Analysis of probabilistic distributions of the output parameters of human systems by the method of equations of state (Both Gaussian and Paretian paradigms adequately represent the behavior of human systems in various situations) [1]

Pavel Barseghyan, PhD

Abstract

From the point of view of mathematical modeling, the qualitative adequacy and quantitative accuracy of the methodology for describing human activities are directly based on the qualitative adequacy and quantitative accuracy of the deterministic patterns of their behavior and the probability distributions of the parameters contained in them.

If we consider the output parameters representing human activity from this point of view, such as, say, the duration of activities performed by people, then it should be said that there are many uncertainties in the choice of the probability distributions of these quantities, which directly affect the quality of simulations.

In particular, this issue is of particular interest from the point of view of the transition from the Gaussian paradigm to the Paretian paradigm.

This work is devoted to the analytical substantiation of the probability distributions of the output parameters of human activity using the method of equations of state of human systems.

Introduction

To ensure progress in various spheres of people's lives, methods of planning their activities, forecasting the future and assessing various associated risks are needed.

In each specific case of any human activity, it is assumed that certain results will be obtained in such conditions when the necessary resources are influenced by random factors, and people, acting in random environments, in turn, have random characteristics.

In such conditions, the organization of human activity and the analysis of its feasibility involves the use of expert systems, which should be based on the basic quantitative laws of

[1] How to cite this paper: Barseghyan, P. (2021). Analysis of probabilistic distributions of the output parameters of human systems by the method of equations of state (Both Gaussian and Paretian paradigms adequately represent the behavior of human systems in various situations); *PM World Journal*, Vol. X, Issue II, February.

human life, behavior and activity of a deterministic nature and the probability distributions of system parameters included in these laws.

Based on this analysis, decisions are made about what needs to be done at the moment, and the future is predicted based on tactical and far-reaching strategic considerations.

It is also clear that tactical goals, decisions and corresponding actions naturally follow from general strategic considerations.

From the point of view of mathematical modeling, the qualitative adequacy and quantitative accuracy of the description of human activity are directly based on the qualitative adequacy and quantitative accuracy of the above-mentioned deterministic relationships and probability distributions of system level parameters of human activity.

If we consider the input and output parameters of the activity of human systems from this point of view, then it should be noted that at present there are many fundamental and practical uncertainties in the choice and substantiation of the probability distributions of these quantities.

First of all, these uncertainties are associated with the current transition from the Gaussian paradigm of the analysis of human activity to the Paretian paradigm, the choice of probability distributions of the parameters of human activity for this analysis and the manipulations performed on them.

Despite the fact that many important studies have been carried out in this direction [1,2,3,4], there are still fundamental and practical difficulties in this area.

Moreover, there are fundamental difficulties both in the field of representing the deterministic laws of human activity, and in the field of choosing the probability distributions of system parameters that characterize their activities, which will be presented below.

As for the difficulties of an applied nature in this direction, they inherently have both objective and subjective features, of which the following three can be distinguished:

1. The limited capabilities of methodologies and the underlying mathematical models, because of which human intuition plays an important role in them, which, in turn, carries the risk of people making arbitrary decisions.
2. These methods do not have a full scientific justification, they mainly have a business focus, which can become a possible reason for the loss of a business in the future.

3. Subjectively, the Gaussian paradigm is more easily understood by people, on the other hand, understanding the Paretian paradigm requires deeper scientific training.

Despite these difficulties, there is a gradual transition from the Gaussian paradigm to the Pareto paradigm, which is dictated by the limited possibilities of the Gaussian approach and is carried out under the pressure of business practice.

However, both of these approaches and the models based on them take place in real life as adequate reflections of the life, behavior and activities of people in various situations.

Observations and scientific analysis of the life, behavior and activity of human systems show that the behavior of the same system under different conditions can be adequately represented in one case by Gaussian models, and in the other by Pareto-type models.

This means that if it is possible to construct a fundamental mathematical theory describing the life, behavior and activity of human systems, then within the framework of this approach it will be possible to present the Gaussian and Paretian paradigms as different concrete examples of the application of the same theory.

In this sense, the mathematical theory of human systems creates the necessary capabilities that, with the help of the mathematical apparatus of the equations of state, allow combining the resources necessary for activity, human properties and the results of human efforts in one quantitative model of human activity [5].

