Method of choice of test signals for automatic intrusive measurement VTQoS

Peter Počta, Martin Vaculík
University of Žilina, FEE, Dept. of Telecommunication, Univerzitná 1, 010 26, Slovakia
pocta@fel.utc.sk , vaculik@fel.utc.sk

Abstract

This document describes the simulations in transmission chain. The aim of simulations was to obtain optimal testing signal determined for automatic intrusive measurement VTQoS (Voice Transmission Quality of Service). Evaluation of simulation signal was based on calculation of mean square measure and correlation coefficient. Suitable signals of beforehand defined group were chosen on the basis of these parameters. Suitable signals will be used to create resulting test sequence.

Keywords

VTQoS, test sequence, correlation coefficient, mean square measure, intrusive measurement

1. Introduction

VTQoS is one the important parts of QoS (Quality of Service). It is very important for the granters and for the users. An increase in the complication and the complexity of networks is vissible, when communication network includes in more and more transmission technologies. The measurement of the voice transmission quality becomes one platform for other platforms, by help of it we can compare simultaneously different transmission technologies and it is such closer to view of the users.

It is possible to measure and evaluate transmission parameters of networks. Evaluation of quality end-to-end provides optimal results according to complexity of network technologies. The participants evaluate it with same way, too.

Voice service is the most wide-spread service, which receiving participant uses filter and predicative abilities of human brain in. It is very important optimal evaluation for this service.

Evaluation of the Quality of voice service exert with intrusive and non-intrusive method, objectively or subjectively.

We only monitor existing dialogue by non-intrusive method. Evaluation algorithm has not got original sample of signal primary. It is a disadvantage of this method. Also it is very difficult to detect some types of interference signal for transmission.

We make a connection for intrusive method, which only serves for transmission of testing voice sample. These methods are known from the beginning of telecommunication technique, when defined sequence vowels-logathoms were transmitted after building-up connection. A receiver had to recognize these logathoms. This way of subjective valuation is used to nowadays (e.g. method MoS).

Recent technique and program equipment provides the objectification of this measurement way with transmitting beforehand defined sound sample, it is receivement on destination side and a comparison of transmission sample and original sample with suitable algorithm, which imitates the way of perception and evaluation of quality transmission opinion by average listener. It is for example E-model defined in ETR-250, or algorithm PSQM (Perceptual Speech Quality Measurement) defined in P.861 ITU-T also PESQ (Perceptual Evaluation of Speech Quality) defined in P.862 ITU-T.

Choice of optimal test sequence is very important for all these methods.

Recommendation P.50 ITU-T is interested in problem of test signals. It is possible to use either real voiced records or special artificial speech as testing signal. The speech consists of voiced and voiceless consonants.

The generator of impulses is used for the generation of voiced consonants, which input parameter is the frequency of the impulses f_0 in. It introduces the height of the voice. We can use a generator of noise for generation voiceless consonants.

Nowadays automatic intrusive measurements VTQoS are realized by real voice records. This method is very objective, if we use only one language. If we want to realize tests much more objectively, we must use speech sample from many languages. We can characterize every language with specific formants. The first three formants are the most important. We can specify a kind of phones by help frequency and width of formants.

These formants are different influenced with transmission chain.

Therefore the aim of this document is to choose suitable signals, which will be used for creation optimal test sequence. The choice of suitable signals for test sequence will be realised from the set of Non-speech-like (fully artificial) signals, which are defined in P.501 ITU-T. These are classical measuring signals. P.501 ITU-T recommendation divides them into deterministic and random signals. Optimal test sequence will provide quality testing without depending on characterstics of concrete speech.

If we use these signals, we can compare results which were realized in other countries simultaneously. For example comparison of networks in given corporation in some countries (Deustche Telecom, Orange, Vodafone,....). [1], [2]

2. Description of simulation model

Test signals are modeled in Matlab as the sequence of digital samples, which pass through the competent type of communication channel (Noise Channel AWGN and Binary Symetric Channel BSC).

