

SelfSimMGI Readme

An OMNET package with an implementation of a useful modification of well-known self-similar traffic model “M/G/ ∞ Input” is presented. The package is designed for analysis of real WAN traffic traces, identification of model parameters and generation of a synthetic traffic, which is close to real one by a number of characteristics simultaneously. These characteristics include traffic value probability distribution, normalized autocorrelation function and Hurst parameter.

Packadge:

1. Traffic generators: **FromFile_Source**, **MGI_Source** and **MGI_LOG Source**
2. Traffic statistics collector: **Traffic_Stats**
3. Others: **Sink**
4. Utilities: **pcap2TL**
5. Network configurations
6. Appendix

1. Traffic generators

FromFile_Source:

It reads traffic trace file in TL or tcp format (as these formats are used in Internet Traffic archive [1]), creates packets with appropriate time of arrival and size, according records in trace file, and sends them to output gate.

MGI_Source:

It implements **Modified Input M/G/ ∞** traffic model. It has two modes of operation:

1. **Source parameter calculation.** In this mode it reads traffic statistics file (this xml file is created by Traffic_Stats module, see below), calculates parameters of appropriate **Modified Input M/G/ ∞** model and writes it into xml profile file. **It does not generate traffic in this mode!** It just calculates parameters!
2. **Run.** In this mode it generates traffic according **Modified Input M/G/ ∞** model and sends it to output gate. Parameters of the model it reads from xml profile file, which was created in first mode. So, usually before start of this module in run mode it is necessary to make start of it in source parameter calculation mode once.

MGI_LOG Source:

Almost the same as **MGI_Source**, but it implements basic **Input M/G/ ∞** model for comparison with Modified one. In basic model all individual sources have rate 1 cell per slot.

2. Traffic_Stats:

It passes traffic through and collects follow statistical information:
probability distribution of traffic value (traffic value in cells per timeslot),
normalized autocorrelation function values,
Hurst parameter and data for its calculation.

All these data it writes into xml file for using by **MGI_Source** and **MGI_LOG Source** modules.

More detailed description of the modules see in Appendix.

3. Sink: This module just destroys arrival packets.

4. To simplify experiments with traffic traces in pcap format, **pcap2TL** program is included in the package. It transforms pcap to TL format and can also cut big TL files into parts of smaller size.

5. Several simple network configurations are provided in **omnetpp.ini** file of this package for test purposes:

FromFile – for trace traffic file of TL or tcp formats processing,

MGI_Parameters_Calculation – for parameters calculation for **Modified Input M/G/∞** model,

MGI_Run – for **Modified Input M/G/∞** model traffic generation

MGI_LOG_Parameters_Calculation – for parameters calculation for **basic Input M/G/∞** model

MGI_LOG_Run - for **basic Input M/G/∞** model traffic generation

The results of processing traffic from CAIDA archive [2] by this package see in [4].

To play with the model, small piece of one CAIDA trace in TL format is included in the package.

APPENDIX:

FromFile_Source

Module's parameters:

filetype – type of file with real traffic trace (TL or tcp is possible);

tracefile – name of file with real traffic trace.

MGI_Source

Module's structure described in file *MGI_Source.ned*. This module consists of three sub modules: **MGI_Parameters_Calculator**, **MGI_Generator** and **MGI_Collector**. Also there is module type **Individual_MGI_Source**. It presents a single traffic source. Exemplars of this module type are created and destroyed dynamically during model run. Source code of these modules is in a files *MGI_Generator.cc*, *MGI_Parameters_Calculator.cc*, *MGI_Collector.cc*, *Individual_MGI_Source.cc*.

Parameters of module **MGI_Source**:

calcRequired – if true, module will start in source parameter calculation mode.
if false, module will start in run mode.

Parameters of module **MGI_Parameters_Calculator**:

data – xml document containing traffic statistics, calculated previously by **Traffic_Stats**;

profile – name of output file with the model parameters calculated.

Tags in xml profile:

- TIMESLOT – time slot duration (in seconds) –value from statistics file of original trace;
- CELLSIZE – cellsize (in bytes) –value from statistics file of original trace;
- TRAFFIC-MEAN – average rate of traffic (in cells per timeslot) – equals to parameter CELLSPERSLOT from statistics file of original trace;
- ACF-SHAPE-PARAMETERS – autocorrelation function shape parameters;
- LIFETIME-MEAN – average lifetime of the individual source (in timeslots);
- SOURCE-BORN-RATE – new individual source born rate (per timeslot);
- MAXIMAL-SPEED –maximal value of informational rate of the source;

- SPEED-DISTRIBUTION – individual source rate probability distribution (rate in cells per timeslot);

Parameters of module **MGI_Generator**:

profile – name of xml file with above described tag structure.

Module **MGI_Collector** has not any parameters.

MGI_LOG Source:

Parameters required for basic Input M/G/ ∞ are the same, except individual source rate probability distribution. In basic model all sources have rate 1 cell per slot.

Traffic_Stats

Parameters of the module:

- timeslot – timeslot for discretization of traffic;
- cellsize – cellsize (in bytes) used for quantization of traffic by value;
- numHistCells – size of histogram for PDF evaluation;
- num_of_points – size of histogram for ACF evaluation;
- num_of_sizes, minSize, maxSize – parameters for evaluation of Hurst parameter;
- outFileName – name of output file to write statistics;
- saveDiscreteTraffic – to save discretized traffic or not.

The module creates a statistics xml-file with the following information:

- TIMESLOT – it is taken from the corresponded parameter of module *Traffic_Stats.ned* and means the same thing;
- CELLSIZE – it is taken from the corresponded parameter of module *Traffic_Stats.ned* and means the same thing;
- CELLSLPER SLOT – average cells per one timeslot;
- BYTESPERSEC – average bytes per second;
- TRAFFIC-VARIANCE – traffic variance;
- TRAFFIC-MAXVALUE – maximal value of traffic measured in cells;
- TRAFFIC-DISTRIBUTION – probability distribution of traffic value (traffic value in cells per timeslot);
- ACF-RANGE – size of calculated part of normalized ACF (points from zero);
- ACF-VALUES – normalized autocorrelation function values;
- HURST and so on – Hurst parameter and data for its calculation.

References:

1. <http://ita.ee.lbl.gov> The Internet Traffic Archive sponsored by ACM SIGCOMM.
2. <http://www.caida.org> The Cooperative Association for Internet Data Analysis (CAIDA).
3. Numerical Recipes in C: The Art of Scientific Computing / William H. Press et al., - 2nd ed. Cambridge University Press ISBN 0-521-43108-5
4. A. Yu. Privalov and A. Tsarev. Analysis and Simulation of WAN Traffic by Self-Similar Traffic Model with OMNET. // report accepted to IWCMC 2014 Conference.