komunikácie
C O M M U N I C A T I O N S

Juraj Smieško *

HYPEREXPONENTIAL MODEL OF TOKEN BUCKET SYSTEM

In the present paper we apply the Theory of Markov Chains for a formal description of a hyperexponential model of Token Bucket System (TBS). This paper extends the topic covered in [3] by replacing Exponential model of human speech by a more precise hyperexponential model.

In the first section we deal with analysis of the Hyperexponential model of On-Off Source of human speech. In the second section we deal with steady-state analysis of Token Bucket system and we solve the balance equations for probabilities of states. In the third section we use real parameters of VoIP and we calculate characteristic of TBS and in the last section the values of probability of packet loss are approximated by an exponential regression function. We can directly compute values of probability of packet loss using this exponential function.

1. Hyperexponential model of On-Off Source

The On-Off Source Model of human speech is composed of two periods, where On period belongs to the active periods in human speech and the Off period belongs to the silence. There are two main states On and Off and the source switches from one state to another.
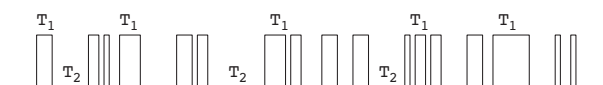

Fig. 1 On-Off model of human speech

Talk-spurt duration is modelled by the variable T_1 and pause duration is modelled by the variable T_2. The probability densities of T_1 and T_2 are modelled by two weighted geometric distribution functions.

Every increment of these variables T_1 and T_2 is equal to 5 ms and then the average talk-spurt duration is $ET_1 = 227$ms and the average pause duration is $ET_2 = 596$ms (see [1]).

$$T_1 \sim f_1(k) = c_1(1 - U_1)U_1^{k-1} + c_2(1 - U_2)U_2^{k-1}$$

$c_1 = 0.60278 \qquad c_2 = 0.39817$
$U_1 = 0.92446 \qquad U_2 = 0.98916$

$$T_2 \sim f_2(k) = d_1(1 - W1)W_1^{k-1} + d_2(1 - W_2)W_2^{k-1}$$

$d_1 = 0.76693 \qquad d_2 = 0.23307$
$W_1 = 0.89700 \qquad W_2 = 0.99791$
for $k = 0, 1, 2, \dots$.

The best approximation of these processes is the Hyperexponential model. The discrete geometrical distributions are approxi-
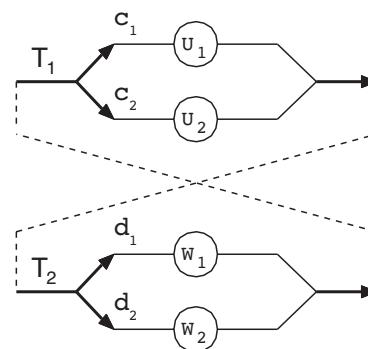
mated by continuous exponential ones and the rest of this model has completely the same nature as the geometrical one. In the new model we can keep the same averages.

ON: $T_1 \sim f_1(t) = c_1\alpha_1 e^{-\alpha_1 t} + c_2\alpha_2 e^{-\alpha_2 t}$
$c_1 = 0.60278 \quad c_2 = 0.39817 \quad \alpha_1 = 0.01511 \quad \alpha_2 = 0.00217$

OFF: $T_2 \sim f_2(t) = d_1\mu_1 e^{-\mu_1 t} + d_2\mu_2 e^{-\mu_2 t}$
$d_1 = 0.76693 \quad d_2 = 0.23307 \quad \mu_1 = 0.02060 \quad \mu_2 = 0.00042$

Fig. 3 and Fig. 4 show approximation of hypergeometrical distributions (points) by hyperexponential distributions for On and Off periods.

