Modelling a router

Ronald Rousseau and Iris Vandecasteele

KHBO, Industrial Sciences and Technology
Zeedijk 101, B-8400 Oostende, Belgium
ronald.rousseau@kh.khbo.be

Abstract. We represent a router's traffic using stochastic modelling techniques.

1 Introduction

Research on digital libraries encompasses a range of intertwined technical, social and political issues. Before studying who is using which part of your digital library, and how digital documents are accessed, it is important to know if your digital library is accessed at all and when. Our department, the department of industrial sciences and technology of the school for higher education in Ostend (Belgium), was interested in monitoring the use of its online resources. As a part of this project we constructed a stochastic model for the institute's router. We are, indeed, convinced that mathematical modelling and analysis continues to play an important role for computer systems and services such as digital libraries. Hence, this contribution focuses on one particular technical aspect within the larger framework of digital libraries.

2 Data

A router is a device that connects the Internet with any number of LANs. In our case this is just one Local Area Network. Over a period of three weeks we collected data on the router's traffic. The signal for which we report the results here, is the hourly average of the lowest and highest incoming traffic (in kbits per second), filtered by a five-point median filter. Fig.1 shows the raw data. This signal clearly exhibits a daily and a weekly pattern.

3 Method

After detrending the signal (subtraction of the mean) we modelled the obtained signal using the auto-regressive (AR) model. Model parameters were obtained by minimizing the mean squared error (MSE-minimization). Practical calculations were done

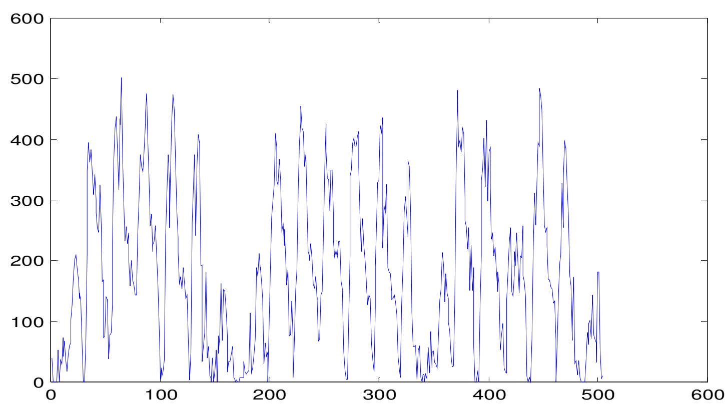

Fig.1 Router traffic: raw data for a three week period

based on Burg's method as implemented in MATLAB ®. Burg's algorithm is an efficient approach which minimizes backward and forward prediction errors. Finally the order of the model was estimated using Akaike's criteria. We performed the calculations for each week separately (results not given here), and for the whole three-week period (a data series of 504 points, or 500 points after filtering).

4 Results

The whole (detrended) 3-week period can be described by the following 7^{th} order model:

$$y(n) - 1.0597\, y(n-1) + 0.0079\, y(n-2) + 0.2920\, y(n-3) - 0.0687\, y(n-4) - 0.0715\, y(n-5) - 0.2015\, y(n-6) + 0.2700\, y(n-7) = w(n)$$

where w(n) is a white-noise signal with zero mean.

A lower order model, however, can describe one-week periods of observed data. This result, being an exercise in stochastic modelling, yields a signal with the same (or at least similar) stochastic behaviour as the data signal. It cannot directly be used for long-term predictions. From this modelled signal all other important parameters can be derived, in particular its power spectral density.

5 Conclusion

Monitoring a digital library includes many aspects. In this contribution we have focused on the technical aspect of representing a router's traffic, using stochastic modelling techniques.

July 2002