

Dependency Test: Portraying Pearson's Correlation Coefficient Targeting Activities in Project Scheduling

Jana Shafi

Department of Computer Science,
Prince Sattam bin Abdul Aziz
University,
Saudi Arabia

Amtul Waheed

Department of Computer Science,
Prince Sattam bin Abdul Aziz
University,
Saudi Arabia

Sumaya Sanober

Department of Computer Science,
Prince Sattam bin Abdul Aziz
University,
Saudi Arabia

Abstract—In this paper, we discuss project scheduling with conflicting activity-resources. Several project activities require same resources but, may be scheduled with the certain lapse of time resulting in repeatedly using the same kind of resources for executing dissimilar activities. Due to the frequent usage of same resources multiple times, expenditure become more expensive and project duration extends. The problem is to find out such kind of activities which are developing implicit relations amid them. , we proposed a solution by introducing TVs (Transparent view of Scheduling) model. First, we analyze and enlists activities according to required resources, categorize them and then we segregate dependent and independent activities by indicating a value. Performing Dependency test on activities by using Pearson's Correlation Coefficient (PCC) to calculate the rate of relations among the ordered activities for similar resources. By using this model we can reschedule activities to avoid confusion and disordering of resources without consumption of time and capital.

Keywords—TVS; Transparent; Dependency; PCC; Activity; Resource; Schedule; Project Introduction

I. INTRODUCTION (MOTIVATION)

With the development of project management system in order to direct the project schedules along with the activities and various resources, project scheduling becomes the helpful tool for standardizing and ordering activities as well resources according to specifications [19, 20]. However, project schedules returned by popular, methods are not satisfactory .Sometimes much of the resources are linked with the activities which have nothing to do with them .It is a remarkable part for the developers to pick resources in order to get an appropriate activity. As we have experience in using project schedules every day, the result set returned by estimation of activity resource set is really too big and mostly and merely useless. The relationship between the activities is obvious to users or managers, while it is not for the project schedules [21].

Late delivery of software projects results in huge loss of manpower, industrial efforts and money which discouraged our software industry less or more for accepting challenges for successful projects in recent years[1].Developers, Engineers, and researchers gave a thought of planning all parameters of software management in an organized manner so that the modules of the project become transparent with due time. This outcome called as Project scheduling which minimizes the failure of projects and encourages workers morale[1][2]. For Complex projects mostly engineering tasks take place in

parallel so one work may be interdependent in another work or task. These interdependencies can be understood by schedules only.

An activity – Must have a clear start and a clear stop – Must have a duration that can be forecasted – May require the completion of other activities before it begins – should have some ‘deliverables’ for ease of monitoring.

Resources- they are used to accomplish the project activities. Resources are classified in vast areas such as manpower, finance, and budgeting, inventories, maintenance and services are required for the completion of the task in limited time.

In order to show where the problem is we input the following activity and resources: Activity1”Questionnaire: Public” Activity2 “Demonstrating new software: Public Resources: Vehicles, Chauffer, Maps, Locations.

TABLE I. EXAMPLE OF ACTIVITY RESOURCE SCHEDULING

S.No	Activity	Resources	Dependents
1	Questionnaire Public	Location,Vehical,Chauffer, Map	Yes
2	Demonstrating New Software Public	Location,Vehical,Chauffer, Map	Yes
3	System Configuration	Software and Hardware	No

We can exemplify from above(table1:) that as both the activities requires some common resources for example Vehicles, Chauffer, Maps, and Locations which set both the activities having invisible relation. Let’s analyze the aforesaid activities. We want resources in order to accomplish activity1” Questionnaire: Public” which requires “Vehicles, Chauffer, Maps, Locations” and activity2 “Demonstrating new software: Public” requires the same kind of resources as activity1 as mention above. However, both the activities are totally different.

Colliding resources with more than one activity leads an invisible relation among activities. Colliding resources is the key to the whole problem in prioritizing the activities.

