Real-time Ranking of Electrical Feeders using Expert Advice*

Hila Becker[1], Marta Arias[2]

[1] Computer Science, Columbia University, New York
[2] Center for Computational Learning Systems, Columbia University, New York

Abstract. We are using machine learning to construct a failure-susceptibility ranking of feeders that supply electricity to the boroughs of New York City. The electricity system is inherently dynamic and driven by environmental conditions and other unpredictable factors, and thus the ability to cope with concept drift in real time is central to our solution. Our approach builds on the ensemble-based notion of learning from expert advice as formulated in the continuous version of the Weighted Majority algorithm [16]. Our method is able to adapt to a changing environment by periodically building and adding new machine learning models (or "experts") based on the latest data, and letting the online learning framework choose what experts to use as predictors based on recent performance. Our system is currently deployed and being tested by New York City's electricity distribution company.

Keywords: Concept Drift, Online Learning, Weighted Majority Algorithm, Ranking

1 Introduction

We are developing a machine learning online system that ranks feeders that supply electricity to the boroughs of New York City according to their susceptibility to impending failure in real-time. Primary feeders constitute a critical part of the distribution grid of New York City; feeder failures put networks, control centers, and field crews under considerable stress, especially during the summer, and cost millions of dollars in Operations and Maintenance expenses annually. Our work is focused on 943 underground primary feeders, distributing electricity to the New York City boroughs of Manhattan, Brooklyn, Queens, and the Bronx. Being able to predict incipient failures in close to real-time could enable crews and operators to take short-term preventative actions thus reducing the risk of failure. More details on this application and an earlier incarnation of our system can be found in [8]. In this paper we present an online machine learning ranking system for the feeders ranking problem that is able to cope with concept drift automatically.

Related Work. The problem of dealing with concept drift in the context of learning from data streams is receiving much attention recently, see e.g. [10, 20, 21, 3, 11, 22, 24, 26]. Most of these algorithms divide the input stream of data into subsets of sequential data (or "data windows"), and repeatedly build predictive models using only one or several contiguous windows of data at a time. These algorithms mainly differ in how a single or a combination of window-specific models are used to make future predictions. We distinguish two broad categories: the ones that maintain a single model and replace it with newer ones to cope with the concept drift [6, 20, 5, 7], and ensemble-based methods that keep a set of models and use combinations of the existing models to make predictions. Ensemble-based algorithms that use averages or weighted averages for future predictions include [21, 22, 24, 3]. All these algorithms are similar in that they use heuristics to estimate the predictive accuracy of the ensemble models and use these to weigh models' predictions. Additionally, the

* This work has been partly supported by a research contract from Consolidated Edison Company of New York