We construct a transmission diagram of Markov model of On-Off Source (loops are omitted):

The rate matrix of this Markovov chain is

$$Q_0 = \begin{pmatrix} -\alpha_1 & 0 & d_1\alpha_1 & d_2\alpha_1 \\ 0 & -\alpha_2 & d_1\alpha_2 & d_2\alpha_2 \\ c_1\mu_1 & c_2\mu_1 & \mu_1 & 0 \\ c_1\mu_2 & c_2\mu_2 & 0 & \mu_2 \end{pmatrix}$$

Fig. 2 Hyperexponential On-Off Source

* **Juraj Smieško**
Department of Mathematical Methods, Faculty of Management Science and Informatics, University of Žlina

Fig. 3 Approximation for On period
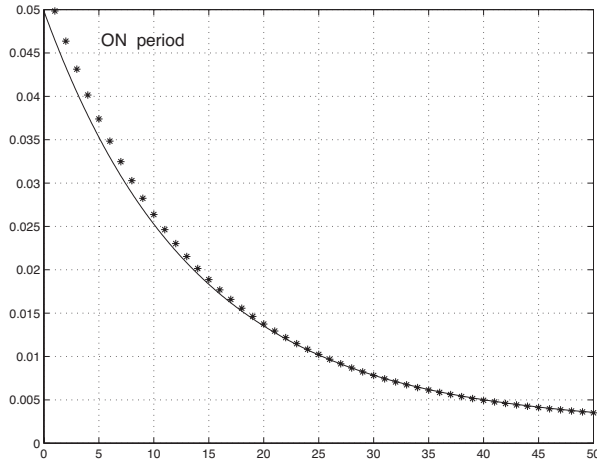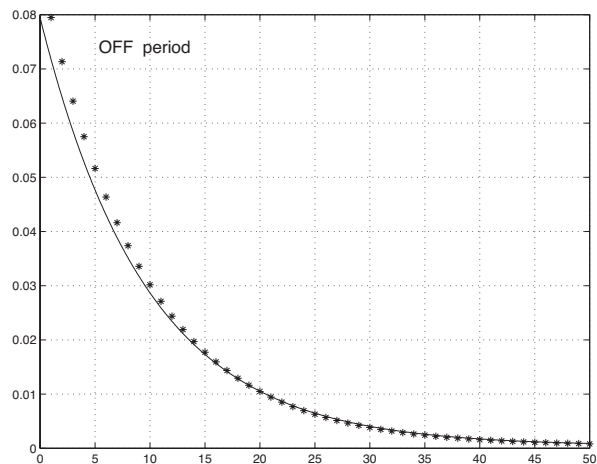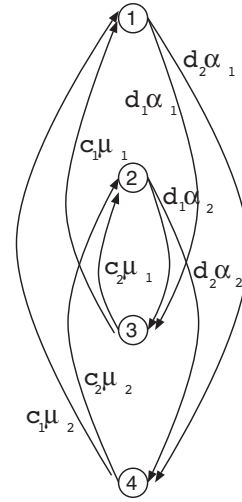

Fig. 5: Markov chain for hyperexponential model

Notice that these probabilities are very similar to those in the Exponential model (see [3], $P_{ON} = 0.27586$ and $P_{OFF} = 0.72414$). But in the Hyperexponential model we have a more detailed structure of On-Off source than in the Exponential model.

2. Steady-state Analysis of Token Bucket System

The Token Bucket System is related to VoIP problems. We have formed the steady-state analysis of VoIP under the Token Bucket Control. Our main problems will be to compute the probability characteristics of TBS and to find relation between probability of packet loss and bucket depth. At first we will analyze the working of TBS in general.

There is a flow of packets entering the TBS. The Token Bucket System is generating tokens (marks) and then it marks each packet with one of them. Only the marked packets will go through network. For practical reasons we can assume a limited bucket with depth "n". When the bucket is empty (there are no tokens), the TBS cannot mark any packet, which is lost then. Let P_{lst} be probabilities of this random phenomenon and we will call it "probability of packet loss".