Nothing in this boundless universe exists independently. All objects are related to other activities in various means. We comprehend this activity from the way it relates to other

activities. Regard activity1 "Questionnaire" and activity2 "Demonstration" one of the relations between them are resources "Vehicles, Chauffer, Maps, Locations" in this context. For better understanding the relationships between activities have to be defining before the developers, a tester understands the semantic of each other. In the project management system, schedules are presented by the activity estimation of resources for example "System Configuration" an activity presented by "IT block" resources they dealt with activity resource combinations. However, it is not always necessary that all activities must be related to each other or with any other kind of activities. Activities which carry out in an individual manner are called isolated activities.

II. RELATED WORK

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

With the development of Project Scheduling, hundreds of methods, ways established algorithms. Project scheduling in the recent years observed along heuristics, constraint-resources, metaheuristics, and resource-based constraints, consistency tests are furnished[4,14,18]. E.L. Demeulemeester and W.S. Herroelen study and display depth knowledge of project scheduling by several algorithms and classify them accordingly for better practical examples are computed [24].

J.Alcaraz focused Genetic Algorithm counting resource allocations[10].Dale.F.Cooper suggested project scheduling a problem with multiple constrained resources with an experimental investigation with a set of the project their characteristics is scheduled by each of these heuristics with a variety of priority rule[8,17]. J-H-Cho and Y-D Kim emerges with another simulated annealing algorithm for resource constrained project scheduling proble[9].Christian Artigues presented the flow network model for static and dynamic resource-constrained project scheduling[13,15].Peter Brucker encompasses notations, classifications, models and methods of project scheduling problems[11,12,16]. And R.Kolisich compiled a survey on deterministic project scheduling remarking net present value maximization and make span minimization[7].The search algorithm for the resource constrained project scheduling problem with an interval is defined with a solution by representing resource flows extending the disjunctive graph model for shop scheduling problems by Poppenborg and Kust in 2016.

Execution of task by generalizing precedence relationships and assigning resources for completion of the task in given deadline is deliberated by Bianco et al. for defining a problem of leveling resources[25].To decrease the resource consumption a method is introduced by author Naffaf et

al.(2016) which consist of three mixed-integer mathematical programs and an adapted satisfiability test for the cumulative constraint.

Pearson correlation coefficient is a formula designated in a statistical test between two variables to determine how strong that relationship is. The value "0" indicates no linear relationship and value "1" indicates a positive linear relationship and "-1" indicates a negative relationship. R denotes Pearson correlation coefficient value [22 , 23].

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n \sum x^2 - (\sum X)^2][n \sum Y^2 - (\sum Y)^2]}} \quad (1)$$

III. PROBLEM STATEMENT

The main issue address in this paper is how to identify the relations between the various activities which carried out in project scheduling later in project development. The resources are viewed as independent or interdependent. Currently, a challenge is to know when prioritizing resources using an activity, how many activities are implicitly interconnected to each other which should be recognized to identify the dependent activities.

IV. PROPOSED SYSTEM

In this paper, we concentrate on the framework of activities along with computation and manipulation in order to identify, sort as well categorize dependent and independent activities.

A. Paper Organization

In the paper Section IV will introduce the system of "Transparent view of scheduling"TVS , depicts its various classification and its formal model.Section V & VI implements TVs and manipulates results via Pearson correlation coefficient.

Schedule's point of view can be understood by TVS Model which is able to provide a transparent view of interdependent and independent tasks. TVS provides a computational framework for network activities. TVS abandons complication at the workplace. Thus in TVS everything (activity) belongs to some or the other category which separates various activities from to be get confused or left behind. As we are using mathematical computation to be more focused on points

With each task efforts and duration of time are allocated and thus a task is a part of a network that aware the software team to meet the product delivery deadline.

Fig.1 proper scheduling is essential for the project which an experienced team can do [1] and must include

- a) Tasks must be created inside the network as shown in.
- b) Efforts and timings are allocated to each task
- c) Interdependencies between tasks must be transparent
- d) Resources must be allocated for the targeted work



Fig. 1. Activity Network

Work done for both optimistic and pessimistic scheduling in order to get more realistic parameters for the project to proceed.

Fig.2 Transparent view of scheduling enables specification viewers as well as programmers and clients to instantly associate their activity in concerning categories to carry out process interdependently or independently according to the requirements and resources.

