AMCIS 2008 Proceedings

Americas Conference on Information Systems (AMCIS)

2008

Leveraging Group Cohesiveness to Form Workflow Teams

Akhilesh Bajaj
University of Tulsa, akhilesh-bajaj@utulsa.edu

Robert Russel
University of Tulsa, rrussel@utulsa.edu

Recommended Citation

Bajaj, Akhilesh and Russel, Robert, "Leveraging Group Cohesiveness to Form Workflow Teams" (2008). *AMCIS 2008 Proceedings*. 347.
http://aisel.aisnet.org/amcis2008/347

Leveraging Group Cohesiveness to Form Workflow Teams

Akhilesh Bajaj
University of Tulsa
akhilesh-bajaj@utulsa.edu

Robert Russell
University of Tulsa
rrussell@utulsa.edu

ABSTRACT

Past literature shows that workflows will be performed with greater efficiency and/or effectiveness if workflow teams have higher group cohesiveness. The major contribution of this work in progress is the creation and implementation of a formal generalized methodology that incorporates ideas from two diverse fields: social network theory and workflow modeling, and allows optimization of work groups along group cohesiveness. In order to implement this model we present newly created algorithms to structure and represent the problem of workflow load representation, possible team sets and social network metric optimization so that standard integer programming solvers can attempt to solve it.

Keywords: Workflows, social networks, cohesiveness

INTRODUCTION

Business processes usually consist of a number of interrelated activities, performed by people with different skill sets or roles. Recently, there has been extensive interest in *workflow management systems* that provide a means of automating the allocation and scheduling of the activities constituting workflows. As pointed out in (Governatori Rotolo and Sadiq 2004), most activities in an organization are executed by a number of agents, possessing different skill-sets. In their description of the dynamic allocation of workflows to agents, (Kumar Aalst and verbeek 2001) point out several shortcomings with current allocation approaches. These limitations primarily stem from an inability to deal with uncertainty, and to incorporate knowledge about the organization into the allocation process. This lack of the incorporation of organizational knowledge is also recognized in other works such as (Casati Castano and Fugin 2001; Governatori et al. 2004; Momotko and Subieta 2002).

Workflow allocation has also been studied in the *job design* literature in human resource management (Corts forthcoming; Hemmer 1995; Holmstrom and Migrom 1991; Itoh 1994; Valsecchi 1996). Typical issues in this area include trade-offs between serial and parallel production, employee motivations and incentives management. However, studies in this area also have not typically considered intra-organizational relationships between agents when allocating workflows, which is the focus of this work.

A rich body of literature dealing with *social networks* (SN) exists in the social sciences. The sociocentric approach to SN (Moreno 1934; Simmel 1955; Wasserman and Faust 1999) typically deals with measuring and quantifying the relationships between individuals in a group. The focus is on measuring the structural patterns of interaction and how these patterns can explain outcomes. Graph based analyses in the area utilize measures like *centrality* and *group cohesiveness* to characterize different groups, based on individual links between members (nodes) of that group. Centrality of each member can be characterized by the number of other nodes to which the member is linked. Cohesiveness reflects how reachable a node is, on average, from any other node in the network.

In this work, we take a first step *towards incorporating ideas from SN to the problem of workflow allocation*. The primary contribution of this work is a methodology that creates workflow teams of human actors within an organization, based on optimizing group or team cohesiveness. Past literature (Austin and Bobko 1985; Evans 1986; Mudrack 1989; Thomas and Griffin 1983) shows that the constituent activities of the workflows will be performed with greater efficiency and/or effectiveness if the social characteristics of the workgroup of actors are optimized along an SN measure, such as group cohesiveness. Our contribution here is a methodology that can used to allocate a generic set of workflows within an organization, to optimize workgroups along group cohesiveness. The rest of this paper is organized as follows. In section 2, we describe earlier work on workflow allocation and social networks. Section 3 consists of a detailed description of our methodology, including the model, algorithms, and an illustrative example. We conclude in section 4.