In particular, this new theory allows the method of the equations of state to be applied to planning and evaluating project behavior in high-risk areas.

A high level of project risk can be caused by various sources and reasons, including a lack of financial, information and other resources, ineffectiveness of the project team, conflict situations, etc.

Thus, in a multidimensional high-risk environment, a consistent transition from the Gaussian paradigm to the Pareto paradigm, which has much wider possibilities, is very important to improve the quality of project planning and execution.

It can be shown that if two different project teams with different performance capabilities perform the same project under the same conditions, it is possible that estimates for a high-performing team are made using a Gaussian approach, and for estimates for a lower-performing team, the Pareto approach.

This is because what is risky for a weak team may not be risky for a strong team.

In other words, design complexity of a project that is non-linear for a weak team can be linear for a strong team.

One of the main goals of this paper is to show that both the Gaussian approach and the Pareto approach are applicable to the analysis of various project situations, and that Pareto-type analysis is adequate in the high-risk area.

The further purpose of this article is to show, using specific examples based on equations of state, that under different conditions, the same mathematical model for the output parameters of human activity can generate various probabilistic distributions, including Gaussian distribution, skewed distributions, and fat tail distributions.

State equations of the activity of human systems and the probability distributions of its output parameters

The characteristic parameters of many created by nature and man-made phenomena and material objects naturally obey the Gaussian law of normal distribution.

These can be the size and weight of people, animals and plants, the characteristics of factory products, and so on.

In contrast to these quantities, which have symmetric distributions of their characteristic parameters that are best approximated by Gaussian law, many of the parameters associated with the operation of human systems have asymmetric distributions, some of which have fat-tailed distributions.

As mentioned above, in various spheres of human activity, including project management, entrepreneurship, public administration, etc., under the pressure of practical requirements, there is a gradual transition from the Gaussian paradigm to the Paretian paradigm.

This is manifested in the fact that various symmetric and asymmetric distributions are introduced into circulation, thus meeting specific requirements for simulating phenomena and processes in various fields of human activity.

Although these studies cover both purely mathematical aspects of the problem and the issues of applied quantitative modeling, they do not solve the main problem, which consists in creating a general concept and theory of probability distributions of the output parameters of the activity of human systems.

More generally, the solution of the problem lies in a unified quantitative approach to describing the life, behavior and activities of human systems, ideally in the form of a unified mathematical theory of the fundamental nature of these systems.

The role of this theory in solving the problem under consideration is that it allows one to obtain the results of human activity as a function of input parameters that characterize the resources necessary for the activity and the characteristics of people.

Having these functional relationships and some empirical data, as well as hypotheses regarding the distributions of input parameters and using analytical means, one can obtain the probability distributions of the output parameters of human activity.

This method allows to obtain probability distributions of the output parameters of human activity by much more reliable analytical methods compared to the choice of these distributions in a hypothetical way.

In this sense, the equations of state of human systems provide ample opportunities for the study of these distributions by analytical methods and their practical applications [5].

The equations of the state of life, behavior and activity of people reflect the idea of the continuity of life of human systems and quantitatively combine the resources necessary for the functioning of human systems, human experience, knowledge and skills, the complexity of life, as well as the results of their activities in the form of quality and level of life, safety of people, progress of society, and so on.

This theory is based on the equations of state of human systems, which are the result of a balance between the sequence of demands of life, presented to people and the actions they take to meet these demands.

The components of this balance are:

- The magnitude, difficulty and complexity of the successive problems that people face in life;
- Biological, social, material, intangible, ideological, moral and other resources necessary for human activity,
- Experience, knowledge and skills of people,
- The results of human activity.

The combination of parametric representations of these components and the conditions of balance between the needs of life and human capabilities give rise to equations representing the state of human systems.

The life, behavior and activity of human systems can be in static or dynamic modes, that is, the parameters that represent their state may not depend or depend on time.

The static state of human systems is represented by algebraic equations, and the dynamic state is represented by differential equations [5,6].

In the simplest case, the static mode of human activity can be represented by the following equation of state

$$N * T * P = W * D \qquad (1)$$

This equation describes the case where a collaborative group of N people, working at performance P during period T, performs an activity of size W and difficulty D.

