TCP Environment: State-space Model and Analysis of Local Behavior

Vladimir Zaborovsky, Aleksander Gorodetsky, Vladimir Muljukha
Politechnical University, St.-Petersburg, Russia;
e-mail: vlad@neva.ru

Abstract. The analysis of network control and resource management in stochastic environment under chaotic conditions (fractal behavior and long range dependence) is of great importance. The choice of models and methods depends on accuracy of identification and measurement of traffic characteristics. But without robust theoretical preconditions network environment modeling will fail to simple simulation results that does not reflects real network features. Therefore Internet as environment of control that characterized by specific features including dynamism, hierarchy and adaptability is a topic of active fundamental researches. In the paper we consider state-space model of TCP environment that reflects various important feature of virtual transport connections. Proposed approach based on analysis of probability of possible TCP states and estimation of available throughput. The received results are confirmed experimentally and verified by modeling of specific local TCP behavior.

Keywords—computer network environment, state-space, transport protocols, analysis; traffic, model

I. Introduction

Internet is a global and commonly used information exchange environment. But models describing processes of information transfer in a form of packet flow qualitatively differ from well known models of weights movement or energy transfer. Till now the methodology of synthesis of such models has no complete scientific substantiation and mainly based on various stochastic approximations. Below we consider specific aspects of modeling of network processes in the form of the packet flow under control of TCP protocol and static routing algorithm. In the modern computer network interaction between nodes or applications are carried out by creating of virtual connections or transport streams of the various natures: audio, video, or digital data. Characteristics of these streams are influenced by network structure and casual behavior of signals in the telecommunication links. Below packet transport processes consider as virtual connections that possesses properties of adaptability to a wide range of characteristics of a computer network protocols and applications. For many control and management purpose it is necessary to have the adequate functional and parametrical description of properties of TCP behavior but due to chaotic nature of network environment we needs to verified proposed models in specific network condition [1,7,8]. In the offered approach, we consider certain factors of packet transmission process which influence directly parameters of TCP connection. We considers computer network as the data transmission environment that studied by the analysis of separate virtual transport connections. Thus it is necessary to notice, that the chosen individual connection cannot render essential influence on a state of whole computer network. At the same time influences of the computer network environment on individual virtual connection can changes the states of TCP protocol. Speaking about these states we should taking into account a states of the standard TCP protocol machine which characterizes a behavior of the congestion window (CW) on a fast recovery (FR), congestion avoidance (CA) slow start (SS) phases when the congestion window grows linearly or equal to receiver window parameter.

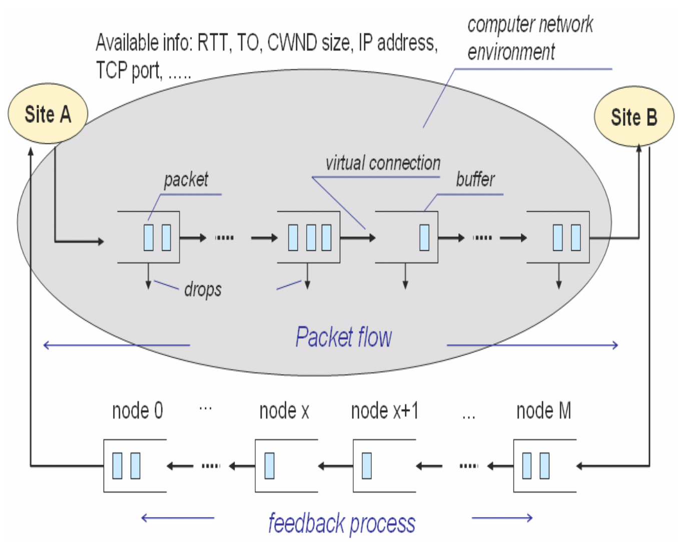


Figure 1. Influence of the computer network environment on individual TCP connection controled by acknowledgement packets

Each of the states is characterized by certain throughput and is represents as the state graph node. Graph arcs between states are characterized by corresponding probabilities. Transitions between states occur under the influence of computer network environment. We suggest to use the special form of approximation of uniform distribution function, namely, uniform function with sliding left border, to model casual character of the environment influence.