The task of this model is to simulate situations in transmission signal through the given transmission chain especially from the point of view of the channel parameters. Main parameters are Error Probability for BSC channel and parameter SNR for AWGN channel. The model is made to occur many errors for transmission. Therefore I did not use any channel coding. We need maximum number of errors, because we want to get the most sensitive signal for these errors. Original source and destination file are recorded by sound card into wav file. The both files are compared simultaneously. This comparison is based on calculation of mean square measurement and correlation coefficient. The determination of suitable signals, which will be used for intrusive measurement was the aim of the simulation.
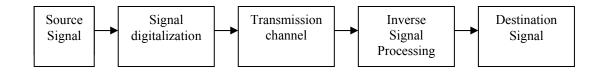

Fig. 1   Principal arrangement of model

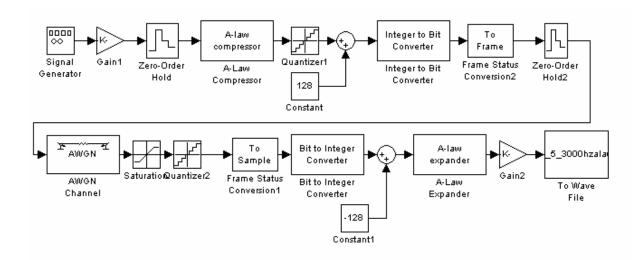**2.1 More detailed description of simulation model**

Fig. 2   Simulation model for AWGN channel

Simulations were done for two models. I only present one simulation model for AWGN channel. The blocks „Zero-Order Hold 2", „Saturation 2", „Quantizer 2" are not used in the simulation model for BSC. Instead of AWGN channel is used BSC channel.

Source signal is generated with the block „Signal Generator". This block provides generation of harmonical signal, impulse or noise signal with beforehand defined parameter.You must specify the Initial seed vector in the simulation. The Initial seed parameter initializes the random number generator that the Gaussian Noise Generator block uses to add noise to the input signal. The probability density function of n-dimensional Gaussian noise is

$$f(x) = \left((2\pi)^n \det[K]\right)^{-1/2} \exp\left(-(x-\mu)^T[K]^{-1}(x-\mu)/2\right) \quad (1)$$

where x is a length-n vector, $[K]$ is the n-by-n covariance matrix, μ is the mean value vector, and the superscript T indicates matrix transpose.

It is possible to modify output signal from the generator in block „Gain", which multiplies the signal by the factor K. y= K*u, where K can be a scalar, vector, matrix.
The signal is sampled simultaneuously (block „Zero-Order Hold"). The block „Zero-Order Hold" provides a mechanism for discreting one or more signals in time or resampling the signal at a different rate. It works on the basis of Shannon and Kotelnik's sample theorem, which says,

$$f_{vz} \geq 2 f_{max} \quad (2)$$

where f_{vz} is sample frequency and f_{max} is maximum frequency sampled signal.
Sampling signal is given in the compression with the compression curve for the A-law or μ-law in the block „A-law (μ-law) Compressor". The formula for modeling of output signal is:

$$y = \frac{A|x|}{1+\log A} \text{sgn}(x) \qquad \text{for } 0 \leq |x| \leq \frac{V}{A} \quad (3)$$

$$y = \frac{V(1+\log(A|x|/V))}{1+\log A} \text{sgn}(x) \quad \text{for } \frac{V}{A} \leq |x| \leq V \quad (4)$$

where A= 87.56 is the parameter of the compressor for the curve A,V is the peak magnitude of x, log is the natural logarithm, and sgn is the signum function (sign in Matlab). This characteristic is used in Europe. For μ curve is the formula similar, where the constant is μ= 255. The formula for modeling of output signal is:

$$y = \frac{V \log(1+\mu|x|/V)}{\log(1+\mu)} \text{sgn}(x) \quad (5)$$

This characteristic is used in USA, Canada, Japan.

Compressed signal is quantized in the block „Quantizer" by help of the round-to-nearest method, which produces an output that is symmetric about zero. Formula is:

$$y = q * \text{round } (u/q) \quad (6)$$

where y is the output, u the input, and q the Quantization interval parameter.

In the next block („Integer to Bit Converter") is mapped each integer at the input vector to a group of bits in the output vector. If M is the number of bits per integer parameter, then input integers must be between 0 and 2^M-1, when first is MSB. As a result, the output vector length is M times the input vector length.

Resulting data flow is set into frames or sequence samples. It determines the block „Frame Status Conversion".