Fig. 4: Approximation for Off period

We solve the balance equations $\pi . Q_0 = 0$, where $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$:

$$\pi_2 = \frac{c_2\alpha_1}{c_1\alpha_1} \pi_1 \qquad \pi_3 = \frac{d_1\alpha_1}{c_1\mu_1} \pi_1 \qquad \pi_4 = \frac{d_2\alpha_1}{c_1\mu_2} \pi_1$$

We compute the probability π_1 from the normalisation condition $\sum_{i=1}^{4} \pi_i = 1$

$$\pi_1 = \frac{c_1\alpha_2\mu_1\mu_2}{(c_1\alpha_2 + c_2\alpha_1)\mu_1\mu_2 + (d_1\mu_2 + d_2\mu_1)\alpha_1\alpha_2}$$

The probabilities for real parameters of VoIP are $\pi = (0.04902, 0.22546, 0.04579, 0.67973)$.

Let P_{ON} and P_{OFF} be probabilities that the Source is in state On or Off:

$$P_{ON} = \pi_1 + \pi_2 = 0.27448 \qquad P_{OFF} = \pi_3 + \pi_4 = 0.72552$$

We will deal with the Hyperexponential model of On-Off Source of human speech (see above). When the Source is in On state it starts generating flow of packets, which represents human speech. The packet rate in usual VoIP systems (using G.729A) is 50 p/s. We will assume that the flow of packets is modelled by Poisson process $N_p(t)$ with rate $\eta = 0.05$ *p/ms*.

In general there can be any token rate. Usually it is the same as an average number of packets entering the TBS. The flow of tokens will be modelled by Poisson process $N_p(t)$ with a rate λ:

$$\lambda = \eta . P_{ON} + 0 . P_{OFF} = 50 \, p/s . 0.27448 = 13.724 \, p/s = $$
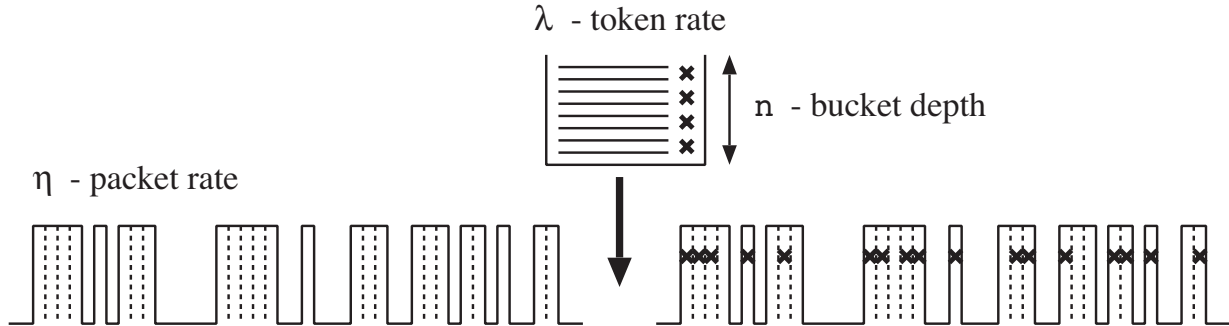$$= 0.01372 \, p/ms$$

Fig. 6: Token Bucket system

In the Token Bucket System we have three elementary random phenomena entering the packet, generating token and switching between On-Off states in Source. We have assumed that each of them is Poisson process, and because of that we can model this TBS by Markov chain.

Probability of an increasing number of tokens in the bucket by one during "short" time Δt is $P(N_T(\Delta t) = 1) = \lambda \Delta t + o(\Delta t)$. If the bucket contains k tokens, probability of a decreasing number of tokens in the bucket by one during "short" time Δt is $P(N_P(\Delta t) = 1) = \eta \Delta t + o(\Delta t)$. If there is an arriving packet and the bucket is empty, we cannot mark it and this packet is lost. Probability of packet loss was named P_{lst}.

