VOL. 11, NO. 21, NOVEMBER 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

A DUO FOLD SYSTEMATIC TOURNAMENT INVESTIGATION ON TEAM FORMATION USING LONGITUDINAL AND QUANTITATIVE PREDICTIVE PRINCIPLES

Krishankumar R. and Ravichandran. K. S.
School of Computing, SASTRA University, Thanjavur, India
E-Mail: krishankumar@sastra.ac.in

ABSTRACT

In software development lifecycle, team formation is playing an indispensable role in developing software application. Standard and ad-hoc principles are available for team formations. Newly adopted key factors are incorporated nowadays for the optimal group selection or team formation. Past records, innovative ideologies and relevant experiences in the field brings real intuition of a person which is essential to form a good team to achieve the ultimate goal of a project or a task set completion. In this article a systematic investigation on team formation is carried out to identify the current region of interest by tapping the unfilled gaps in the team formation agenda and to understand the effects of teams in product development scenarios. A Selective Search (SS) mechanism is adopted to collect 100 papers for survey. A systematic investigation on teams is carried out using cohort study and regression approaches for a tournament based analysis. The validation of the model is done using test and measures from statistical themes. The intuitive inference that is obtained from the analysis is that, team formation has made an exhaustive effort on qualitative modeling but a subtle contribution towards quantitative model development. Also the concept of optimizing the quantified parameters adhering to teams using traditional or non-traditional optimization techniques is discussed only on a narrow lane. It is concluded that the current hot spots for effective team formation lies in the design of an optimized quantitative model. It also infers from the analysis that there is an effective need for both the cohort study and regression mechanism to understand the missing region in team formation by setting up a theoretic framework and wrapping it up using a mathematical principle.

Keywords: team formation, multi linear regression model, least square estimate, two tier architecture, selective search mechanism.

1. INTRODUCTION

In the recent years, software development is reaching its highest peak in the research arena. Team coordination, team formation plays a vital role in building an effective software application. Technically, team formation is one which literally brings up the product to the market in time. Time factor is playing an important role in software development. Work break down structure (WBS) and Team formation (TF) are the bull eyes for the software development. Among the above said, the latter nowadays gears up well in the research academy. Business management people know the organizational behaviour of the employees and their skill set. Knowledge and wisdom give more impact in forming a team. Team members should not be disturbed at any point of time, so as to finish the project within stipulated time. People with a higher rate of team spirit and team patriotism tend to form an effective team. Groups and teams are two different terms governing companies. Teams work with a common objective in mind and groups work with individual objectives.

The idea of team formation relies completely on the Tuckman Model (Adair, 2004)namely i) Forming - boundaries and goals are to be framed well in advance. ii) Storming - ideas and approaches should be in a innovative manner. iii) Norming - rules and regulations are formulated in black and white with a rigid manner. iv) Performing- growing and maturity of the team and the products will improve gradually. The above said are the four key factors for any good team formation. Inevitable seven steps are also there for long run stability towards the team formation. Cohesiveness - functional activities within the module, Communication - between employees and employer-employees, Group think - every member in the team should agree to work with a single agenda, Homogeneity - every piece of work should be in the same order without any deviation, Role identity- it is an important feature that identifies the skill set of an employ and it tries to assign the right job to the right person, Stability - workers confidence and their ability towards the work gives high success to the project, Team size - the size of the team should be so concise and effective to achieve the task. Other factors also have greater dominance. Some of the daily activities bring up the sheerness towards the work .Kickoff - some initial meetings should be conducted in the day starting, Team agreement - represents effectiveness and efficiency, Delivery process - models like agile ,waterfall, incremental and spiral, Conflict - it should be avoided between the co-workers, Personality assessment - reward them in the right sense, Team building events - extra-curricular activities will make people work together, Social events - to sense the outside world to generate more knowledge, Celebration - At end of the project, some small parties are arranged to encourage the worker and make them do their jobs in a powerful and also in a stronger manner. The above said quality attributes put a greater effort in team formation. Two decades ago, the approach used for team formation or team building varied widely from managers to managers, from organizations to organizations and from leaders to leaders. The computing experts wanted to share their hands in helping

organizations to form an effective team. Efforts then emerged using computing and statistical techniques to optimize the concept of team building and team forming. Several approaches came into existence from then on. These approaches or strategies drove members into effective teams. In this article efforts are made to survey such strategies for team formation using a novel survey mechanism based on quantitative predictive principle and are segregated majorly into three categories namely Models, Methods, Concepts (MMC), Tools, Software, Algorithms (TSA) and Optimize. The concept of team is more prevalent not just in software organizations but in each and every place where there is involvement of both man and machine. Team work is a key concept of knowledge discovery and military zones. When teams are introduced there is migration from individual device centric model to group device centric. Key elements are the 44 drivers(Brennen, Strong, Ryder, Blendell, & Molloy, 2007; Ghobadi, 2015). A team is said to be effective if its Information technology (IT), interaction and performance factor is well close to the optimal threshold. The relationship between IT, organizational structure, cross interaction and performance is discussed (C. J. Chen, 2007; Isenberg & Carpendale, 2007)For teams the most preferred network structure is direct link but it is not feasible in the long run. To attend the global problems use synchronous and asynchronous ICTs (Chan, Jiang, & Klein, 2008; Fruchter, Swaminathan, Matsumura, & Ohsawa, 2008). The concept of team forming is basically an NP-Hard problem and so no traditional approach exist for this problem to each optimality. So many non-traditional techniques, tools and algorithms have helped in forming a better team (Alavi & Dillenbourg, 2012; W.-N. Chen & Zhang, 2012; Doherty & O'Riordan, 2009; Park, Pattipati, An, & Kleinman, 2012).

The remainder of the section covers Literature Survey and its Analysis in section 2, followed by An Innovative Investigation of Teams using Multi Linear Regression Model in section 3, A Comparative Study on Team Formation in Section 4, followed by Results and Discussion in Section 5 and Conclusion in Section 6.

2. LITERATURE SURVEY AND ITS ANALYSIS

2.1 Cohort study: a starter mode

The term TEAM exemplifies Together Everyone Accomplish More. As the acronym is spelled it proves to be a rudimentary concept of any software company. The Selective Search (SS) scheme is followed rather than an exhaustive scheme as SS is feasible. Around 100 articles from various esteemed journals are being surveyed to obtain the current region of focus for the problem in hand. The survey is focused on tapping the gap in the literature and finding efforts to fill the same. As the initial step the articles for review are being gathered(Kitchenham & Charters, 2007). This is done based on a twofold approach with the first fold contributing to the simple Google Scholar and Scopus search and the second fold contributes to a more specific journal search. Around 1332 articles were initially gathered. These article were scrutinized based on their relevancy to the domain of interest. The relevancy was checked with respect to both domain aspect and method aspect. Keywords were gathered from experts for the same. Some sample keywords are team, coordination, group, cooperation, interaction etc for domain and fuzzy logics, neural network, genetic algorithm, aggregation operators, ant colony optimization, particle swarm optimization, adaptive neuro fuzzy inference system etc for methods respectively. A SS mechanism is then followed for filtering articles pertaining to these keywords. Additionally factors like citations, article scope, author proficiency in the current domain etc were also considered. Based on these factors 100 articles were selected from a pool of 1332 articles for the investigation. The Table-1 depicts the Longitudinal Survey on Team Formation from 2007-2015.

Table-1. Longitudinal survey on team formation from 2007-2015.

Year of publication	Number of articles	Techniques adopted	Number of techniques
2007	11	Self assignment, Empirical model, theoretic framework, principal agent theory, Regression analysis, qualitative inductive analysis, device centric model, action research method, shared analysis information visualization	9
2008	10	Optimization approach, empirical model, KB model. Game theory, analytical framework, Process assessment model, pair programming, direct link model, Two fold approach model, theoretic model, Influence Diffusion model	11
2009	8	Analysis model, agile methods, declarative and procedural memory, XP method, SGD model, research model, Genetic programming	7
2010	7	Dynamic capability theory, agile methods, empirical study, visualization and external representation, VTManager, research model, trust and goal congruence, pair programming, meta analysis, Big Five, multilevel model	11
2011	11	Portfolio design, DSIMDP, Quality content analysis, Grounded Theory, quality based model, socio-material bricolage, formal model, Stochastic automata network, common ground building, conceptual model, Re-representation, max-min algorithm, Factor Graph, DPOMDPs,	13
2012	20	Memetic algorithm, GA and Simulated Annealing, Grounded theory, empirical model, global teaming, theoretic model, competence model, work place model, agent based negotiation model, quantitative model, agent based model, team work benefit aware model, reference andassessment model, structural equation and partial least square, Lantern tool, distributed algorithm	16
2014	7	Cloud based tool- ARCA, Quantitative model, Two stage clustering, resource estimation method, conceptual model, operational replication, Five factor model, replicated quasi experiment	8
2015	13	Greedy, Hill Climb, Random search, Brute force, uninorm aggregate operator, expert weighting method, analytical network process model, virtual field, behaviour blending, situation awareness, event interest model, math model, meta model, coalition model, empirical model, Big5, perception, preference, lean&agile, theoretic framework	20

In the lime light of the Table-1 it is significant that during the year 2012 major contribution on team formation is being made followed by 2013 and 2015 respectively. To validate the above inferences some instances from politics, E-communication, IT sectors, agile techniques, optimization are depicted (Bella et al., 2013; Dayan, Elbanna, & Di Benedetto, 2012; Gider, Likar, Kern, & Miklavcic, 2012; Stryker, Santoro, & Farris, 2012). Unfolding of community is key aspect in community network analysis. Node community is found rather than the whole graph community using greedy optimization. Ensemble methods like ensemble ranking and clustering and naive combine and rank is used to select local modularity function. Results show that ensemble ranking approach proves better in finding ego centred communities (Janz & Prasarnphanich, 2009; Kanawati, 2015). Optimization is one method for improving team coordination. The composition approach with heuristic search is used. Methods like Genetic Algorithm (GA), Hill Climbing (HC) etc are compared with experts' advice using three project instances from an oil company in Brazil. Results prove that GA is best for larger cases and other methods are competitive for smaller cases with a limit of data confidentiality. It also provides

collaboration management solution (COMPOOTIM) that provides method to introduce and manage collaboration (Magdaleno, de Oliveira Barros, Werner, de Araujo, and Batista, 2014).

The Table-2 shows the Strategy Based Classification of 100 Articles. As a part of knowledge that is mined from the Table-2, significant attention is given to Model, Method or Concept (MMC) section followed by Tools Software or Algorithms (TSA) and Optimize sections. To justify the inference some instances are highlighted. Case study and qualitative research mechanism were adopted to examine developers' interaction data for free open source software project; organizational pattern to software projects; precursors incorporations and teams behaviour in the organization.(Crowston, Howison, Masango, and Eseryel, 2007; Gonçalves, Ferreira, Gonçalves, Putnik, & Cruz-Cunha, 2014; Iorio & Taylor, 2014; Savelsbergh, Poell, & van der Heijden, 2015; Wale-Kolade, 2015). Traditional process activities are different from modern activities. To analyze the dynamics an empirical model, three empirical axes are taken and tested over five teams of six students each involved in three projects. The results show that there is high variability of actual process and assessment of it is weaker in nature(Germain & Robillard, 2008). Global Software Development (GSD) teams are tough to create (NP-Hard) and manage. A new method called virtual team (VT) manager is developed to manage virtual teams (Guzmán, Ramos, Seco, & Esteban, 2010).