Taking into account that the difficulty of human activity D has an upper limit of feasibility $\overline{D}$ then the equation of state will look like this [7]

$$N * T * P = W * \frac{\overline{D} * \mathrm{D}}{\overline{D} - \mathrm{D}} \qquad (2)$$

Some of the parameters included in these equations are input parameters, the second part represends itself the process of human activity, and the third part is the output parameters of the same activity.

In particular, as an output parameter, one can consider the duration of human activity T.

All parameters included in the equations of state, including those representing resources, the complexity of life requirements, human capabilities and results of human activity, as mentioned above, also have random components that can have arbitrary probability distributions depending on the situation or other circumstances.

As for the considered symmetric, asymmetric and fat-tailed distributions, they can, depending on the situation and circumstances, describe the behavior of the input and output parameters of human systems.

The meaning and role of the probabilistic interpretation of these equations is that they transform the distributions of the input parameters into the distributions of the output parameters.

In other words, in an analytical way, that is, using the equations of state, functional dependences of the output parameters on the input parameters of the human system are obtained, and, therefore, the output parameters are considered as functions of random arguments, that is, input parameters.

Let's use this approach to study the probability distributions of duration T, considering it as one of the output parameters of human activity.

Duration of human activity as a function of random arguments

From the above equations of state (1) և (2), two expressions can be obtained for the duration T of human activity, the first of which contains a linear relationship between this value and the difficulty of activity D, and the second contains a nonlinear relationship $T(D)$, which becomes linear in areas of small difficulties (Fig.1):

In the case of a linear relationship for the duration of human activity, one can get:

$$T = \frac{W*\mathrm{D}}{N*P}\,, \qquad (3)$$

and in the case of a nonlinear relationship, we get the following expression for the same value

$$T = \frac{W}{N*P} * \frac{\bar{D}*\mathrm{D}}{\bar{D}-\mathrm{D}} \qquad (4)$$

All quantities on the right side of these expressions in various life situations can have both deterministic and random behavior with corresponding probability distributions.

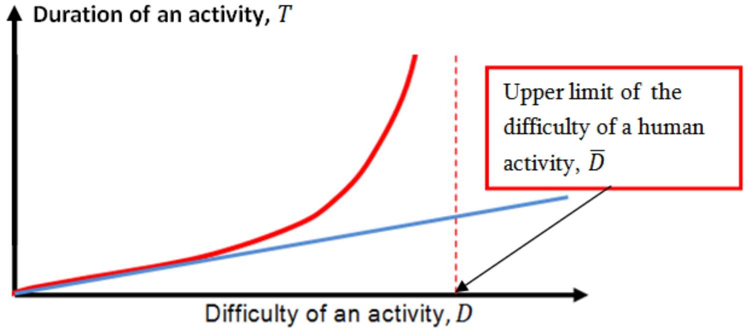


Fig.1 Graphical representation of linear and nonlinear mathematical models of functional relationship $T(D)$. In the area of small difficulties, the two models coincide.

Depending on which distributions have input parameters and in which mode a given human system operates, linear or nonlinear, the duration of activity will have: 1. normal or other symmetric distribution, 2. asymmetric distribution, and 3. fat-tail distribution.

Let's consider some typical examples of them as special cases.

1. All characteristic parameters of the human system described by the equation of state (3) and operating in a linear mode are constant values with the exception of the difficulty D, which has a normal distribution.

In this case, the duration of activity T will have a normal distribution as a result of a linear transformation of the input normal distribution of difficulty D (red curve of the distribution density function in Fig. 2).

2. All characteristic parameters of the human system described by the equation of state (4) and operating in a nonlinear mode are constant values with the exception of the difficulty D, which has a normal distribution.

In this case, the duration of activity T will have an asymmetric distribution due to the nonlinearity of the dependence T (D) (brown curve of the distribution density function in Fig. 2).

In such a situation, the closer the mean value D_{av2} of the distribution function $\psi_2(D)$ to the upper feasibility limit $\overline{D}$, the greater the spread of the asymmetric output function, which will eventually become a fat-tailed distribution.

3. All characteristic parameters of the human system, described by the equation of state (3) and operating in a linear mode, are constant values, with the exception of the difficulty D and productivity P, which have normal distributions.

This particular case is of interest in view of the serious delays that unexpectedly occur in the normal course of human activity.

