Teaching Computer Science Using Visualization of the Evaluation Process

Oleg V. Stukach
TPU, 30 Lenin Avenue, Tomsk, 634050, Russia

Abstract – **In this paper, a new methodology of training of the software testers is offered. The methodology is based on the modified method of expert evaluations. Control quality charts are used for processing and visualization of results. Features of expert testing during training and revealing of distinctions are considered. Advantages of the developed methodology are shown on example.**

Index Terms – **control quality chart, software tester, evaluation methodology, Statistica software**

I. Problem of Computer Sciences Training

Quality of training of the computer science experts is the basic factor of competitiveness of the software production organization. The expert should be able to orientation in techniques of software testing, correctly and quickly to choose necessary ones. It is known that concept of quality is connected to expectations of the consumer. Thus, in many cases quality can be estimated only by the expert way i.e. from positions of any characteristic set of preferences. Therefore, the problem of numerical techniques development of the quality software estimation and simultaneously quality of training of testers is rather urgent, especially in the market competition conditions. The traditional ways of estimation connected with summarized rating are not appropriate to the software developer teaching. Partially knowledge is dangerous because any error in a program may refuse one despite decorating of the program interface.

For improvement of the program organization of computer systems the continuous training and self-improvement of computer experts and students are necessary. At realization of this training it is necessary to combine forms of traditional training with opportunities of independent work, distance learning, practical development of projects and realizations of experimental researches. The modern methods of statistical and expert evaluation of quality of the computer system program organization play not the last role at training. In this sense the application of mathematical and numerical methods allows to increase quality of testing and finally, quality of all projects.

There is a feature of student training on specialty "IT engineer". The student should not only have the appropriate knowledge on the specialty but also be able to work in a team. Modern program complexes are created by the large number of developers actively cooperating among themselves. Therefore the average level of a group is more important that knowledge and skills of each developer during project performance. Besides, the problem of the software estimation quality has no authentic right solution. Thus, the ordinary problems frequently are not appropriate for modern training software developers.

II. The Modified Expert Evaluation Method

Use of the expert evaluation method allows to solve a triad of problem:

- To estimate the knowledge level of each expert;
- To show of non-standard creative abilities;
- To show how expert work in group.

The new methodology of testing and simultaneous evaluation score consists in the following. Specialist in the expert role is offered to answer some questions, evaluating some probable alternatives. For example, it is possible to offer to evaluation three source codes of program by possible criteria: brevity of code, originality of algorithm, minimal memory size, minimal time of operation, absolute autonomy etc. Evaluation is carried out in a non-numerical scale to exclude possible mistakes. The most suitable grades are following: excellent, very good, good, satisfactory, unsuccessful, and bad [1, 2]. Screening of works the non-numerical scale is replaced to numerical as in Delphi method [2]: 1 (excellent), 0.75 (very good), 0.625 (good), 0.5 (satisfactory), 0.25 (unsuccessful), 0 (bad). Further average mean and dispersion of grades of each expert was calculated.

Work in a group means that the problem is solved by all members. Therefore from grades of each expert the "average expert" was deduced, which grades represented average grades and dispersions of all experts. In this case we modeled a classical evaluation expert method. It is clearly that the average expert not always satisfies to necessary requirements. In most cases the poor-quality grades may included in the average expert ones. The average expert only allows to see tendencies in evaluations of all expert team, but its grades is not a guarantee of true.

It is known from the formal theory of measurement that the validity of measurements is determined by comparison with standard. There are no standard in the expert evaluation method because there is no absolutely true result of measurement. In our case the standard is teacher which knowledge corresponds for high degree what should be

achieved by each expert and hence, the average expert. It is the reference expert.

Mean and dispersion for grades of each expert, average, and standard experts were calculated. It allows to compare a knowledge level of experts from each other and with the average expert. The researches result for one group is shown in Fig. 1.

