Behavioral pattern recognition and knowledge extraction for decision-making in software project management

Arturo Peralta
Dept. of Information Technologies and Systems.
University of Castilla La Mancha.
Ciudad Real, Spain.
arturo.peralta@alu.uclm.es

Eduardo Conde
Dept. of Business Organization.
National University of Distance Education
Madrid, Spain.
jconde@ind.uned.es

Manuel García
Dept. of Manufacturing Engineering
National University of Distance Education
Madrid, Spain.
mggarcía@ind.uned.es

Francisco P. Romero
Dept. of Information Technologies and Systems.
University of Castilla La Mancha.
Ciudad Real, Spain.
francisco.romero@uclm.es

Short Paper. CSCI-ISPC

***Abstract*— The PSP (Personal Software Process) is a set of practices with the aim of improving the work of software engineers from the point of view of schedule management and personal productivity. Both managers and employees of software development companies can take advantage of these techniques with the aim of improve the efficiency of the company and its personal work. This paper presents several experiences in using the specified time record described by the Personal Software Process. The analysis carried out for both employees and managers provide a set of observations and recommendations concerning the type of work and schedule management made for each type of worker for various situations.**

Keywords— Project planning, Estimation tasks, Schedule Management, Behavioral Pattern Recognition

I. INTRODUCTION

Project requires gathering and processing an incessant stream of information and knowledge. Moreover, in the software development context, priorities and resources are generally not well defined and are changed constantly. In this context, one of the main problems facing both managers and employees is to make an efficient schedule management. On one hand, the employee, besides its scheduled supervised activities, have to specify their hours for autonomous work. On the other hand, the managers must estimate the necessary time for preparing meetings, reports, presentations, etc.

The problem of schedule management can be considered as a cause of stress, and thus, it can be placed within the field of psychology. Cognitive Psychology [1] provide us a framework to conceptualize the problem of stress and schedule management based on activity. In this context, the inability to enumerate and record data concerning the tasks done is a symptom of disorganization. For this reason the cognitive therapy recommends a methodical and systematic observation over some aspects of the problem. This examination process increases the control over one's own time and awareness of what we are doing.

Several techniques have been developed within the software engineering area with the aim of improving work organization. The use of these techniques can be extended to regular activities of any officer or employee with the same purpose. One of the most important tools is the Personal Software Process (PSP) [2]. The PSP provides a structured framework of forms, guidelines, and procedures for software development in order to help software engineers to be more productive and produce higher-quality software. Both PSP and TSP (Team Software Process) [3] (its extension to team management) are based on the collection of metrics with a high level of detail. Once the data has been collected, statistical and intelligent data analysis allow a realistic planning, tracking, prediction, and control of projects and software products.

For this purpose, there have been some experiments with students of Software Engineering o Artificial Intelligence courses. Some of the most important are those made in Mexico [4] and Thailand [5]. These studies analyze trends in controlled situations under a strong set of constraints. Also noteworthy, other works, for example [6] and [7], presents restricted experiments with different sets of software engineers are made by analyzing the impact of using PSP on their ability to estimate and plan their work, the quality of the software they develop and their productivity.

Results showed that the use of all levels of PSP is not applicable in a real-world software development environment because most developers abandon PSP practices after the first use if it is not mandatory. Therefore, only the PSP0 level (time-recording log) is applicable for improving productivity [8].

Using time records in this paper are described different experiences in order to improve the estimation and time management in business, both from the point of view of newly hired employees, those with several years of experience, and finally the manager responsible for overseeing the development of their respective projects.

978-1-5090-5510-4/16 $31.00 © 2016 IEEE
DOI 10.1109/CSCI.2016.167

The rest of the paper is organized as follows. In the following sections the experiences with employees in training period, with experienced employees' members of a development team for a particular project, and project management software are presented. For this study has been used time record books provided by a group of employees (different roles) of a real software development company, located in Ciudad Real (Spain). Finally, the paper closes with the conclusions that can be drawn from the study and future lines of work.