Table-2. Strategy based classification of 100 articles.

Strategies	Number of articles
MMC	83
OPTIMIZE	7
TSA	10

2.2 Role of model method and concept (MMC) in team formation

The classical proverb states that Unity is Strength. All software companies adhere to this proverb. The teams are thus formed to attain prosperity. In this paper efforts are made to understand the region of research in such effective team building. Also contribution is made to that region to make teams more effective and efficient for the software companies. The survey has paved way for three major sections of discussion as depicted in a nut shell in Table 2. The elaboration of Table 2 is the key aspect of this section. The three major categories are the MMC, TSA and Optimize. The third column of Table 1 is the overall classification that signifies that, in the current year (2015) around 20 versatile strategies have been performed for an effective team formation. This seeks for the emergence of a fine tuned research in this area of study which is explicit from the inference gained. The problem of team formation is thereby a demanding research topic that is expecting greater efficacy in the next decade. The Figure-1 depicts the Scatter Plot for Team Formation Survey which is a 3-D amalgamation of year of publication, number of articles published and techniques adopted each year.

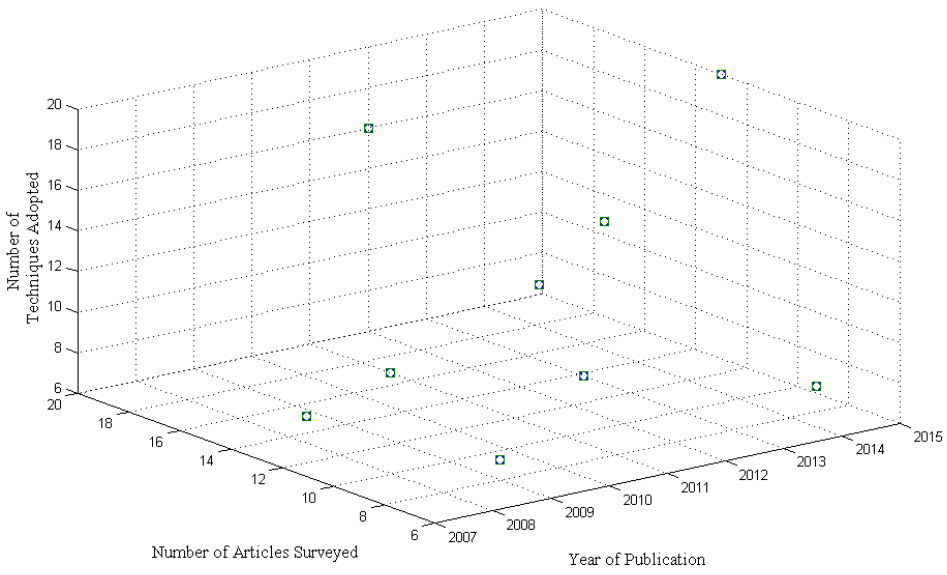

Figure-1. Scatter plot for team formation survey.

The major amount of contribution for team building or for team formation is made in this section. The Table-2 signifies the above statement. GSD is a big problem for information system research and practice of face to face communications. New knowledge based model and theoretic framework is set for this purpose to

achieve efficacy (Fruchter & Courtier, 2011; Giuffrida & Dittrich, 2015; Johri, 2011; Kotlarsky, van Fenema, & Willcocks, 2008; Lee, Espinosa, & Delone, 2013). Maintaining GSD teams' performance in a dynamically changing environment is tough. To achieve this Lean and Agile environment is used with thematic case study by interviewing 16 practitioners from five different organizations and obtained 33 performance factors. These are improved by soft factors like team identity, spirit etc that revolve around social and behavioural sciences (Assudani, 2011; Brookes, Morton, Grossman, Joesbury, & Varnes, 2007; Fagerholm et al., 2015; Richardson, Casey, McCaffery, Burton, & Beecham, 2012; Tuffley, 2012). Social relationship among the global organizations is important. Traditional methods don't concentrate on this aspect. So a new Analytic Network Process Model is proposed to integrate social with analytic networks and it is tested for water and food security research in Mekong River Basin. Results prove that companies can prioritize network goals and increase their impact and reach(Grady, He, & Peeta, 2015; Nguyen-Duc, Cruzes, and Conradi, 2014). Teams abide by different policies and factors that are framed as a part of organizations' protocols which must be embedded in students project works. Teams that do software projects are viewed as a set of collective action problems with main grip on what policies to be enacted for the teams' success. The institutional analysis is used to set theoretical principles and analytical framework used in game theory is also used over a case study that shows how to apply framework, understand principles and an example resembling real time scenario (André, Baldoquín, & Acuña, 2011; Colomo-Palacios, Casado-Lumbreras, Soto-Acosta, García-Peñalvo, & Tovar-Caro, 2012; Dubielewicz & Hnatkowska, 2008; Tenenberg, 2008). Applications involving higher security concerns and uncertainty seek a team effort. Architecture of software is a key concept. Analysis model called the Architecture Requirements Engineering Error and Accuracy (AREA-TEAM) is incorporated into banking system and results prove to be beneficial and effective(Gaubatz, Lytra, & Zdun, 2015; Georgiev, Kabamba, & Tilbury, 2008; Smari, Clemente, & Lalande, 2014; Sulayman, Riaz, Naqvi, & Aman, 2009). Creating coordination in uncertain environments is tough. A new Dec-Sparse Interaction Markov Decision Process (SIMDP) model is devised to promote local interaction in multi agents and it is compared with modified MDP and Dec- MDPs. A solution with error bounds are set for this model. A reinforcement learning algorithm is designed to allow multi agents learn when to coordinate and to act individually and is tested for navigation scenarios (Melo & Veloso, 2012; Sun, Lin, & Xu, 2015)

Teams have problem learning things in coordination. In teams new software methodology can cause failures as the teams may lack learning. Agile methods like single and double loop learning is used for learning and also triple loop learning is used to test ineffectiveness in teams and how power factors like power inherent in desire for conformity and cohesion affects learning and results prove that desire for cohesion affects team learning(Melo & Veloso, 2012; Sun et al., 2015; Toffolon & Dakhli, 2009).Agile practices in companies promote intra team knowledge sharing. A new conceptual model is devised and ground theory is applied for data from four Brazilian organizations and an expert advice is taken for intra team knowledge sharing that promotes organization and seeks its cope with the enterprise agility aspects(Santos, Goldman, & de Souza, 2014; Tessem, 2014). In software companies considering team as a single memory unit is a key aspect. Software development needs team memory. Declarative and procedural memory is taken for 67 projects in 38 firms and using partial least square(PLS) it is found that customer and innovation orientation, social responsibility and market were positively mapped to declarative memory; customer and innovation, systematic management control were positive for procedural memory (Giblin, Brennan, & Exton, 2010; Hoda, Noble, & Marshall, 2013; Keskin, 2009). The team coordination is motivated by physical appearance, communication and psychometric traits.. The relationships between team processes, personality, task characteristics, quality and satisfaction in teams are analyzed, XP method is used over 35 teams with 105 members and results indicate that the job satisfaction for agreeable and conscientiousness personality groups are high with cohesion being low due to fight amongst teams and positive relation between quality and extraversion factor (Acuña, Gómez, & Juristo, 2009; Hannay, Arisholm, Engvik, & Sj¿berg, 2010; Yuan, Zhang, Chen, Vogel, & Chu, 2009). The study deals with how personality and climate affect team development and product quality. Replicated Quasi experiments are conducted on real time data to improve coordination and quality among teams(Acuña, Gómez, Hannay, Juristo, & Pfahl, 2015; Gómez & Acuña, 2014; Kosti, Feldt, & Angelis, 2014; Mishra, Mishra, & Ostrovska, 2012).

Complex products will optimize the structure using new product development. Based on the concept of overlapping ratio and relative coordination a quantitative model is devised that reduces coordination time(Cataldo, Llc, & Herbsleb, 2012; Giordano & George, 2013; Yang, Yao, Lu, & Zhang, 2014). The relationship between effectiveness and abstract, personal trust is being analyzed for Virtual Teams (VTs) using theoretical frameworks and qualitative analysis that is used to improve the collocation, mobility and centralization in the presence of dominance(Baskerville & Nandhakumar, 2007; Crowston, Li, Wei, Eseryel, & Howison, 2007; Lavrač et al., 2007; Ocker, 2007). The team and its flexibility are vital in software companies. The product quality depends on ability to respond to changes. The dynamic capability theory is used to construct a model that depends on flexibility which in turn depends on reactive and capabilities which is tested via questionnaire set. The results indicate strong links for reactive and mixed links for capability with comprehensive and efficient response to improve quality(Basri & O'Connor, 2011; Li, Chang, Chen, & Jiang, 2010; Sarker, Sarker, & Schneider, 2009; Sutanto, Kankanhalli, & Tan, 2011). The analysis of High Performance Teams (HPTs) are done using empirical

studies taken from medical and company perspectives (Fernandes, Sales, Santos, & Webber, 2011; Garcia, Maria Balmaceda, Schiaffino, & Amandi, 2013; Petre, 2010; Philip, Wende, & Schwabe, 2012).

Teams involve multi levels, multi disciplines and multi agents. Some typical teams are self interested agents(Brafman & Domshlak, 2013; Griffith & Sawyer, 2010; Hayano, Hamano, & Sugawara, 2013; Sanchez-Anguix, Julian, Botti, & Garcia-Fornes, 2012).Intra and inter composition of teams is essential. Internal composition of work team is a key aspect. Iterative mix and replication methods are used to recruit people for teams. Qualitative research is done by interviewing managers, do cross sectional survey to find correlation; systematic mapping, replication survey is performed and it shows that personality and behaviour owe better correlation(Da Silva et al., 2013; Liberatore & Luo, 2010; Quesada, Palomares, & Martínez, 2015). Team building with regards to robots is a hotspot area. Collective movements of robots are managed by using virtual fields, behaviour blending and situation awareness. It is intended for local navigation. Robots are anonymous and it is decentralized with movement governed by rules. It is tested for real time and simulation and results prove to be effective(Arsenyan, Büyüközkan, & Feyzioğlu, 2011; Cifuentes, Girón-sierra, & Jiménez, 2015; Gunna & Anderson, 2013). There are a variety of external factors that affect team building. Critical technologies like wrong tool selection, limited internet access etc are key factors affecting teams' performance. To validate, 15 teams were observed for two year period and Grounded Theory (GT) approach was adopted. Results provide knowledge and guidelines for virtual teams and also prove that lack of internet availability, bandwidth, tool experience etc can affect teams' performance (Oakley, Hanna, Kuzmyn, and Felder, 2007; Saldaña-Ramos, Sanz-Esteban, García-Guzmán, and Amescua, 2012; Weimann, Pollock, Elsje, and Brown, 2013). Knowledge sharing is very vital in companies. Knowledge among Information System Development (ISD) teams must be shared. The sharing is promoted by using antecedent characteristics deployed in a research model to improve team performance(Crowder, Robinson, Hughes, and Sim, 2012; Janz and Prasarnphanich, 2009; Koh and Lim, 2012; Sha and Chang, 2012).When teams are involved software design becomes a social activity. A simple Teamwork Benefit Awareness model is set and validated for junior IT professionals to measure performance. Results are compared with nationality and gender sets and interesting inference is obtained for two Hofstede's country (Fernández-Sanz and Misra, 2012; Tessem, 2014). There are two questions addressed in regards to E-media variety and Team performance. Based on structural equation model and partial least square method around 290 teams from 66 organizations are analyzed and results analysed via wrap PLS 2.0 show that role of media is significant for teams(Kock and Lynn, 2012)