In the paper the description of states of TCP-connection is given in a form of creative hypothesis about stationary probabilities of staying of TCP-connection in the chosen state. Stationary requirement is experimentally confirmed therefore parametric description of transition between TCP states proposed to use to analysis of throughput of virtual TCP connection.

Realistic estimation can be achieved by the analysis of probability of TCP connection stay in each of possible states taking into account their individual throughput.

The paper is organized as follows. In section 2 the model of the network environment is shown and analysis the factors which influence TCP connection including the changes of CW parameters. In section 3 the TCP connection model in state-space is presented. The graph of transition between states is used to define influence of the network environment. The description of the spitted state is entered and used to estimates of various probabilities of TCP conditions. Model of complexities for TCP process in the state-space are resulted and experimental confirming by offered description. In section 4 there are presented the mathematical model of a congestion window (CWND), and distribution function of time intervals between the moments of loss/delay of the packets, which influence a state of TCP-connection. Section 5 is the conclusion in which the received results are listed and offers directions of future works.

II. TROUGHPUT MODEL OF THE NETWORK ENVIRONMENT

The increase of throughput is the main objective of development of new generation of computer networks and protocols. This goal can be reached applying new methods of processing and routing information. But for working out these methods it is necessary to know, what is model of computer network, what characteristics are processing, and how is the cross-state influencing increase productivity.

To answer these questions it is necessary to create adequate model of basic network processes. Attempts to create a mathematical model of TCP flows and traffic behavior in terms of measurement characteristics were undertaken for a long time. But many open questions in the understanding and implementation of modelling concepts are remains unexplored especially in the case when fluctuation factors have stochastic or even chaotic natures. Complexity of modeling task stem from the fact that we cannot make independent measurement traffic parameters: all such measures are correlated and this fact reduce efficacy of feed-back control and restricted accuracy of standard identification methods which can provides asymptotic parameter estimations only in steady-state stochastic condition.

The object of our research is the model TCP connection that used for transfer packets through a computer network environment. To provide correct description of virtual connection behavior we use the state space model of TCP protocol. Each states of TCP protocol are modeled as the nodes that connected by arcs of probability transitions. Below we consider only those states which directly influence current values of virtual connection throughput and observed, that allows model identification by current traffic measurements. The considered states are slow start, congestion avoidance, and fast recovery. We also analysis the several "events" along with "states", namely: a time-out (TO) and fast retransmit (FR) take place. The sequence of these events can concern one state. On fig. 2 is shown typical influence process of network environment in the form of loss/delay moments, changes of throughput, and dynamics of a congestion window.

