

8th CIRP Conference on Intelligent Computation in Manufacturing Engineering

Wavelet Based De-noising in Manufacturing and in Business

Gábor Bényász^{a*}, László Cser^{a*}

^aCorvinus Univ. of Budapest, Budapest 1092 Fővám tér 8, Computing Science Department, Hungary

* Corresponding author. Tel.: +36-1-482-7461; fax: +36-1-482-7463. E-mail address: laszlo.cser@uni-corvinus.hu.

Abstract

Last two decades a powerful new tool has appeared in the quality assurance and quality control: the data mining. It enables discovering the hidden connections and confluences in big data systems, collected by on-line measuring and control systems in manufacturing processes. The main difficulties appear in distinguishing the valuable signal and the noises in measurements, as well as the time dependent, relatively slow changes indicating the inherent tendency of changes inside the process.

It has been shown ([1], [2] etc.) that the Self Organizing Maps can be used effectively in analyzing the manufacturing process (e.g. hot rolling) in order of predicting the expected quality. Source of the analysis is the big amount of data collected by the hierarchically structured control system of the hot rolling mill. Since the sampling frequencies in different measurement devices were different, the problem of the possible noise became vital for the reliability of quality prediction.

Recent paper intends to illustrate the efficiency of the wavelet analysis for noise removal in different areas of data mining, including the thermal processes in hot rolling, and in another area, very far from it – in financial analysis.

Technically, the wavelet is based on the usage of orthogonal functions. Unlike the similar and mathematically closest Fourier analysis - wavelets conserve the entirely time dependent information, where the temporal identification of the signals frequency content is also important.

© 2012 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of CIRP ICME

Keywords: Analysis, Measurement, Data mining

1. Wavelet and de-noising

Recently, the data mining is getting more and more frequently used in quality assurance and in quality control: according to its basic features, it makes possible to examine the latent information content in large data systems, produced by on-line measurement- and control systems (e.g. in manufacturing). The major challenge is in the separating the real signal from the noise of measurements, especially in the case, when the examined time series of values contain time dependent, relatively slow changes indicating some tendencies of changes inside the examined phenomena (e.g. tool wear or development of fatigue crack in cold forging tools, or roll wear, as well as growing surface damage in hot rolling etc.).

In some former papers ([1], [2] etc.) has been shown that the Self Organizing Maps represent suitable tool in analyzing the process of hot rolling in order of predicting the expected quality (e.g. accuracy or microstructure).

Data sources in the analysis were large datasets collected by the hierarchically structured control system of the hot rolling mill. Since the sampling frequencies and sources of noises of the devices were different, the problem of noise became the central point for reliability of the quality control.

Recent paper is an attempt to apply wavelet methods in data mining process, not only manufacturing (e.g. the thermal processes in hot rolling), but in the economy/cs as well.

Although wavelets have become widely accepted only in the last two decades, its mathematical bases and formulas had been already invented in the first decade of the last century ([3]). Technically, the wavelet bases on the usage of orthogonal functions. Unlike the similar and mathematically closest Fourier analysis - wavelets conserve entirely time information, a feature that enabling this technique in applications, where the temporal identification of the signals frequency content is also important. One of the most successful fields of

wavelet application is the signal and time-series de-noising (e.g. [4]-[6]).

1.1. Wavelets

Although the wavelets application became widely accepted only in the last two decades, its mathematical background has been invented and published by the Hungarian mathematician Alfred Haar in 1909 ([3]). Later on – catalyzed by with the extensive spread of its application – its mathematical background has been refocused, developed and adapted to the given application domain ([7], [8]).

Technically the wavelet bases on the usage of orthogonal functions. Unlike to the mathematically closest Fourier analysis, wavelets conserve entirely time information.

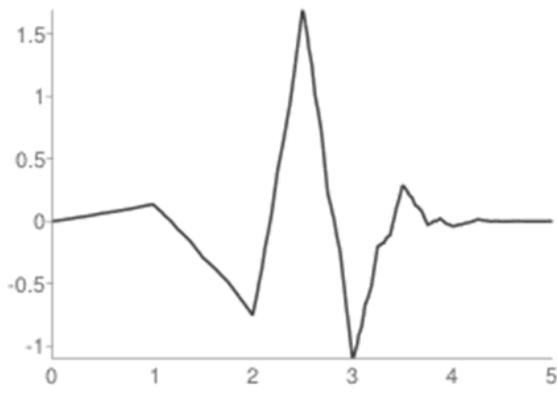


Figure 1. - Daubechies 3 (db3) Wavelet function ([8])

Itself the wavelet is a wave whose energy is concentrated in a specific location and its analytical function meets the related mathematical constrains (see in [9]). Based on this feature wavelet is capable to capture the properties of signals (more broadly: time series) simultaneously in time and frequency domain.