TVS includes

a) Outsource Activities-Those activities which include all external resources,external components,external behavior and resources other than usual which are or will affect our activities in near or far future.For Example:Power-shortage,Politics,Rate ofInflation etc

b) Insider-Activities which can be completed within organization including manpower, resources, and Coordination.(internal resources)

c) Dependent-Some insider and outsource activities cannot be completed without each other. These are the proportional activities which change with the variance of each other.

For example:-Detail Marketing Plans sales tax, Bond insurance etc

For example:Clients,Distributors,advertisement etc.

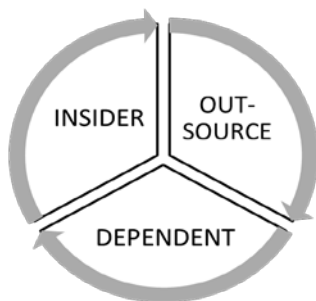


Fig. 2. Transparent view of scheduling

V. MODALITIES WITH EXAMPLE

A. As we already know that we have TVS classes as Inside,Outsider,Dependent which we abbreviated for our convenience in the following way.

- Inside- I_n ,
- OutSource- O_s ,
- Dependent- D_e

B. Formal Model

Figure.3 portrays how the activities are implicitly dependent on each other for resources.

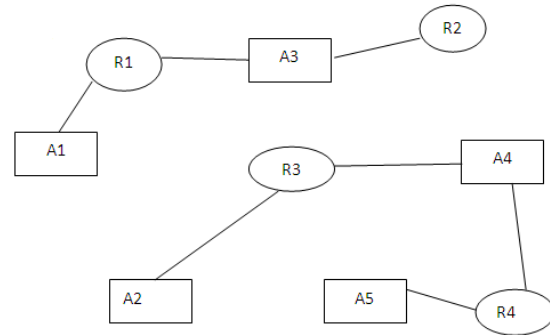


Fig. 3. Resource-Activity Relationship Model

Definition 1: A keyword 'I_s' is a set of activities which Include internal resources.

Definition 2: A concept of activities C in a given schedule is Presented by a square vertex of the graph.

Definition 3: A relation collection in a given domain R is a Set of related activities. It represented by arch of the graph.

Definition 4: A graph of an activity relation G is a set of Vectors of the form (A, R) where A is activity Set and is a relation.

Definition 5: An activity relation subgraph G_p is a subgraph of G.

Definition 6: Keyword 'O_s' is a set of activities belongs to outside resources.

Definition 7: An activity-resource pair set $RksetI_sO_s$ is a set of activities of the form($Kis, Kos, Risos$) where Kis and Kos where Kis refers to Insider, Kos -Outsource $Rqisos$ -is a relation between Ki and Ko .

Definition 8:An activity-resource pair candidate set $CRksetp$ is a set an activity-resource pair set and every $Rqisos$ presents an arc of G_p .

Definition 9:A result set $U(Kis, Kos, Rqisos$ is the returned Dependency value when we correlate an the activity of $CRksetp$ into TVS.

C. Demonstration

We are going to demonstrate how to classify activities and to get a correlation coefficient of 'I_n' and 'O_s'.In figure 3,

-0.933	- 0.200	0.871	0.040	0.187
-0.933	- 0.200	0.871	0.040	0.187
M _X : -0.933	M _Y : - 0.200	SUM: 1.867	SUM: 4.800	SUM: 0.400

B. Key

- a) I_s: I_s Values
- b) O_s: O_s Values
- c) M_I: Mean of I_s Values
- d) M_O: Mean of O_s Values
- e) I_s – M_I & O_s – M_O: Deviation scores
- f) (I_s – M_I)² & (O_s – M_O)²: Deviation Squared
- g) (I_s – M_I)(O_s – M_O): Product of Deviation Scores