Transmission channels
Block „AWGN channel" adds Gaussian white noise to real or complex input signal. When the input signal is real, this block adds real Gaussian noise and produces a real output signal. When the input signal is complex, this block adds complex Gaussian noise and produces a complex output signal.
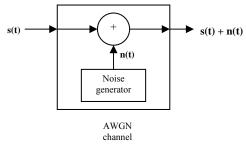

AWGN channel

Fig. 3 AWGN channel model

Block „BSC" introduces binary errors in transmission signal. Binary signal is transmitted on the input port. This block processes each vector element independently, and introduces an error bit in a given spot with probability Error probability. The first output port is binary signal that has passed through the channel. The second output port is the vector of errors that were

introduced. Condition acceptance characteristics of y_j symbol provided that x_i was sent out. They are given with these formulas:

$$p(y_j / x_i) = \frac{p_b}{s-1} \quad \text{for} \quad i \neq j \qquad (7)$$

$$p(y_j / x_i) = 1 - p_b \quad \text{for} \quad i = j \qquad (8)$$

where :
p_b is probability wrong acceptance of symbol
$1-p_b$ is probability right acceptance of symbol
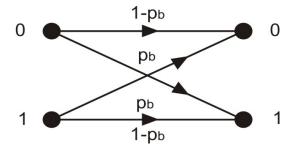s is number of states.

Fig. 4 Transition characteristic

Transmission channel has got a function of the limiter, when the signal oversteps adjusted values, then limitation sets in.

The signal is converted with inverse way, when it passes the line. At first sequence of bits changes to parallel word with length of M bits in the block „Bit to Integer Converter", then the block „A-law (µ-law) Expander" restores data, which were compressed with a compressor. The formula for the expander is inverted to compressed function.

For A-law is:

$$x = \frac{y(1 + \log A)}{A}$$
$$\text{for} \quad 0 \leq |y| \leq \frac{V}{1 + \log A} \qquad (9)$$

$$x = \exp(|y|(1 + \log A)/V - 1)\frac{V}{A}\operatorname{sgn}(y)$$
$$\text{for} \quad \frac{V}{1 + \log A} \leq |y| \leq V \qquad (10)$$

For µ-law is:

$$x = \frac{V}{\mu}\left(e^{|y|\log(1+\mu)/V} - 1\right)\operatorname{sgn}(y) \qquad (11)$$

Block „To Wave File" writes audio data to a Microsoft Wave (.wav) file in the uncompressed pulse code modulation (PCM) format. It is important to use standard sample frequencies (8000, 11025, 22050, 44100 Hz), the amplitude of input signal should be in the range of nominal values ±1 . [3], [6]

3. Methods of comparison of signals

Correlation coefficient serves for comparison of signals in time area. Mean square measure is used in spectrum (frequency) area.

3.1 Correlation coefficient

This coefficient determines measure of linear dependence between two arbitrary quantities X a Y. If $\rho_{xy} = 1$ it means, that Y is linear depended on X. The case $\rho_{xy} = -1$ means inverse linear dependence between Y and X. Value ρ_{xy} is defined in range from -1 to +1. It represents statistic dependence.
It is:

$$\rho_{xy} = \frac{Cov(X,Y)}{\delta_x \delta_y} \quad \text{where} \quad -1 \leq \rho_{xy} \leq 1 \qquad (12)$$

and

$$Cov(X,Y) = \frac{1}{M}\sum_{i=1}^{M}(x_i - \mu_x)(y_i - \mu_y) \qquad (13)$$

$$\delta_x = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(x_i - \mu_x)^2} \qquad (14)$$

$$\mu_x = \frac{1}{M}\sum_{i=1}^{M}x_i \text{ (it is same for y)} \qquad (15)$$

M is number of samples
Covariance (COV) means measure of linear dependence between X and Y. Matlab corr2 function was used for the calculation of correlation coefficient in Matlab. This function realizes calculation by help with this formula:

$$r = \frac{\sum_m \sum_n (A_{mn} - \overline{A})(B_{mn} - \overline{B})}{\sqrt{\left(\sum_m \sum_n (A_{mn} - \overline{A})^2\right)\left(\sum_m \sum_n (B_{mn} - \overline{B})^2\right)}} \qquad (16)$$

where $\overline{A}$ = mean2(A) , and $\overline{B}$ =mean2(B).
Matlab function mean2 realize calculation of mean value.