2.3 Role of tools software and algorithms (TSA) in team formation

Teams have a greater impact in software development. Managers and leaders strive hard to build an effective team by adopting certain specialized Tools, Software and Algorithms. These are made of computer logics, OR techniques or statistical methodologies. These Tools, Software and Algorithms are made viable by either making it commercial or by making it as Free/Libre Open Source Software (FLOSS).Agile computing techniques are becoming so prevalent in effective team building(Bipp, Lepper, & Schmedding, 2008). Distributed teams can't have face-face root cause analysis. Anew cloud based tool**,** Adaptive Root Cause Analysis (ARCA),is built and compared with 35 other RCA tools by incorporating them over four agile teams and two international companies and ARCA proves much effective for RCA, but has the problem of software support for creation of cause-effect diagram. The concept of decentralized coordination is essential for a good team structure(Lehtinen, Virtanen, Viljanen, Mäntylä, & Lassenius, 2014).A novel approach to decentralized coordination called factor graph is introduced. This forms a tree structure by eliminating dependency between functions and variables that contribute least to quality. A max-sum algorithm along with two pruning method is developed and tested over mobile sensor domain and results show two percent closeness to optimal value and 92% reduction in search space(Rogers, Farinelli, Stranders, & Jennings, 2011). Autonomous agents play important role in cooperation**.** Agents are designed to allow collaboration using two aspects namely teaching and leading for proper decision making. The ad-hoc agents are more knowledgeable and give optimal possible joint utility (Stone, Kaminka, Kraus, Rosenschein, & Agmon, 2013). The key challenges in decentralized operation are the collaboration which is handled using DEC-POMDPs algorithm using the identical condition plans. To keep it online, follow several local policies using new and fast local search method using linear programming. Results prove that it works best for large problems, when lines go imperfect and improve scalability of decision theoretic planning (Wu, Zilberstein, & Chen, 2011)

2.4 Role of optimization in team formation

The efficacy of team structure and its formation depends on the pattern of optimization. Many soft computing and evolutionary computing techniques are being incorporated to achieve optimization. Memetic algorithm is used to help professors in setting collaborative teams for software engineering. The algorithm is designed with many alternatives, each corresponding to different criteria. These criteria are mapped to different roles of the student and to test the algorithm's efficacy it is tested over 8 data sets and results indicate that for all datasets quality improves via this algorithm (Yannibelli & Amandi, 2012).Recruiting of decrecruiting of staffs to stabilize the network is a key aspect(Dorn, Skopik, Schall, & Dustdar, 2011). Staffing for any software project is not an easy task. This staffing

problem is viewed as a constraint satisfaction problem which is solved via an optimization approach that uses many utility functions that are maximized and minimized as of need. Some parameters considered are characteristics of project, human resource and development constraints (Barreto, Barros, & Werner, 2008).

2.5 Cohort study: a finisher mode

In this section attempts are made to bring out the overall theme of the analysis conducted so far with regards to effective team formation. The theme thus concentrated, constitutes the research problem for study. The research problem under consideration is to construct an optimized model driven approach for effective team formation in diversified domains. The Table 3 depicts the Strategies Contribution in Team Formation. The three major classifications of strategies have been broken down to eleven detailed sub categories and their percentage of contribution in effective team formation is estimated via a longitudinal investigation accompanied by probability theory. The formula to calculate the percentage of strategy contribution is given in Eq. (1)

$$\text{Percentage of Strategy Contribution } P_{SC} = \left(\sum_{i=1}^{n} A_i / N\right) \times 100 \quad (1)$$

where $\left(\sum_{i=1}^{n} A_i / N\right)$ is the probability of finding an article in a specified i^{th} sub category.

The Figure-2 depicts the Estimation of Contributors Percentages in Team Formation. The graph consists of the sub categories in the X Axis and the Percentage of Strategy Contribution in Y Axis. The peak points in the graph are the critical regions that are visualized and analyzed using intuitive learning principles to gain research inferences.

Table-3. Strategies contribution in team formation.

Strategies adopted	**Model method concept** (MMC) in (%)	**Tool software algorithm** (TSA) in (%)	**Optimize** in (%)
Evolutionary Optimization			85
Greedy Optimization			29
Operational Research Optimization			14
Empirical Model	64		
Math Model	10		
Statistical Model	2		
Analytical Model	3		
Research Model	20		
Algorithm Driven (Distributed, Markov Chain, Max/Min, Agent Based)		40	
Tool Driven (ARCA, Lantern, Virtual Team Manager)		30	
Software Program Driven (Pair Program)		30	

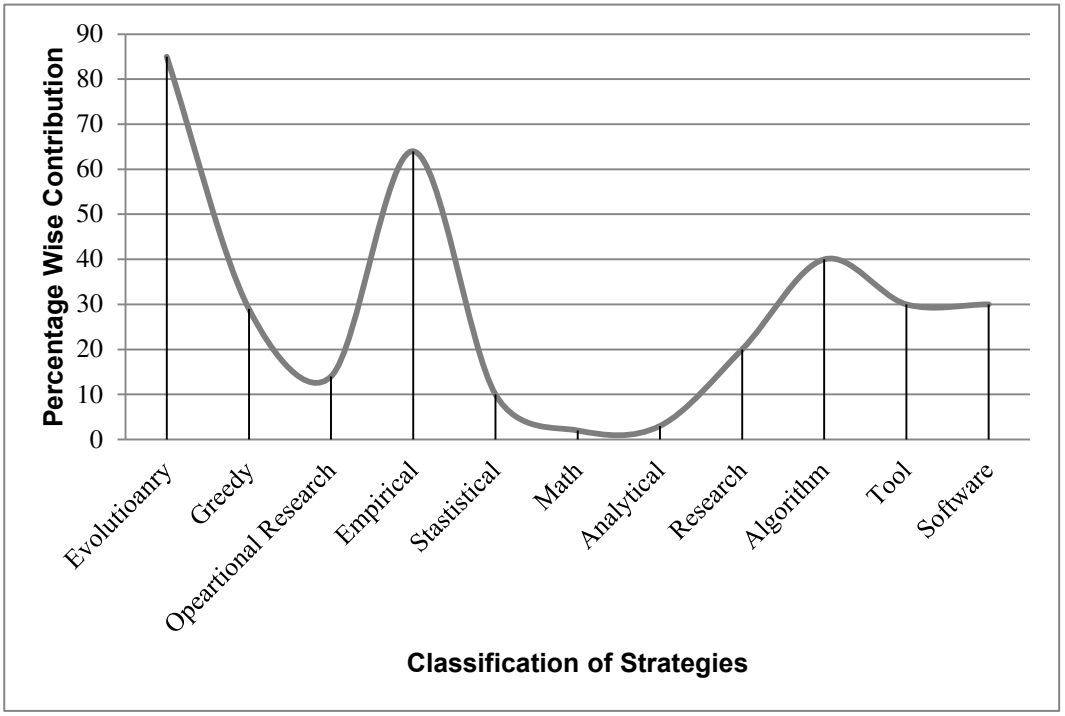

Figure-2. Estimation of contributors percentages in team formation.

3. AN INNOVATIVE INVESTIGATION OF TEAMS USING MULTI LINEAR REGRESSION MODEL

3.1 Generic regression model for team formation

The main objective of this article is divided into two folds. The first fold covers the strategy incorporated to identify the hot trend in the past with regards to the team formation problem and to figure out the cutting gap that needs to be resolved. This is done by designing a Multi Linear Regression Model (MLRM) that is being wrapped by Least Square Estimate (LSE) approach. The second fold involves construction of measures to evaluate the proposed model for its trustworthiness, reliability and accuracy. The idea, nuances and the concept of multi linear regression has been deployed in this article(Faraway, 2002).

The MLRM has more than one input called the regressors and an output called the response. Each of the input is accompanied by a coefficient and the response by an error rate. The general form of MLRM is given in Eq. (2) and the model describes a set of coefficients corresponding to the input that are calculated using LSE approach. The value of X_0 is taken as unity for the regression model.

$$M = \sum_{i=0}^{n} \beta_{i\,N_i} \quad (2)$$

where M is response of the MLRM, β_i is the coefficient corresponding to a particular regressor, N_i is the regressor and n is the number of columns. The Least Square Estimate is an approach which is predominantly used for regression model. The LSE uses the matrix method to solve for the regression coefficients. The general matrix form for the regression model is given in Eq. (3)

$$M = N\beta + \varepsilon \quad (3)$$

where N is the input matrix of order pxq, Y is response matrix of the order px1, β is the regression coefficient of order qx1 and ε is the error rate considered to be 0.5 for manually developed data set. The Least Square fit has the main objective of calculating regression coefficients and then evaluating the expected response thus calculated using residual schemes. The formula to calculate β the regression coefficient is given in Equation (4)

$$\beta = (N'N)^{-1}(N'M) \quad (4)$$

where N' is the transpose of N and (*) is the matrix multiplication.

3.2 Components of the generic regression model

The Figure-3 depicts An Empirical Two Tier Multi Linear Regression Model. This model is proposed as an empirical predictive model for judging the future scope of research in team formation. There are two distinctive tiers considered namely the top tier comprising of the Design Layer and the bottom layer comprising of the Evaluation Layer. These two layers are formed using various components. These components have been broadly classified into three categories namely:

Formation category

This is the primary category of the empirical predictive model that is used for a longitudinal survey of cohort study on effective team formation. In this category the model is designed using inputs taken from various parameters that constitute an empirical data set. The inputs

are termed as regressors and they are accompanied by a set of constants called regression coefficient that remain static for a model taken under consideration. The model is designed with these two parameters using Eq. (1). This designed model is further used for predicting new sample set. The regression coefficient remain static and the regressor alone changes based on the sample set. Most often a newer sample is not calculated as a crisp value but a range is fixed for the newer sample with crisp lower and upper bounds. Before prediction a systematic training of the proposed multi linear regression model is necessary which is dealt in detail in next category.

Formulation category

This category deals with the mathematical deployment for the proposed model. The expected response is calculated for each instance of the regressor and the residual is calculated for each of the expected response. These residuals are plotted using a special type of graph called the residual plot. This plot is the benchmark for standardizing the need for linear (or) nonlinear type of regression. The results of this category is fed into the next category for evaluating the quality metrics like reliability, accuracy, adequacy etc.

Findings category

In this category the proposed model is evaluated using several metrics like t-Test, ANOVA (F-test, p value), R^2 and R^2_{adj} measure and Root Mean Square Error (RMSE). Each of the measure is used to validate a specific quality metric namely; the two tests are used for testing the adequacy of the individual and the model as a whole respectively. The two measures are used to test the reliability of the model and the RMSE is used to test the accuracy of prediction. Only after satisfying the significant threshold values for each of the measures, the proposed model is used for predicting newer (unknown) sample set.