II. ANALYSIS OF WORKING TIME OF EMPLOYEES DURING TRAINING PERIOD.

In the case study of this work, the self-monitoring activities used were performed using a form of time based on the proposal of the Personal Software Process. It is included for each workday the actual time spent on each task calculated from the start time and end time, discounting interruptions.

Table 1 shows the fields used for the self-registration for the employee in training period. The field "Day" means the day's work; "Start" and "End" indicate the start and end of the session. "Disp" the time spent on breaks, o interruptions, in minutes. "Tm" indicates the uptime of the session (calculated from previous data). Finally appear the data associated with each task: "Tec" indicates the associated technology used for performing the task of training (Java, PHP, .Net, etc.), "Type" indicates the type of task performed (assistance of course, doing exercise, meeting for monitoring, etc.) and finally "Des" is the proper description of the task.

TABLE 1. SELF-REGISTRATION TABLE FOR EMPLOYEES IN TRAINING PERIOD

Day	Start	End	Disp	Tm	Tec	Type	Des

This experiment was conducted with programmers employed in training period, in which they are learning about the use of certain technologies that will be necessary for the development of their work within the company. These employees were hired by the company, and were formed over a period of three months in different technologies.

During this period, they were offered the possibility of making part of the training at home, by the self-study of the provided materials. Moreover, as introduction were offered a seminar about time management of two hours in which were explained the tools that they should use for planning their weekly work (eg Google Calendar) and the used system for registering their working time.

The activity consisted on deliver every Monday, and during three weeks, a schedule of the weekly planning training for all required technology and the time spend during the week for each learning task. Weekly schedules must be submitted the Monday of the week in question and the records the following Monday at the end of the week.

During the seminar they were stressed about the importance of recording the time they had interrupted during the training period. Also they were told we had to have working sessions and no travel time or rest. Also it was recommended to update the time recording the most frequently possible. Some relevant qualitative results of this experiment were the following:

• There was a not very high percentage of employees in training (22%) that gave the same plan for every week and the self-reports were very similar, containing inconsistencies as indicating the study of certain technologies that really had not been required. All these data were discarded.

• In the analysis of calendars radical change between the first and the second weekly calendar was found. One possible explanation is the effect of time recording after his first week of work.

Some quantitative results can be seen in Table 2. Note that the activity was conducted during the first weeks of training, which explains certain results.

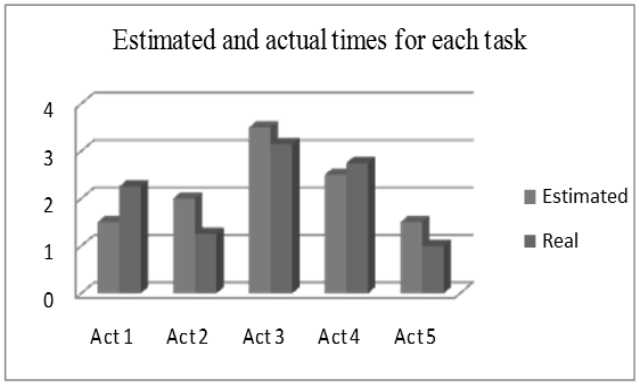
TABLE 2. QUANTITATIVE RESULTS FROM EXPERIENCE

Hours of daily study during the week	8 hours
Classroom activity VS non-contact activity	63% VS 37%
Daily disruptions	1 hour
Percentage of study during the weekend	3%

The study of the time spent for autonomous learning tasks, as Java programming language, can be seen in Figure 1.

Conclusions extracted as the first task, though simple, was hard to novice employees because it was the first task of training performed within the company. The same happened with activities which involved a longer research (eg, activity number 4).

Fig 1. Estimated and real times for each task



At the end of the activity the employees complete an anonymous survey results.

Among his responses can be highlighted the following aspects:

• Employees indicated that time recording helped them to be aware of their use of time and especially disruptions.

• The time to complete the self-registration is very short if it is made every day. If it is done at the end of the week it is heavy and slow, generating many mistakes.