3.2 Mean square measure

Measure of differences of 2 short-term segments speech signal is very important, when we compare speech signals. We mark x, y vector of symptoms of microsegment speech, then we mark d(x, y) measure of difference. Measure of difference must have these characterisctics:

1. symmetry d(x,y) = d (y,x)
2. positive definition
   d(x, y)> 0 for $x \neq y$
   d(x, y) = 0 for x = y

This measure is based on a comparison of the spectrums of test microsegment with referenced microsegment. The most common norm is L_2-norm, which is defined like this:

$$d_2(t, r) = \left[\sum_{i=1}^{N} |y_{ti} - y_{ri}|^2 \right]^{1/2} \qquad (17)$$

where:
N is number of points FFT in given microsegment
y_{ti} is absolute value of i-th coefficient FFT of test microsegment
y_{ri} is absolute value of i-th coefficient FFT of reference microsegment.

It is used for expedition of calculation alternative, which use applications L_1-norm, which is given with this formula:

$$d_1(t,r) = \left[\sum_{i=1}^{N} |y_{ti} - y_{ri}| \right] \qquad (18)$$

FFT function implemented in Matlab is used for the calculation of coefficient FFT. The formula for FFT is:

$$X(k) = \sum_{j=1}^{N} x(j) \omega_N^{(j-1)(k-1)} \qquad (19)$$

$$x(j) = \frac{1}{N} \sum_{k=1}^{N} X(k) \omega_N^{-(j-1)(k-1)} \qquad (20)$$

where

$$\omega_N = e^{(-2\pi i)/N} \qquad (21)$$

N is number of points FFT
[4], [5]

4. Presentation of results

Principle of simulations is based on recording wav files and their simultaneous offline comparison. It is based on calculation of mean square measure by L_2-norm. By help of it we can compare signals with spectrum. Another comparative parameter is correlation coefficient, which provides a determination of measure of linear dependence between 2 signals.

We can compare given signals more expressively and more complex when we use 2 parameters for the evaluation of the signals. Nowadays the comparison is provided by help of correlation coefficient.

If we want to discover the differences in spectrum area, we will add next parameter. By help of correlation coefficient and mean square measure we can find the signal, which is the most sensitive for noise influences, which arise on communication channels.

Algorithm of comparison:
1. Reading of source and reference signal
2. Segmentation of signals in n intervals with 8000 samples.
   These parameters are calculated in each interval:
   1. Correlation coefficient r
   2. Coefficients FFT
   3. Mean square measure d
3. Calculation of average value:

$$\bar{r} = \frac{\sum_{i=1}^{n} r_i}{n} \quad (22) \ , \qquad \bar{d} = \frac{\sum_{i=1}^{n} d_i}{n} \quad (23)$$

n is number of intervals
r_i is correlation coefficient of i-th interval
d_i is mean square measure of i-th interval

We will get much more precise results, if we segment the signals in n intervals and then we calculate parameters.
These types of signal were researched:
1. Sinusoidal signal of the frequencies 300, 800, 1200, 2400, 3000 Hz
2. Bipolar signal of square course and sawtooth signal of frequencies 300, 800, 1200, 2400, 3000 Hz
3. Random signal of frequencies 300, 800, 1200, 2400, 3000
4. Gaussian white noise (mean value= 0, variance= 1)
5. Unipolar signal of square course with the width of the impulses 50 %

We can create every speech by help with 2 generators- the generator of impulse and the generator of noise, according to synthesis of speech signal. The signals on the basis of these facts were chosen and they were used for simulation. These simulations were realized by the usage of the signals displayed in upper part of document for compressed characteristics according to the curves A and μ and for the both types of the channels (AWGN, BSC).

The examples of simulation results for the compression by the curve A are summarized for next tables.

Values listed in tables were acquired by frequency 300 Hz (Sinusoidal signal, Random signal, Sawtooth signal, Bipolar signal of square course).