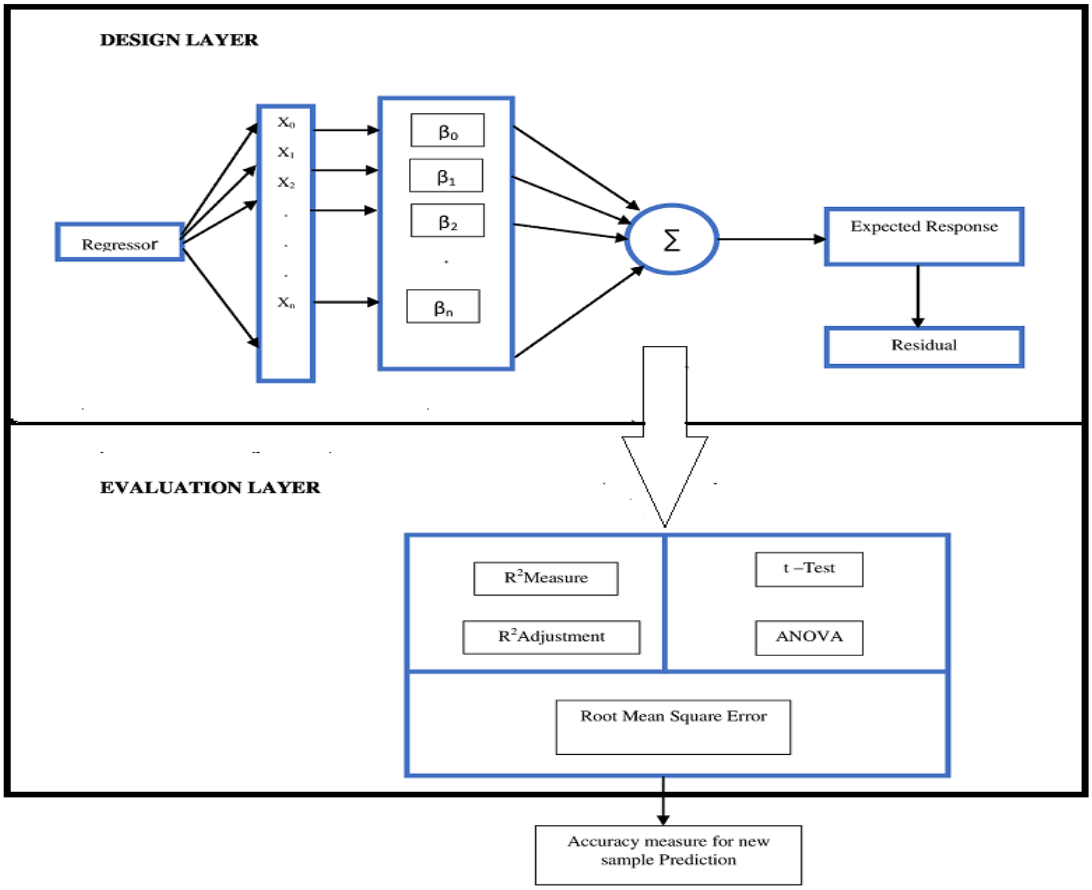

Figure-3. An empirical two tier multi linear regression model.

3.3 Mathematical background for the proposed model

The Table-4 depicts the Regression Model of Survey on Team Formation. It is a 9x5 matrix with last column being the actual response and all the preceding columns are all inputs. The Table 5 depicts the Estimation of Residuals for the Regression Model. In this table residual is being calculated that tells whether the model is feasible with regards to linear approach or not by using a residual plot. The Residual plot for the Regression Model is given in Figure-4. The plot must be randomly distributed for linear regression approach to be feasible for the given model. The Table-5 shows the Estimation of Regression Coefficients for the proposed MLRM and the LSE fit for the given model.

Table-4. Regression model of survey on team formation.

Regressor (N1)	Regressor (N2)	Regressor (N3)	Regressor (N4)	Response (M)
2007	0	11	0	82.5
2008	1	8	1	99.5
2009	1	7	0	89.5
2010	0	6	1	92.5
2011	1	8	2	99.5
2012	0	17	3	92.5
2013	2	9	2	99.5
2014	0	6	1	92.5
2015	2	11	0	89.5

N1 - Year of Publication, N2 -OPTIMIZE, N3 - Model Method and Concept (MMC), N4 - Tool Software and Algorithm (TSA), M - Response with manual error considered.

The residual is the difference between the actual response and the expected response. This is the measure of linearity in a regression model. The random distribution of residual points in the space signifies the linearity property. The formula to find residual is given in Equation (5).

$$\text{Residual } e_i = M_i - M_i'' \quad (5)$$

where M_i is the Actual Response and M_i'' is the Expected Response (or) Calculated Response from the MLRM.

The formula in Equation (5) is used in Table-5 for estimating the residual and plotting those points in the residual plot in Figure-3 to determine the feasibility of the model. The Table-5 depicts the Estimation of Residuals for Regression Model. The calculation of residual is an important parameter for judging the nature of the proposed multi linear regression model. The residual is generally the difference between the actual response and the expected response. Most regression model considers the residual of case order type while a few regression models consider the residual of fitness and probability. The random distribution of residuals in the residual plot signifies that the linear model chosen for the application under consideration is valid. If the plot forms a sequential pattern then the non linear model is chosen for the application. In the case of application pertaining to business, finance and economics fitness residual or probability residuals is used.

Table-5. Estimation of residuals for regression model.

Actual response (M)	**Expected response** (M'')	**Residuals** (e)
83	84.53	-1.53
100	96.03	3.97
90	91.38	-1.38
93	93.66	-0.66
100	100.25	-0.25
93	92.72	0.28
100	102.19	-2.19
93	92.24	0.76
90	88.96	1.04

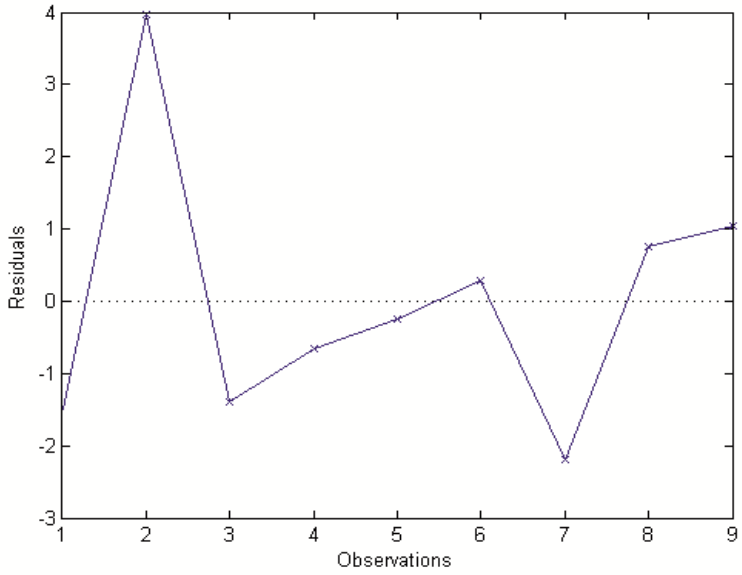

Figure-4. Residual plot for regression model.

The Table-6 represents the coefficients of the MLRM that are estimated using LSE approach. Once the coefficients are calculated the model can be used to predict unknown samples. The detailed concept of evaluation of MLRM is discussed in the next section.

Table-6. Estimation of regression coefficient for the proposed MLRM.

Regression coefficient (β_i)	Values for β	Equation of LSE fit
β_0	806.69	
β_1	-0.35	$M = 806.7N_0 - 0.35N_1 + 3.63N_2 -$
β_2	3.63	$0.98N_3 + 5.28N_4$
β_3	-0.98	
β_4	5.28	

LSE - Least Square Estimate

3.4 Evaluation of designated regression model

The MLRM is constructed as a part of first fold to identify the potential hot trend in the literature with respect to team formation and to tap the gap that is to be filled for a successful outcome. The efficacy and reliability of the proposed model is evaluated using several statistical methods. These statistical methods describe the model's complete trait. These traits include the quality of service measures. The statistical methods that are chosen for the evaluation are t-Test, Analysis Of Variance (ANOVA) test, R^2 measure, R^2_{adj} (adjusted R^2) measure and root mean square error.

These evaluations are carried out based on the hypothesis that is formulated for the proposed MLRM. The hypothesis is split into Null hypothesis (H_0) and Alternate hypothesis (H_1). The Eq.(6) states this as

$$H_0: \beta_0 = \beta_1 = \ldots = \beta_n$$
$$H_1: \beta_i \neq 0 \text{ for at least one } i = 0, 1 \ldots n \quad (6)$$

The t-Test is a test conducted on individual regressor to understand its significance with respect to the proposed model. This test is basically a weaker form of testing as it considers the individuals and pays subtle attention to the model as a whole. In reality model as a whole has greater significance in prediction. As the count for input increases the sum of square and the variance and decreases the residual sum of square. The formula to perform t-Test is given in Eq. (7)

$$t_{stat} = \frac{\beta_i}{\sqrt{D_{ii}\sigma^2}} \quad (7)$$

Where β_i is the regression coefficient, σ^2 is the variance and D_{ii} is the diagonal elements of $(X'X)^{-1}$

The t-Test is sometimes referred to as a marginal test and the Null hypothesis is rejected only if $|t_{stat}| > t_{\frac{\alpha}{2},n-k-1}$ is true. Here α is the confidence of 95%, so set it to 0.05, n is total number of records and k is the number of parameters

The ANOVA test also called the F-test is a test for understanding the significance of the model as a whole. The F-test also checks for the hypothesis from Eq. (5). The procedure to calculate ANOVA is depicted in Table 7 ANOVA Calculation Procedure.

Table-7. ANOVA calculation procedure.

Variations	Sum of squares (SoS)	Degree of freedom	Mean square	F_{stat} (ANOVA)
Regression	SS_{Reg}	K	MS_{Reg}	
Residual	SS_{Res}	n-k-1	MS_{Res}	MS_{Reg}/MS_{Res}
Overall	SS_O	n-1		

$MS_{Res}$is equal to variance (σ^2) – Residual Mean Square, MS_{Reg} is Regression Mean Square

The formula to calculate regression sum of square, residual sum of square and overall sum of square is given by Eq. (8)

$$\text{Regression SoS } SS_{Reg} = \beta' N' M - \frac{(\sum_{i=1}^{n} Mi)^2}{n}$$
$$\text{Residual SoS } SS_{Res} = M'M - \beta' N' M \quad (8)$$

$$\text{Overall SoSSS}_O = M'M - \frac{(\sum_{i=1}^{n} Mi)^2}{n}$$

where (*)' is the transpose of the (*) and n is number of rows.

The ANOVA is called the complete test. The Null hypothesis is rejected when $|F_{stat}|>F_{\alpha,k,n-k-1}$ is true and once the null hypothesis is rejected it improvises the stability and feasibility of the model. Here α.is taken to be 0.05 and k is the total parameters and n is the number of records.