The wavelet decomposition (or wavelet expansion) can be defined by using a special wavelet function (Figure 1). This expansion is itself the representation of combinations of orthogonal real-value functions, which are generated by scaling and shifting transformations of the base wavelet. This base wavelet from which all the transformed functions are generated, called “mother” wavelet.

For the function $f(t)$ to be analyzed the wavelet decomposition, equation looks like:

$$f(t) = \sum_i a_i \psi_i(t) \quad (1)$$

where $\psi_i(t)$ are the orthogonal basis functions, whilst the a_i are the wavelet coefficients, which can be computed as the inner product of the original $f(t)$

function and $\psi_i(t)$ wavelet base functions:

$$a_i = \langle a_i, \psi_i(t) \rangle = \int f(t) \psi_i(t) \quad (2)$$

This equation shows that the wavelet coefficients are produced as ortho-normal decompositions. By replacing equations (2) in equation (1), the general form of wavelet equations is:

$$f(t) = \sum_j \sum_k a_{jk} \psi_{jk}(t) \quad (3)$$

where the a_{jk} are the coefficients of the discrete wavelet transformation (DWT).

The daughter functions (3) are generated from the mother wavelet by scaling and translation (product with $2^{j/2}$) transformation (shift with $-k$):

$$\psi_{jk}(t) = 2^{j/2} \psi(2^j \cdot t - k) \quad (4)$$

1.2. Wavelet de-noising

Based on the results of wavelet based applications, the most successful fields of wavelets is data analysis. More closely, it has been proved that wavelets outperform over the traditional methods in removal of noise in signals. Shrinkage methods for noise removal ([4]) enabled wide set of approaches using wavelets with probabilistic methods, developing new and more effective de-noising techniques.

The wavelet based signal de-noising includes the following basic steps:

- Decomposition: choosing a wavelet and a level N , than computing the wavelet decomposition of the signal at level N .
- Finding threshold detail coefficients: for each level from 1 to N , selecting a threshold, applying soft thresholds to the detail coefficients.
- Reconstruction: computing wavelet reconstruction using the original approximation coefficients of level N and the modified detail coefficients of levels from 1 to N .

The first step creates a wavelet decomposition of the original signal using given wavelets at a given level (in the example it is the Daubechies wavelet function, shown in Figure 1, the level of decomposition is 3). The second step concludes the filtering by dropping those wavelet atoms, whose wavelets coefficients are under the threshold value.

There are two basic kinds of the applicable thresholds; hard and soft. Hard threshold is simpler, but the soft threshold method has advantageous mathematical properties ([4]). In the example, broken down in equation (5)-(6), the heuristic version of Stein's

Unbiased Risk Estimate method is used, as its threshold selection rules are conservative enough and would be more convenient, when small details of the signal are near the noise range. It is a mixture of the universal threshold-selecting scheme and the scale dependent adaptive threshold-selecting rule. It provides the best estimation results in the sense of optimal minimum mean-square error [5], an outperforming good estimation, when the inherent function is not known [13].

In this approach, level-dependent thresholds has been chosen by regarding different resolution levels (different j) for the wavelet transformation, as for independent multivariate normal estimation problems.

Within one level (scale - j) is

$$y_{jk} = \psi_{jk} + \varepsilon z_{jk} \quad k=0..2^j-1 \quad (5)$$

and it is necessary to estimate the

$$(\psi_{jk})_{k=0}^{2^j-1} \quad k = 0 \dots 2^j - 1 \quad (6)$$

Stein's Unbiased Estimate of Risk is:

$$\hat{\Theta}_k^{(t)} = \eta_t(\psi_{jk}) \quad (7)$$

where ε depicts the Gaussian white noise with i.i.d. distribution $N(0,1)$.

It gives an approach of the risk (understood as optimal minimum mean-square error [13]) for a particular threshold value t . Minimizing it in t gives a selection of the threshold level for level j [4]. (A slight modification of this model, employed in case, when the data vector had a very small L2 norm, in which case the unbiased risk estimate is very noisy and a fixed threshold is employed).