C. Result Details & Calculation

I_s Values

$$\sum = 28$$

$$\text{Mean} = 0.933$$

$$\sum(I_s - M_I)^2 = SSI = 1.867$$

O_s Values

$$\sum = 6$$

$$\text{Mean} = 0.2$$

$$\sum(O_s - M_O)^2 = SSO = 4.8$$

I_s and O_s Combined

$$N = 30$$

$$\sum(I_s - M_I)(O_s - M_O) = 0.4$$

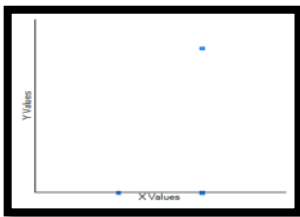
R Calculation

$$r = \frac{\sum((I_s - M_I)(O_s - M_O))}{\sqrt{((SSI)(SSO))}}$$

$$r = 0.4 / \sqrt{((1.867)(4.8))} = 0.1336$$

Meta Numeric (cross-check)

$$r = 0.1336$$



Depicting Correlation Graph

Result clarifies that insider is not dependent on outsource And vice-versa.

The value of R is 0.1336. Although technically a positive correlation, the relationship between variables is weak (nb. the nearer the value is to zero, the weaker the relationship).The value of R², the coefficient of determination, is 0.0178.

D. Performance Analysis

a) First, the five text boxes represent the calculations that would be required if you were to calculate the R value in stages.

b) Second, there is more than one way to calculate the R value, but these are all mathematically equivalent.

c) Third, in the "Result Details & Calculations" box, you'll find what we've called a cross-check value, which is the R value calculated using an algorithm supplied by the [Meta Numeric](#) statistical library. This should be identical to the value that we've calculated.

d) Forth, In the graph Depicting Correlation also we showed you that as the result is zero means the Dependability between Insider and Outsider activities is not proved hence no correlation.

e) Hence this experiment proved that how we can weight our activities as well classify them as dependent or independent

VII. APPLICATION

The concept can be useful to accurately correlate activities according to their nature which gives perfect scheduling/sequence to carry out in an organization.as well activities can be assigned independently or dependently to the team or individuals according to the requirements.

PCC Result directly classifies the existence of activities correlation by computing values positive , negative and neutral.PCC result is represented in graphical view for enhanced perceptive of determining dependent or independent activities.

TVs model is efficient in communication among team leaders who are responsible for project modules .This also helps in controlling cost as usage of resources is clearly defined.The concept is useful in managing changes in schedules which become an easy job and it can adjust according to the environment, requirements and time.Project Schedule ,module's responsibilities are simply assigned and allotted with a time limit for carrying out the project on time. TVs serves as a virtual model to track progress according to plan , measure quality , cost control in real-time schedules. Project schedules can be structured in a hierarchical form which enables in easily identifying the relationship among activities. The model portrays flow of activities and resources simultaneously which assist in debugging errors.

VIII. SUMMARY & CONTRIBUTION

In the project scheduling, estimation of activity-resource is recorded and prioritize. We call the activity defining the resource and its relations among various activities.Here the question arises is how to define the activity resource? Whether the relation should be activity -activity- resource? The answer is no.

The paper proposes "TVS" activity-resource-activity correlation base scheduling since it takes an advantage of activity-resource estimation and achieves whole correlation value of activities. Defining schedule of activity resource can be standardized by correlating them and weighing them.

Initially, we explore and procure activities according to essential resources, sort out and segregate Insider (dependent) and outsource(independent)activities by denoting with a value 0 for existence and 1 for non-existence.

Performing Dependency test on activities by using Pearson's Correlation Coefficient (PCC) to calculate the rate of relations among the ordered activities for similar resources.

The Key idea in Project scheduling is the sequence of events according to their category Insider(I_s),Outsider(O_s) and Dependent(De).This paper presents a close view of complex, time-taking Project Scheduling activities , an approach through which many activities can be weighed accordingly and we can easily correlate the dependency.

IX. FURTHER FUTURE WORK

Project development, scheduling and various accurate operations can be improve and manipulate in future work. The transparent view of scheduling can lead to an era of automatic generalizing events and sequences of Project scheduling engrave with computations for accuracy.