• As proposed improvements they recommended that the process were completely anonymous and they could receive some feedback about their behavior.

III. TIME ANALYSIS OF WORK OF A SOFTWARE DEVELOPMENT MANAGER ANALYSIS OF WORKING TIME OF EMPLOYEES DURING THE DEVELOPMENT OF A SOFTWARE PROJECT

For this study we have used the data obtained by Personal Soft-ware Process of 5 employees (programmers) provided while participating in five Software projects. In addition, to be able comparative studies and get some conclusions, the 5 projects that were selected were with the same duration, estimated on 125 hours, more or less.

Table 3 shows the registration of time used by each employee during his work in a determinate software development project. As a distinctive feature it contains the identification of the phase of the life cycle at which the task (analysis, design, programming, documentation, etc.) and the associated task description.

TABLE 3. SELF-REGISTRATION TABLE FOR EMPLOYEES

Day	Start	End	Disp.	Tm	Phase	Des.

After a detailed study of the data recorded in these projects should be noted that the level of hours finally dedicated to the realization of each project was actually between 200 and 250 hours, something much higher than the 125 hours that was estimated. This is usual because of the inevitable "human factor" and the consequent error in the initial estimates and the number of unforeseen and difficulties during project development and unknown in an initial assessment.

In addition, it is also possible to observe that the human and technical characteristics of each employee are a key factor for realistic planning of the project. This is reflected in the percentages allocated by each type of employees to different phases. Studying these percentages we can distinguish two types of workers, some who have difficulty establishing contact with technology (spending more than a third of the time in these phases) and on the other hand those who have no problems with technology but are less comfortable with the completion of documentation or, in general, less technical tasks (spend more than 35% at work).

One recommendation that can be extracted from the analysis of the records is to base the performance of any software project development in long sessions, which is not always easy. In the cases studied the sessions have an average of 3 hours 15 minutes, which has made that more than 30% of sessions has been reached end.

By contrast, the great length of sessions makes them more prone to interruptions. Thus, it is observed that the percentage of sessions with interruption has been around 80%, with average of 14 minutes per disruption.

We can also see that most errors appear in the tasks performed with a larger number of disruptions.

As for the weekly distribution of work employees who do not attend training courses or participate in other work, they perform their duties during the days of the week, always on days like Monday and Thursday and only sometimes extend their working day during the weekend for the conclusion of a very specific activity.

By contrast, the employees who attend training to proposed courses by the company, or perform other additional activity, often they are forced to continue their business professionals tasks outside their working hours including weekends.

All these aspects are essential to make realistic planning before to the beginning and completion of software projects.

IV. TIME ANALYSIS OF WORK OF A SOFTWARE DEVELOPMENT MANAGER

The experience with a manager of software development projects has been extended longer than the study made with employees' with programmer role. In this case had been used time records from January to September 2015, including the following distribution on the number of projects aimed at programmers and the number of employees involved in them.

TABLE 4. PROJECTS AND WORKERS MANAGED

Period	Projects	Workers
01/01/2015-31/03/2015	5	16
01/04/2015-30/06/2015	6	21
01/07/2015-30/09/2015	3	15

The register used by the manager has a singularity; a description of the task using two fields "Type" (Management, Training, Meeting, Research) and description.

TABLE 5. SELF-REGISTRATION TABLE FOR THE MANAGER

Day	Start	End	Disp.	Tm	Type	Des.

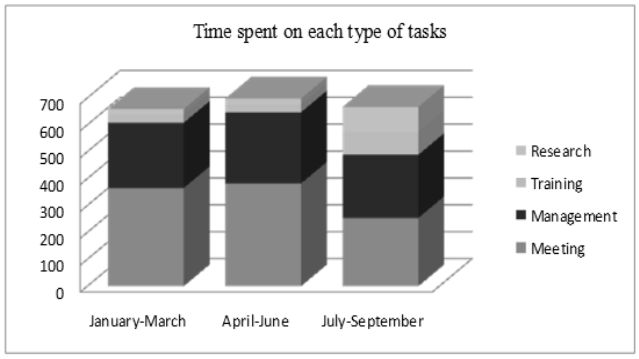
The results of aggregate data from these records is shown in Fig 2. The total working time in each quarters is constant.

