops boarding passengers, it is essential to be able to re-parameterise and re-run models for different data sets from different months of the year.

It is also necessary to be able to re-run an analysis based on historical measurement data if timetables change, or the key definitions used in the evaluation of regulatory