REFERENCES

- [1] Software Engineering: A Practitioner's Approach, 6/e Roger S Pressman, R.S. Pressman and Associates
- [2] Software Engineering:-Ivan Marsic,Rutgers,the state University of New Jersey,September 10,2012
- [3] Pearson product-moment correlation coefficient from Wikipedia, the free encyclopedia
- [4] V. Ahuja V. Thiruvengadam, (2004),"Project scheduling and monitoring: current research status", Construction Innovation, Vol. 4 Iss 1 pp. 19 – 31
- [5] Interpretation of the Correlation Coefficient: A Basic Review RICHARD TAYLOR, EDD, RDCS JDMS 1:35-39,January/February 1990
- [6] A survey of variants and extensions of the resource-constrained project scheduling problem, Sönke Hartmann,1,DirkBriskomb,doi:10.1016/j.ejor.2009.11.005
- [7] An integrated survey of deterministic project scheduling, Volume 29, Issue 3, June 2001, Pages 249–272 R Kolischa, , R Padmanb, , .
- [8] Heuristics for Scheduling Resource-Constrained Projects: An Experimental Investigation Dale F. Cooper Royal Holloway College (University of London),Published Online: July 1, 1976 Page Range: 1186 – 1194Journal of the Operational Research Society(1997)48,736–744.
- [9] A simulated annealing algorithm for resource constrained project scheduling problemsJ-H Choi and Y-D Kim2February 2001, Volume 102, Issue 1-4, pp 83-109
- [10] A Robust Genetic Algorithm for Resource Allocation in Project Scheduling J. Alcaraz, C. Maroto
- [11] Resource-constrained project scheduling: Notation, classification, models, and methods Peter Bruckera, 1, 1, Andreas Drexlb, , Rolf Möhringc, 2,2,Klaus Neumannd, 3, 3, Erwin Pesche, 4, 4
- [12] Bianco L, Caramia M (2011) A new lower bound for the resource-constrained project scheduling problem with generalized precedence relations. *Comput Oper Res* 38:14–20
- [13] Cavalcante CCB, de Souza CC, Savelsbergh MWP, Wang Y, Wolsey LA (2001) Scheduling projects with labor constraints. *Discrete Appl Math* 112(1–3):27–52
- [14] de Reyck B, Demeulemeester EL, Herroelen WS (1999) Algorithms for scheduling projects with generalized precedence relations. In Węglarz, pp 77–106
- [15] Drezet L-E, Billaut J-C (2008) A project scheduling problem with labour constraints and time-dependent activities requirements. *Eur J Oper Res* 112((1):217–225
- [16] Hartmann S (1999) Project scheduling under limited resources: models, methods, and applications. Number 478 in *lecture notes in economics and mathematical systems*. Springer, Berlin
- [17] Hartmann S, Briskorn D (2010) A survey of variants and extensions of the resource-constrained project scheduling problem. *Eur J Oper Res* 207:1–14
- [18] Klein R (2000) Project scheduling with time-varying resource constraints. *Int J Prod Res* 38:3937–3952
- [19] Węglarz, J (eds) (1999) Project scheduling: recent models, algorithms, and applications. Kluwer, Dordrecht
- [20] Editorial “Project Management and Scheduling” Rainer Kolisch1 · Erik Demeulemeester2 · Rubén Ruiz Garcia3 · Vincent T Kindt4 · Jan Węglarz5 Published online: 3 March 2016 © Springer-Verlag Berlin Heidelberg 2016
- [21] J. Józefowska, M. Mika, R. Różycki, G. Waligóra, and J. Węglarz, “A heuristic approach to allocating the continuous resource in discrete-continuous scheduling problems to minimize the makespan”, *Journal of Scheduling* 5 (6), 487–499 (2002).
- [22] Park, E., and Lee, Y. J., 2001, Estimates of the standard deviation of Spearman's rank correlation coefficients with dependent observations: *Comm. Statist. Simul.*, v.30, no.1, p. 129-142.
- [23] Robinson, P. M., 1977, Estimation of a time series model from unequally spaced data: *Stochast. Proc. Appl.*, v.6, no.1, p. 9-24.
- [24] E.L. Demeulemeester and W.S. Herroelen, *Project Scheduling – A Research Handbook*, Kluwer, Boston, 2002
- [25] Bianco L, Caramia M, Giordani S (2016) Resource leveling in project scheduling with generalized precedence relationships and variable execution intensities. doi:10.1007/s00291-016-0435-1 (this issue)