Some preliminary results of memory cache analysis with the use of non-extensive

Dominik Strzałka

Department of Powerengineering, Powerelectronics and Complex Systems
Faculty of Electrical and Computer Engineering, Rzeszów University of Technology
Al. Powstańców Warszawy 12, 35-959 Rzeszów, Poland
strzalka@prz.edu.pl

Abstract—**The problem of modeling different parts of computer systems requires accurate statistical tools. Cache memory systems is an inherent part of nowadays computer systems, where the memory hierarchical structure plays a key point role in behavior and performance of the whole system. In the case of Windows operating systems, cache memory is a place in memory subsystem where the I/O system puts recently used data from disk. In paper some preliminary results about statistical behavior of one selected system counter behavior are presented. Obtained results shown that the real phenomena, which have appeared during human-computer interaction, can be expressed in terms of non-extensive statistics that is related to Tsallis proposal of new entropy definition**

Keywords—complex systems; cache memory; system counters;Tsallis entropy

I. INTRODUCTION

The description of phenomena that appeared in nowadays computer systems selected parts (like disks, memory, network, etc.) seems to be very important topic in computer engineering. This is mainly due to the fact that computer systems consist of many interdependent subsystems and it is very important to understand how they work and what kind of interactions exist between them. Usually this topic is described in terms of queueing theory, but we can see that modern computer systems have very high level of complexity. Here, we understand the term complexity as a feature that relates this machines to complex systems, where the whole is more that sum of its parts [1]. The paradigm change in computer systems perception (towards complex systems) is not a new idea. For the first time it has appeared in the seventies of XXth century and was proposed by Dijkstra [2] and then by M. Gell-Man [3].

Nowadays computer systems (even personal computers) offer multitask processing, very user friendly access interfaces and high performance; this is done thanks to operating systems and sophisticated electronic integrated circuits. However, this is not easy taking into account the fact that in these complex systems the amount of available resources is always limited [4]. This is especially well-visible in the case of memory that, in contradiction to Turing machines properties, is limited, heterogeneous and consists of many technological, physical and logical solutions, which in fact make the whole memory subsystem complex and very hard to be understood. The multitude of assumed solutions is in fact a result of expectations and technological development done in order to achieve the highest possible performance, but with assumption about limitation of solutions costs. In this paper we show that in the case of cache memory behavior measured only by one system counter, the complexity existed in time series can be expressed in terms of non-extensive statistics. Obtained results presents preliminary outcomes for personal computers working under Windows operating systems.

The paper is organized as follows: after preliminaries in Introduction, we show, in Section 2, theoretical background for non-extensive statistics including the problem of power laws and long-range dependencies. This background will be used in Section 4 to explain obtained results, which were possible to be achieved thanks to the experiment shortly described in Section 3. Paper is crowned in Section 5.

II. THEORETICAL BACKGROUND

In 1988 C. Tsallis published his famous paper [5] about the concept of new definition of non-extensive entropy. Since then a lot of papers have been published in order to critique, support, develop and discuss this proposal. The whole concept is based on a simple q-generalization of logarithm:

$$\ln_q(x) = \frac{x^{1-q}-1}{1-q} \qquad q\in R, \ \ln_1(x) = \ln(x) \qquad (1)$$

with inverse function given by:

$$e_q^x = [1 + (1-q)x]^{\frac{1}{1-q}} \qquad \text{with } e_1^x = e^x \qquad (2)$$

The equation (1) leads to the extension (or generalization) of classical definition of Boltzmann-Gibbs entropy given by

$$S_{BG} = -k \ \ln(W) \qquad (3)$$

where k is Boltzmann constant and W is the number of system microstates, to the form:

$$S_q = -k \ln_q(W) \qquad (4)$$

The development of non-extensive entropy concept introduce its continues form for the set of probabilities given by the function $p(x)$:

$$S_q = -k \int_{-\infty}^{+\infty} p(x) \ln_q p(x) dx \qquad (5)$$

The extremization of BG entropy in continues form

$$S_{BG} = -k \int_{-\infty}^{+\infty} p(x) \ln p(x) dx \qquad (6)$$

with the constrains $\int_{-\infty}^{+\infty} p(x)dx = 1$, $<x> \equiv \int_{-\infty}^{+\infty} xp(x)dx = 0, <x^2> \equiv \int_{-\infty}^{+\infty} x^2 p(x)dx = \sigma^2$, leads to the Gaussian probability distribution given as:

