RELATIONSHIP BETWEEN TEAM SIZE AND EFFORT REQUIRED DEVELOPING AN INFORMATION SYSTEMS: AN EXPLORATORY STUDY

Nafisseh Heiat, College of Business, Montana State University-Billings, 1500 University Drive, Billings, MT 59101. 406 657-2224, nheiat@msubillings.edu

ABSTRACT

The objective of this paper is to investigate the relationship between teams, and summary work effort in software development based on data from the ISBSG repository. It includes a discussion of team dynamics and properties, a thorough examination of the ISBSG repository, and explanations of the models and statistics used to evaluate the data. Three sets of hypothesis are tested by applying regression analysis to the sample data. The results are then analyzed to determine if a relationship exists between summary work effort, and maximum team size.

Introduction

Over time numerous models and theories have been made in attempt to estimate the effort required to develop a project in Information Systems. Finishing projects late can be very costly and detrimental to a company. It is to a company's advantage to accurately forecast completion time. Many research studies have been conducted to identify factors that may affect effort. Some of these factors include, among others, models, language, platform, and team size. All of these factors have been proven to be related to effort in various studies. Team size is an interesting comparison because while optimum teams may exist, it is difficult to contribute effort to team size only. Teams are subjected to a variety of immeasurable factors. Thus while one can form a hypothesis, the weight given to it is difficult to assign given the dynamics of teams.

The ISBSG repository contains information on hundreds of software projects. It attempts to be a stable reflection of the IT Industry. The objective of the repository, and book published with it, is to provide insight on the factors mentioned above that affect effort. The data should help managers make better decision in the future, from knowledge gained today. Based on data from the ISBSG repository, the purpose of this paper is to explore the relationship that exists between summary work effort, and maximum team size. For this paper three sets of hypothesis are being tested.

1. **H0=Max Team Size is not an indicator of Summary Work Effort.**
   H1=Max Team Size is an indicator of Summary Work Effort.
2. **H0=Function Points are not an indicator of Summary Work Effort.**
   H1-Function Points are an indicator of Summary Work Effort.
3. **H0=Function Points +Max Team Size is not an indicator of Summary Work Effort.**
   H1=Function Points + Max Team Size is an indicator of Summary Work Effort.

Regression analysis will be applied to the data samples. The results will then be analyzed to determine if a relationship exists

Teams

In order to test the relationship between teams and summary work effort in software development, it is necessary to have an understanding of team dynamics. This information will aid in analyzing the final data. Teams have many variables. As a result it is difficult to use them for a stable variable measure. Results based on team size can never be fully reliable because there are so many factors that affect team dynamics. In general individuals are affected by four factors:

- Profile of their personalities

- Mental images resulting from their perceptual process
- Willingness and ability to overcome degrees of uncertainty
- Influence of their subconscious minds

While teams are considered optimal for tackling complicated problems they do have some weaknesses in the social-psychological area:

- Teams, or groups, can affect the performance of a highly motivated individual.
- When working with others there can be a loss of self-awareness.
- In a team, or group, there may be pressure to achieve consensus among members.
- Within a group individuals may make riskier decisions because it is harder to trace the outcome back to one person.
- When groups become too large, there is an increase of having social loafers.

These factors are important because they subconsciously affect team dynamics. To date there have been many studies seeking to prove that team size affects overall group performance. While many studies have been done, minimal literature has been published in regards to determining the optimum team size. This is in part due to all of the variables that affect teams. There are many factors that are involved in teams, which could directly or indirectly affect effort. In this study we had only access to the team size. However, it is clear that many factors can affect team effort in addition to team size.