After having filtered out the coefficients, using the Stein's Unbiased Risk Estimate, thresholds for the data are reconstructed by the inverse wavelet, decomposing step 1.

2. Wavelets as tools in data mining

2.1. Temperatures in hot rolling process

In the rolling mill, temperatures are measured in many places; like in the Figure 2 (arrows are indicating the place of pyrometers and diagrams measured by them).

Pyrometers are built into the line and they are measuring the temperature of the running strip. Basis of the sampling is the measuring frequency, and all

measurement results appear in the co-ordinates "Temperature vs. Time". Pyrometers are fully independent on each other and their results are used and stored by different computers (with different clock-signal). Such kinds of measurements are very useful for temperature control, but they cannot be used for the prediction of the expected quality, since it depends on the cooling and deformation "history" in each elementary volume of the strip. Based on the additional speed measurement, diagrams shown in Figure 2 can be transformed into the length co-ordinates of the strip.

Figure 3 shows the measured temperatures vs. strip length and the elongation of the strip during the consequent passes. All diagrams are transformed to the same initial reference point. However, when it is necessary to transform the measured data into the length co-ordinates of the strip, the beginning of the measurement should be distinguished from the moment, when the strip exactly arrives the measurement device. The main problem in transforming the time-dependent signals into the natural (length) co-ordinates of the strip is caused by the different sampling frequencies and by the different noise levels.

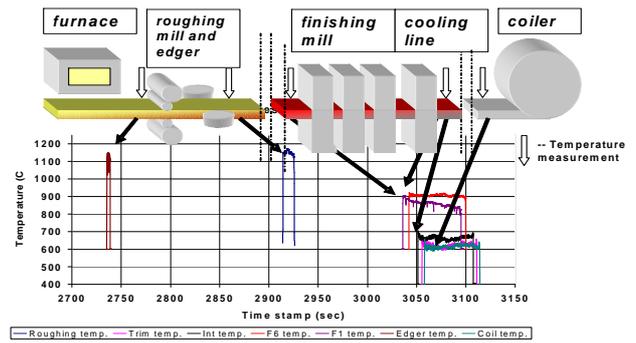


Figure 2: Schematic structure of the measurements in a hot rolling mill ([2])

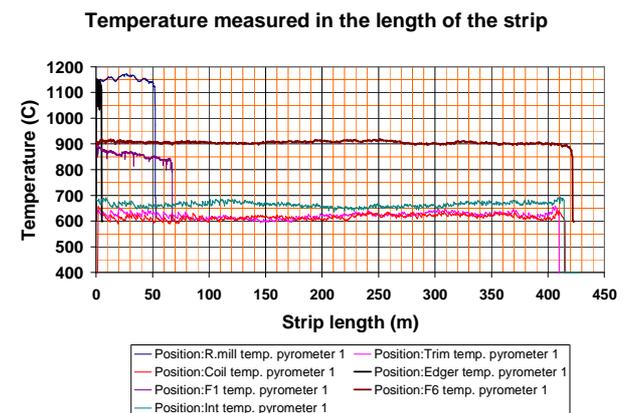


Figure 3: Temperatures transformed into length co-ordinates ([2])

Fig. 3 shows an example: temperatures and speeds are measured at different locations of the mill, but in time

co-ordinates (vector $T(t)$ with components of temperatures measured in different locations, as well as $v(t)$ analogously). Using these two vectors, it is necessary to estimate the temperatures in natural co-ordinates of the strip $T(s)$. Because of the elongation of the strip, it is necessary to transform the strip length into the interval (0,1). The accuracy of the vector components after some transformations with noisy arguments ($v(t) \rightarrow s(t) \rightarrow t(s) \rightarrow T(s) \rightarrow T(l)$) can be very low, jeopardizing the results of the data mining.

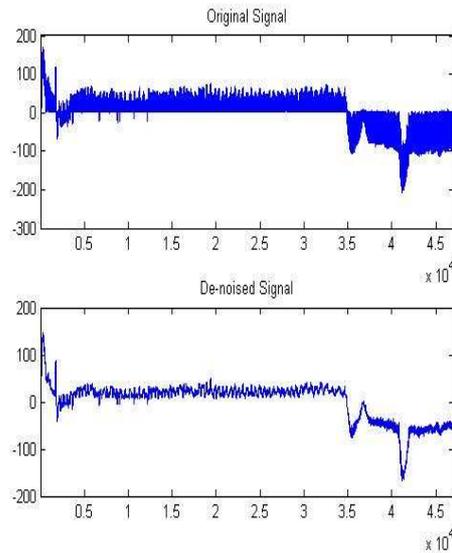


Figure 4: Measured signals and the real values

Since the scatter of temperature and other parameters influencing the thermo-mechanical process can play an essential role in the evolution of microstructure, it is very important to separate the real signal from the noise and perform the transformations only with the signals. That's why the wavelet transformation opens new perspectives in such kinds of manufacturing data.