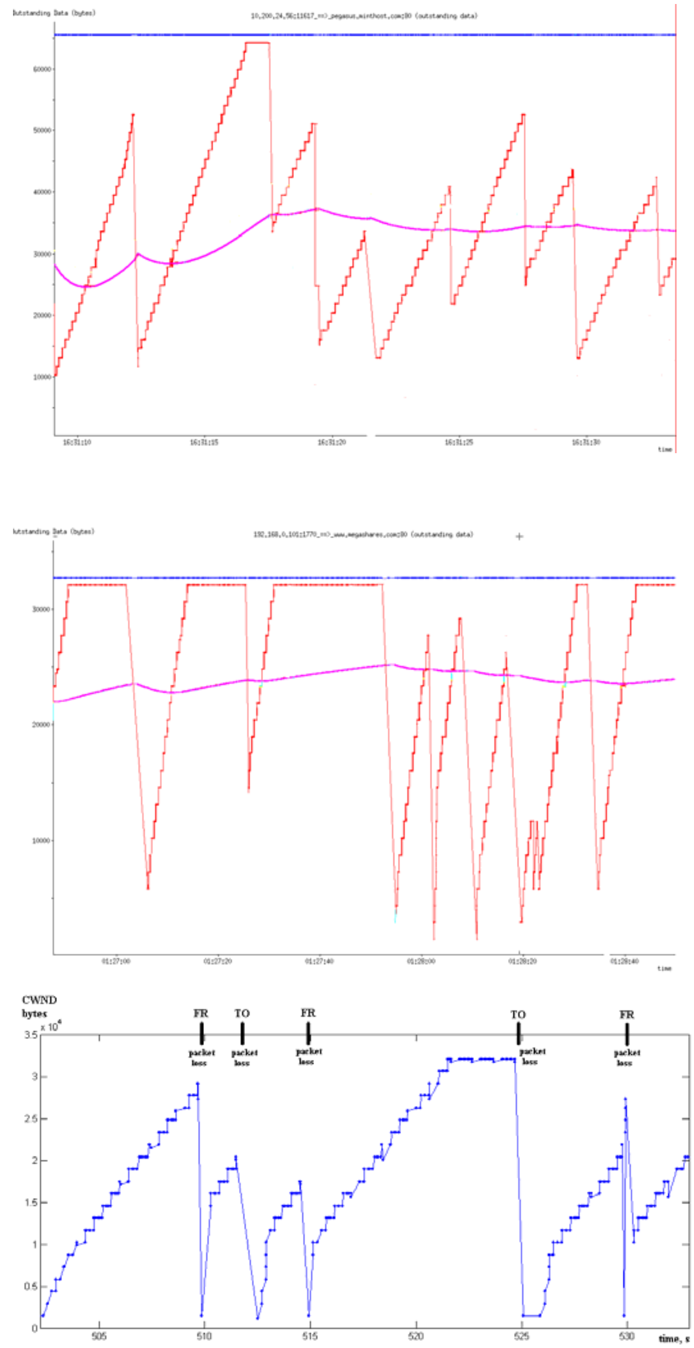


Figure 2. Influence of the computer network environment by means of loss/delay of packets on the size of the congestion window of TCP-connection.

On the figure 2 is drawn the plot of throughput that was calculated by the formula

$$P = \frac{\sum_{j=1}^{N} P_j t_j}{\sum_{j=1}^{N} t_j},$$

where P_j is throughput in the time interval j, t_j - is the duration of this interval, and N is the total number of these intervals in measurement period.

III. TCP STATE-SPACE MODEL

In recent years the modeling of the TCP behavior has received considerable attention [1,2,3,4]. But well-specified RFC models do not allow having a clear understanding concerning dynamic properties of the network traffic. Following to the describing above hypothesis we present a new approach to modeling TCP connection throughput. Therefore throughput of all TCP-connection can be define as throughput in a certain state increased by probability of reach specific state condition:

$$y = \sum_{ij} P_{ij} U_i(k_j)$$

where P_{ij} – probability of staying in U_i state with k_j throughput; the probability calculates as a part of time when TCP is staying in the state characterized by specific measure of throughput.

Based on TCP model in state-space it is possible to increase gradually of throughput estimation by means of an iterative changed the number of states. As a result the total model complexity and corresponding accuracy is characterized by numbers of states, namely:

- Model with one state - slow start (SS) (transitions occur through event TO). The given model is elementary because of all proceeding processes are statistically independent. The interval of loss of the packet is fixed and moves on a time base. (Fig. 3);

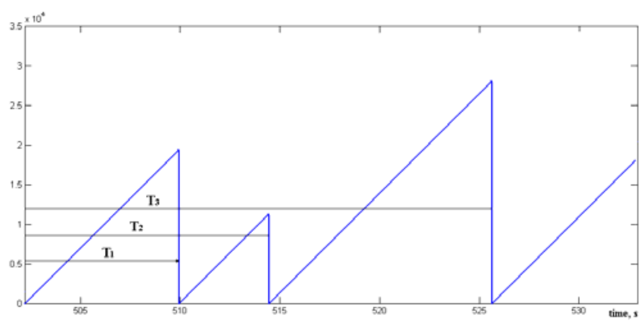


Figure 3. An elementary model of behaviour of the congestion window of TCP-connection with only SS states.

- State of retransmit (transitions occur through a state FR). The model is more adequate since the individual period between loss/delay intervals are correlated with all previous intervals of loss/delay events. The right border of the interval is defined by the moment of achievement the value of destination node congestion parameter. The left border changes depending on value of a congestion window at the moment of the previous loss/delay interval (Fig. 4).