The $R^2$and R^2_{adj} are two measures that are used to check for the total adequacy of the proposed model. Out of these two the best is R^2_{adj} because it can be used for variable selection and to avoid over-fitting. The formula for these two measures are given in Eq. (9)

$$R^3 = \frac{SS_{Reg}}{SS_O}$$
$$R^2_{adj} = 1 - \frac{SS_{Res}/(n-p)}{SS_O/(n-1)} \quad (9)$$

where $SS_{Res}/(n-p)$ is the in the numerator and is the residual mean square, $SS_O/(n-1)$ is the denominator and is a constant. The Mean Square Error (MSE) is another measure to test the reliability of the model. It is found using the formula given in Eq. (10). The model sets a threshold of 5% as a tolerance to error. This means that error rate within this threshold is acceptable and beyond this is not. The MSE magnifies the outliers and makes it easy for the user to understand the stability of the system. The square root of MSE is Root MSE.

$$\text{Mean Square Error (MSE)} = \frac{\sum_{i=1}^{n}(M_i - M_i'')^2}{r}$$
$$\text{Root MSE} = \sqrt{MSE} \quad (10)$$

where r is number of rows in the data set and M_i-M_i'' difference between the actual and expected. The Table 8 shows the Overall Estimation of Model Efficiency. This table conveys the desired quality of service which a model poses for being effective. This table summarizes all the test results and it proves that the proposed model is feasible and stable.

A walkthrough over Table-8 gives the complete efficacy chart of the model which is used for investigating teams and its formation for a successful project completion. Two standard test metric one for individual and other for holistic analysis has been used along with two standard measures which are used for determining the appropriate fitness of the model into the application.

Table-8. Overall estimation of model efficiency.

Regression	t-Test Result	R^2 Result	R^2_{adj} Result	Root Mean Square Error (RMSE)	ANOVA Result (F_{stat} and p value)	Range of new sample
Proposed Model	β_0 =1.04 β_1= -0.92 β_2= 2.95 β_3= -3.31 β_4= 5.18	0.897	0.794	2.60	8.7 0.029	84.35<=M_{new}<=97.30 90<=M_{new}<=115

The prediction of an unknown sample using the regression model is one of the essential features of the MLRM. The proposed model is also developed to predict a new sample. The prediction in regression is generally based on range values. The last column of Table 5 shows the range of value for the response, based on the input pattern that is fed to the model. The Equation (11) is used to estimate the range of the new response.

$$M'' - t_{\frac{\alpha}{2},n-k-1}\sqrt{\sigma^2(1 + N'_{new}(N'N)^{-1}N_{new})} \leq M_{new} \leq M'' + t_{\frac{\alpha}{2},n-k-1}\sqrt{\sigma^2(1 + N'_{new}(N'N)^{-1}N_{new}} \quad (11)$$

The Figure-5 depicts the Comparison Chart for Team Vs Product. The comparison is between Quality of

Service (QoS) and the Average Percentage Value (APV). The APV is calculated using Equation (12)

$$\text{Average Percentage Value (APV)} = \frac{\sum_{i=1}^{n} P_i}{n} \quad (12)$$

If QoS is in the form of non linguistic (numerical) then use P_i values as it is given.

Else if QoS is in the form of linguistic terms (Categorical) then set threshold values for P_i based on heuristic intuition and confirm the threshold with experts.

where P_i is the percentage value of a corresponding QoS in the i^{th} article and n is the total number articles that pertain to the respective QoS.

The Table-9 shows the Estimation of Average Percentage Values for Teams and Products. There are 11 QoS that are taken for consideration and APV for team and product is calculated using Equation (12). These calculated values are then plotted over a comparison graph shown in Figure-5. The APV that is obtained in the table for teams and products is an intuitive guidance; indicating the behaviour of each attribute or parameter taken for consideration.

Table-9. Estimation of average percentage values for teams and products.

Quality of Service (QoS)		Average Percentage Values (APV)	
	Teams	Proportionality	Products
Reliable	75	Direct	80
Scalable	60	Direct	75
Trust	80	Direct	80
Cohesion	85	Inverse	40
Accuracy	80	Direct	85
Adapt	70	Direct	80
Coupling	45	Inverse	85
Stability	70	Direct	75
Durable	80	Inverse	45
Consistent	80	Direct	85
Elapse Time	20	Direct	15

This table depicts the relationship between teams and products with regards to the 11 QoS. The proportionality column signifies the nature of proportionality that exists among the products and teams with respect to the considered QoS. The word Direct signifies direct proportionality and Inverse signifies inverse proportionality. The definition for the QoS are as follows (Pressman, 2009): Reliable - Self Heal, Scalable - Expand, Trust - Belief, Cohesion - Intra Bond, Accuracy - Focus and Preciseness, Adapt - Accustom to environment, Coupling - Inter Bond, Durable - Lifespan, Consistent - Recursive Success, Elapse Time - For teams :Time between decomposing and regrouping; For products: Time between design and delivery.

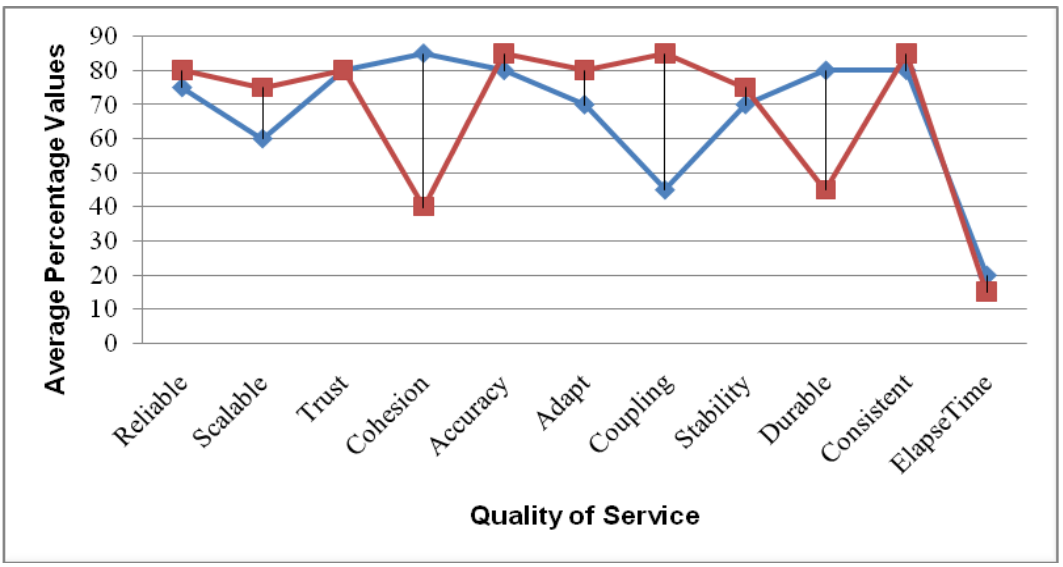

Figure-5. Comparison chart for team vs product.

3.5 Merits and demerits of least square estimate

The approach of LSE is used to wrap the MLRM. It is an elegant method for prediction with following advantages (i) Guarantees maximum likelihood solution (ii) Highly consistent and asymptomatically normal (iii) LSE line is the best linear unbiased estimator according the Guass Markov theorem. The LSE has certain limiting factors also namely, (i) LSE minimises the squared function thereby making the outliers negligible in certain applications. (ii) It requires the transpose of the sample covariance matrix which may not be possible in all cases; in such scenario use LASSO or Ridge regression.

4. COMPARATIVE INVESTIGATION ON TEAM FORMATION

4.1 Research theme under study

This section deals with the research theme that is common to both the theoretic as well as the predictive model. As mentioned early in the article; the research theme is construction of an optimized model driven approach for team formation. This theme has been taken for evaluation by both the models and inferences are thereby made.

4.2 Theoretic way of evaluating research theme

This section deals with the standard way for surveying any form of application. A longitudinal survey mechanism has been adopted to investigate the research trope. A selective search based duo fold approach is adopted. According to this, articles from 2007-2015 are first collected, filtered based on relevance (100 articles) and analyzed using Eq. (1). The graph in Figure-2 provides an informative inference by adhering the research motif. The critical regions in the graph indicates that the contribution of those strategies in team formation is effective and the troughs indicate that these strategies have a scope into team formation which are unturned. The evolutionary technique and the empirical model show peak contributions to team formation while the math & statistical model and the operational research methods show trough contributions.

4.3 Predictive way of evaluating research theme

In this section a quantitative predictive model, MLRM is used to evaluate the analysis done on team formation. This approach has a mathematical background to support the predictions made on behalf of the future era. The prediction is approved only after the model evaluation. This predictive model also adheres to the research theme with a strong mathematical support. The t-Test values from Table 8 drives the inference in support of the theme taken under consideration.

4.4 Final words on formulated inferences

In this section a thorough understanding of the two models are made. Though these two models adhere to the research theme the effective adherence is from the predictive model and not the theoretic model because (a) Predictive model has a strong mathematical background to support the result than the theoretic model. (b) The quantitative predictive model consumes less time to predict but more time to train. (c) The predictive model is highly scalable with regards to a specific application. To illustrate these facts about these two models Table 10 has been constructed and is being supplemented by Figure-6.

The Table-10 is the Linguistic Comparative Study of Two Models that is used to analyze the five key factors and its impact on these two models (theoretic and predictive). The inferences gained are categorical in nature thereby giving a linguistic (intuitive) knowledge of predictive model's superiority over theoretic model.

Table-10. Linguistic comparative study of two models.

Comparison factors	Theoretic model	Predictive model
Mathematical Support	Weak Support	Strong Support
Training Time	Less Time	More Time
Predicting Time	More Time	Less Time
Nature of Prediction	Intuitive Prediction	Mathematical Prediction
Scalability	Less Scalable	Highly Scalable

The Table-11 depicts the Numerical Comparison of Two Models Based on Normalized Time Stamp Values. In this table the comparative analysis of cohort and regression model is done using time stamp values which is normalized using decimal point normalization (10^j) where j denotes the range of normalizing factor. There are four parameters that are considered for time stamps namely the theoretic training and predicting followed by predictive training and predicting. The time stamps are calculated for each of the proposed classifications.

Table-11. Numerical comparison of two models based on normalized time stamp values.

Time stamps vs classification	MMC	TSA	Optimize
Theoretic training	0.6	0.4	0.4
Theoretic predicting	0.9	0.7	0.6
Predictive training	0.8	0.6	0.5
Predictive predicting	0.5	0.3	0.3

The Figure-6 depicts the Comparison Chart for Training and Prediction Time Stamps which considers four novel parameters namely Theoretic Training and Prediction, Predictive Training and Prediction for the investigation of efficacy among the two models. This pictorial theme gives a wider scope for understanding the tournament that has been conducted. The inference obtained from the figure is based on the normalized time stamp values (decimal point normalization) which indicate that the theoretic training takes less time compared to predictive training and predictive prediction takes less time compared to theoretic predictions thereby concluding with a mathematical support that the predictive model has a superior edge over the theoretic model. The green color rhombus stands for MMC, blue color for Optimize and purple color for TSA respectively. The graph shown in Figure-6 is a radar graph that uses the four axis principle of north, south, east and west. Each of these axes is corresponding to a specific strategy adopted for investigating the guidelines and principles for effective team formation. As noted from the radar, theoretic prediction takes a longer time (0.9) since the process of driving to a significant inference deals with drawing intuitive parallels and comprehending those parallels into justifiable results. On the other hand the process of predictive prediction takes only very less time (0.5) since it is a mathematical driven technique and can be easily comprehended for valid inferences using mathematical formulae. This notion of prediction changes in terms with the training aspect as the training phase takes more time with respect to predictive training (0.8) as the model has to be configured for future prediction by making modularization in terms with the parameters governing the prediction. A trio-fold iteration has to be incorporated before identifying the feasibility of the model, so that the model evolves as a valid candidate for future prediction. While the time taken to train the theoretic principle is less comparatively (0.6) as this training is an intuitive training and cannot be used as a determinant factor for future prediction until the exact future pattern is not known. So once the future pattern is identified time drifts more into drawing parallels to match the proximity of the pattern which comes under the theoretic prediction class. One implicit inference that is also evolved from the radar is that the time for theoretic prediction is not only greater than the predictive prediction but also predictive training.