AWGN channel (A-law)

SNR	Sinus	Noise	Uni. Square	Bip. Square	Random	Sawtooth
5	4067,2	1231,4	4301,2	5606,9	2421,2	3372,0
10	2583,5	784,7	2769,7	3585,3	1671,6	2120,6
15	1331,8	194,4	1340,0	1867,8	1054,5	1088,0
25	1243,7	87,0	1237,4	1750,0	1010,4	1019,4
50	1243,7	87,0	1237,4	1750,0	1010,4	1019,4

Table 1 Mean square measure for AWGN channel and A-law compressed characteristic

SNR	Sinus	Noise	Uni. Square	Bip. Square	Random	Sawtooth
5	0,7024	0,5850	0,5835	0,7406	0,6833	0,6807
10	0,9079	0,8519	0,8123	0,9203	0,9026	0,9016
15	0,9959	0,9927	0,9894	0,9967	0,9955	0,9959
25	0,9999	0,9998	1	1	0,9999	0,9999
50	0,9999	0,9998	1	1	0,9999	0,9999

Table 2 Correlation coefficient for AWGN channel and A-law compressed characteristic
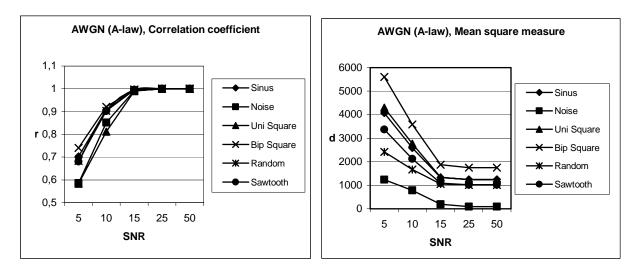

Fig. 5 Graphs for AWGN channel and A-law compressed characteristic

BSC (A-law)

	Sinus	Noise	Uni. Square	Bip. Square	Random	Sawtooth
0,00001	417,9	88,2	1237,7	626,7	295,4	299,4
0,0001	426,8	93,4	1241,0	638,5	300,9	308,4
0,001	517,3	139,1	1277,8	751,4	346,2	391,2
0,01	1097,1	362,7	1637,0	1495,6	666,2	865,7
0,1	3158,6	986,7	3418,1	4180,2	1760,7	2493,6

Table 3 Mean square measure for BSC channel and A-law compressed characteristic

	Sinus	Noise	Uni. Square	Bip. Square	Random	Sawtooth
0,00001	0,9999	0,9998	1,0000	1,0000	0,9999	0,9999
0,0001	0,9998	0,9996	0,9996	0,9999	0,9997	0,9997
0,001	0,9983	0,9971	0,9958	0,9987	0,9984	0,9983
0,01	0,9829	0,9705	0,9569	0,9859	0,9834	0,9839
0,1	0,8300	0,7529	0,7220	0,8614	0,8403	0,8408

Table 4 Correlation coefficient for BSC channel and A-law compressed characteristic
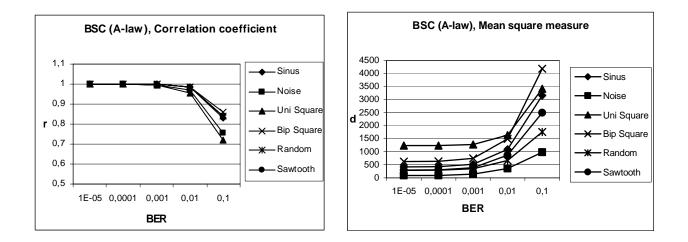

Fig. 6 Graphs for BSC channel and A-law compressed characteristic

5. Conclusion

The principle of choice of signal for automatic measurement VTQoS is based on simple rule. This rule is, that the signals are used for measurement, which have maximum difference of values competent boundary parameters.

Calculation is realized on the basis of below given formulas, then are chosen 2 best signals for each coefficient (corelation coefficient, mean square measure) on each channel. We get 8 signals and we choose 4 the best of them.

It is for AWGN channel:

$$\delta_r = \overline{r}_{(SNR\ 50)} - \overline{r}_{(SNR\ 5)} \qquad (24)$$

$$\delta_d = \overline{d}_{(SNR\ 5)} - \overline{d}_{(SNR\ 50)} \qquad (25)$$

It is for BSC channel:

$$\delta_r = \overline{r}_{(BER\ 0,00001)} - \overline{r}_{(BER\ 0,1)} \qquad (26)$$

$$\delta_d = \overline{d}_{(BER\ 0,1)} - \overline{d}_{(BER\ 0,00001)} \qquad (27)$$

where:

δ_r is the difference of average values of boundary parameters of correlation coefficient

δ_d is the difference of average values of boundary parameters of mean square measure

$\overline{r}_{(SNR\ 50)}$ is average value of correlation coefficient for competent value of boundary parameter for AWGN channel

$\overline{d}_{(SNR\ 50)}$ is average value of mean square measure for competent value of boundary parameter for AWGN channel

$\overline{r}_{(BER\ 0,1)}$ is average value of correlation coefficient for competent value of boundary parameter for BSC

$\overline{d}_{(BER\ 0,00001)}$ is average value of mean square measure for competent value of boundary parameter for BSC

Boundary parameters are SNR for the value of 50 dB and for the value of 5 dB on AWGN channel. It is parameter BER for the values 0,00001 and 0,1, when we use BSC channel.