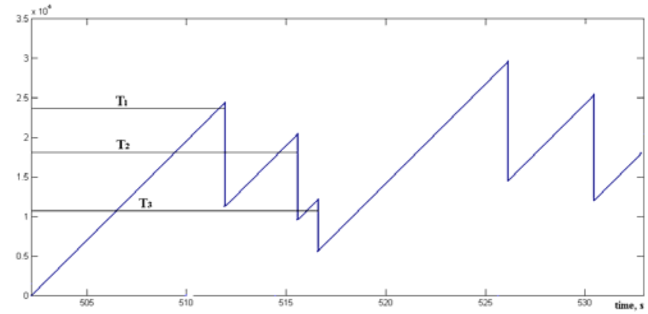


Figure 4. Model of behaviour of the congestion window of TCP-connection with the state of FR

- Model with two states and consecutive alternation of the SS and FR states (Fig. 5).

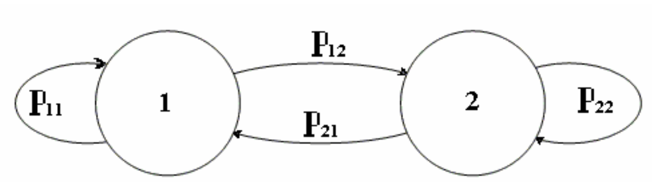


Figure 5. State graph of a congestion window of TCP-connection with SS and FR states.

On Fig. 5: "1" is a state of fast retransmit (FR), "2" is a state of slow start (SS), P_{ij} is the probability of transition from i to j state. For example, experimentally received intervals for this probability are: $P_{12}+P_{22} = 0.17\text{-}0.23$, $P_{11}+P_{21} = 0.77\text{-}0.83$.

On fig. 6 the fragment of CWND is presented. We calculate the relation of quantity FR to the total number of losses/delays of packets to receive probabilities P_{11}, P_{21}, P_{12} and P_{22}, and also we analyze the quantity of the same transitions following one after another.

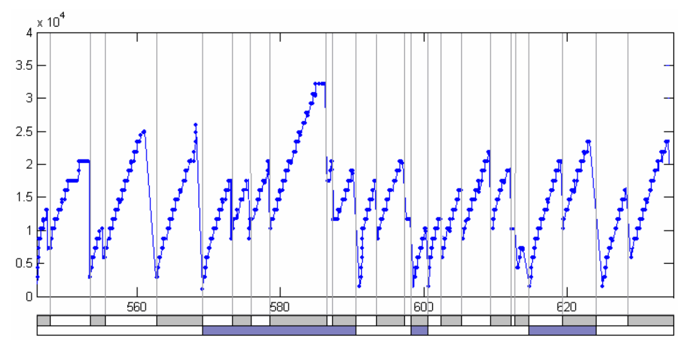


Figure 6. Example of CWND for calculation of probabilities of transition P_{11}, P_{21}, P_{12} and P_{22} .

The proposed model with two states (Fig. 5) can be considered as an enclosure of one simple model into another. In this case the used indexation of arcs probabilities is based on light rule - event TO is the first index, event FR – the second one. Thus the second index at approach of the moment TO is zero, because since that time there is no correlation between the sizes of a congestion window in two segments (Fig. 7).

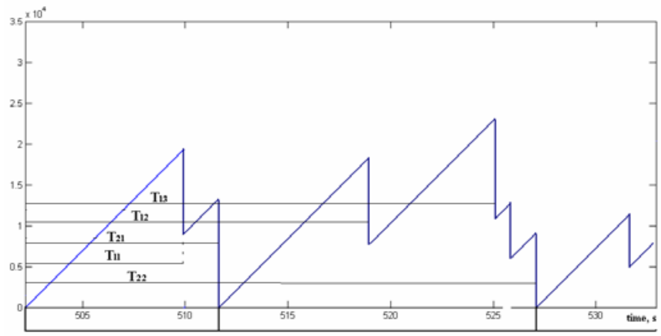


Figure 7. Ccongestion window behaviour o with double indexation for the two states model of SS and FR.