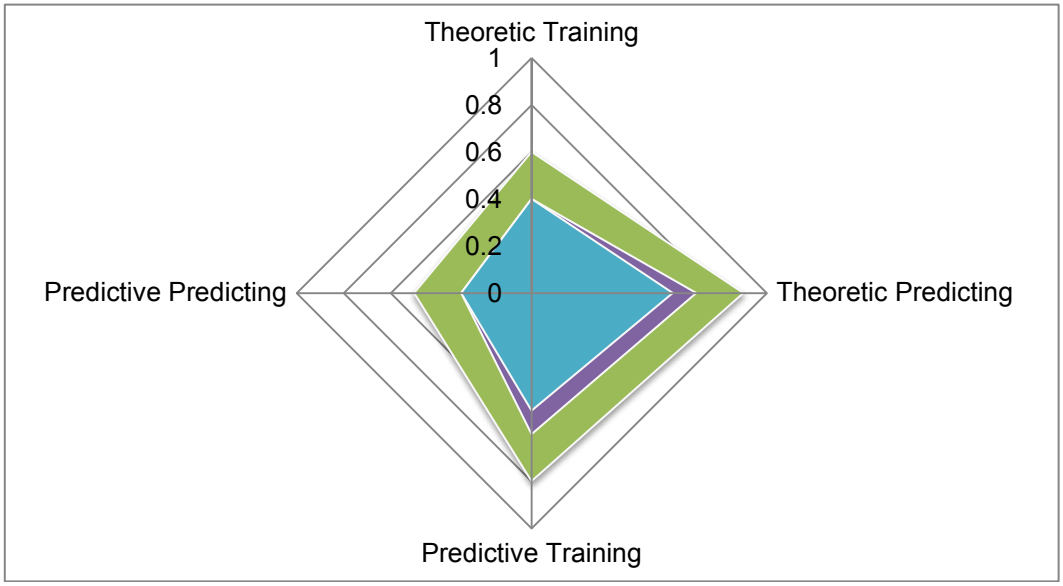

Figure-6. Comparison chart for training and prediction time stamps.

5. RESULTS AND DISCUSSIONS

Team activities have now become a part and parcel of companies and academics. Teams are involved in all places where activity by a single person becomes ordeal. In this article a novel two tier MLRM is proposed to construct a survey based on three heuristic classification set namely the MMC, OPTIMIZE and TSA. The pre-requisites for design of this survey model are drawn from Table-1 and Table-2. The second tier validates the proposed model using statistical measures and verifies the accuracy, reliability and adequacy. As pictures speak louder than text, graphical representations are also considered. In Figure-1 that depicts Scatter Plot for Team Formation Survey, it is clear that the need for study in this area is imperative as the year 2012 marks a peak value followed by 2013 and 2015. It is self explained from this graph that the current year (2015) has an exponential scope for team formation and effective strategies to supplement them. Around 20 versatile strategies for team formation in the year 2015 is validated via an exponential trend line. As a part of future scope the gap in the research study is to be figured out and effective & novel strategies are to be incorporated. These graphical intuitions are taken as a rudimentary element for the construction of the novel survey mechanism using regression model. The validation process begins with a t-Test measure which is a test for individual regressor. This tells the role of each regressor in that particular model. The value for each of input is measured using Eq. (7) and compared with t-distribution threshold (2.776). The inference that is gained out of this measure is that in the nine years of survey from 2007 to 2015 there has been an enormous amount of contribution in the field of MMC accompanied with TSA (t-Test criteria satisfied); but subtle contribution in the field MMC accompanied by OPTIMIZE (t-Test criteria not satisfied). The second measure is the F-test supported by a p-value in correspondence to the ANOVA technique. This is used to test the adequacy of the model as a whole. The designed model is validated using Table-5 supplemented by Eq.(7). The measured value is compared with a threshold (6.39) from F-distribution table. The inference is that the proposed regression model is adequate and is reliable. The p-value (0.02) that is calculated from the ANOVA also infers better model stability as the calculated value is less than the threshold of 0.05.

The two additional measures namely R square and Adjusted R square are used to test the fitness of the model. The adjusted R square is taken as a yard stick for fitness evaluation due to direct proportionality problem of R square measure with number of regressor in the model. The value of 79% infers a significant fit of the proposed model for the present survey application. This fitness measure is also strongly supported by the graph in Fig.4. This is a residual plot and the randomness of the plot validates that the linear regression model is the most effective for the application under consideration. The final phase of the model in the two tier architecture is to predict the new entry level sample.

The model is trained primarily with a known set of samples (responses known) and then set ready for prediction. The new sample (Present Year (2015), Six MMC and two OPTIMIZE) is fed as input to the model making sure that the sample is free off extrapolation problem and the range of response value (Table 8 last column first value) for the model is provided using Eq. (11). The inference gained from the range is that the article count for team formation will improve significantly thereby making a benchmark in the efficacy of team formation. This type of an interval based response is more reliable than a single valued response in terms with prediction models. This is strongly supported using RMSE measure that infers the accuracy rate of the regression model. The value of 2.6% indicates lower error rate in prediction compared to the threshold of 5%. After proving the model's stability for an extrapolation free test sample; a test sample with extrapolation (Present Year (2015), Six

MMC and Six OPTIMIZE) was fed to the model and the interval range (Table 7 last column second value) justified the same fact of benchmark trait in the efficacy of team formation. The Fig.5 is the comparison of QoS values for a given team and the end product that is built by that team. It shows the proportionality relationship between teams and product pertaining to a given QoS. The square marked red lines in the graph depict the product to QoS line and the diamond marked blue lines depict the team to QoS line. The inference gained from the chart is that the Cohesion, Durability and Elapse Time for a team must be lower so that the end product flourishes. These inferences were cross validated using experts opinion (Business Consultant, Management Staffs and Senior Project Manager) and the results proved to be justifiable and valid. These intuitions derived from the survey helps in better and effective team formation.

6. CONCLUSIONS

Teams in software companies are an inevitable resource factor. The success of any company depends on its human resource. A selective effort has been taken to survey 100 papers from esteemed journals for over nine years and better understanding of teams, their structures and formations. The potential hotspot for research are being tapped using a tournament survey mechaism wherein a comparison is made in terms with efficacy of investigation based the cohort study versus the innovative mechanism of two tier MLRM. The intuitive inference thus gained is that MLRM has a cutting edge in terms with performance over cohort study while investigating effective team formation because of its strong mathematical support.The advantages of this review are (i) It serves as a better platform for future surveys conducted in this field of study. (ii) This acts as a centralized hub for knowing the current trend in team fornation. (iii) It uses the principle tournament survey by making a comparative study with theoritic and predictive measures. (iv) A predictive model with strong mathematical background is proposed for investigating team formation. (v) The current trend that is discovered with respect to team formation are validated both theoritically and mathematically for intutive justifications and reasoning purposes.

The articles from current trend and also state of art articles are being collected and analyzed based on a duo fold search scheme. The collected 100 articles are analyzed to the core, using a regression mechanism based on LSE approach and the results of the analysis are validated suing statistical principles. The limitation of this survey is that, it has constructed a dedicated model for review on teams which cannot be extended for other empirical applications.The survey based on regression mechanism on effective team formation finally proves that the current trend is moving towards an optimized model driven team formation strategy. As a part of future scope; efforts must be made to construct a holistic model for survey. As a final word for take home message; an equi-propotion amalgamation for both theoreric (Cohort) and predictive principles (Regression) are needed for a successful tournament survey to be constructed pertaining to any heterogeneous diversified application.

ACKNOWLEDGEMENTS

We, the authors would like to thank the funding agencies, University Grants Commission for their financial support for the research from Rajiv Gandhi National Fellowship of Grant no: F./2015-17/RGNF-2015-17-TAM-83. We also thank the SASTRA University for providing an excellent infrastructure for performing our research work. We also thank the interview participants and the anonymous reviewers for their effort in making the research a successful work.

REFERENCES

Acuña S. T., Gómez M. and Juristo N. 2009. How do personality, team processes and task characteristics relate to job satisfaction and software quality? Information and Software Technology. 51(3): 627-639. http://doi.org/10.1016/j.infsof.2008.08.006.

Acuña S. T., Gómez M. N., Hannay J. E., Juristo N. and Pfahl D. 2015. Are team personality and climate related to satisfaction and software quality ? Aggregating results from a twice replicated experiment. Information and Software Technology. 57: 141-156. http://doi.org/10.1016/j.infsof.2014.09.002.

Adair J. 2004. The John Adair Handbook of management and Leadership. Leadership, 242.

Alavi H. S. and Dillenbourg P. 2012. An ambient awareness tool for supporting supervised collaborative problem solving. IEEE Transactions on Learning Technologies. 5(3): 264-274. http://doi.org/10.1109/TLT.2012.7

André M., Baldoquín M. G. and Acuña S. T. 2011. Formal model for assigning human resources to teams in software projects. Information and Software Technology. 53(3): 259-275. http://doi.org/10.1016/j.infsof.2010.11.011.

Arsenyan J., Büyüközkan G. and Feyzioğlu O. 2011. Modelling Collaboration Formation with a Game Theory Approach. World Congress on Engineering 2011, 1, 2073-2085. http://doi.org/10.1016/j.eswa.2014.10.010.

Assudani R. H. 2011. Role of familiarity in affecting knowledge gaps in geographically dispersed work. IEEE Transactions on Professional Communication. 54(3): 314-332. http://doi.org/10.1109/TPC.2011.2161805.

Barreto A., Barros M. deO and Werner C. M. L. 2008. Staffing a software project: A constraint satisfaction and optimization-based approach. Computers and Operations Research. 35(10): 3073-3089. http://doi.org/10.1016/j.cor.2007.01.010.

Baskerville R. and Nandhakumar J. 2007. Activating and perpetuating virtual teams: Now that we're mobile, where

do we go? IEEE Transactions on Professional Communication. 50(1): 17-34. http://doi.org/10.1109/TPC.2006.890849.

Basri S. and O'Connor R. V. 2011. A study of software development team dynamics in SPI. Communications in Computer and Information Science, 172, 143-154. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84864269413&partnerID=40&md5=b7cc3d9810d790b7810f1fc7bb70addc.

Bella E., Fronza I., Phaphoom N., Sillitti A., Succi G. and Vlasenko J. 2013. Pair Programming and Software Defects - A Large, Industrial Case Study. 39(7): 930-953.