$$p_{BG}(x) = \frac{e^{-\beta x^2}}{\int_{-\infty}^{\infty} e^{-\beta x^2} dx}, \qquad (7)$$

with $\beta > 0$ as the Lagrange parameter determined by σ^2. The equation (6) is considered as a ground state of the Central Limit Theorem (CLT) [6]. For non-extensive entropy (5) we can use the same way, which gives the stationary $p_q(x)$ distribution called q-Gaussians:

$$p_q(x) = \frac{e_q^{-\beta x^2}}{\int_{-\infty}^{+\infty} e^{-\beta x^2} dx} \ (q < 3; \beta > 0), \qquad (8)$$

where e_q is defined by (2). For a set of N such random where e_q is defined by (2). For a set of N such random variables their sum for $q < 5/3$ and $N \to \infty$ converges to a (stable) Gaussian distribution, whereas when $q > 5/3$ we can obtain a set of α-stable Lévy distributions [7]. Such distributions are used in statistics to generalize CLT for distributions with power laws. The α (stable) parameter is described as:

$$\propto = \frac{3-q}{q-1} \text{ with } q \in \left(\frac{5}{3}, 3\right) \qquad (9)$$

with finite variance distributions for q < 5/3 and infinite variance distributions for q ≥ 5/3.

The lack of finite variance is usually related to the existence of power-laws for analyzed system with following consequences:

- the existence of long-term and long-range dependencies;

- difficulties of calculations and interpretation of some statistical properties;

- scaling phenomena with complex dynamics;

- manifestation of non-extensive statistics.

In paper [8] it has been postulated that between the non-extensive q parameter and the Hurst exponent H exists a direct relation expressed as $H = 1/(3-q)$.

The template is used to format your paper and style the text All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

III. EXPERIMENT

In order to show the complex behavior of computer memory system, a simple experiment was performed. Thanks to the perfmon tool, an useful administrator program that is included into Windows systems, we collect a set of data that are related to cache memory. Perfmon is a computer program usually used by system administrators to trace, monitor and manage different parts of computer system (Fig. 1).
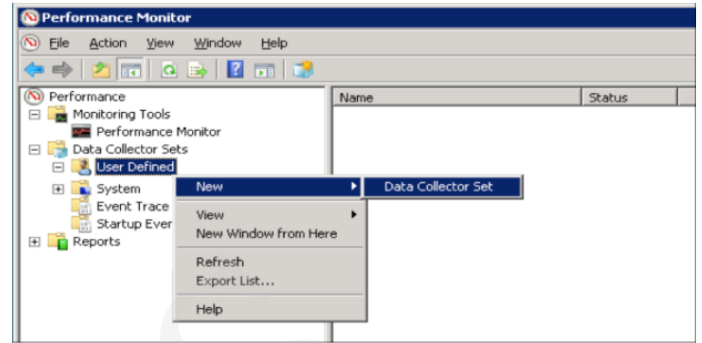

Fig. 1 Perfmon main window. A new data collector set is going to be defined

It helps to show the overview of the system and gives reports that can be used for system performance optimization. Its possibilities are based on different system counters that are normally traced by Windows system kernel, but thanks to the perfmon they also can be shown to computer user with different resolutions. This is very important, because perfmon almost doesn't generate any additional workload (let's note that this is normal computer program that is processed, so it must influence system performance). In experiment we trace 5 computer systems with different configurations; obtained data have at least 500k records. Details about systems configurations and traced records are presented in Table I.

TABLE 1 CONFIGURATIONS OF TRACED SYSTEMS AND SOME DETAILS ABOUT TIME SERIES

Config	CPU	RAM	HDD	Operating system	Number of collected records
1	AMD Athlon 64 X2 Dual-Core TK-55, 1.80 GHz	3.0 GB DDR2	Hitachi HTS542512K9SA00 ATA Device	Windows 7	697047
2	Intel Core i5 650 3.20GHz	4.0 GB DDR3	Unknown type, 500GB	Windows 7	519969
3	DualCore Intel Core 2 Duo T6600, 2.2 GHz	8.0 GB DDR3	ST9500325 ASG ATA Device	Windows 7	684499
4	Intel Pentium Dual CPU T322 2GHz	2.0 GB DDR2	Hitachi HTS543225L9A300 ATA Device	Windows 7	947020
5	Intel DualCore Celeron M T2130 2x 1.86GHz	2.0 GB DDR2	Segate ST9120822A SATA	Windows 7	1033098