The accuracy of the proposed model can be increase based on model with «the spitted states». For example the needs for split state arises in a case when the considered state is observed at various traffic intensity or overload of the computer network environment, that directly influences values of probability of transitions between states (Fig. 8).

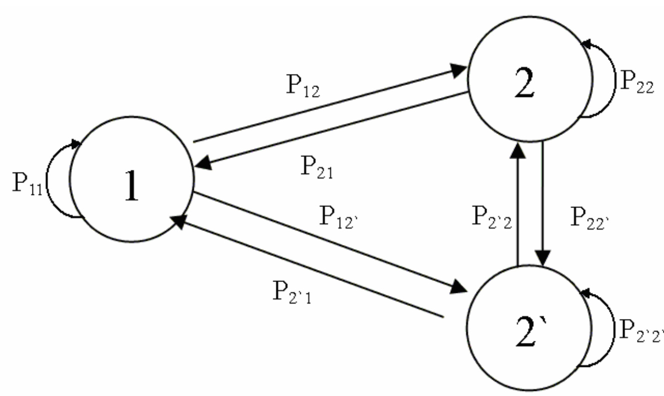


Figure 8. Model of behaviour of a congestion window of TCP-connection with the split state for occurrence of an overload of a computer network.

It is possible to increase the number of the split states and by this way increase accuracy of throughput estimation of individual TCP connection. But taking into account the parametric restriction of the offered model, we will also limit the model complexity by two spitted states. In this case probabilities of transition for the given model are strongly depended on congestion condition of a network environment. It is easy to understand that in case without congestion the TCP model is reduced to the state-space model with two states. In congestion network environment the "2" or FR state is spitted into several sub states, namely ($P_{12`}$, $P_{22`}$, $P_{2`3`}$…) and total model complexity is increase.

Based on value of P_{ij} and model parameter k_i –traffic throughput in "i" state , it is possible to calculate throughput of TCP connection:

$$P = \sum_{i=1}^{N} (\sum_{j=1}^{N} P_{ij}) k_i \ ,$$

where N is total number of states of TCP connection.

IV. The analysis of local behaviour of a congestion window of TCP

From the previous analysis lies the fact that key contribution to the traffic fluctuation is brought from packet layer caused by TCP dynamics. Fractal character of the traffic is caused by losses/delays factors under the influence of congestion of network environment. To create a formal model of this influence we proposal to use specific approximation of distributions function of time interval between model states: parametric uniform and exponential distribution functions (Fig. 9).

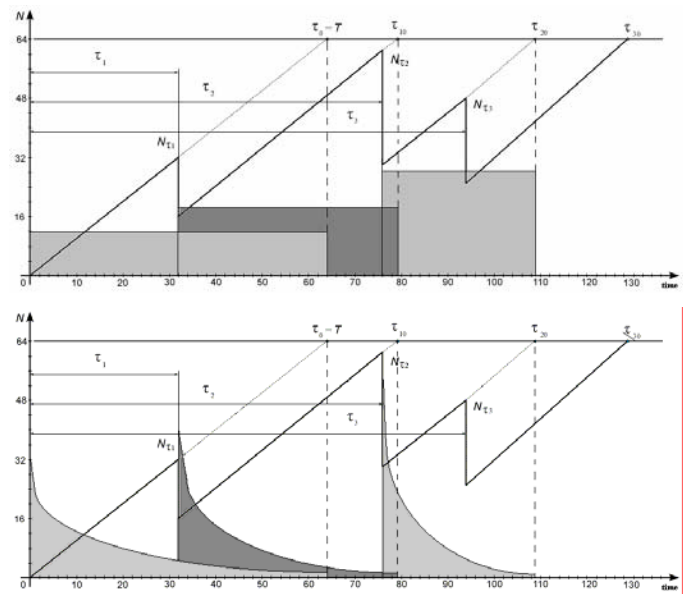
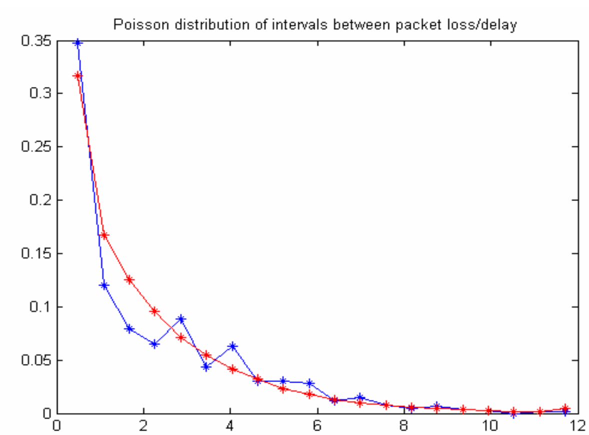