Bipp T., Lepper A. and Schmedding D. 2008. Pair programming in software development teams - An empirical study of its benefits. Information and Software Technology. 50(3): 231-240. http://doi.org/10.1016/j.infsof.2007.05.006.

Brafman R. I. and Domshlak C. 2013. On the complexity of planning for agent teams and its implications for single agent planning. Artificial Intelligence. 198, 52-71. http://doi.org/10.1016/j.artint.2012.08.005.

Brennen S. D., Strong R. J., Ryder C. J., Blendell C. and Molloy J. J. 2007. Teams, computer modeling, and design. IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans. 37(6): 995-1004. http://doi.org/10.1109/TSMCA.2007.904747.

Brookes N. J., Morton S. C., Grossman S., Joesbury P. and Varnes D. 2007. Analyzing social capital to improve product development team performance: Action-research investigations in the aerospace industry with TRW and GKN. IEEE Transactions on Engineering Management. 54(4): 814–830. http://doi.org/10.1109/TEM.2007.906859.

Cataldo M., Llc R. B. and Herbsleb J. D. 2012. Coordination Breakdowns and their Impact on Development Productivity and Software Failures Coordination Breakdowns and their Impact on Development Productivity and Software Failures. 39(3): 343-360.

Chan C. L., Jiang J. J. and Klein G. 2008. Team task skills as a facilitator for application and development skills. IEEE Transactions on Engineering Management. 55(3): 434-441. http://doi.org/10.1109/TEM.2008.922633.

Chen C. J. 2007. Information technology, organizational structure, and new product development-the mediating effect of cross - Functional team interaction. IEEE Transactions on Engineering Management. 54(4): 687-698. http://doi.org/10.1109/tem.2007.906831.

Chen W.-N. and Zhang J. 2012. Ant Colony Optimization for Software Project Scheduling and Staffing with an Event-Based Scheduler. IEEE Transactions on Software Engineering. 39(1): 1-17. http://doi.org/10.1109/TSE.2012.17.

Cifuentes S., Girón-sierra J. M. and Jiménez J. 2015. Expert Systems with Applications Virtual fields and behaviour blending for the coordinated navigation of robot teams : Some experimental results. Expert Systems with Applications. 42(10): 4778-4796. http://doi.org/10.1016/j.eswa.2015.02.008.

Colomo-Palacios R., Casado-Lumbreras C., Soto-Acosta P., García-Peñalvo F. J. and Tovar-Caro E. 2012. Competence gaps in software personnel: A multi-organizational study. Computers in Human Behavior. 29(2): 456–461. http://doi.org/10.1016/j.chb.2012.04.021.

Crowder R. M., Robinson M. a., Hughes H. P. N. and Sim Y. W. 2012. The development of an agent-based modeling framework for simulating engineering team work. IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans. 42(6): 1425-1439. http://doi.org/10.1109/TSMCA.2012.2199304.

Crowston K., Howison J., Masango C. and Eseryel U. Y. 2007. The role of face-to-face meetings in technology-supported self-organizing distributed teams. IEEE Transactions on Professional Communication. 50(3): 185-203. http://doi.org/10.1109/TPC.2007.902654.

Crowston K., Li Q., Wei K., Eseryel U. Y. and Howison J. 2007. Self-organization of teams for free/libre open source software development. Information and Software Technology. 49(6): 564-575. http://doi.org/10.1016/j.infsof.2007.02.004.

Da Silva F. Q. B., França a. C. C., Suassuna M., De Sousa Mariz L. M. R., Rossiley I., De Miranda R. C. G., Espindola E. 2013. Team building criteria in software projects: A mix-method replicated study. Information and Software Technology. 55(7): 1316-1340. http://doi.org/10.1016/j.infsof.2012.11.006.

Dayan M., Elbanna S. and Di Benedetto a. 2012. Antecedents and Consequences of Political Behavior in New Product Development Teams. IEEE Transactions on Engineering Management. 59(3): 470-482. http://doi.org/10.1109/TEM.2011.2166078.

Doherty D. and O'Riordan C. 2009. Effects of shared perception on the evolution of squad behaviors. IEEE Transactions on Computational Intelligence and AI in Games. 1(1): 50-62. http://doi.org/10.1109/TCIAIG.2009.2018701.

Dorn C., Skopik F., Schall D. and Dustdar S. 2011. Interaction mining and skill-dependent recommendations for multi-objective team composition. Data and Knowledge Engineering. 70(10): 866-891. http://doi.org/10.1016/j.datak.2011.06.004.

Dubielewicz I. and Hnatkowska B. 2008. Improving software development process implemented in team project course. Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5102 LNCS (PART 2), 687-696. http://doi.org/10.1007/978-3-540-69387-1_80.

Fagerholm F., Ikonen M., Kettunen P., Münch J., Roto V. and Abrahamsson P. 2015. Performance Alignment Work: How software developers experience the continuous adaptation of team performance in Lean and Agile environments. Information and Software Technology. http://doi.org/10.1016/j.infsof.2015.01.010.

Faraway J. J. 2002. Practical Regression and Anova using R. Reproduction, 21(July), 212. http://doi.org/10.1016/0360-1315 (91)90006-D.

Fernandes P., Sales A., Santos A. R. and Webber T. 2011. Performance Evaluation of Software Development Teams: A Practical Case Study. Electronic Notes in Theoretical Computer Science. 275(1): 73-92. http://doi.org/10.1016/j.entcs.2011.09.006.

Fernández-Sanz L. and Misra S. 2012. Analysis of cultural and gender influences on teamwork performance for software requirements analysis in multinational environments. IET Software, 6(3): 167. http://doi.org/10.1049/iet-sen.2011.0070.

Fruchter R. and Courtier R. 2011. Building common ground in global teamwork through re-representation. AI and Society, 26(3): 233-245. http://doi.org/10.1007/s00146-010-0303-5.

Fruchter R., Swaminathan S., Matsumura N. and Ohsawa Y. 2008. Interaction Dynamics in AEC Global Teamwork. Tsinghua Science and Technology, 13(SUPPL. 1): 145-150. http://doi.org/10.1016/S1007-0214 (08)70141-X.

Garcia P., Maria Balmaceda J., Schiaffino S. and Amandi A. 2013. Automatic detection of team roles in computer supported collaborative work. IEEE Latin America Transactions, 11(4): 1060-1065. http://doi.org/10.1109/TLA.2013.6601751.

Gaubatz P., Lytra I. and Zdun U. 2015. Automatic enforcement of constraints in real-time collaborative architectural decision making. Journal of Systems and Software, 103, 128-149. http://doi.org/10.1016/j.jss.2015.01.056.

Georgiev D., Kabamba P. T. and Tilbury D. M. 2008. A new model for team optimization: The effects of uncertainty on interaction. IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans, 38(6) 1234-1247. http://doi.org/10.1109/TSMCA.2008.2001072.

Germain É. and Robillard P. N. 2008. Towards software process patterns: An empirical analysis of the behavior of student teams. Information and Software Technology. 50(11): 1088-1097. http://doi.org/10.1016/j.infsof.2007.10.018.

Ghobadi S. 2015. What drives knowledge sharing in software development teams: A literature review and classification framework? Information and Management, 52(1): 82-97. http://doi.org/10.1016/j.im.2014.10.008.

Giblin M., Brennan P. and Exton, C. 2010. Introducing Agile Methods in a Large Software Development Team: The Developers Changing Perspective. Agile Processes in Software Engineering and Extreme Programming, 184-189. Retrieved from http://www.springerlink.com/index/RU72H8210206K1N3.pdf.

Gider F., Likar B., Kern T. and Miklavcic D. 2012. Implementation of a multidisciplinary professional skills course at an electrical engineering school. IEEE Transactions on Education, 55(3): 332-340. http://doi.org/10.1109/TE.2011.2174238.

Giordano G. and George J. F. 2013. The effects of task complexity and group member experience on computer-mediated groups facing deception. IEEE Transactions on Professional Communication. 56(3): 210-225. http://doi.org/10.1109/TPC.2013.2273817.

Giuffrida R. and Dittrich Y. 2015. A conceptual framework to study the role of communication through social software for coordination in globally-distributed software teams. Information and Software Technology, 63, 11-30. http://doi.org/10.1016/j.infsof.2015.02.013.

Gómez M. N. and Acuña S. T. 2014. A replicated quasi-experimental study on the influence of personality and team climate in software development. Empirical Software Engineering. 19(2): 343-377. http://doi.org/10.1007/s10664-013-9265-9.

Gonçalves P., Ferreira L., Gonçalves J., Putnik G. D. and Cruz-Cunha M. M. 2014. Direct Communication versus Virtual Communication in Virtual Teams. Procedia Technology, 16, 3-10. http://doi.org/10.1016/j.protcy.2014.10.062.

Grady C. a., He X. and Peeta S. 2015. Integrating social network analysis with analytic network process for international development project selection. Expert Systems with Applications, 42(12): 5128-5138. http://doi.org/10.1016/j.eswa.2015.02.039.

Griffith T. L. and Sawyer J. E. 2010. Research Team Design and Management for Centralized R&D. IEEE Transactions on Engineering Management. 57(2): 211-224. http://doi.org/10.1109/TEM.2009.2023132.

Gunna T. and Anderson J. 2013. Dynamic heterogeneous team formation for robotic urban search and rescue. Procedia Computer Science. 19(3): 22-31. http://doi.org/10.1016/j.procs.2013.06.009.

Guzmán J. G., Ramos J. S., Seco A. A. and Esteban A. S. 2010. How to get mature global virtual teams: A framework to improve team process management in distributed software teams. Software Quality Journal, 18(4): 409-435. http://doi.org/10.1007/s11219-010-9096-5.

Hannay J. E., Arisholm E., Engvik H. and Sj¿berg D. I. K. 2010. Effects of Personality on Pair Programming. IEEE Transactions on Software Engineering. 36(1): 61-80. http://doi.org/10.1109/TSE.2009.41.

Hayano M., Hamano D. and Sugawara T. 2013. Role and member selection in team formation using resource estimation. Frontiers in Artificial Intelligence and Applications, 252, 125–136. http://doi.org/10.3233/978-1-61499-254-7-125.

Hoda R., Noble J. and Marshall S. 2013. Self-organizing roles on agile software development teams. IEEE Transactions on Software Engineering. 39(3): 422-444. http://doi.org/10.1109/TSE.2012.30.

Iorio J. and Taylor J. E. 2014. Precursors to engaged leaders in virtual project teams. International Journal of Project Management. 33(0) -. http://doi.org/http://dx.doi.org/10.1016/j.ijproman.2014.06.007.

Isenberg P. and Carpendale S. 2007. Interactive tree comparison for Co-located collaborative information visualization. IEEE Transactions on Visualization and Computer Graphics. 13(6): 1232-1239. http://doi.org/10.1109/TVCG.2007.70568

Janz B. D. and Prasarnphanich P. 2009. Freedom to Cooperate: Gaining Clarity Into Knowledge Integration in Information Systems Development Teams. IEEE Transactions on Engineering Management. 56(4): 621-635. http://doi.org/10.1109/TEM.2009.2023451.