This fact was a base for derivation of this rule, because we need the signals, which will be very sensitive for every influences, which we can see on real networks. When we reach maximum difference for same conditions, then given signal is much more sensitive than the other signals for interference influences which arise on given channel. It is possible to state on the base of these facts, that this signal is more advantageous for intrusive measurement VTQoS.

δ_r	AWGN (A-law)	BSC (A-law)
Sinus	0,2975	0,1699
Noise	0,4148	0,2469
Uni. Square	0,4165	0,2780
Bip. Square	0,2594	0,1386
Random	0,2790	0,1596
Sawtooth	0,3192	0,1591

Table 5 Diference δ_r for Correlation coefficient

δ_d	AWGN (A-law)	BSC (A-law)
Sinus	2823,5	2740,7
Noise	1144,4	898,5
Uni. Square	3063,8	2180,4
Bip. Square	3856,9	3553,5
Random	1407,4	1465,3
Sawtooth	2352,6	2194,2

Table 6 Diference δ_d for Mean square measure

The results of simulations have finished according to theoretical supposition. These types of the interference influence expressively on stochastic signals with Gaussian distribution from the point of view of correlation coeficient (e.g. Gaussian white noise). The entropy of stochastic signals with Gaussian distribution is bigger than the entropy of deterministic signals. It is supposed, that the transmission of stochastic signal with Gaussian distribution and deterministic signal on transmission channel with same probability of error causes loss of bigger quantity of information of stochastic signal, because this signal has bigger entropy. Lowering of quantity of transmission information causes lowering of value of correlation coefficient in the

case of Gaussian white noise, when we compare it with correlation coefficient of deterministic signal for these same conditions of transmission. These types of interference from the point of view of mean square measure have an expressive influence on the signals, which have upper harmonical components (e.g. bipolar signal of square course). The interference shows for bigger number of harmonical components, because this difference arises on many spectrum components in the spectrum. It causes total growth of the value of mean square measure. It is followed from these results, that these signals can be used in test sequence: Unipolar signal of square course, Bipolar signal of square course, Gaussian white noise, Sinusoidal signal.

However sequence will be used for automatic intrusive measurement VTQoS on a real telephone channel. This channel uses the frequency band from 300 Hz to 3400 Hz, therefore does not transmit DC component. For this result final sequence should only contain the signals: Bipolar signal of square course, Gaussian white noise, Sinusoidal signal. We must not forget on this fact, that bipolar signal of square course only contains odd harmonical components. We must count with it, that first harmonical component will be only transfer, when we use these signal with high frequencies over 1700 Hz also we only obtain sinusoidal signal. That is why we can only choose the signal with the frequency, so the first three harmonical components (1, 3, 5 harmonical component) can be transfered in the zone of telephone channel. Thereby we secure the highest correlation between ideal and real course. It follows, that we can use bipolar signal of square course with the frequencies from 300 to 670 Hz for transmission on telephone channel.

6. Acknowledgement

7. References

[1] Křenek, J. , Holub, J. : Meření kvality hlasového přenosu v telekomunikačných sítích, In: ST 5/2004, pp. 6-8.

[2] Holub, J. , Šmíd, R. : Hodnocení hlasových přenosů v telekomunikačných sítích,In: ST 6/2001, pp 3-5.

[3] Matlab Help.

[4] Konvit, M. : Teória oznamovania, ALFA, 1989, 274 p.

[5] Psutka, J. :Komunikace s počítačem mluvenou rečí, ACADEMIA, 1995, 287 p.

[6] Franeková, M. : Modelovanie komunikačných systémov v prostredí Matlab, Simulink a Communications Toolbox, EDIS, 2003, 129 p.