The most important thing about carried experiment is fact that obtained records are based on normal work of computer user. This means that during experiment we don't use any benchmarks or computer programs that generate any artificial workload; records represent typical work of different computer users who use office programs, internet browsers, mail clients, games, multimedia, etc. The weakness of this approach is a fact that there is no possibility to repeat unique behavior of each user, but on the other hand, in the case of personal computers they usually don't process the extreme workload that is generated by benchmarks. Our approach shows computer system behavior during man-machine interaction. Obtained data were collected with resolution 1 s (the highest possible in perfmon) thus during 1 hour we were able to have 3.6k observations. Taking into account data presented in Table I it is clear that we trace our systems at least for 130 hours (or more). In our experiment we focused on one system counter, which helps evaluating memory and especially cache usage: Cache Bytes. It shows the amount of resident pages allocated in RAM that the Kernel threads can address without causing a page fault [9] consisting of:

- Memory\System Cache Resident Bytes as the size (in bytes) of the pageable operating system code in the file system cache (includes only current physical pages and does not include any virtual memory pages not currently resident),
- Memory\System Driver Resident Bytes as the size (in bytes) of pageable physical memory that is used by different device drivers,
- Memory\System Code Resident Bytes as the size (in bytes) of operating system code located in physical memory (it can be written to disk when is not used),
- Memory\Pool Paged Resident Bytes as the size (in bytes) of the paged pool. Space used by the paged and nonpaged pools is taken from physical memory, so a pool that is too large denies memory space to processes.

IV. RESULTS OF EXPERIMENT

Figures 1a-e show behavior of counter Cache Bytes for 5 different computer configurations. The first interesting, very well-visible fact is the behavior of this counter represented by high jumps (up and down) in cache memory usage. These jumps are related to the fact that during processing computer user runs different computer programs that require memory (especially visible by Pool Paged Resident Bytes).
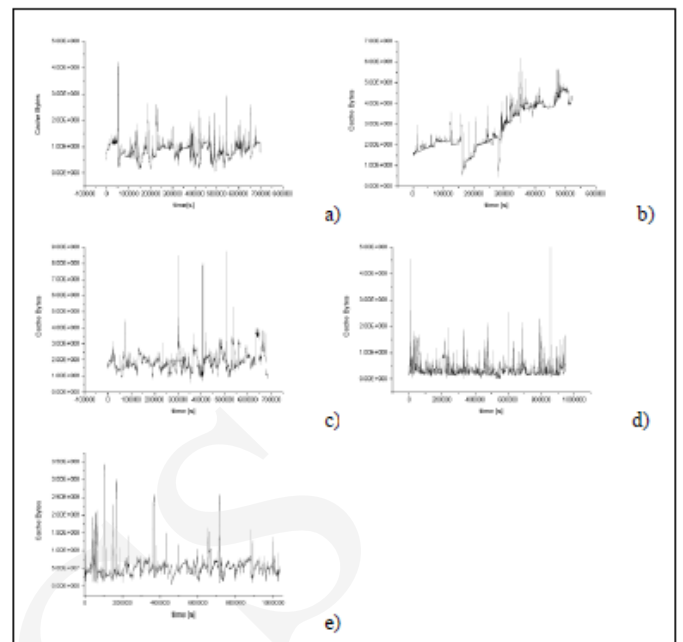

Fig. 2 Behavior of counter Cache Bytes during experiments for five hardware configurations from Tab. 1. Subfigures: a) computer Id. 1, b) computer Id. 2, c) computer Id. 3, d) computer Id. 4, e) computer Id. 5.

Because we cannot be sure whether these time series are stationary or not (especially in Fig. 2c, where we have well-visible increasing trend), the simplest way is to analyze their increments - figures 3a-e show details.

