

EFFORT ESTIMATION BASED ON STORY POINTS IN AGILE APPROACHES: A SYSTEMATIC LITERATURE REVIEW

Ch. Prasada Rao¹, Dr.P.Siva Kumar², Dr. Rama Sree. S³ and S N S V S C. Ramesh⁴

Abstract- Software effort estimation is an essential activity in the process of software development. Estimates are the basis for bidding, planning, budgeting and scheduling for the software development. Most of the projects failed due to over run, over budget or both. According to standard surveys, 30% to 40% of software projects failed due to inaccuracy of software effort predictions. There has been a lot of research done on effort estimation in conventional process models for the last three to four decades; however these models have not given accurate predictions. Effort predictions can be done with different parameters like Lines of Code, Function Point Analysis, Object Points, Use Case Points etc. None of them have given accurate effort estimation for acquiring high success rate of software projects in traditional process models. To improve the accuracy and efficiency of effort estimation, in recent days an Agile Process Model is introduced. 70% of the software industries have adapted the Agile model where effort is estimated using Story Points. The objective of this study is to provide a detailed overview of the effort estimation techniques in Agile Software Development based on Story Points.

Keywords – Agile Methodologies, SCRUM, Story Points, Effort Estimation

I. INTRODUCTION

In software industries, the main goal of software management is to develop a successful software project. Software management focuses on four P's for software development: People, Process, Project and Product. The first P which indicates People is who gets the job done. These people are considered based on skill set, ability and reliability for developing the software successfully. Process is an important framework to organize the different activities for developing the software. It also provides structure and context for estimating the required effort. Third P indicates Project, which generally deals with different people with different roles and responsibilities where they are going to achieve the common goal. Project is the combination of people and process to produce the desired product has shown in Fig.1. Product is termed as Artifact in the software development. An Artifact is delivered after the completion of each phase in software development life cycle (SDLC). The main goal of SDLC in software engineering is to deliver a high quality product within time and budget. Different process models for developing the software product like Waterfall model, Spiral Model, Incremental and Prototype Model etc exists in the literature but each model has its own pros and cons. There has been a lot of research been done on effort estimation in traditional process models for the last three to four decades. All the models, techniques, planning, activities, tools etc used in traditional process models have not given accurate predictions. To overcome the problem of inaccurate prediction, the top industrial experts have identified & framed a new process model called Agile Process Model.

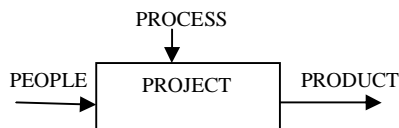


Figure.1 FOUR P's

¹ K L University Vaddeswaran, Guntur Dist

² Dept. of CSE K L University Vaddeswaran, Guntur Dist

³ Dept of CSE Aditya Engg College Surampalem, E.GDist

⁴ Dept. of CSE Aditya College of Engg & Tech Surampalem, E.GDist

Now-a-days, Agile software development is very popular and it enhances the success rate of software projects. Since its inception in 2001, most of the software industries have adapted the Agile approach and it substituted the conventional methods of software development [3]. Agile approaches include Extreme Programming (XP), Scrum, Dynamic System Development Method (DSDM), Feature Driven Development (FDD) and Adaptive Software Development (ASD). Most of the researchers and practitioners have considered XP and Scrum for their studies and implementation. Agile process model is both Incremental and Iterative where completion of tasks is very faster and quicker. The advantages of agile process model is to deliver the products in short cycles, simple design, less documentation, continuous customer involvement to give better feedback, continuous rigorous testing, refining, refactoring, robust coding principles, contiguous integration etc [2]. Using, either Traditional Approaches or Agile Approaches, the main goal is to improve the success rate of software projects with accurate software effort predictions. The literature shows that software effort estimation using agile process gives accurate results than the traditional process models. A few of the common effort estimation techniques that can be used in Agile Software Development are Expert Opinion, Analogy, Disaggregation and Planning Poker [1]. Planning Poker, introduced by Greening [1], is a technique for effort estimation in Agile which combines the elements of Expert Opinion, Analogy and Disaggregation.

Functional points, Use case points, Object points, LOC and Story points etc are the key parameters for effort prediction in Agile process. But story points have been proved to provide more accurate results than the others specified. In Agile effort estimation, requirements are represented using User Stories and these are estimated using Story Points [3]. Story Points are a predominant means of sizing in Agile methods. Each iteration is termed as "Sprint" in Agile Methodologies. For each Sprint there may be "n" number of story points. Usually one story point takes one ideal working day for its development. For example if there are 6 story points in one sprint then it may take 6 to 7 working day to complete. The total number of story points that a team can construct and deploy in a single sprint for a project is denoted as "Team Velocity" or "Story Points per Sprint" [3]. The total story points and the team velocity are the main input factors to forecast the software development effort in agile methodology. The main goal of this study is to understand and provide an overview of different technologies used for estimating effort in agile approaches based on story points.