Figure 9. Congestion window of TCP-connection and distribution function of intervals between loss/delay events.

Validations of proposal approximation are following from experimental results observed in various network environment conditions (Fig 10).



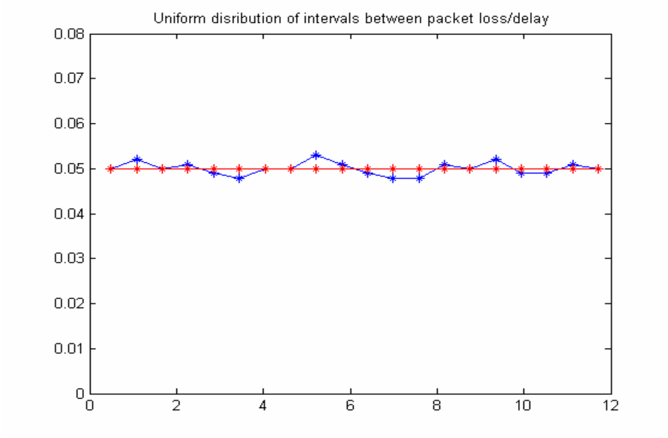


Figure 10. Schedules of distribution of intervals between repeated transfers of packets for networks with high (a) and low (b) loading.

Calculations for parameters of the pure uniform distribution have been resulted earlier [5,6,9]. Proposed methods based on second order estimation and relying on a creative hypothesis of the parametric form of distribution function of time interval between FR phases. Local approximation of each phase of TCP connection during the time interval "k" can be described by formulas (Fig. 8):

$$y_{k+1} = y_k + 1; \quad \text{if} \quad \psi(\xi) = 0$$
$$y_{k+1} = y_k / 2; \quad \text{if} \quad \psi(\xi) = 1,$$

Where $\psi(\xi)$ is function of the indicator of an overload; ξ is a stochastic variable with exponential distribution function with left sliding border. The right border of distribution function is always in a point of possible achievement of value of a receiver window. Thus the area under a distribution curve (S) is less than 1, and the difference 1-S characterizes probability of transition in a state of congestion and directly depends on congestion condition of network environment.

Statistical features of considered model, as well as in a case with pure uniform distribution, can be described by following expressions:

$$\tau_{10} = T + \frac{\tau_1}{2},$$
$$\tau_{30} = T + \frac{\tau_1}{8} + \frac{\tau_2}{4} + \frac{\tau_3}{2},$$

Where τ_1 is the variable with an exponential distribution of density of probability, namely: $\lambda e^{-\lambda\tau}$, $0 \le \tau \le T$,(a=0, *b*=*T*); τ_2 is the variable with distribution: $\lambda e^{-\lambda(\tau-\tau_1)}$, $\tau_1 \le \tau \le \tau_{10}$ (a=τ_1, *b*=τ_{10}); τ_3 is the variable with экспоненциальным distribution: $\lambda e^{-\lambda(\tau-\tau_2)}$, $\tau_2 \le \tau \le \tau_{20}$ (a=τ_2, *b*=τ_{20}).

The moments of the second order of a congestion window based on this model can be expressed by ratios:

$$D(\tau_{20}) = M_2(\tau_{20}) - M_1^2(\tau_{20});$$
$$D(\tau_{30}) = M_2(\tau_{30}) - M_1^2(\tau_{30}).$$

Time interval between FR can be approximated by the function:

$$D(t) = K(t-64)^{1+\alpha},$$

where, for uniform distribution K = $2*10^{-7}$; α = 4, and for exponential - K = $7*10^{-7}$; α = 4.176 (Fig. 11). Therefore TCP traffic as dynamic process possesses scale invariance which obviously opens the internal nature of network interaction.