Increments presented in Fig. 3 seems to be typical for processes that usually are described by Lévy □-stable processes [10], which we described in Section 2. Figs 4a-e) show probability distributions of increments presented in Figs. 3 – each of them is done in log-lin scale with the fit of normal distribution in order to show the behavior of distribution tails.
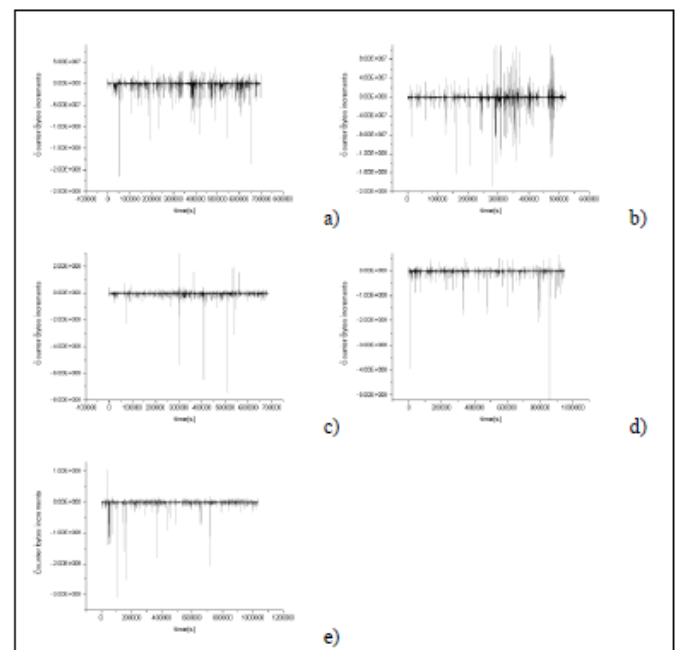

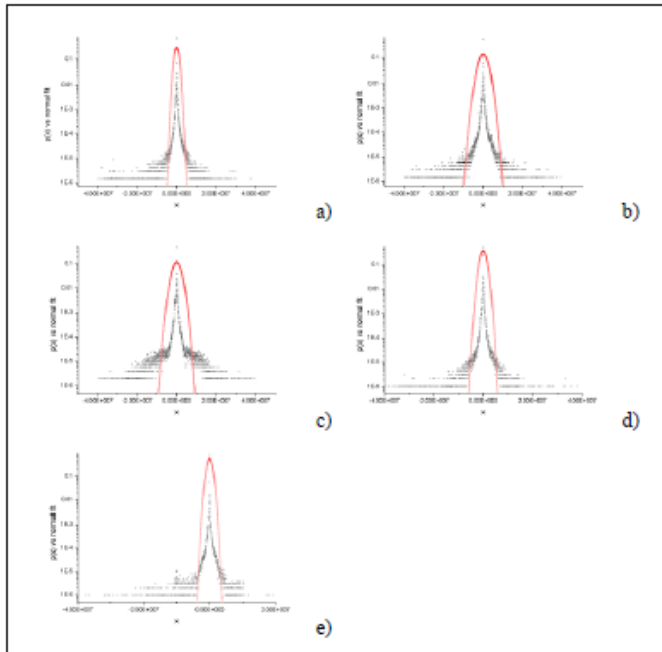Fig. 3 Increments of time series presented in Figures 2a-e.

Fig. 4 Probability distributions of increments of time series presented in Figures 3a-e.

From Figs. 4a-e) it is clear that each probability distribution (open squares) has slowly vanishing tail in comparison to normal distribution (open triangles). Figs. 5a-e) show the right tail of each distribution with linear fit that exhibits the possible existence of power laws (i.e., heavy tails).
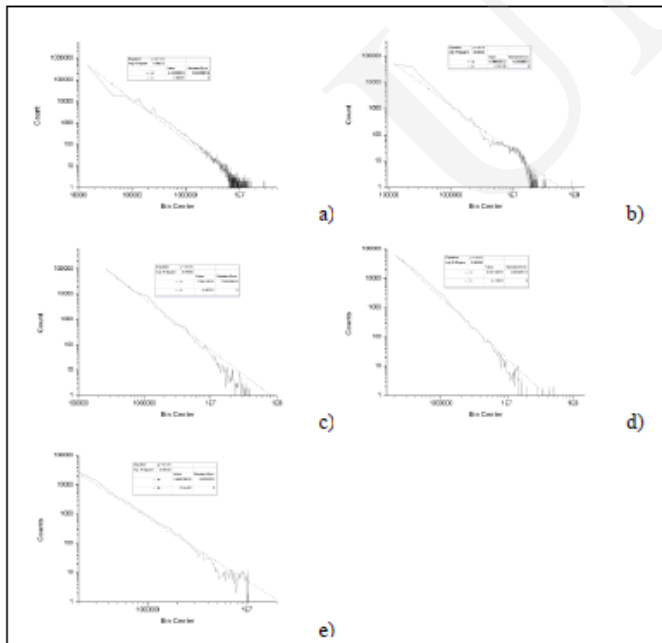

Fig. 5 Linear fit for right tails of probability distributions from Figures 4a-e.