II. SYSTEMATIC LITERATURE REVIEW

A. Research Questions

In this study, the following research questions have been framed to conduct the SLR:

- RQ1: What are the different techniques that have been used for effort estimation in agile approaches based on story points?
 - o RQ1a: What are the performance metrics to be evaluated using these techniques?
 - o RQ1b: What accuracy level has been achieved by these techniques?
- RQ2: What are the different datasets considered to predict the software effort in agile process based on story points?

III. RESULTS

This section provides the results captured for the research questions specified in Section 2.1. The primary investigation of this SLR is to analyze different techniques for estimating the effort in agile approaches based on story points. In the process of searching for the papers from different organizations like IEEE, Science Direct, ACM, Springer Link, etc. 12 papers are identified in the software effort estimation based on story points in agile software development. This section presents the systematic review results by referring these 12 papers.

Alaa et al. [3], proposed an algorithmic model considering Technical Factors (TF) and Environmental Factors (EF) to calculate Adjusted Story Points (ASP) and implemented the model using Multi-Layer Neural Network. In this four industry projects are considered and the accuracy metrics like MER, Size estimation are evaluated. The results show that the MER decreased from 34% to 12% and size estimation error decreased from 28% to 5.9%

Aditi Panda et al. [4], used different types of neural networks like General Regression Neural Network (GRNN), Probabilistic Neural Network (PNN), GMDH & Cascade Correlation Neural Network (CCNN) to predict the estimation. The different performance metrics like MSE, R2, MMRE and PRED are evaluated using 21 projects dataset. It was stated that the CCNN outperformed the other networks.

Satapathy et al. [5], evaluated and compared MMRE & PRED with various Support Vector Regression (SVR) Kernel methods for effort estimation. They have taken 21 projects dataset and evaluated them using MMRE and PRED. SVR-RBF outperformed the other models

Morakot et al. [6], have proposed a deep learning model which is a novel combination of two powerful techniques: Long Short-Term Memory (LSTM) and Recurrent Highway Network. The new model was evaluated and compared the accuracy metrics like MAE & SA using Long-Deep Recurrent Network (LD-RNN), Long Short-Term Memory (LSTM), Bag-Of-Word (BOW) and Random Forest Techniques with 16 projects dataset which includes 23,313 issues. For the performance measure MAE, the model with lower value is the better and for SA, the model with higher value is the better. Using wilcoxon test analysis, LD-RNN has given the better result with respect to MAE & SA.

Christophe et al. [7], estimated the effort with Story Points and COSMIC FP. COSMIC FP provided better MMRE & R2 than Planning Poker estimations. They have used 24 industrial projects dataset. To the best of our knowledge, estimating the effort based on FP is at entire project level but based on story points is at issue level. He only stated that the story points are widely adapted and used in the agile approaches. May be for these projects, COSMIC FP have given better accuracy.

Kayhan et al. [8], to predict the accurate effort provided an automated tool called as “Auto Estimate” which integrates the features of J48 and manual Planning Poker techniques. J48 is an open source Java tool in WEKA to produce decision trees. The required data for the work like story points, estimated data, actual effort etc is collected from 10 different teams of IBM Rational Team Concert. They evaluated MMRE with different supervised learning techniques and planning poker technique. J48 and Planning Poker together outperformed the others.

Suman et al. [9], used story points for not only predicting development effort, size and cost but also for predicting maintenance effort. He calculated the size of maintenance and duration of the maintenance using Adjusted Story Points (ASP) by an algorithmic model.

Rashmi et al. [10], proposed an algorithmic model to estimate the effort in agile environments based on story points. They have considered various resistance factors, project and people related factors. The model computes the Total Estimated Effort (TEE) and Total Estimated Cost (TEC) using Unadjusted Value of story points (UV), Velocity Factor (VF) and Complexity Factors.

Hind et al. [11], introduced an Algorithmic model to adjust the story points in Scrum environments using three adjustment factors: Priority Factor (PF), Story Size Factor (SF), Complexity Factor (CF). Velocity is also improved by using two factors: Friction Factors (FR) and Dynamic Factors (DF). They designed a model by considering all these factors to calculate the total time needed to complete the Scrum project.

Yves et al. [12], proposed a unified-model to elicitate and represent the User Stories in agile process. User Story template may help researchers and practitioners to predict the accurate development size, cost and maintenance effort.

Porru et al. [13], have given a broad way of thinking in effort estimation based on story points in agile environments. They evaluated the performance metric (MMRE) based on 8 open source projects using Support Vector Machine (SVM), Naïve Bayes (NB), K-Nearest Neighbor (KNN) and Decision Trees (DT). SVM outperformed than others.

Tanveer et al. [14], suggested to implement a tool incorporating developer’s knowledge, experience, complexity and impact of changes. They also advised to consider expert’s knowledge and explicit consideration of cost drivers. They developed a case study based on the report given by German Multinational Software Corporations, SAP SE.

The essence of the above research publications is tabulated for quick reference in Table 1.