Figure 4 illustrates the efficiency of the noise removal from a measured data set (the real value is hardly to see in the original signal). The speed in the first stand shows clear tendencies (Figure 5), but the accuracy is not enough for estimation of the real position of the strip. The de-noised signal measured in drive shows the real scatter of the mandrel current (Figure 6).

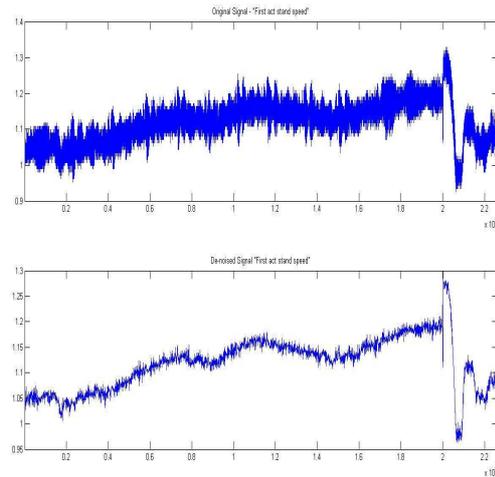


Figure 5: Measured and de-noised speed in the first stand

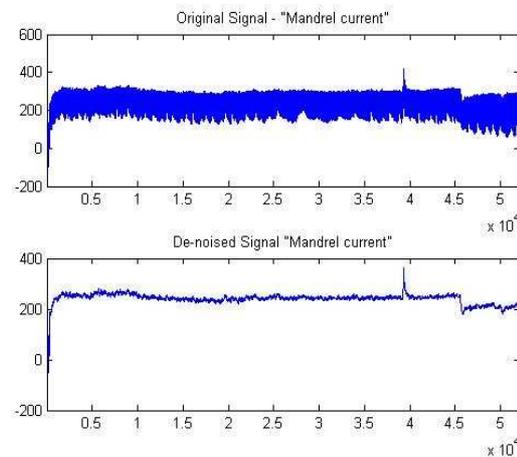


Figure 6: Mandrel current

2.2. Finance domain: analysis of stock prices

Similarly to the previous application in quality assurance of manufacturing, the same method is utilized to the time series of stock prices in order to de-noise the Gaussian white/noise component from the stock prices. Just in order of the comparison, unity and presentation the same wavelet functions have been utilized. Of course, according to the application domain, different method could be accepted, as optimal due to the different nature of the analyzed phenomena. The test data in this domain are the stock (closing) prices listed on the New York Stock Exchange on the time interval 04/01/2002 and 04/01/2012. The analogy of signal de-noising and the volatility-analysis is only in the mathematical approach, since the volatility of the stock prices is an important characteristic of the stock operations. The examined stock are the General Electric co., Microsoft - these are less volatile based on VIX - and Google - is more volatile - (YAHOO Finance/www.finance.yahoo.com).

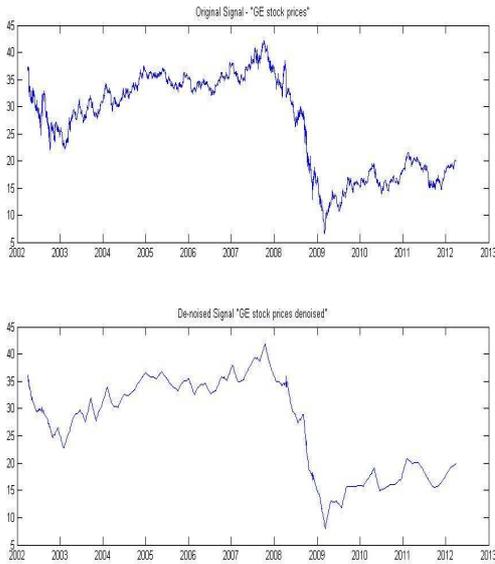


Figure 7: GE stock prices

As from the time series can be seen, the motion of the Google stock prices is more volatile, than the others two (Figure 7-9). The de-noised series and the original signal differ much more in the case of Google, than for the other examined stock prices. It contains bigger noise component as the others, and the de-noising algorithm eliminates this component from the original time series. This feature can be captured by the SNR (Signal to Noise Ratio) value of the different stock prices. The Google’s SRN is significantly lower than of the Microsoft and GE (shown in Figure 10), what means that it resulted naturally in a smaller SNR value. The SNR computed as follows:

