MAP tele DOCTORAL PROGRAMME IN TELECOMMUNICATIONS

PROJECT TITTLE:

UNDERSTANDING USER BEHAVIOUR BASED ON THEIR USAGE PATTERNS FROM 802.11 LOG DATA

STUDENT:

DOSSA MASSA

MAP TELE 2008/09

SUPERVISOR:

PROF. RICARDO SANTOS MORLA

Abstract

Wireless LANs (802.11 networks) are increasingly common with most Universities, organizations, cities already using them or planning to deploy such infrastructure, thus it is increasingly important to understand trends in the usage pattern of these networks . In this research project, we aim to study user behavior in terms of their usage patterns, from the traces of eduroam FEUP 802.11 logs by focusing on the behavior of access points (APs), and intend to devise a way to classify access points and group them based on their similarity. In addition we intend to predict future behavior of access point from the limited 802.11 log data. Understanding APs behavior and prediction of future behavior of APs is crucial for workload analysis of the network, hence can lead to better capacity planning and potential network optimizations, also could be useful in the management of the physical infrastructure e.g. at FEUP. In addition understanding APs behavior and its future behavior is important for many applications such as traffic engineering for APs and resource provisioning for QoS sensitive applications.

Literature review showed that most papers addressed issues related to users i.e. users movement, users location predictions and users classifications, few have addressed APs behavior while none addressed prediction issues (associations, events and traffic) related to AP. Works that addressed AP behavior focuses on the features that leaves out massive data which could be useful in characterizing APs behavior, thus we intend to use more fine grained usage information.

In this study we employ the use of probabilistic graphical model, in trying to predict future APs behavior, thus having extracted the daily usage information of all APs in the FEUP hotspot for the 86 days i.e. daily number of unique users and the associated number of events for each AP. Therefore, knowing number of events for each AP in the past three days and by employing Infer.NET a tool for performing inference in probabilistic graphical models, we have been able to predict the number of events for the next day.