Table.1 Search Results

S.No	Authors	Methodology/Techniques Used	Accuracy Metrics evaluated & achieved	Dataset Used
1	Alaa et al	Multi Layer Neural Network	MAE=12% & Size Estimation=5.9%	4 industry projects
2	Aditi Panda et al	GRNN, PNN, GMDA & CCNN	MSE=0.0059, R ² =0.9303, MMRE=0.1486 and PRED=94.76	21 Project dataset
3	Satapathy et al	Support Vector Regression (SVR) Kernel methods: SVR Linear Kernel,	MMRE=0.0747 & PRED= 95.90	21 Project dataset

		Polynomial, RBF, Sigmoid		
4	Morakot et al	Deep Learning Model: LD-RNN, LSTM, BOW	MAE= 2.09 & SA=52.66	16 project dataset & 23,313 issues
5	Kayhan et al	Auto Estimate: J48+Planning Poker	MMRE=91.75	10 project data of IBM
6	Rashmi et al	Enhanced Algorithmic Model	Calculated Time & Cost using ASP	Sample Data
7	Hind et al	Enhanced Algorithmic Model	Calculated Time & Cost using ASP	Sample Data
8	Porru et al	SVM, NB, KNN and DT	MMRE is ranging from 0.16 to 0.61	8 open source projects

IV. CONCLUSION

The objective of this study is to perform a systematic literature review on effort predictions based on story points in agile approaches. Agile methodologies are very popular and can deliver the software projects quickly and fastly. In Agile approaches there has been a variety of estimation techniques applied on story points and velocity for predicting effort. Based on the systematic study, the techniques are ranging from algorithmic models to soft computing techniques. Establishing and enhancing an algorithmic model is also very essential for estimating the effort. It is observed that, now-a-days most of the researchers focused on machine learning techniques to acquire accurate effort estimations. There is a scope to generate a novel model by integrating expert judgments with machine learning techniques for effort estimation. Though FP, LOC, UCP, Story Points etc are the popular size oriented metrics, the Story Points are widely recognized and used in the agile environments. The datasets (ISBSG, Promise, NASA, UCEEDB etc) on FP, LOC, Object Points and UCP are available publicly but there is no dataset available on story points publicly. Extracting the story points from various industries, academics, organizations, cross-companies etc is very essential. Few authors provided Story point datasets related to industrial and academic environment. A systematic literature survey is carried out taking 12 papers based on story points used in effort estimation in agile environments. The results of the review are also presented in detail. Based on the results acquired from this Systematic Literature Review, it is recommended to use sound engineering techniques and tools to estimate the effort accurately.

REFERENCES

- [1] Mohammed Usman., Emilia.M, Francila.W, Ricardo.B, "Effort Estimation in Agile Software Development:A Systematic Literature Review", ACM,2014.
- [2] Andreas.S, Martin .K, Reiner.D, "Effort estimation for Agile Software Development Projects", Proceedings Fifth Software Measurement European Forum, Melan,2008.
- [3] Alaa-El-deen Hamouda , "Using Agile story points as Estimation Techniques in CMMI Organization" , Agile Conference 2014.(p1).
- [4] Aditi Panda , Satapathy S.M , Santanu Kumar .R , "Empirical Validation of Neural Network Models for Agile" , Science Direct,Procedia Computer science 2015.
- [5] Satapathy S.M , Aditi Panda , Santanu Kumar .R , "Story Point Approach based Agile Software Effort Estimation using various SVR Kernel methods".
- [6] Morakot Choetkiertikul, Hoa Khanh Dam, Truyen Trany, Trang Phamy, Aditya Ghose and Tim Menzies , "A Deep Learning Model for Estimating Story Points".,[CS.SE] 6 sep 2016.
- [7] Christophe C,Alain Abran,Rachida .D, "Effort Estimation with Story Points and COSMIC Function Point- An Industry Case Study", Software Measurement News, Vol .21, No.1,2016. Pages 25-36.
- [8] Kayhan..M, Alhal .V.S, Ramanadhan J, Raju Ramnath , "Cost-Effective Supervised Learning Models for Software Effort Estimation in Agile Environments", IEEE 40th Annual Computer Software and Applications Conference,2016.
- [9] Jitender .C, Dr.Ugrasen Suman, "Story Point based Effort Estimation Model for Software Maintenance", Science Direct,Procedia Technology for (2012) 761-765.
- [10] Rashmi.P, Naresh .C, "Estimation in Agile Environment using Resistance Factor",IEEE 2014.
- [11] Hindh.Z, Mohammed..A, "Adjusting Story Point Calculation in Scrum Effort and Time Estimation", IEEE 2015.
- [12] Yves.W, Samedi.H, Manuel.K, Isabelle.M, Stephen.P, "Billing a Rationale Diagram for Evaluating User Story Search" .,IEEE 2016.
- [13] Simon.P,Alessandro.M,Serge.D, Michele.M,Roberto.T,"Estimating Story Points from Issue Reports",ACM,2016.
- [14] Binish .T, Liliانا.G, Martin.U.L.F, "Understanding and improving Effort Estimation in Agile Software Development,-An Industry Case Study",ACM,2016.
- [15] Kitchenham.B, Charters.S, "Guidelines for Performing Systematic Literature Reviews in Software Engineering"..version 2.3, Technical Report, Keele University and University of Durham, 2007.