In all cases we have results that indicate the possible existence of power-law behavior for tails distribution. The estimated slopes of fitted data are respectively: b = -1.905, b =

-1.761, b = -2.007, b = -2.1, b = -2.144. Because we use probability distribution function, the estimation of tail values in log-log scale can be done in relation to the Eq. (10):

$$\lim_{|x|\to\infty} p(x) \approx \frac{\alpha C^\alpha}{|x|^{1+\alpha}} \approx |x|^{-1-\alpha}, \qquad (10)$$

where C is a positive constant, however this is a very rough approximation giving us stability index $\alpha = -b+1$: $\alpha = 0.905$, $\alpha = 0.761$, $\alpha = 1.007$, $\alpha = 1.1$, $\alpha = 1.144$. But this method works satisfactory only when we have ideal (simulated, not skewed) probability distributions. Here, in order to calculate the stability index we need to focus on different approach based on relation:

$$\lim_{|x|\to\infty} (1 - F(x)) \approx \frac{\alpha C^\alpha}{|x|^{1+\alpha}} \approx |x|^{-\alpha}, \qquad (11)$$

where $F(x)$ is cumulative distribution function. We obtain respectively: $\alpha = 1.825$, $\alpha = 1.651$, $\alpha = 1.781$, $\alpha = 1.93$, $\alpha = 1.69$, however again basing on graphical approach where we plot 1-$F(x)$ vs x on a log-log scale. In order to obtain more accurate results we need to use methods based on Hill estimator or maximum likelihood estimator.

V. CONCLUSIONS

In this paper we show behavior of one system memory counter (obtained from five different computer systems). Despite the fact, that our findings (power law behavior in analyzed time series) were done basing on graphical methods and without any parametrical statistical tests, we were able to discover some interesting features of Cache Bytes counter. Obtained results make a strong connection between Tsallis entropy, its thermodynamical background and computers. It seems that the existence of Lévy stable processes in the case of computer systems are not widely known in the case of computer systems analysis (behavior of different system parts), but on the other hand they are showing new interesting experiment possibilities. The so far obtained results need to be developed relating them to the more detailed study of time series statistical properties (for example towards statistical self-similarity in time domain [11]) and to be compared with results for other different hardware systems configurations. This task is not the easy one, however proposed approach seems to confirm that computer systems work in states that are far from equilibrium states that can be described and analyzed by the concept of Tsallis entropy. Let's note that this thermodynamical background, which at first glance seems to be proper for considerations of combustion engines, is firmly seated in computer systems, which in fact are also machines that perform (electrical) energy transformations into useful work (calculations).

REFERENCES

[1] Aristotle, Metaphysics (from The Complete Works of Aristotle: The Revised Oxford Translation), ed. J. Barnes, Princeton, 1984.

[2] P. Wegner, Research paradigms in computer science, Proc. of the 2nd Int. Conf. on Soft. Eng., San Francisco, California, pp. 322-330, 1976.

[3] M. M. Waldrop, Complexity: The Emerging Science at the Edge of Order and Chaos, Simon and Schuster, USA, 1992.

[4] F. Grabowski, Nonextensive model of self-organizing systems, Complexity, vol. 18, no. 5, pp. 28-36, 2013.

[5] C. Tsallis, Possible generalization of Boltzmann-Gibbs statistics, J. Stat. Phys., vol. 52, pp. 479–487, 1988.

[6] C. Tsallis, F. Baldovin, R. Cerbino, P. Pierobon, Introduction to Nonextensive Statistical Mechanics and Thermodynamics, in The Physics of Complex Systems (New Advances and Perspectives), Eds. F. Mallamace, H.E. Stanley, IOS Press, 2004.

[7] C. Tsallis, Nonextensive statistical mechanics, anomalous diffusion and central limit theorems, Milan J. Math., Vol. 73, pp. 145–176, 2005.

[8] D. O. Cajueiro, B. M. Tabak, "Is the expression H = 1/(3-q) valid for real financial data?", Physica A, vol. 373, pp. 593-602, 2007

[9] https://msdn.microsoft.com/en-us/library/ms804008.aspx

[10] A. Weron, K. Burnecki, S. Mercik and K. Weron, Complete description of all self-similar models driven by Lévy stable noise. Phys. Rev. E, 71 p. 016113, 2005.

[11] P. Dymora, M. Mazurek, Network Anomaly Detection Based on the Statistical Self-similarity Factor, Lecture Notes in Electrical Engineering vol. 324(1), pp. 271-287, 2015.