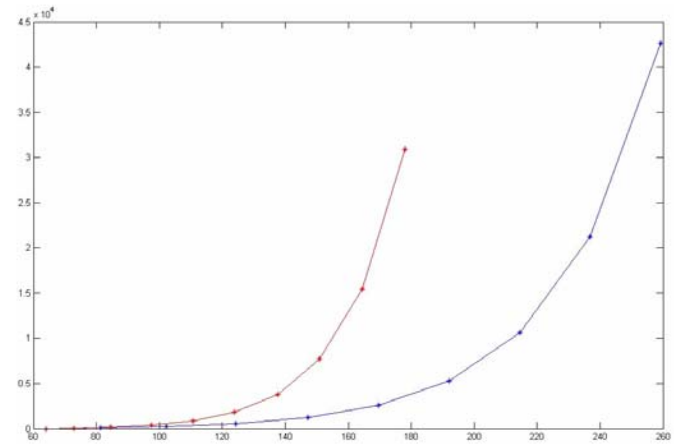


Figure 11. Schedules of the moments of the second order exponential (red) and uniform (blue) distributions

The basic reasons of such behavior are statistical correlation (dependence) between 1) transmission media and network application; 2) TCP capacity and router buffers. This correlation may cause specific burst in packet counts and deform spectral or statistical properties of network traffic. Attractive feature of proposed model is that all mentioned above characteristics including parameter α can be defined in terms of standard network measurement framework.

Proposal model considerer TCP flow as the transition process between several protocol states i.e. as a set of corrected transition from one state to another state which reflects congestion condition and influences characteristics. So, from the formal point of view TCP protocols can be representative as a nonlinear operator that define transition rules and phases between different states of TCP connection.

V. THE CONCLUSION.

1. The necessary tool of research is the formal mathematical model of investigated object. Proposal TCP model with spitted state-space can be used to define complicated network condition and suitable for modeling and management purposes.

2. Each of considered states is characterized by some probabilistic measure, transitions rules and parameters n which define important network characteristics and TCP throughput

in "the spitted states" conditions and "returnable" probabilities.

3. On each observable time interval parameters of all TCP states are defined as average value of current throughput of TCP data flow.

4. The further development of the proposal model is required to define optimal number of the split states and parameters characterizing a connection throughput in different congestion conditions.

References

[1] Floyd S. , Paxon V., Wide-area traffic: The failure of Poisson modeling, IEEE/ACM transaction on networking, 3(3), p.226-244, June 1995.

[2] W.Willinger, R.Govindan, S.Jamin, V.Paxon "Scaling phenomens in the Internet: Critically examing criticality", in Proc.Nat.Acad. Sci USA, vol. 99, Feb.2001 pp. 2573-2580.

[3] Vladimir Zaborovsky Multiscale Network Processes: Fractal and p-Adic analysis //Proceedings of ICT 2003, v 2, pp. 835-840.

[4] J. Cao, W.S.Cleveland, D.Lin. amd D.X.Sun, "On the Nonstationary of Internet traffic", ACM SIGMETRICS, pp. 102-112, 2001.

[5] Gorodetsky A. Structure of the differential equations and fractal natured processes in dynamic systems // Scientific and technical sheets SPbSPU. 2004. № 4.

[6] Gorodetsky A. The system methods of definition of statistical characteristics of the network traffic // Scientific and technical sheets SPbSPU. 2005. № 1.

[7] Zaborovsky V., Meylanov R. Peer-to peer fractal models: new approach to describe multiscale network processes // Proceedings of ICT 2002, Beijing, China, v.1, pp1095-1100.

[8] V.Zaborovsky Network Complexity:Cross-Layer Models and Characteristics. AICT 2007, Morne, Mauritius.

[9] Zaborovsky V., Ygorov S., Fedorov N. TCP flows in stochastic network environment: models and characteristics. ICN 2008