PREVIOUS WORK

Workflow Allocation

Workflow allocation has been investigated the automated **workflow management systems** (WFMS) area, where attention has been paid to the dynamic allocation of activities. (Governatori et al. 2004) point out that most WFMSs refer to underlying organizational role lists in order to allocate activities to machines accessible by agents (which could be a human or an information system) who can perform these roles. (Kumar et al. 2001) provide several shortcomings in the activity allocation methods of WFMSs, many of which can be attributed to a lack of organizational knowledge on the part of the WFMS. One of the pioneering attempts to overcome these limitations is presented in (Kumar et al. 2001), with the use of object constraint language (OCL) to model teams of agents and their relationships in an organization. A limitation of this approach is that OCL does not support concepts that are usually used to characterize organizational relationships. Similarly, (Momotko et al. 2002) use an object-oriented language to model organizational constraints, with the same limitations. A third approach in (Casati et al. 2001) uses the event-condition-action framework to model organizational constraints. Finally, (Russell Aalst Hofstede and Edmond 2005) describe various patterns of resource allocation in workflows such as escalation and deallocation. Our methodology presented here complements these approaches in that it is a static methodology which can be implemented alongside a dynamic selection scheme in a WFMS. Further, we explicitly utilize social network metrics, which traditionally have closer ties to organizational modeling.

Social Network Measures

A social network is a structure whose nodes represent members in a social context and whose edges represent links such as interaction, collaboration or influence between the nodes (Liben-Nowell and Kleinberg 2003). SN analysis has attracted considerable interest from social and behavioral scientists over the last few decades (Moreno 1934; Wasserman et al. 1999). Recently, researchers in artificial science and data mining have also recognized that an organization can benefit from the interactions within the informal social network amongst its members that can often supplement the official hierarchy imposed by the organizational chart (Kautz Selman and Shah 1997; Raghavan 2002). Social networks can be represented using graphs (with the members as nodes and the interaction links as edges) or socio-matrices, where the sending members are the rows and the receiving members are the columns. In this work, we utilize a sociometric notation, which is the most widely used today (Wasserman et al. 1999).

Several measures have been used in the SN literature to characterize a network, from the perspective of either one actor, or a group. Measures from a member standpoint include the centrality and the prestige of the actor in a SN, with finer definitions including degree centrality, closeness centrality and betweenness centrality (Freeman 1979). Group level measures include the overall group (or team) cohesiveness, which was initially defined as the “forces that act on members to stay in the group” (Festinger 1950).

There is considerable support in sociological theory that network cohesion or cohesiveness is an important explanatory variable in studying the emergence of consensus among members of a group (Collins 1988; Friedkin 1984). Group cohesiveness is thought to drive an array of variables including group homogeneity and influence of group norms.

Group cohesiveness has also been shown in numerous studies to have a positive impact on performance in the workplace. Examples of earlier work include (Thomas et al. 1983) who concluded that increased cohesiveness fosters better task redesign. (Austin et al. 1985) demonstrated that participative goal setting is more effective in cohesive groups. (Evans 1986; Mudrack 1989) investigated how increased cohesiveness reduces the negative effects of situational constraints on organizational behavior. (Mudrack 1989) summarizes several studies that show how group cohesiveness leads to improved productivity in teams where the cohesiveness is based on task focus, as opposed to social integration.

More recent studies further strengthen the positive link between team cohesiveness and workplace performance. (Seers Petty and Cashman 1995) report a study of 103 manufacturing workers that found a positive association between job satisfaction, internal work motivation and team cohesiveness. (Sanders and Nauta 2004) showed how increased cohesiveness results in lower employee absenteeism. In (Steinmark 2002) low cohesiveness was found to negatively affect creative group work such as brainstorming, as well as more routine tasks. In (Jordan Feild and Armenakis 2002) multiple tasks were used on 50 army teams to determine the effects of social cohesion, group potency and team-member exchange on team performance. Social cohesion was found to have a positive effect on physical performance, mental performance and the assessment of the team by the commander.

The cohesiveness of subgroups within a larger SN can be measured based on four different factors: a) the degree of mutuality of the links between the members in the subgroup, *i.e.* to what extent do the members of the subgroup choose each other?, b) the reachability from one member to another in the subgroup, c) the frequency of ties between members which measures the degree to which members have ties to one another within a subgroup against a theoretical maximum of ties to all members in the subgroup, and d) the relative frequency of ties which measures the frequency of ties of the subgroup relative to the ties in the larger SN.