We will construct a transition diagram of the whole system. Component columns of the diagram refer to a number of tokens in the token bucket. We assigned this as "k-level" for $k = 0, \dots n$. For easy reading loops in the transition diagram are omitted:

Let X_k and Y_k be probabilities that in the TBS are k tokens and Source is in On states and let Z_k and W_k be probabilities that in the TBS are also k tokens, but Source is in Off states. Let P_k be a probability of k-level. We see that $P_k = X_k + Y_k + Z_k + W_k$ and $P_{lst} = X_0 + Y_0$. Taking from the Theory of Markov Chains [2] we

can write the following balance equations for state probabilities of steady-state Markov chain:

for $k = 0$

$$0 = -(\lambda + \alpha_1)X_0 + c_1\mu_1 Z_0 + c_1\mu_2 W_0 + \eta X_1$$
$$0 = -(\lambda + \alpha_2)Y_0 + c_2\mu_1 Z_0 + c_2\mu_2 W_0 + \eta Y_1$$
$$0 = d_1\alpha_1 X_0 + d_1\alpha_2 Y_0 - (\lambda + \mu_1)Z_0$$
$$0 = d_2\alpha_1 X_0 + d_2\alpha_2 Y_0 - (\lambda + \mu_2)W_0$$

for $k = 1, \dots, n-1$

$$0 = \lambda X_{k-1} - (\lambda + \alpha_1 + \eta)X_k + c_1\mu_1 Z_k + c_1\mu_2 W_k + \eta X_{k+1}$$
$$0 = \lambda Y_{k-1} - (\lambda + \alpha_2 + \eta)Y_k + c_2\mu_1 Z_k + c_2\mu_2 W_k + \eta Y_{k+1}$$
$$0 = \lambda Z_{k-1} + d_1\alpha_1 X_1 + d_1\alpha_2 Y_k - (\lambda + \mu_1)Z_k$$
$$0 = \lambda W_{k-1} + d_2\alpha_1 X_1 + d_2\alpha_2 Y_k - (\lambda + \mu_2)W_k$$

for $k = n$

$$0 = \lambda X_{k-1} - (\alpha_1 + \eta)X_n + c_1\mu_1 Z_n + c_1\mu_2 W_n$$
$$0 = \lambda X_{k-1} - (\alpha_2 + \eta)Y_n + c_2\mu_1 Z_n + c_2\mu_2 W_n$$
$$0 = \lambda Z_{k-1} + d_1\alpha_1 X_n + d_1\alpha_2 y_n - \mu_1 Z_n$$
$$0 = \lambda W_{k-1} + d_2\alpha_1 X_n + d_2\alpha_2 Y_n - \mu_2 W_n$$

We want to solve the balance equations $p \cdot Q = 0$, where $p = (X_0, Y_0, Z_0, W_0 \dots X_n, Y_n, Z_n, W_n)$. The rate matrix Q is more complicated, so we will write it to the block matrix:
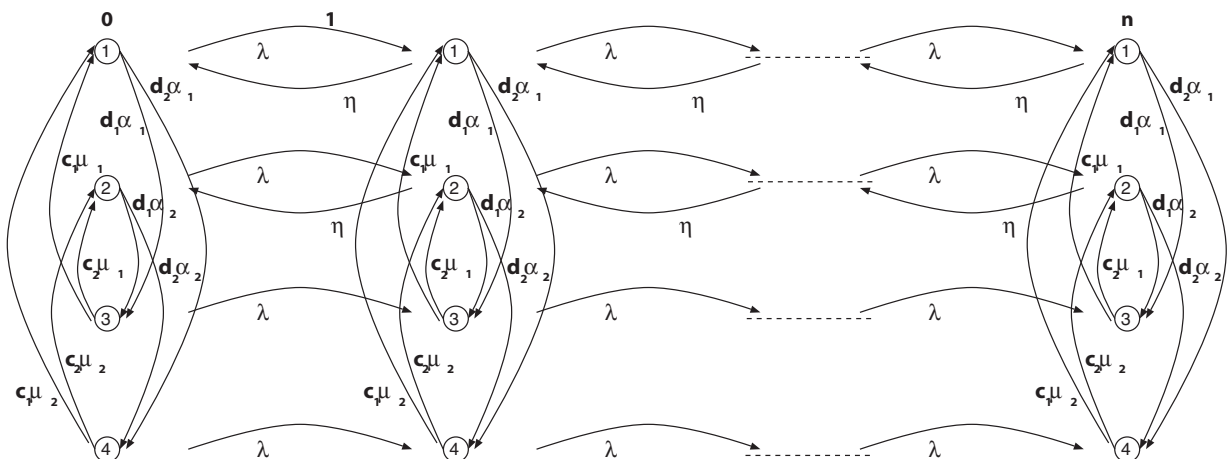