$$SNR = \frac{\bar{P}}{\bar{I}}, \quad \bar{P} = \frac{1}{N} \sum_i Fd_i,$$

$$\bar{I} = \frac{1}{N} \sum_i |Fd_i - Fo_i| \quad (8)$$

where Fd is the de-noised signal, and Fo the original time series of stock prices.

Besides the de-noised GE signal the wavelet power spectrum (WPS) graph is also presented in Figure 8. This map shows the wavelet coefficients in the time-scales space. It graphically shows the different long-term and short-term phenomena in the time-series (or signal), as it shows the effect of different scales (power of 2 – marked on y-axis on the graph labeled with “level”) in the time-series by its coefficients (the greater values of coefficients are indicated by lighter colors).

The application of wavelets for signal processing in economics and statistics is relatively new. The novelty in time-series analysis (more precisely data cleansing in the

example shown in Fig. 7-9) basically differs from the traditional statistics in the area of economics and statistics by the fact that it applies a “top-down” approach, whilst the traditional statistics applies the “bottom-up” approach (it captures tendencies first and handles errors as residual). The wavelets enable multi-scale analysis at the same time.

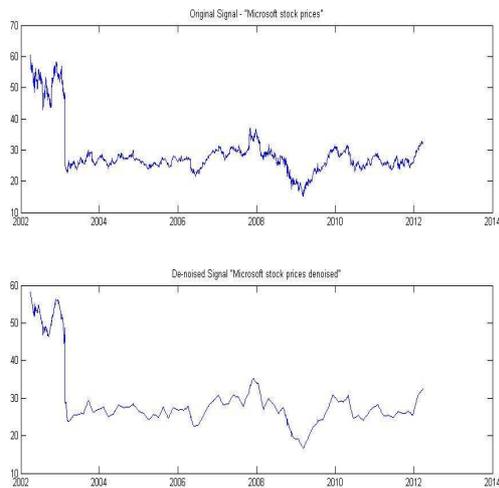
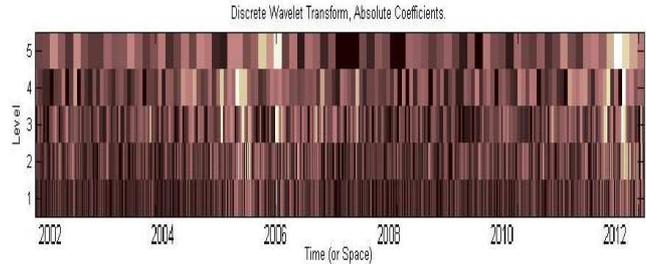