Measures of subgroup cohesiveness reflect one of these four factors. For example, (Bock and Husain 1950) proposed a ratio of the average strength of ties in a subgroup to the average strength of ties of members in the subgroup to SN members outside the subgroup, *i.e.* utilizing factor d) in the discussion above. (Alba 1973) viewed just the average strength of ties in the subgroup as the “centripetal” force holding the subgroup together, while the strength of ties of members in the subgroup to members outside the subgroup was viewed as a “centrifugal” force. Excellent summaries of different subgroup cohesiveness measures can be found in (Sailer and Gaulin 1984) and more recently in (Wasserman et al. 1999).

In this work, we leverage the notion of SN metrics such as cohesiveness in forming workflow groups and take a first step in proposing a methodology that allows for the optimal formation of workflow groups. The optimizing criterion used in this work is a measure of subgroup cohesiveness of the workflow subgroup, *though other measures for cohesiveness may also be used, without loss of generality*. Next, we describe our methodology.

CREATING SUBGROUPS BASED ON OPTIMIZED COHESIVENESS

We first present a general model of a workflow load, member roles or skill sets, and a depiction of all the possible member subgroups (or teams) for each workflow in the load. Next, we present algorithms to formulate the problem of optimizing a social network metric as an integer program that can fed to solvers such as CPLEX (www.ilog.com).

Generalized Workflow Model

Figure 1 summarizes the notation used in our formulation.

M: The set of |M| = m members, each element shown as m_i

W: The set of |W| = w workflow instances, each element shown as w_j

R: The set of |R| = r roles, each element shown as r_k

X: A relation between M and R, depicting if member m_i can perform role r_k , each element shown as x_{ik}

Y: A relation between R and W, depicting the relative amount of time expended by each role in a workflow, each element shown as y_{kj}

t: The total time that will be taken to execute all workflow instances under consideration

Figure 1. Summary of notation used in the formulation

In an organization, let there be a set of members:

$M = \{m_i, i = 1,..,m\}$

and a set of workflow_instances:

$W = \{w_j, j = 1, .. ,w\}$

A workflow_instance[1] requires a set of roles that perform various activities within the workflow.

Let there be the set of roles in the organization: $R = \{ r_k, k = 1,..,r\}$

We allow each member m_i the ability to perform multiple roles, shown by the relation X:

$X = \{x_{ik}$ such that $i = 1, ..,m$ and $k=1,..,r$ and $x_{ik} = 1$ if member i can perform role k, and 0 otherwise$\}$

We capture the relative amount of time spent by each role r_k in accomplishing a workflow w_j with the following relation :

$Y=\{y_{kj}$, $k = 1,..,r$ and $j = 1,..,w$ and $\sum_{k=1}^{r} y_{kj} = 1$ for all $j = 1,..,w\}$

Let each w_j have associated a time $t_j \varepsilon R^+$ (set of positive real numbers) that can be taken to complete it, including all slack.

For each role r_k let there be c_k members who can perform that role. Then

$c_k = \sum_{i=1}^{m} X_{ik}$, $k = 1, \dots r$.

Size of Solution Space:

For each workflow_instance w_j, the solution space size is:

$$s_j = \prod_{k=1}^{r} Z_k \binom{c_k}{1} \text{ subgroups.}$$

$$= \prod_{k=1}^{r} Z_k c_k \qquad (1)$$

Where $j = 1,\dots w$ and

$Z_k = 1/c_k$ if $Y_{kj} =0$ and

$Z_k = 1$ if $Y_{kj} > 0$ and

() represents the combination symbol.

The total solution space size $s = \sum_{j=1}^{W} s_j$ subgroups

Decision Variables:

The number of decision variables = s.

We need to select one subgroup g_{jq} for each workflow_instance w_j,

Where $j = 1,\dots w$ and $q = 1, .. s_j$. (The total number of subgroups that need to be picked = w subgroups).