Fig. 7: Markov chain for Hyperexponencial TBS

$$Q = \begin{pmatrix} A_0 & \Lambda & 0 & \ldots & 0 & 0 & 0 \\ \Gamma & A & \Lambda & \ldots & 0 & 0 & 0 \\ 0 & \Gamma & A & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & A & \Lambda & 0 \\ 0 & 0 & 0 & \ldots & \Gamma & A & \Lambda \\ 0 & 0 & 0 & \ldots & 0 & \Gamma & A_\nu \end{pmatrix} = \begin{pmatrix} A+\Gamma & \Lambda & 0 & \ldots & 0 & 0 & 0 \\ \Gamma & A & \Lambda & \ldots & 0 & 0 & 0 \\ 0 & \Gamma & A & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & A & \Lambda & 0 \\ 0 & 0 & 0 & \ldots & \Gamma & A & \Lambda \\ 0 & 0 & 0 & \ldots & 0 & \Gamma & A+\Lambda \end{pmatrix}$$

$$A_0 = \begin{pmatrix} -(\lambda+\alpha_1) & 0 & d_1\alpha_1 & d_2\alpha_1 \\ 0 & -(\lambda+\alpha_2) & d_1\alpha_2 & d_2\alpha_2 \\ c_1\mu_1 & c_2\mu_1 & -(\lambda+\mu_1) & 0 \\ c_1\mu_2 & c_2\mu_2 & 0 & -(\lambda+\mu_2) \end{pmatrix} \qquad A_0 = \begin{pmatrix} -(\alpha_1+\eta) & 0 & d_1\alpha_1 & d_2\alpha_1 \\ 0 & -(\alpha_2+\eta) & d_1\alpha_2 & d_2\alpha_2 \\ c_1\mu_1 & c_2\mu_1 & -\mu_1 & 0 \\ c_1\mu_2 & c_2\mu_2 & 0 & -\mu_2 \end{pmatrix}$$

$$A = \begin{pmatrix} -(\lambda+\alpha_1+\eta) & 0 & d_1\alpha_1 & d_2\alpha_1 \\ 0 & -(\lambda+\alpha_2+\eta) & d_1\alpha_2 & d_2\alpha_2 \\ c_1\mu_1 & c_2\mu_1 & -(\lambda+\mu_1) & 0 \\ c_1\mu_2 & c_2\mu_2 & 0 & -(\lambda+\mu_2) \end{pmatrix} \quad \Gamma = \begin{pmatrix} \eta & 0 & 0 & 0 \\ 0 & \eta & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \Lambda = \begin{pmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

We assigned Q_0 as rate matrix of the Hyperexponential On-Off Source. It can be easily seen that:

$$A_0 = A + \Gamma, \ A_n = A + \Lambda \ \text{and} \ Q_0 = \Gamma + A + \Lambda$$