Figure 8: Microsoft stock prices

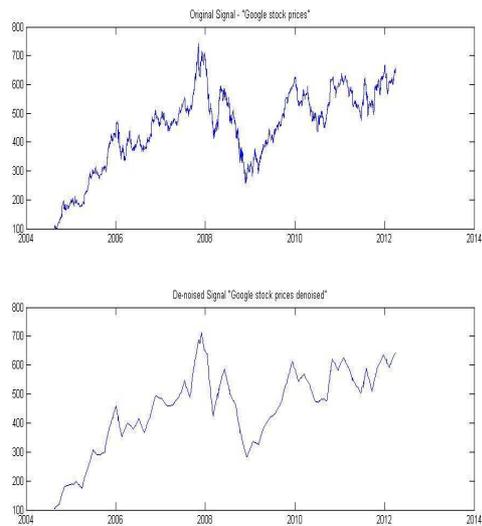


Figure 9: Google stock prices

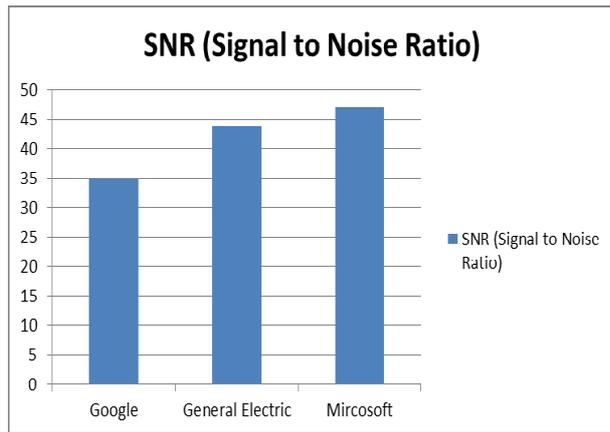


Figure 10: Signal to Noise Ratio

3. Conclusions

Two different areas have been shown, as illustration of the wavelets, as an efficient new method in the data cleansing for data mining. The strength of wavelet transformation over the traditional filtering and denoising methods (e.g. Fourier-transformation) is in the clean separation of high frequency scatter and slow tendencies in the measured (collected) big data systems.

In the quality assurance, as well as in the quality prediction of hot rolling products, the “thermo-mechanical history” of deformation inside the strip in each elementary volume is the main factor. Since the different measurement systems, controlling the rolling mill use different sampling frequencies, transformation of the measured values into unified time co-ordinates is connected with smoothing or interpolation of measured values. In both cases the influence of the accuracy in separation of noise from the real measured values is essential. The comparison with own former results (e.g. [1]) shows that the introduction of wavelets opened new perspectives in the industrial data mining.

The reason of scatter in financial data (e.g. stock prices) has been resulted by the volatility and differs in information content very strongly from the noise of on-line measuring systems. The wavelet transformation enables separation of the volatility from the long-term tendencies in stock prices.

The big computation need, connected with the wavelet transformation requires high performance computers for the big databases. The CUDA architecture enables data mining also in that case by commonly used PCs as well ([10]).

Acknowledgement

Financial support from TÁMOP-4.2.2./B-/1-2010-0023 is gratefully acknowledged.

References

- [1] L.Cser, O. Simula, P.Mäntylä, P.Ruha, J.Larkiola: Self Organising Maps (Som) in Hot Rolling, Intelligent Computation in Manufacturing Engineering (Ed. Roberto Teti), Proc. of the ICME 2000, 21/23 June 2000, Capri, Italy, pp.291-297
- [2] L. Cser, J. Paavola, D. Martin: Intelligent Transfer of Laboratory Data to Industrial Scale Hot Rolling, Intelligent Computation in Manufacturing Engineering (Ed. Roberto Teti), Proc. of the ICME'04, Sorrento, Italy, June 30-July 2,
- [3] Haar, A., Zur Theorie der orthogonalen Funktionensysteme, Math. Ann. 69 (1910), 331-371
- [4] Donoho, et. al; Adapting to Unknown Smoothness via Wavelet Shrinkage. Journal of the American Statistical Associations, (1995.12.), 1200-1224.
- [5] Chang, S.G.; Bin Yu; Vetterli, M.; Adaptive wavelet thresholding for image denoising and compression, Image Processing, IEEE Transactions , Sep 2000 , Volume: 9 Issue:9 ,1532 – 1546ü
- [6] Portilla, J.; Strela, V.; Wainwright, M.J.; Simoncelli, E.P.; Adaptive wavelet thresholding for image denoising and compression, Image Processing, IEEE Transactions , Nov. 2003, Volume: 12 Issue:11 , 1338 - 1351
- [7] Mallat, S.G.; A theory for multiresolution signal decomposition: the wavelet representation, Pattern Analysis and Machine Intelligence, IEEE Transactions on, Jul 1989, Volume: 11 Issue:7 , 674 - 693
- [8] Daubechies, I; The wavelet transform, time-frequency localization and signal analysis, Information Theory, IEEE Transactions on , Sep 1990 , Volume: 36 Issue:5 , 961 – 1005
- [9] Percival, D. B., Walden, A. T., Wavelet Methods for Time Series Analysis, October 2010, Cambridge, 41-65
- [10] Benyasz, G., Cser, L. Financial Time Series Clustering on CUDA, Conference of PhD Students in Computer Science, Proceeding, June 2010
- [11] Sun, E. W., Meinel, T. , A new wavelet-based denoising algorithm for high-frequency financial data mining, European Journal of Operational Research 217 (2012) 589–599
- [12] He,K., Xie, C., Lai, K., Estimating real estate value-at-risk using wavelet denoising and time series model, Computational Science–ICCS 2008, 494–503
- [13] C. Stein, Estimation of the mean of a multivariate normal distribution, Annals of Statistics,1981, Volume: 9, Issue: 6, 1135-1151