The decision variables are:

$$V_{jq} = \begin{cases} 1 & \text{if the subgroup is selected and} \\ 0 & \text{otherwise} \end{cases}$$

---

[1] We consider workflow instances, as opposed to workflow types, in order to enable multiple instances of the **same workflow type** to be allocated to **different groups**. For example, 500 instances of a workflow type such as handling an insurance claim, are handled as 500 different workflow instances, allocated to different possible groups.

for $j = 1, \dots w$, $q = 1, \dots s_j$

Objective Function:

Let the cohesiveness of a subgroup g_{jq} be coh_{jq}

We define the cohesiveness as the number of links between members of the subgroups, divided by the theoretical maximum number of links possible between the members. *We note that our model formulation does not depend on this definition, and an alternate definition of cohesiveness may be used here.*

The objective function we use is to **maximize** the global cohesiveness given by $\sum_{j=1}^{w}\sum_{q=1}^{s_j} V_{jq} coh_{jq}$

(2)

Constraints:

1. Each workflow_instance w_j can be assigned only one possible subgroup out of the relevant s_j .

2. Each member must not overflow his/her work time for the workload (time capacity). To model constraint 1:

$$\sum_{q=1}^{s_j} V_{jq} = 1 \text{ for all } j = 1, \dots, w \tag{3}$$

To model constraint 2:

Total time for steady state workload = $\sum_{j=1}^{w} t_j = t$

Without loss of generality, we assume each member works for a maximum of t/m time units. If we want to provide for overtime, it is straightforward to add slack to the time for each member by increasing the overall time t.

$$\text{Let } U_{kji} = \begin{cases} 1 & \text{if member } m_i \text{ is allocated role } r_k \text{ in workflow\_instance } w_j \\ 0 & \text{otherwise} \end{cases}$$

U_{kji} can be computed from V_{jq}, which would be an output of the optimizer.

$$\text{Then we have: } \sum_{k=1}^{k} \sum_{j=1}^{w} U_{kji} * t_j * Y_{kj} <= t/m \text{ For all } i = 1,..,m \tag{4}$$

Expressions (2) and (3) comprise a standard set-partitioning model. The objective is to select the optimum combination of entities to "cover" the specified workflow instances. Constraint (4) is an additional requirement to force compliance with member workload capacities. Thus, the workflow allocation formulation is a type of set-partitioning model with additional constraints.

Having formulated the 0-1 integer programming optimization model, we next present the representation and solution method.

Representation of the Model and Workflow Allocation

Our model representation requires four input matrices:

$WR=\{wr_{ij} \mid i= 1..w \wedge j = 1..r \wedge \forall\ i, j: 0 <= wr_{ij} <=1\}$

$MR=\{mr_{ij} \mid i= 1..m \wedge j = 1..r \wedge \forall\ i, j: mr_{ij}=1 \vee mr_{ij}=0\}$

$MM=\{mm_{ij} \mid i= 1..m \wedge j = 1..m \wedge \forall\ i, j: 0 <= mm_{ij} <=1\}$

$WT= \{wt_i \mid i= 1..w \wedge \forall\ i: 0 <= wr_{ij} <= \infty\}$

WR consists of W rows and R columns, and each element describes the extent of contribution of a role in that workflow instance. MR is a binary matrix that consists of M rows and R columns, and each element describes if a member is eligible to perform that role or not. MM has M rows and M columns and captures the actual links between the members in the SN. The WT column vector consists of W rows, and each element represents the time units consumed by that workflow. The overall representation and solution method is shown in figure 2.

$$\left.\begin{matrix} WR \\ MR \\ MM \\ WT \end{matrix}\right\} \rightarrow \text{ Solution Space Representation (WI) } \rightarrow \text{ Integer Program Solver(CPLEX)}$$

Figure 2. Overall Representation and Solution Method

For each workflow *j*, the solution space matrix WI_j represents all the possible member allocations for performing the workflow, based on the input matrices. Each row represents one possible allocation of members that satisfies the input criteria. If a role is not utilized in the workflow, then all values in that column are set to -1. Each element in $WI_j = -1$ or a member number of a member that can potentially perform that role, based on information in MR. The solution space representation for each workflow is a generated matrix $WI_j = \{wi_{ab} \mid a = 1..s_j \wedge b = 1..r \wedge \forall\ a,b: -1 <= wr_{ab} <= m\}$