The weekly working time for each of the quarters is 44 hours, more than 40 hours of a working day stipulated, considering also that this are real working hours, so we must keep in mind that in most cases to get 40 hours of effective work, workdays between 45 and 50 hours would be needed.

In turn, observing the distribution of tasks by type, it has been proven as the preparation of meetings and other management tasks, requiring almost the entire working time spent by management.

By contrast, training tasks, courses, or research, usually do not represent more than 15% of their working time

Fig 2. Time spent on each type of task



In addition, analyzing the data in greater detail we can extract useful conclusions for the organization of the manager's work.

For example, we can identify the influence of the dedicated time or meetings with other workers (in this case very high, around 50%) divided between contact and non-face meetings with tools based on internet use. In this case it is divided equally between contact and non-face meetings.

In addition, we can identify the increase in time spent for meetings during the months in which the manager has to manage a larger number of projects and people (first and third quarter). We consider this relationship could be important, because spending as high time for preparing meetings could have troubling implications for other activities (Research, Training and Management).

Another important point is to establish objectively which peaks of manager's work in order to be able to organize the rest of their tasks. In this case we have identified that the second fort-night of March and June as points of maximum dedication to management tasks, coinciding with the period in which the number of targeted projects was higher.

Finally, is studied the influence of some variables such as the existence of holidays or displacements in the manager's work.

V. CONCLUSIONS

In this paper has been presented certain experiences using the specified time record described by the Personal Software Process. Specifically, we have studied cases of both employees and without experience as the manager responsible for monitoring and project management. To do it, has been analyzed notebooks time records provided by these professionals belonging to a real software development company.

Following this study, we have completed a set of conclusions and recommendations concerning the type of work and time management made for each type of worker to different situations and time periods, which could eventually help to improve their own management.

As future work we propose building a web application that allow to employees to record their worked time, with a mechanism to provide feedback obtained through intelligent and automatic analysis of the provided data . Thus, in addition to facilitate monitoring and use of Personal Software Process on your first level, we will reduce the error rate in recording time, improving quality, and being able to be used with better guarantees for further process analysis and knowledge extraction.

REFERENCES

[1] Beck, J (2009). Terapia cognitiva: conceptos básicos y profundación. GEDISA

[2] Humphrey, W. (2005). PSP, A Self-Improvement Process for Software Engineers. Addison-Wesley.

[3] Humphrey, W. (2006): TSP, Leading a Development Team. Addison Wesley.

[4] Gómez O., Aguileta A., Gómez G., Aguilar R. (2014). Estudio del Proceso Software Personal (PSP) en un entorno académico. Revista Electrónica de Computación, Informática, Biomédica y Electronica (ReCIBE). Universidad de Medellín.

[5] Ramingwong, S., Ramingwong, K. (2012). Implementing a Personal Software Process (PSPSM) Course: A Case Study. Journal of Software Engineering and Applications. Scientific Research Publishing.

[6] Grazioli, F., Vallespir D., Pérez L., Moreno S. (2014). The Impact of the PSP on Software Quality: Eliminating the Learning Effect Threat through a Controlled Experiment. Advances in Software Engineering. Elsevier.

[7] Vallespir D., Grazioli F., Pérez L., Moreno S. (2013). Demonstrating the impact of the PSP on software quality and effort: eliminating the programming learning. TSP Symposium, Software Engineering Institute, Carnegie Mellon University USA.

[8] Valverde C., Grazioli F., Vallespir D. (2012). Un estudio de la calidad de los datos recolectados durante el uso del personal software process. Proceedings of the 9th Jornadas Iberoamericanas de Ingeniería de Software e Ingeniería del Conocimiento (JIISIC '12)