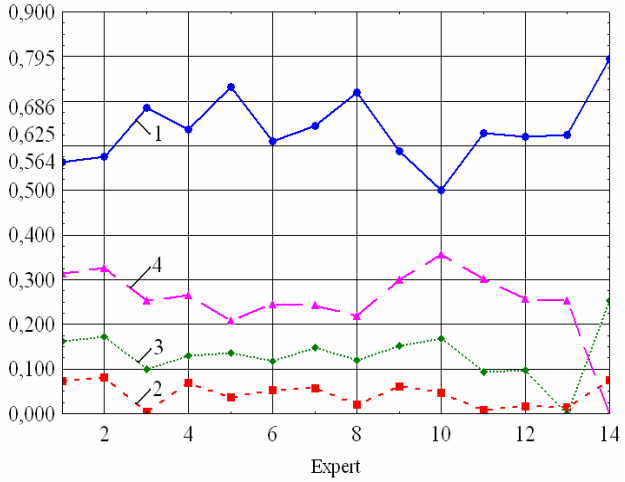

Fig. 1. 1 – mean values of each experts; 2 – dispersions; 3 – deviation of the expert grade from average expert; 4 – deviation of the expert grade from the reference expert

On X-axis the expert number, average (13) and the reference (14) experts are marked. It is possible to observe a deviation of expert grades from reference and average. From diagrams of differences it is possible to determine, as far as the deviation of grades of each expert from average (graph 3) and from reference (graph 4) is much more. Grades of 3, 11, and 12-th experts more closely to average (13), and 2, 9 and 10 are rather far in the evaluations. The grades of average expert are not good approximation to the true value, i.e. to grades of the reference expert (14). Its average evaluation on the graph 1 essentially differs both from evaluation of the average expert, and from everyone. Thus, knowledge of the reference expert essentially highest that average one.

In Fig. 1 the mean grade of 5 and 7 experts is most appropriate to reference. Their knowledge is the highest in group. Tenth expert is most differed from group which grades are far both from average and reference experts. It may be recognized as poor expert. It is justified also by dispersion (graph 2): 10 experts have the biggest dispersion grades.

All questions seem different for experts, so it is possible to judge concern differences through dispersion. The small dispersion marks insignificance of distinctions that is bad in our case. It is possible to judge through dispersion about capability of experts to work in a group. At strong scattering of dispersion it is possible to conclude that solved common decision by the expert group may be no correct though some experts may be absolutely exact in their evaluations.

The uniform dispersion testifies that strong contradictions in group are not observed concerning the done work. Investigations allow to solve the discrimination problem of the most productive experts from group. On this basis it was offered to form separate groups of experts from allocated in several ones. On the dispersion we judged how many these experts are compatible in uniform decisions. We create uniform team having achieved a closely dispersion.

III. Control Charts

Total Quality Measurement is widely used in many technological processes in modern industry and service [3]. The system of quality measurement is necessary to continuously supervise the process conditions and to guarantee quality of production or service. If it is necessary the system alloy to control of process at its deviations from normal mode and to prevent thus losses from release of poor-quality production. Facilities control quality may be used almost in any production for determine and choice of a problem, diagnostics of possible reasons of problem occurrence and make decision and recommendations for elimination of the problem.

It is offered to apply concepts of quality management in the training system of information technology experts. For this purpose the statistical methods of quality management for monitoring, estimations and developments of decisions on management of educational process were used on the basis of control charts [3]. These charts were carried out in the statistical system “Statistica” [4].

An example of control X-R chart with expert grades is shown in Fig. 2. Every three points on the chart corresponds to each expert that is grades for three variants of problem evaluation. Grades of the average expert are not designed in the chart, i.e. 14-th reference expert follows directly after 12. The chart consists of two histograms, X-chart of deviations and R-chart of ranges. It is possible to estimate does the distribution of grades correspond to the Gaussian law from the X-chart histogram. If the distribution law is far from normal, the sample volume should be increased.

Estimations of each expert are presented as a varied curve on the means and ranges chart. As it seems from diagrams, changes may be in the limits so-called “control limits” (shaped lines). These control limits automatically calculated from a “six sigma” rule. Also it can be designed compulsorily if limits of allowable changes of grades are known. The average line represents a population mean of all estimations. Deviations from average and output for control limits give an evident picture of the decision of problem all of expert group. On X-chart it is possible to see the points which are overstepping the control limits. For quality measurement of production it means that it is

necessary to make a correction to the process as it becomes uncontrollable. In our case it may mean that a number of experts show higher knowledge, than the average expert. It is desirable that these points became as much as possible. It is clearly that than knowledge of the reference expert is higher than at any of participants of group. Therefore in the example (Fig. 2) first and third points of the reference expert appeared behind control limits. It is important that grades of each of problems essentially differed. It is not observed at 3 and 11 experts. From R-chart it is visible that 11 expert has identical average range, that puts its work under doubt.