We can determine recurrent formulas for the probabilities W_i and Z_i:

$$(\lambda + \mu_2)W_0 = d_2\alpha_2 Y_0 + d_2\alpha_1 X_0$$
$$(\lambda + \mu_2)W_k = d_2\alpha_2 Y_k + d_2\alpha_1 X_k + \lambda W_{k-1} \quad k = 1, ..., n-1$$
$$\mu_2 W_n = d_2\alpha_2 Y_n + d_2\alpha_1 X_n + \lambda W_{n-1}$$

$$(\lambda + \mu_1)Z_0 = d_1\alpha_2 Y_0 + d_1\alpha_1 X_0$$
$$(\lambda + \mu_1)Z_k = d_1\alpha_2 Y_k + d_2\alpha_1 X_k + \lambda Z_{k-1} \quad k = 1, ..., n-1$$
$$\mu_1 Z_n = d_1\alpha_2 Y_n + d_1\alpha_1 X_n + \lambda Z_{n-1}$$

But the equations for probabilities Y_i and X_i are very complicated:

$$\eta_1 X_1 + c_1\mu_2 W_0 + c_1\mu_1 Z_0 = (\lambda + \alpha_1)X_0$$
$$\eta_1 X_{k+1} + c_1\mu_2 W_k + c_1\mu_1 Z_k + \eta Y_k = (\lambda + \alpha_1)X_k +$$
$$+ \lambda W_{k-1} + \lambda Z_{k-1} + \lambda Y_{k-1} \quad\quad k = 1, ..., n-1$$
$$c_1\mu_2 W_n + c_1\mu_1 Z_n + \eta Y_n = \alpha_1 X_n + \lambda W_{n-1} \ \lambda Z_{n-1} + \lambda Y_{n-1}$$

$$\eta_1 X_k + \eta_1 X_k = \lambda W_{k-1} + \lambda Z_{k-1} + \lambda Y_{k-1} + \lambda X_{k-1} +$$
$$k = 1, ..., n$$

If we assign $p_k^{on} = X_k + Y_k$ and $p_k^{off} = Z_k + W_k$ we have these interesting relations:

$$p_k^{on} = \frac{\lambda}{\eta} \cdot p_{k-1}^{off} + \frac{\lambda}{\eta} p_{k-1}^{on} \quad k = 1, ..., n$$

3. TBS with real parameters of VoIP

To deduce an explicit formula for probability of packet loss $P_{lst} = P_{lst}(n) = X_0 + Y_0$ is unreal, but there is no problem to compute values of P_{lst} for real parameters of our Token Bucket System with the Exponential On-Off Source:

$c_1 = 0.60278$	$c_2 = 0.39817$
$\alpha_1 = 0.01511 \ \text{ms}^{-1}$	$\alpha_2 = 0.00217 \ \text{ms}^{-1}$
$d_1 = 0.76693$	$d_2 = 0.23307$
$\mu_1 = 0.02060 \ \text{ms}^{-1}$	$\mu_1 = 0.00042 \ \text{ms}^{-1}$
$\lambda = 0.01372 \ \text{p/ms}$	$\eta = 0.05 \ \text{p/ms}$

For example the reader can see for himself the difference between $n = 4$ and $n = 5$:

n	X_k	Y_k	Z_k	W_k	P_k
0	0.0237	0.1507	0.0153	0.0113	0.2010
1	0.0098	0.0453	0.0116	0.0150	0.0817
2	0.0060	0.0164	0.0075	0.0167	0.0466
3	0.0048	0.0079	0.0050	0.0177	0.0354
4	0.0046	0.0051	0.0064	0.6191	0.6352

n	X_k	Y_k	Z_k	W_k	P_k
0	0.0222	0.1475	0.0146	0.0108	0.1951
1	0.0091	0.0444	0.0111	0.0143	0.0789
2	0.0054	0.0162	0.0070	0.0158	0.0444
3	0.0042	0.0080	0.0046	0.0167	0.0335
4	0.0040	0.0052	0.0034	0.0174	0.0300
5	0.0041	0.0041	0.0049	0.6047	0.6178

For real use it is enough to have the bucket depth $n = 1, ..., 10$. Now we will increase the bucket depth n and calculate characteristics of models:

$P_{lst}(n)$ – probability of packet loss or probability of an empty bucket in time of arriving packet

$\lambda \cdot P_{lst}(n)$ – average number of lost packets

P_n – probability of full bucket (bucket will refuse tokens)