ISBSG Data

The data for this project comes from the ISBSG repository. The dataset contains values for 789 projects. While the projects attempt to represent the IT industry, the ISBSG repository only uses projects, which utilize Function Point Analysis. The majority of the projects in the repository are less than 2000 function points. In addition, work effort must be recorded to be included in the repository. The data provided comes from over 20 countries and from a variety of industries. When published in 2000, all of the projects were less than 5 years old. The projects used three platforms; mainframe, midrange, and pc. The bulk of the projects, or 61.7%, came from a mainframe platform. There are four main languages included for each project; 2GL, 3GL, 4GL, and Application Generator. The projects could be of three types, new development, re-development, or enhancement. The reasons for including this information are that all of these factors have been statistically linked to summary work effort. In addition to this, a multitude of other subjects was gathered including programming language, application type, and development platform among others. However, they are not critical to the understanding of this project. The three areas of interest for this research are summary work effort, maximum team size, and function points. Function Points is one method used to determine effort in system development. The purpose of FP is to measure the functional requirements of a proposed system. The first step to finding the FP Value involves identifying the type of each required function in terms of inputs, outputs, inquiries, internal files, and external interface files. Once the functions have been identified for each subject, it is multiplied by a complexity weight. By summing the FP's and multiplying it by the technical complexity factor or TCF, the required effort is calculated. FP is a popular method because it is not dependent on any programming language or development tool. The ISBG definitions for these titles were as follows:

- **Function Points**-The adjusted function point count number.
- **Maximum Team Size**- The maximum number of people that worked at any time on the project
- **Summary Work Effort**- Provides the total effort in hours recorded against the project by the development organization.

Once the data was obtained, it needed to be cleansed. The maximum team size column contained many null values. The function point column also contained null values, though not as many as maximum

team size. The cleansed data contained 500 samples. 450 samples were used in estimation and 50 samples used for out of sample testing of the models.

Methodology

Three regression analyses were conducted to determine if a relationship existed between team size and project elapsed time. The first two used simple linear regression. The last one used multivariate equation. The results necessary for discussion can be seen in the following equations:

Effort=-289+686*Max Team Size (1)

T = -0.638 **R-SQ** = 28.3% **MAPE** = 10%

Effort=1080+803*Function Points (2)

T = 2.17 **R-SQ** = 36.4% **MAPE** = 33%

Effort=-2326+6.59(FP)+512(Max Team Size) (3)

T = -4.40 (FP) and 0.41 (Max Team Size) **R-SQ** = 51% **MAPE** = 17%

Discussion

For equation (1) the R squared indicates that only 28.3% of the variation in summary work effort can be explained by maximum team size. A weak relationship between the variables can also be seen by looking at the graph in appendix.

The result for the second linear regression is better than the first, but it is still skewed by the large number of outliers. In this analysis T calculated was equal to 2.17. Therefore, we can accept the statistical significance of Function Points in determining the required effort. The R squared is higher, with a value of 36.4%. However, it is still very low. The plot graph looks similar to the previous graph. The data is all clustered in the lower left-hand corner. There are also many outliers visible. (See Appendix 1).

The result for this equation is better than the other results. The T value of –4.47 for function points again indicates that its' significance, while the T value of 0.41 indicates that the team size is not statistically significant. The R squared for has a value of 51%. This indicates that 51% of the variance for summary work effort is contributed to Function Points and team size.

Conclusion & Limitations

In conclusion, It seems from this analysis that the team size is not a significant contributing factor to estimating the required effort. However, it has been reported in other studies that team size affects effort. There may be several reasons for this discrepancy. First, outliers have a strong effect on regression analysis. The high number of outliers in this project could in part have contributed to the erratic nature of this data. This research should be conducted again, but without the outliers. Perhaps then, the final results will be better.

Secondly, the ISBSG repository did find that max team size did affect summary work effort. However, they did not find this until they removed several other factors. The ISBSG book used a different method know as "Residual Analysis. By comparing the methodology used in the ISBSG research and the methodology used here it becomes apparent that there are some differences. Based on the data and methodology used in this paper, there was not statistically significant relation ship found between team size and summary work effort. It is clear that this project needs to be reevaluated, and redone to find more accurate results. Though the results were not as expected, the overall process was still insightful to understanding the factors that may affect effort in information systems development.

Appendix 1 and references are available upon request.