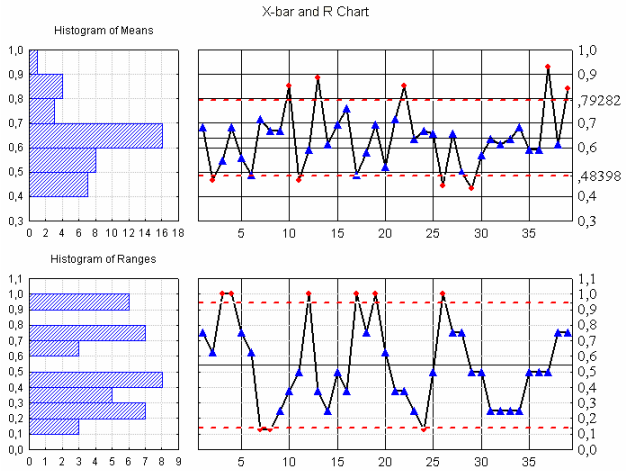

Fig. 2. Example of X-R chart for visualization of expert grades of software developer group

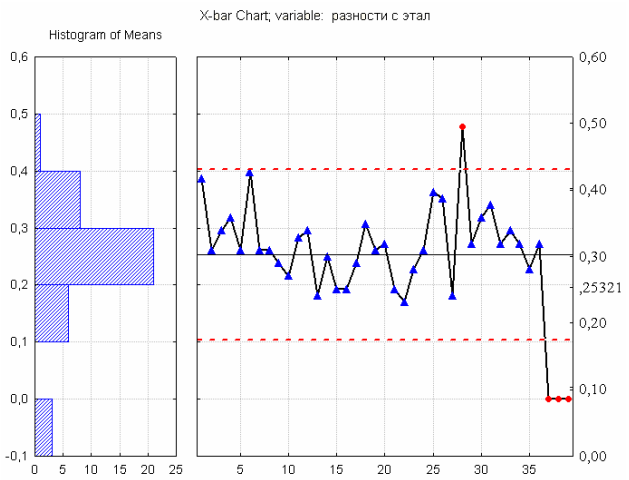

Fig. 3. Example of control X-chart of deviations of experts grades from grades of the reference expert

The control X-chart of deviations for average grades of experts from reference expert is presented in Fig. 3. For comparison the reference expert grades are included also in a chart (three last points). The control limits designed by the six-sigma rule shown as far as the reference expert grades are far from estimations of all group. It is visible also from chart that grades of 10 (poor) experts are far from grades of all group and reference expert. Grades of 5 and 7 experts are much more closed to the standard. It is visible from all constructed diagrams in Fig. 1–3. Hence, a level of knowledge of 5 and 7 experts is the highest in group. In the knowledge process more and more points will come near to average line. It will enable to carry out stratification that is search of the group having identical average grades that will mean improvement of quality of expert work in the group.

Thus, analysis of the control charts allows to correct process of training of the software developers and gives the information concern quality of acquired teaching material. Inclusion in a data set of the reference expert grades allows to estimate a degree of readiness trained and to determine a knowledge level each of them.

IV. Conclusion

Thus, in the paper it is shown that the modified expert evaluation method allows to carry out demonstrative selection of experts and allows to determine, as far as they are able to work as one team. The control chart method allows to estimate quality of training directly in its process and, hence, allows to correct it in time. Fast growth of IT sector and growing competition demand of the organizations of qualitative decisions more and more complicated problems in deadlines. The use of statistical quality monitoring in modern conditions is the most accessible and demonstrative way of evaluation of a level of software developers training.

References

[1] V.A. Silich, M.P. Silich, “The System Analysis and Research of Operations: Textbook”. Tomsk: TPU, 2000. P. 97.

[2] V.S. Anfilatov, A.A. Emelyanov, A.A. Kukushkin, “System Analysis in Management: Textbook”. Ed. A.A. Emelyanov. Moscow: Finance and Statistics, 2002. P. 368.

[3] G. Bengu, W. Swart, “A Computer-Aided, Total Quality Approach to Manufacturing Education in Engineering”. IEEE Transactions on Education. Vol. 39, P. 415-422, August 1996.

[4] “The Electronic Textbook on Statistics”. Moscow: StatSoft, Inc. 2001. http://www.statsoft.ru/home/textbook/default.htm.