EK – average bucket depth

$H = EK/n$ – token bucket usage

n	$P_{lst}(n)$	$\lambda \cdot P_{lst}(n)$	P_n	EK	χ
1	0.2079	2.852 p/s	0.7574	0.7574	75.5%
2	0.1892	2.596 p/s	0.6893	1.4696	73.5%
3	0.1803	2.474 p/s	0.6567	2.1543	71.8%
4	0.1744	2.393 p/s	0.6351	2.8218	70.6%
5	0.1697	2.328 p/s	0.6178	3.4779	69.6%
6	0.1655	2.271 p/s	0.6027	4.1256	68.8%
7	0.1617	2.219 p/s	0.5888	4.7664	68.1%
8	0.1582	2.170 p/s	0.5758	5.4015	67.5%
9	0.1548	2.124 p/s	0.5636	6.0314	67.0%
10	0.1517	2.081 p/s	0.5519	6.6566	66.6%

4. Relation between Probability of packet loss and Token Bucket depth

The most interesting characteristic is probability of packet loss $P_{lst}(n)$. In Fig.7 we compared the values of probability of packet loss $P_{lst}(n)$ for the Exponential model and Hyperexponential model (see [3]).
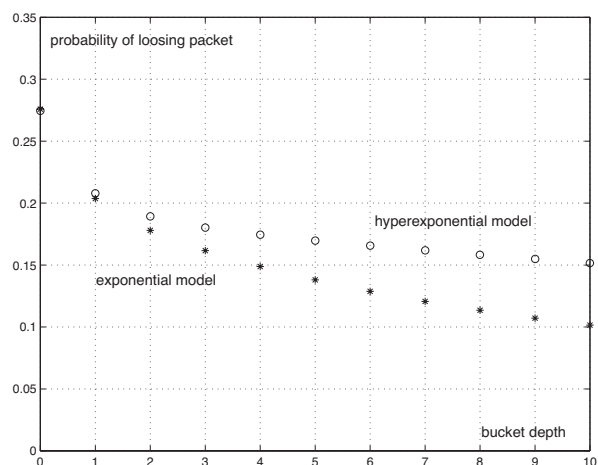

Fig. 8: Compared between exponential and hyperexponencial model

The approximation by hyperexponencial distributions is the best approximation and therefore the Hyperexponential Model of TBS is more accurate in modelling the real TBS than the Expo-nential model, but for computation of state probabilities we had to use numerical methods.

The approximation by exponential distributions is not so "good", but for this model we have gained recurrent formulas for state probabilities and values of characteristics that we get from the Exponential model can be used as lower estimation of characteristics of the real TBS.

Now we use approximation by the regression function $f(n)$ by the least squares method to find relation between Probability of packet loss $P_{lst}(n)$ and the Token Bucket depth n. The Regression function must satisfy these conditions to gain solid approximation:

$$\lim_{n \to \infty} f(n) = 0, \quad \nexists n_0; f(n_0) = 0 \quad \text{and} \quad \forall n_1 < n_2; f(n_1) > f(n_2)$$

If we want few parameters of regression function, there is an ideal exponential function $f_t(n) = ae^{bn}$.

We will measure the quality of approximation by the square root of sum residuals:

