Project	**IEEE 802.16 Broadband Wireless Access Working Group <http://ieee802.org/16>**
Title	**Generalizing 4IPP Traffic Model for IEEE 802.16.3**
Date Submitted	**2000-12-12**
Source(s)	Jun Huang DragonWave Inc 202-535 Legget Dr. Kanata ON K2K 3B8 Canada Voice: 613 599 9991 ext 232 Fax: 613 599 4225 mailto:jun.huang@dragonwaveinc.com
Re:	This contribution is in response to call for contribution, from IEEE 802.16.3 Traffic Model Ad-hoc Committee, sent out 2000-11-21, with a subject of 802.16.3c-00-51 (4IPP Model).
Abstract	The contribution defines a generic traffic model – n Interrupted Renew Process (nIRP), it provides both self-similar traffic and non-self-similar traffic modeling for Broadband Wireless Access (BWA) applications. The model can be used to accurately characterize measured voice, video and data traffic. It is a backward compatible extension of 4IPP model. It is forward extendable as well. The contribution offers a system level method suitable for simulation of MAC/PHY proposals from traffic and performance perspectives.
Purpose	For 802.16.3 to consider the input of the nIRP model for evaluating different MAC/PHY combinations. This contribution is made for Session #11 in Ottawa.
Notice	
Release	
Patent Policy and Procedures	

Generalizing 4IPP Traffic Model for IEEE802.16.3

Jun Huang
DragonWave Inc.

Introduction:

Broadband Wireless Access is expected to accommodate not only data traffic but also voice and video traffic. The anticipated user/node models can range from Small Office Home Office (SOHO) to Small and Medium Enterprise (SME), or from Multi Dwelling Unit (MDU) to Multi Business Unit (MBU) of different sizes.

The applications that are expected from these nodes can vary widely from traditional voice to the latest on-line gamer, including numerous multimedia applications. A highly versatile, universal mathematical model (yet not too complicated for simulation) is thus needed to accommodate these different node traffic profiles.

To this end, we have decided to extend the already proposed 4IPP traffic model [1] (IEEE802.16.3c-00/51 contribution made by Dr. Baugh), rather than to introduce other radically different self-similar traffic models, such as Fractional Brownian Model (used in ATM Forum), Descriptive Model (used in TCP simulations), Chaotic Maps Model, etc. The effort is also made to extend the 4IPP model in such a way to include some traditional non-self-similar models, like ON-OFF model [2] at the same time.

The "beauty" of the IPP traffic model [3] is

- It characterizes the correlation (not just variance) of an aggregated traffic with a minimum number of parameters, i.e. only 3 parameters implicitly related to average, variance and correlation somehow.
- It is also analytically solvable when fed into a queuing system (that models MAC protocol) by using the matrix geometric method.

The disadvantage of IPP, however, is

- It is not suitable for modeling individual traffic (such as a single Voice)

The following extension will overcome above disadvantage, and meanwhile introduce further freedom for engaging more advanced self-similar models.

Description for Generalized Model:

We extend 4IPP model in 3 steps to reach 3 levels of generality.

Step 1. Just Varying the "size":

Interrupted Poisson Process (IPP) is a process obtained by turning on and off a Poisson process alternatively. The on time and off time obey the (negative) exponential distribution. 4IPP is a superposition stream of 4 independent IPP streams. (See Dr. Baugh's contribution for details)

Self-similar traffic is a traffic that has the burstiness over different time scales, the bursty pattern of larger scale resembles that of the smaller of itself recursively. Mathematically speaking the process has an Index of Dispersion for Counts (IDC) increasing linearly with the observation window, or measurement length in terms of time. The IDC curve for IPP traffic increases with the observing window initially, but eventually it flattens out at a constant value. As a consequence, a number of IPP components are needed to ramp the synthesized traffic to the desired IDC value. If measured traffic has steep and high IDC value for the interested time scales, more IPP components might be needed. This can be the case for MDU users, where different video-content and irregular usage-hours make the node traffic burstier.

As such we suggest extending the 4IPP model to nIPP model in the first step, and the number "n" is to be obtained by matching the measurement to the calculated IDC. A method for matching IDC can be found in [4]. "n" can be 1 if non-self-similar traffic is modeled. N can be 2 or 3 if "less self-similar" traffic is encountered, and N can be 5 or higher if "more self-similar" traffic is coped with.

This slight extension is to introduce some flexibility for system designers. They can now choose the different "size" of model based on their traffic measurement profile for their target users and system requirements. Nothing prevents you from using 4IPP.

Step 2. Extending the inter-arrival distribution:

For some applications like traditional voice, the inter-arrival time between the packets during on time (corresponds to talk spurt) is constant; during off time (corresponds to silent period), no packet is generated. If we allow the inter-arrival time between packets to be a general distribution, we then have a notion called Interrupted Renewal Process (**IRP**) in queuing theory. IRP degenerates into IPP when the general distribution becomes exponential. When the distribution is deterministic with a constant value, the IRP becomes Interrupted Deterministic Process (IDP), which is essentially the traditional ON-OFF model.

That solves the disadvantage of IPP; with this slight extension, we can use IDP (ON-OFF) to model individual voice. And nIDP to model n number of voices, where n can be a small number, this is the case for a SOHO node.

On the other hand, according to reference [5], we can use a number (say n=16) of ON-OFF mini-sources to characterize the video traffic. Suggested n is ranged from 10 to 20, we choose 16. Loosely speaking, we can divide a video screen into 4 by 4 blocks, the activity of each block is associated with corresponding mini-source. If there is a major change between consecutive video frames for the particular block, that mini on-off source is interpreted as in "on" state, because the information update has to be sent out; in contrast, if that block looks like still picture, little update needs transmitted, the mini-source is claimed as in "off" state.