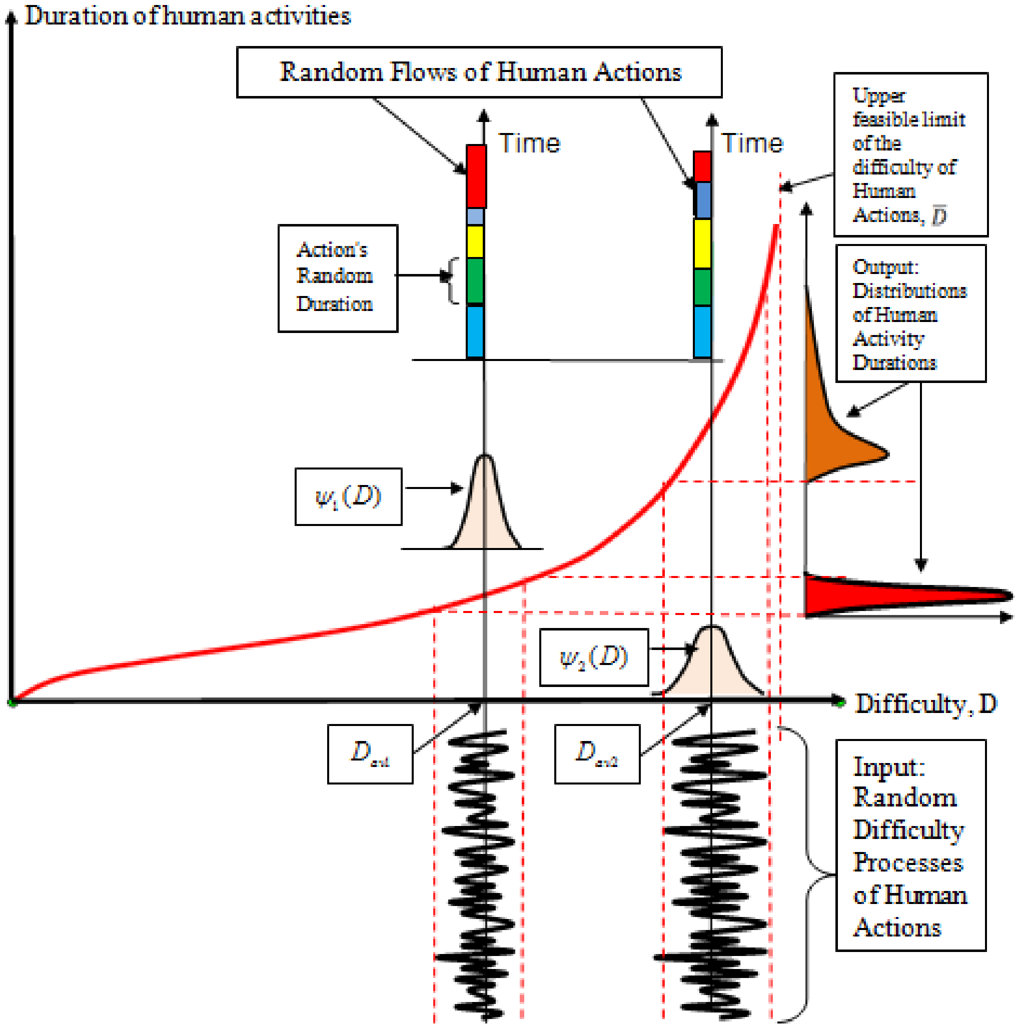


Fig. 2 Two normal input processes representing a probabilistic picture of human activity with distribution densitiy functions of their difficulties ψ_1(D) and ψ_2(D) and average values D_{av1} and D_{av2}

This case does not imply large risks of delay in the relatively small spread of difficulty D and productivity P, which is characteristic of simple human activity.

But if the spread of these two parameters increases depending on the nature of human activity and the state of people, this can cause significant delays in their activity, up to its failure.

This means that the density distribution function of the duration T of human activity will have a large skewness towards large delays.

Qualitatively, this phenomenon can be explained by the fact that accidentally big difficulties in the process of human activity in time can coincide with the low productivity of people.

Quantitatively or mathematically, this phenomenon is explained by the fact that the distribution of the duration T of human activity is the result of the ratio of two normal distributions, which is the Cauchy distribution with a fat tail [8,9].

4. With regard to the fact that the input parameters themselves can have distributions with fat tails, in this sense the parameter of difficulty D, which people face during their activity and the efficiency of their activity P are of particular interest.

These quantities, depending on the scale of the human system, the nature and organization of its activities, can have asymmetric distributions with a fat tail.