$$\varepsilon = \sqrt{\sum_{k=1}^{n} [P_{lst}(k) - f(k)]^2}$$

For real use it is enough to have the bucket depth $n = 1, ..., 10$. Then the exponential regression function is $f(n) = 0.2025 \cdot e^{-0.0312n}$ with $\varepsilon = 0.0144$ and with the mean error $\varepsilon/n = 0.0014$.
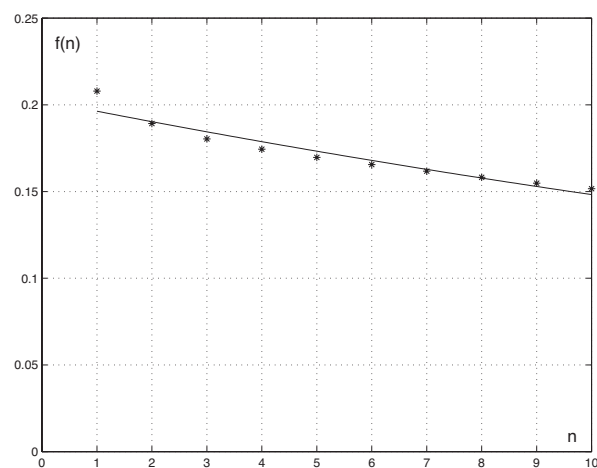
n	$f(n)$	$P_{lst}(n) - f(n)$	n	$f(n)$	$P_{lst}(n) - f(n)$
1	0.1963	0.0116	6	0.1680	−0.0024
2	0.1903	−0.0010	7	0.1628	−0.0011
3	0.1844	−0.0041	8	0.1578	0.0004
4	0.1788	−0.0044	9	0.1529	0.0019
5	0.1733	−0.0036	10	0.1482	0.0034

Fig. 9: Approximation by exponential function for n = 1, ..., 10

The exponential function is approximation (except $n = 1$) with a maximum error $4.4 \cdot 10^{-3}$. If we are satisfied with this precision, we can exchange $P_{lst}(n)$ for $f(n)$ for $n = 2, ...,10$.
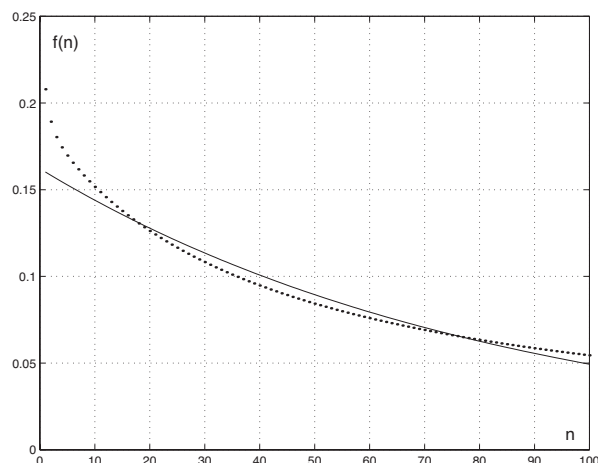

Fig. 10: Approximation by exponential function for n = 1, ..., 100

If we use the bucket depth $n = 10, ..., 100$, the exponential regression function is $f_{lst}(n) = 0.1621.c^{0.0119n}$ with $\varepsilon = 0.0807$ and with the mean error $\varepsilon/n = 0.0008$.

This approximation is very "good" for the depth $n = 11, ..., 100$ with a maximum error $6.4 \cdot 10^{-3}$.

For real use we are satisfied with the function:

$$f(n) = \begin{cases} 0.2745 & \text{for} \quad n = 0 \\ 0.2079 & \text{for} \quad n = 1 \\ 0.2025 \cdot e^{-0.0312n} & \text{for} \quad n = 2, ..., 10 \\ 0.1621 \cdot e^{-0.0119n} & \text{for} \quad n = 11, ..., 100 \end{cases}$$

5. Conclusion

We executed steady-state analysis of Hyperexponential model of Token Bucket System and created balance equations for states of Markov chain. Analytical solution led to complicated recurrent formulas for state probabilities. Therefore we used numerical methods to compute probability of a losing packet. We approxiamted these results using the exponential function. Finally we obtained the function for direct calculation of this probability.

References:

[1] ITU-T Recommendation P. 59, *Artificial Conversational Speech,* 1993.
[2] PEŠKO & SMIEŠKO, J.: *Stochastics Models of Operation Analysis* (in Slovak), Žilinská univerzita, Žilina 1999.
[3] SMIEŠKO, J.: *Exponential Model of Token Bucket System,* Communications – Scientific Letters of the University of Zilina, 4/2003, 66–70.