Johri A. 2011. Sociomaterial bricolage: The creation of location-spanning work practices by global software developers. Information and Software Technology. 53(9): 955-968. http://doi.org/10.1016/j.infsof.2011.01.014

Kanawati, R. (2015). Empirical evaluation of applying ensemble methods to ego-centred community identification in complex networks. Neurocomputing, 150, 417-427. http://doi.org/10.1016/j.neucom.2014.09.042

Keskin H. 2009. Antecedents and consequences of team memory in software development projects. Information and Management, 46(7): 388-396. http://doi.org/10.1016/j.im.2009.07.003.

Kitchenham B. and Charters S. 2007. Guidelines for performing Systematic Literature Reviews in Software Engineering. Engineering, 2, 1051. http://doi.org/10.1145/1134285.1134500.

Kock N. and Lynn G. S. 2012. Electronic media variety and virtual team performance: The mediating role of task complexity coping mechanisms. IEEE Transactions on Professional Communication. 55(3): 325–344. http://doi.org/10.1109/TPC.2012.2208393.

Koh E. and Lim J. 2012. Too early, too bad: Uncovering and understanding the initial participation paradox in technology-mediated learning teams. IEEE Transactions on Professional Communication, 55(1): 55-84. http://doi.org/10.1109/TPC.2011.2172122.

Kosti M. V., Feldt, R., & Angelis L. 2014. Personality, emotional intelligence and work preferences in software engineering: An empirical study. Information and Software Technology, 56(8): 973-990. http://doi.org/10.1016/j.infsof.2014.03.004.

Kotlarsky J., van Fenema P. C. and Willcocks L. P. 2008. Developing a knowledge-based perspective on coordination: The case of global software projects. Information and Management. 45(2): 96-108. http://doi.org/10.1016/j.im.2008.01.001.

Lavrač N., Ljubič P., Urbanič T., Papa G., Jermol M. and Bollhalter S. 2007. Trust modeling for networked organizations using reputation and collaboration estimates. IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews. 37(3), 429–439. http://doi.org/10.1109/TSMCC.2006.889531.

Lee, G., Espinosa, J. A., & Delone, W. H. (2013). Task environment complexity, global team dispersion, process capabilities, and coordination in software development. IEEE Transactions on Software Engineering. 39(12): 1753-1771. http://doi.org/10.1109/TSE.2013.40.

Lehtinen T. O. a, Virtanen R., Viljanen J. O., Mäntylä M. V. and Lassenius C. 2014. A tool supporting root cause analysis for synchronous retrospectives in distributed software teams. Information and Software Technology. 56(4): 408-437. http://doi.org/10.1016/j.infsof.2014.01.004.

Li Y., Chang K. C., Chen H. G. and Jiang J. J. 2010. Software development team flexibility antecedents. Journal of Systems and Software, 83(10): 1726-1734. http://doi.org/10.1016/j.jss.2010.04.077.

Liberatore, M. J., & Luo, W. 2010. Coordination in consultant-assisted IS projects: An agency theory perspective. IEEE Transactions on Engineering Management. 57(2): 255-269. http://doi.org/10.1109/TEM.2009.2013838.

Magdaleno, A. M., de Oliveira Barros, M., Werner, C. M. L., de Araujo, R. M., & Batista, C. F. A. (2014). Collaboration optimization in software process composition. Journal of Systems and Software, 103, 452-466. http://doi.org/10.1016/j.jss.2014.11.036.

Melo F. S. and Veloso M. 2012. Heuristic planning for decentralized MDPs with sparse interactions. Springer Tracts in Advanced Robotics, 83 STAR. (11): 329-343. http://doi.org/10.1007/978-3-642-32723-0_24.

Mishra D., Mishra A. and Ostrovska S. 2012. Impact of physical ambiance on communication, collaboration and coordination in agile software development: An empirical evaluation. Information and Software Technology, 54(10), 1067–1078. http://doi.org/10.1016/j.infsof.2012.04.002.

Nguyen-Duc A., Cruzes D. S. and Conradi R. 2014. The impact of global dispersion on coordination, team performance and software quality - A systematic literature review. Information and Software Technology. 57, 277-294. http://doi.org/10.1016/j.infsof.2014.06.002.

Oakley B. a., Hanna D. M., Kuzmyn Z. and Felder R. M. 2007. Best practices involving teamwork in the classroom: Results from a survey of 6435 engineering student respondents. IEEE Transactions on Education, 50(3): 266-272. http://doi.org/10.1109/TE.2007.901982.

Ocker R. J. 2007. A balancing act: The interplay of status effects on dominance in virtual teams. IEEE Transactions on Professional Communication. 50(3): 204-218. http://doi.org/10.1109/TPC.2007.902656.

Park C., Pattipati K. R., An, W. and Kleinman D. L. 2012. Quantifying the Impact of Information and Organizational Point-to-point Communication Structure. 42(1): 68-86.

Petre M. 2010. Mental imagery and software visualization in high-performance software development teams. Journal of Visual Languages & Computing, 21(3): 171-183. http://doi.org/10.1016/j.jvlc.2009.11.001.

Philip T., Wende E. and Schwabe G. 2012. Exploring failures at the team level in offshore-outsourced software development projects. Lecture Notes in Business Information Processing, 130 LNBIP, 194-211. http://doi.org/10.1007/978-3-642-33920-2_12.

Pressman R. S. 2009. Software Engineering: A Practioner's Approach.

Quesada F. J., Palomares I. and Martínez L. 2015. Managing experts behavior in large-scale consensus reaching processes with uninorm aggregation operators. Applied Soft Computing. http://doi.org/10.1016/j.asoc.2015.02.040.

Richardson I., Casey V., McCaffery F., Burton J. and Beecham S. 2012. A process framework for global software engineering teams. Information and Software Technology. 54(11): 1175-1191. http://doi.org/10.1016/j.infsof.2012.05.002.

Rogers a., Farinelli a., Stranders R. and Jennings N. R. 2011. Bounded approximate decentralised coordination via the max-sum algorithm. Artificial Intelligence. 175(2): 730-759. http://doi.org/10.1016/j.artint.2010.11.001

Saldaña-Ramos J., Sanz-Esteban a., García-Guzmán J. and Amescua a. 2012. Design of a competence model for testing teams. IET Software. 6(5): 405. http://doi.org/10.1049/iet-sen.2011.0182.

Sanchez-Anguix V., Julian V., Botti V. and Garcia-Fornes A. 2012. Reaching unanimous agreements within agent-based negotiation teams with linear and monotonic utility functions. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 42(3): 778-792. http://doi.org/10.1109/TSMCB.2011.2177658.

Santos V., Goldman A. and de Souza C. R. B. 2014. Fostering effective inter-team knowledge sharing in agile software development. Empirical Software Engineering. http://doi.org/10.1007/s10664-014-9307-y.

Sarker S., Sarker S. and Schneider C. 2009. Seeing remote team members as leaders: A study of US-Scandinavian teams. IEEE Transactions on Professional Communication. 52(1): 75-94. http://doi.org/10.1109/TPC.2008.2007871.

Savelsbergh C. M. J. H., Poell R. F. and van der Heijden B. I. J. M. 2015. Does team stability mediate the relationship between leadership and team learning? An empirical study among Dutch project teams. International Journal of Project Management, 33(2): 406-418. http://doi.org/10.1016/j.ijproman.2014.08.008.

Sha X. and Chang K. T. 2012. Research article the role of leadership and contextualization on citizenship behaviors in distributed teams: A relational capital perspective. IEEE Transactions on Professional Communication, 55(4): 310-324. http://doi.org/10.1109/TPC.2012.2188595.

Smari W. W., Clemente P. and Lalande J.-F. 2014. An extended attribute based access control model with trust and privacy: Application to a collaborative crisis management system. Future Generation Computer Systems. 31, 147-168. http://doi.org/10.1016/j.future.2013.05.010.

Stone P., Kaminka G. a., Kraus S., Rosenschein J. S. and Agmon N. 2013. Teaching and leading an ad hoc teammate: Collaboration without pre-coordination. Artificial Intelligence, 203, 35-65. http://doi.org/10.1016/j.artint.2013.07.003.

Stryker J. B., Santoro M. D. and Farris G. F. 2012. Creating collaboration opportunity: Designing the physical

workplace to promote high-tech team communication. IEEE Transactions on Engineering Management. 59(4): 609-620. http://doi.org/10.1109/TEM.2011.2170995.

Sulayman M., Riaz M., Naqvi H. and Aman M. A. 2009. Using area-team to evaluate the software architecture for an online banking system: A case study. Communications in Computer and Information Science, 59 CCIS, 111-118. http://doi.org/10.1007/978-3-642-10619-4_14.

Sun X., Lin H. and Xu K. 2015. A social network model driven by events and interests. Expert Systems with Applications. 42(9): 4229-4238. http://doi.org/10.1016/j.eswa.2015.01.020.

Sutanto J., Kankanhalli A. and Tan B. C. Y. 2011. Deriving IT-mediated task coordination portfolios for global virtual teams. IEEE Transactions on Professional Communication. 54(2): 133-151. http://doi.org/10.1109/TPC.2011.2121630.

Tenenberg J. 2008. An institutional analysis of software teams. International Journal of Human-Computer Studies. 66(7): 484-494. http://doi.org/10.1016/j.ijhcs.2007.08.002.
Tessem B. 2014. Individual empowerment of agile and non-agile software developers in small teams. Information and Software Technology. 56, 873-889. http://doi.org/10.1016/j.infsof.2014.02.005.

Toffolon C. and Dakhli S. 2009. KNOC: A Knowledge-Oriented Cooperative Software Development Process. Information Systems Development. http://doi.org/10.1007/978-0-387-68772-8_38.

Tuffley D. 2012. Optimising virtual team leadership in Global Software Development. IET Software. 6(3): 176. http://doi.org/10.1049/iet-sen.2011.0044.

Wale-Kolade A. Y. 2015. Integrating usability work into a large inter-organisational agile development project: Tactics developed by usability designers. Journal of Systems and Software. 100, 54-66. http://doi.org/10.1016/j.jss.2014.10.036.

Weimann P., Pollock M., Elsje S. and Brown I. 2013. Enhancing Team Performance through Tool Use : How Critical Technology-Related Issues In fl uence the Performance of Virtual Project Teams. IEEE Transactions on Professional Communication. 56(4): 332-353.

Wu F., Zilberstein S. and Chen X. 2011. Online planning for multi-agent systems with bounded communication. Artificial Intelligence. 175(2): 487-511. http://doi.org/10.1016/j.artint.2010.09.008.

Yang Q., Yao T., Lu T. and Zhang B. 2014. An Overlapping-Based Design Structure Matrix for Measuring Interaction Strength and Clustering Analysis in Product Development Project. IEEE Transactions on Engineering Management. 61(1): 159-170. http://doi.org/10.1109/tem.2013.2267779.

Yannibelli V. and Amandi a. 2012. A memetic algorithm for collaborative learning team formation in the context of software engineering courses. … in New Technologies, Interactive Interfaces and …, 7547 LNCS, 92-103. http://doi.org/10.1007/978-3-642-34010-9_9.

Yuan M., Zhang X., Chen Z., Vogel D. R. and Chu X. 2009. Antecedents of coordination effectiveness of software developer dyads from interacting teams: An empirical investigation. IEEE Transactions on Engineering Management, 56(3): 494-507. http://doi.org/10.1109/TEM.2008.927819.