In more general cases, the expected distribution function of the duration of human activity can be obtained analytically, considering it as the result of the product and ratios of random variables with given distributions.

Continuation of analysis

The probabilistic analysis of the duration of human activity can be further enriched and concretized with the use of more ramified and more representative equations of state.

In this sense, the following equation of state of the activity of human systems is of particular interest.

$$N * T * M * I_{Max} = (1 + at_a) * (W - W_{0)}) * D, \qquad (5)$$

where new parameters were added, such as the relatively easily measurable intensity of people's activity I_{Max}, the intensity of their errors a and the average time t_a spent on correcting one error, as well as the motivation of people M and their previous experience W_0 [5].

Similar to probabilistic analysis of the duration of the activity of human systems, it is also possible to analyze other output parameters of human activity, including its results and associated risks.

Particular attention should be paid to the probabilistic representation of the complexities encountered in human systems in the process of life, complexities that can be many times greater than the capabilities of human systems that implement them.

Among such cases are the so-called “black swans” in the life of large-scale human systems, which are rare and difficult to predict, but are largely explainable if they occur [10].

If we consider the complexities and difficulties that people face in life, then a person very rarely encounters extreme complexities and difficulties in everyday life, mainly dealing with simple actions.

This fact suggests that the complexities and difficulties encountered in people's lives can be represented by power laws.

Conclusions

1. The current paradigm of probabilistic analysis of success and risks of failure associated with human activities is based on hypothetical distributions of the considered key parameters, be they symmetric or asymmetric, triangular, Gaussian or lognormal, etc.

2. There is no rational statistical basis for choosing any type of hypothetical distribution for assessment of the results of human activities, and the accuracy and quality of solutions based on such hypotheses can be characterized by the well-known approach “garbage in, garbage out”.

3. This circumstance speaks of the need to search for fundamentally new ways of solving the problem, of which the top-down approach of an analytical or theoretical nature should be especially noted, especially taking into account the fundamental impossibility of having reliable statistical data in this direction.

4. Under these conditions, the only acceptable way to solve the problem is to create a reliable theory of the quantitative nature of human activities, based on a parametric representation of various aspects of human life.

5. The main goal in this direction is the development of parametric representations or mathematical models of various aspects of human activity with relatively easily measurable parameters.

6. The combination of these mathematical models within the framework of a general concept or principle that reflects the life, behavior and activities of human systems, allows to build mathematical models of these systems of a fundamental nature.

7. Mathematical models of human systems are in the form of equations of state, which contain three types of parameters: input parameters, parameters representing the essence of systems and output parameters.

8. Input parameters in the equations of state of human activity usually represent financial, informational, material, intangible and other resources necessary for the activity.

9. In the same equations, the parameters reflecting the essence of systems represent the characteristics of people and the characteristics of the activities performed.

10. The output parameters of human activity reflect the quantity and quality of the results of the activity, the probability of its successful completion or the risks associated with failure, as well as the duration of the activity.

11. Given that the first two types of parameters are random in nature, but also measurable, it is possible to get an idea of their probability distributions by direct measurements or expert methods.

12. On the other hand, mathematical models of human systems make it possible to present the results of their activities, including the duration of activities, as functions of random parameters in an analytical way in the form of formulas.

13. This means that the probability distributions of the output parameters of human activity can be obtained analytically, using the mathematical apparatus of functions of random arguments of probability theory, instead of being hypothesized.

14. Even if the parameters representing the components of human activity have a symmetric, say, Gaussian distribution, the result of their activity may have a fat-tailed distribution due to the mathematical form of the equations of state.

References

1. Joel A. C. Baum and Bill McKelvey. (2006). **ANALYSIS OF EXTREMES IN MANAGEMENT STUDIES**. Research Methodology in Strategy and Management, Volume 3, 123–196 Copyright r 2006 by Elsevier Ltd.
2. Crawford, G. Christopher and Aguinis, Herman and Lichtenstein, Benyamin and Davidsson, Per and McKelvey, Bill, **Power Law Distributions in Entrepreneurship: Implications for Theory and Research** (January 2, 2015). Journal of Business Venturing 30(5) 696-713 doi:10.1016/j.jbusvent.2015.01.001, Kelley School of Business Research Paper No. 15-25, Available at SSRN: https://ssrn.com/abstract=2578029
3. Andriani, P., & McKelvey, B. (2009). **From Gaussian to Paretian Thinking: Causes and Implications of Power Laws in Organizations**. Organization Science, Vol. 20, No. 6,

