PhD Proposal Title: Internet Traffic Forecasting using Data Mining Techniques

Motivation and Objectives:

Internet traffic forecasting is an important issue for any medium/large network provider that has received few attention from the Computer Networks research community. The aim of this PhD is to study novel solutions for Internet traffic forecasting tasks, which may be used in distinct contexts such as:

- Forecasting of Internet TCP/IP traffic, given a network topology with several nodes and links;
- Forecasting of specific Internet traffic aggregates (e.g. crucial flow aggregates, classes of service, etc.)
- TCP/IP end-to-end throughputs predictions
- Forecasting scenarios where only incomplete information in available
- Improving Internet Traffic identification/classification tasks

These are challenging tasks, in particular when few information is available. The intention is to develop advanced data mining techniques (e.g. Neural Networks or Support Vector Machines) that can provide timely and accurate Internet traffic forecasts. Given this context, the results of this research are expected to effectively contribute to enhance several Internet management related tasks such as: general resource management procedures, improve anomaly detection efforts in the Internet (e.g. spread of virus detection), increase the Quality of Service levels of Internet Providers, assist intelligent optimization tools for configuration and planning of TCP/IP networks, among many others.
Using the proposed forecasting techniques, prototypes of intelligent, user-configurable and versatile Internet traffic forecasting engines could be implemented in order to provide useful outputs for other Internet related management tools.

Additional notes: Although data collection often requires a huge effort/time, real-world data is/will be available (e.g. UKERNA TCP/IP traffic - bytes/s of every link in the British academic network).

Supervisors:

- **Paulo Cortez**, pcortez@dsi.uminho.pt, Department of Information Systems/Algoritmi R&D Centre, Univ. Minho, Guimarães, Portugal (http://www3.dsi.uminho.pt/pcortez)
- **Pedro Sousa**, pns@di.uminho.pt, Department of Informatics/CCTC R&D, Univ. Minho, Braga, Portugal (http://marco.uminho.pt/~pns)

Bibliography (previous work performed in this area):

- P. Cortez, M. Rio, M. Rocha and P. Sousa. Multiscale Internet Traffic Forecasting using Neural Networks and Time Series Methods. In Expert Systems, Wiley-Blackwell, In press.
- P. Cortez, M. Rio, P. Sousa and M. Rocha. Forecasting Internet Traffic by Neural Networks Under Univariate and Multivariate Strategies. In K. Madani (Ed.), Proceedings of 4th International Workshop on Artificial Neural Networks and Intelligent Information Processing (ANNIIP 2008), pp. 61-71, Funchal, Portugal, May, 2008. INSTICC Press (http://hdl.handle.net/1822/8027).
- P. Cortez, M. Rio, P. Sousa and M. Rocha. Topology Aware Internet Traffic Forecasting using
- Neural Networks. In J. de Sá et al. (Eds.), Artificial Neural Networks - ICANN, 17th International Conference, Lecture Notes in Computer Science 4669, pp. 445-454, Porto, Portugal, September, 2007, Springer (http://hdl.handle.net/1822/7634).
- P. Cortez, M. Rio, M. Rocha and P. Sousa. Internet Traffic Forecasting using Neural Networks. In Proceedings of the 2006 IEEE International Joint Conference on Neural Networks, pp. 4942-4949, Vancouver, Canada, July, 2006 (http://hdl.handle.net/1822/6581).