As a consequence, the extended nIRP is good for characterizing the voice, video and data, all at one "shot". By allowing the arbitrary inter-arrival distribution, we also leave ourselves some headroom for future possible development of more advanced video models such as heterogeneous mini-source model. The above model family is still solvable when attempting to predict queuing behavior under semi-Markovian assumptions (the on time and off time is exponentially distributed and the "switch-over" may only fall on the moment of packet arriving).

Step 3. Stretching the sojourning time:

For some application like LAN [6], the "on" time or "off" time (sojourning times) is distributed following Pareto or Weibull distribution rather than Exponential distribution. As a consequence, we can relax the restriction for sojourning time of being exponential (like IPP), to include Pareto and Weibull as well (for IRP). This last step of extension destroyed analytical beauty of IPP in queuing analysis. But as to the simulation (with IRP), it incurs the same amount of modeling complexity.

We'd like to propose this extension because the Pareto centric models have already been extensively used in 3G wireless designs, LMCS/LMDS MAC evaluation and advanced satellite communication planning. Examples are: (1) Pareto ON-OFF model has been used in 3G designs; (2) WWW (Pareto) traffic model has been utilized in LMDS study [7]; and (3) Pareto IPP [8] has been employed in advanced satellite studies. The commercial tools are available for all these models.

In summary, we call the extended model nIRP, the inter-arrival time can be constantly or exponentially distributed (or any distribution in between), the sojourning time can be Pareto or exponentially distributed (or any distribution in between); and by definition, it is completely backward compatible with 4IPP, however it offers much freedom for including voice, video and data models suited for various node traffic profiles.

Simulation Method:

When simulation is the concern for obtaining the performance of the proposed MAC protocols, the nIRP has similar structure and complexity as comparing with 4IPP. The difference is just between the different distributions for variables involved. The nIRP model can be easily realized in any commercial network simulation package. The minimum change needed is: instead of calling a subroutine for an exponential function, now we call that of a Pareto distribution or a Weibull distribution etc., when generating traffic from applications of a node. Here is a summary of the model family.

Model Attributes			Model Characters		
Number of "Mini" sources	Inter-arrival time	Sojourning time	Type of Applications	Model's "Nickname"	Example Applications
4	Exponential	Exponential	Data	4IPP	WAN
1	Constant	Pareto	Data	Pareto ON-OFF	LAN
1	Weibull	Pareto	Data	WWW	PAN
1	Constant	Exponential	Voice	ON-OFF	Phone
1	Exponential	Pareto	Video	Pareto IPP	DTV
16	Constant	Exponential	Video	Mini ON-OFF	Video Conf.

Table 1. Summary of members in nIRP model family

Further Considerations:

This contribution is a "guild line" for simulation. More detail results will be reported later. In terms of the system test scenarios and performance metrics for evaluating different MAC/PHY combinations, we refer to Dr. Baugh's contribution.

The additional point we want to bring up is that the performance such as the overall access delay can be sensitive to the traffic profile. For example, it was identified in [7] that the DOCSIS MAC offers less delay than DAVIC MAC when WWW traffic is injected, but the difference starts diminishing when batch Poisson traffic is resorted to. The advantage of batch Poisson model is that the packet size distribution is accurately captured.

As a consequence, using single model to observe the performance metric might not be enough. We are using the family of models testing various MAC/PHY combinations. We will extend nIRP model further to include batch Poisson.

Even that may not enough to draw conclusions, it is practically important to establish node level models for different types of users (SOHO, SME, MDU, MBU) based on the generic traffic models (for application) and measured statistics of the traffic combination (of applications) for each node. And then conduct the evaluations to see if they mimic field deployment.

References:

1) C.R. Baugh, "4IPP Traffic Model for IEEE802.16.3," IEEE802.16.3c-00/51, Meeting #10, Tampa, Oct. 2000.

2) P.T. Brady, "A Statistical Analysis of ON-OFF Patterns in 16 Conversations," Bell Syst. Tech. J., Vol. 47, No.1, Jan. 1968.

3) A. Kuczura, "The Interrupted Poisson Process as an Overflow Process," Bell Syst. Tech. J., Vol. 52, No.3, March 1973.

4) J. Huang, T. Le-Ngoc, J.F. Hayes, "Broadband Satcom System for Multimedia Services", IEEE ICC'96, Dallas, June 1996.

5) B. Maglaris, D. Anastassiou, P. Sen, G. Karlsson, J.D. Robbins, "Performance Models of Statistical Multiplexing in Packet Video Communications," IEEE Trans. on Comm., Vol. 36, July 1988.

6) W.E. Leland and D.V. Wilson, "High Time-resolution Measurement and Analysis of LAN Traffic: Implementations for LAN Interconnection", Proc. IEEE Infocom'91, Bal Harbour, April 1991.

7) M.T. Ali, R. Grover, G. Stamatelos, D.D. Falconer "Performance Evaluation of Candidate MAC Protocols For LMCS/LMDS Networks," IEEE JSAC, Vol. 18, No. 7, July 2000.

8) S.N. Subramanian, T. Le-Ngoc, "Traffic Modeling in a Multi-Media Environment", CCECE'95, Montreal, Sept. 1995.

Acknowledgement:

Thanks go to Mr. Rod McMullin (DragonWave Inc) and Prof. Oliver Yang (Ottawa University) for offering detail comments.