November-December 2009, pp. 1053-1071. http://leeds-faculty.colorado.edu/dahe7472/Adriani%2025614715.pdf.
4. Hilbert, M. (2013), **Scale-free power-laws as interaction between progress and diffusion**. Complexity. doi: 10.1002/cplx.21485; http://onlinelibrary.wiley.com/doi/10.1002/cplx.21485/abstract. http://martinhilbert.net/Powerlaw_ProgressDiffusion_Hilbert.pdf
5. Pavel Barseghyan (2017) **"Elements of the Mathematical Theory of Human Systems Part 1: Assessment of the Results of Human Actions and Activities Based on the Method of State Equations "** *PM World Journal Volume 6, Issue 12 December 2017* – 13 pages. http://pmworldjournal.net/wp-content/uploads/2017/12/pmwj65-Dec2017-Barsegyan-mathematical-equation-of-human-systems-part1-featured-paper.pdf
6. Barseghyan, P. (2019). **Elements of the Mathematical Theory of Human Systems – Part 6: Human Life as a Project (The differential equation of human life);** *PM World Journal*, Vol. VIII, Issue II (February). Available online at https://pmworldjournal.net/wp-content/uploads/2019/02/pmwj79-Feb2019-Barseghyan-the-differential-equation-of-human-life-featured-paper.pdf .
7. Barseghyan, P. (2020). **Quantitative Models of Emergence in Light of Mathematical Theory of Human Systems - Part 2: Nonlinear Mathematical Models of New Quality Emergence**; PM World Journal, Vol. IX, Issue IV, April. https://pmworldlibrary.net/wp-content/uploads/2020/03/pmwj92-Apr2020-Barseghyan-Quantitative-Models-of-Emergence-Part2-nonlinear-models.pdf
8. Continuous Univariate Distributions, Vol. 2 (Wiley Series in Probability and Statistics). (1994), by Norman L. Johnson, Samuel Kotz, N. Balakrishnan.

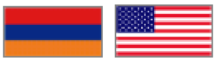
9. Ratio distribution. Wikipedia.
https://www.google.com/search?q=ratio+normal+distributions&oq=ratio+distributions&aqs=c hrome.2.69i57j0i19i22i30l7.23980j0j15&sourceid=chrome&ie=UTF-8

10. Taleb, Nassim Nicholas. **The Black Swan: The Impact of the Highly Improbable**. New York: Random House, 2007.

About the Author

Pavel Barseghyan, PhD

Yerevan, Armenia
Plano, Texas, USA



Dr. Pavel Barseghyan is a consultant in the field of quantitative project management, project data mining and organizational science. Has over 45 years' experience in academia, the electronics industry, the EDA industry and Project Management Research and tools development. During the period of 1999-2010 he was the Vice President of Research for Numetrics Management Systems. Prior to joining Numetrics, Dr. Barseghyan worked as an R&D manager at Infinite Technology Corp. in Texas. He was also a founder and the president of an EDA start-up company, DAN Technologies, Ltd. that focused on high-level chip design planning and RTL structural floor planning technologies. Before joining ITC, Dr. Barseghyan was head of the Electronic Design and CAD department at the State Engineering University of Armenia, focusing on development of the Theory of Massively Interconnected Systems and its applications to electronic design. During the period of 1975-1990, he was also a member of the University Educational Policy Commission for Electronic Design and CAD Direction in the Higher Education Ministry of the former USSR. Earlier in his career he was a senior researcher in Yerevan Research and Development Institute of Mathematical Machines (Armenia). He is an author of nine monographs and textbooks and more than 100 scientific articles in the area of quantitative project management, mathematical theory of human work, electronic design and EDA methodologies, and tools development. More than 10 Ph.D. degrees have been awarded under his supervision. Dr. Barseghyan holds an MS in Electrical Engineering (1967) and Ph.D. (1972) and Doctor of Technical Sciences (1990) in Computer Engineering from Yerevan Polytechnic Institute (Armenia).

Pavel's publications can be found here: http://www.scribd.com/pbarseghyan and here: http://pavelbarseghyan.wordpress.com/. Pavel can be contacted at terbpl@gmail.com