Packet-Loss Prediction Model Based on Historical Symbolic Time-Series Forecasting

by

Hooman Homayounfard

A dissertation submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in the Faculty of Engineering and Information Technology

UNIVERSITY OF TECHNOLOGY, SYDNEY

October 2013

© Copyright 2013

by

Hooman Homayounfard

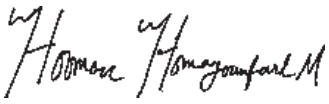
© All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.

CERTIFICATE OF ORIGINAL AUTHORSHIP

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Candidate



Acknowledgements

During my PhD program, I have received support and encouragement from a great number of individuals. I express my deep gratitude to my principal supervisor, Dr Paul Kennedy, for his excellent guidance in writing this thesis and the papers which preceded chapters. Dr Paul's professional advice and unyielding support led me to accomplishing this task.

I would like to thank my co-supervisors Prof John Debenham, Prof Robin Braun and Prof Simeon Simoff for advising me on the work of this thesis. I should also acknowledge Prof Barry Jay, Prof Didar Zowghi, Prof Massimo Piccardi, Prof Longbing Cao, Prof Jie Lu, Dr Ante Prodan, Dr Vinod Mirchandani, Dr Priyadarsi Nanda, Dr Masoud Talebian, Andrew Litchfield, Jose Vergara, Nima Ramzani, Debi Taylor and Vahid Behbood. These are a few of many people who had a positive impact on my achievements.

I am in debt to the whole UTS community for challenging my mind over the past years. I do thank Mrs Eilagh Rurenga for polishing chapters in my thesis. I am beyond grateful for the help from Craig Shuard and Phyllis Agius, the FEIT research officers.

Dear family and friends, thank you for your incredible support. I would have been lost without you in this process. My special regard to my wife and mother, Narges and Mahin, for their sincere love, patience and support. The thesis is dedicated to my dear Leila.

Publications

1. **Hooman Homayounfard**, Paul Kennedy, and Robbin Braun, *NARGES: Prediction Model for Informed Routing in a Communications Network*, In J. Pei et al., editor, *LNAI*, volume 7818, pages 327-338. Springer Berlin Heidelberg, 2013.

2. **Hooman Homayounfard** and Paul Kennedy, *HDAX: Historical Symbolic Modelling of Delay Time Series in a Communications Network*, In P. J. Kennedy, K. Ong, and P. Christen,editors, *AusDM09*, volume 101 of CRPIT, pages 129-138, Melbourne, Australia, 2009.

Abstract

Rapid growth of Internet users and services has prompted researchers to contemplate smart models of supporting applications with the required Quality of Service (QoS). By prioritising Internet traffic and the core network more efficiently, QoS and Traffic Engineering (TE) functions can address performance issues related to emerging Internet applications. Consequently, software agents are expected to become key tools for the development of future software in distributed telecommunication environments. A major problem with the current routing mechanisms is that they generate routing tables that do not reflect the real-time state of the network and ignore factors like local congestion.

The uncertainty in making routing decisions may be reduced by using information extracted from the knowledge base for packet transmissions. Many parameters have an impact on routing decision-making such as link transmission rate, data throughput, number of hops between two communicating peer end nodes, and time of day. There are also other certain performance parameters like delay, jitter and packet-loss, which are decision factors for online QoS traffic routing.

The work of this thesis addresses the issue of defining a Data Mining (DM) model for packet switching in the communications network. In particular, the focus is on decision-making for smart routing management, which is based on the knowledge provided by DM informed agents. The main idea behind this work and related research projects is that time-series of network performance parameters, with periodical patterns, can be

used as anomaly and failure detectors in the network. This project finds frequent patterns on delay and jitter time-series, which are useful in real-time packet-loss predictions.

The thesis proposes two models for approximation of delay and jitter time-series, and prediction of packet-loss time-series – namely the Historical Symbolic Delay Approximation Model (HDAX) and the Data Mining Model for Smart Routing in Communications Networks (NARGES). The models are evaluated using two kinds of datasets. The datasets for the experiments are generated using: (i) the Distributed Internet Traffic Generator (D-ITG) and (ii) the OPNET Modeller (OPNET) datasets.

HDAX forecasting module approximates current delay and jitter values based on the previous values and trends of the corresponding delay and jitter time-series. The prediction module, a Multilayer Perceptron (MLP), within the NARGES model uses the inputs obtained from HDAX. That is, the HDAX forecasted delay and jitter values are used by NARGES to estimate the future packet-loss value.

The contributions of this thesis are (i) a real time Data Mining (DM) model called HDAX; (ii) a hybrid DM model called NARGES; (iii) model evaluation with D-ITG datasets; and (iv) model evaluation with OPNET datasets.

In terms of the model results, NARGES and HDAX are evaluated with offline heterogeneous QoS traces. The results are compared to Autoregressive Moving Average (ARMA) model. HDAX model shows better speed and accuracy compared to ARMA and its forecasts are more correlated with target values than ARMA. NARGES demonstrates better correlation with target values than ARMA and more accuracy of the results, but it is slower than ARMA.

Contents

List of Tables

List of Figures