Figure 3 depicts the algorithm used to generate each WI_j matrix.

```
For each workflow $W_j$

    Compute $s_j = \prod_{k=1}^{R} Z_r \binom{c_r}{1}$ subgroups.

        Where j = 1,…w and

        $Z_r = 1/\binom{c_r}{1}$ if $Y_{rj} = 0$ and

        $Z_r = 1$ if $Y_{rj} > 0$

    Allocate matrix $WI_j$ with $s_j$ rows and |R| columns
    Initialize all elements in $WI_j$ to -1
    For each column k in $W_j$ ( k = 1 to |R|)
        If WR(j,k) >0 (role is utilized in $W_j$) then
            Sequentially write the id of all members that can provide
                that role down the kth column of $WI_j$ (one id per  element),
                repeating the list down the column until all rows of WI
                $_j$ have been filled in the kth column
            Sort the $WI_j$ matrix along the kth column, in ascending order
                utilizing a standard sort procedure such as Quick sort
        End if
    End For k
End For j
```

Figure 3. Algorithm for Generating the WI_j matrix for workflow W_j

Illustrative Example

The representation and formulation presented so far is generalizable to any organization that has workflow instances that need to be amongst members with different skill sets, where a social network metric can be optimized. The example we present next illustrates the usage of our methodology in a simple scenario, and promotes an understanding of the methodology. In this scenario, there are five members in a workgroup: Andrew, Bob, Charlene, Jane and Mary, who work at an insurance office. There are four organizational roles supported by the five group members: receptionist, sales person, office manager and claims inspector.

	Receptionist	**Sales Person**	**Office Manager**	**Claims Inspector**
Andrew	1	1	1	0
Bob	0	0	0	1
Charlene	1	0	0	0
Jane	1	1	0	0
Mary	0	0	1	1

Figure 4. MR: MxR Role Capability Matrix

Each member can perform one or more of these roles, as shown in the MR matrix in figure 4. A value of 1 in row <i,j>indicates that a member m_i can perform role r_j. There are 20 workflow_instances that need to be performed by this group, over a period of time, representing two workflow_types: handling a customer's insurance claim, and signing up a customer for an insurance policy.

Handling a customer claim requires 20% of time from a receptionist, 60% of time from a claims inspector and 20% time from an office manager. Each instance of this workflow type takes 3 hours, and there are 8 instances in the workload under consideration. Signing up a customer for a policy requires 20% of time from a receptionist, 60% of time from a sales person and 20% time from an office manager. Each instance of this workflow type takes 4 hours, and there are 12 instances in the workload under consideration. The problem the office manager faces is how to form teams to complete the workflows in this workload, so as to maximize the overall cohesiveness for the workload. For the 5 members, m_1, … m_5, we have a sociomatrix **MM** of size 5x5, as shown in figure 5.

	Andrew	**Bob**	**Charlene**	**Jane**	**Mary**
Andrew	1				
Bob	1	1			
Charlene	1	0	1		
Jane	0	1	0	1	
Mary	0	1	0	0	1

Figure 5. MM: MxM Non-directional Sociomatrix

Each cell in the matrix has a value 1 if the two members cross a pre-determined threshold that enhances cohesiveness (for example: mutual respect or "liking" for each other), otherwise it has a value zero. A binary, non-directional representation of a social network is well accepted in the literature. However, continuous values from 0 → 1 can also be used here, with no change in the methodology.

We wish to compute the optimal allocation for a steady state organizational workload of eight instances of the first workflow type and 12 instances of the second workflow type.

The total time units worked by the total of all members is:

$(8t_1+12t_2) = (8*3 + 12*4) = 72$ hours.

The **WR** matrix is shown in figure 6 for the 20 workflow_instances. The first eight represent instances of the first workflow type, and the remaining 12 represent instances of the second type.

	Receptionist	**Sales Person**	**Office Manager**	**Claims Inspector**
W1	0.2	0.0	0.2	0.6
W2	0.2	0.0	0.2	0.6
W3	0.2	0.0	0.2	0.6
W4	0.2	0.0	0.2	0.6
W5	0.2	0.0	0.2	0.6
W6	0.2	0.0	0.2	0.6
W7	0.2	0.0	0.2	0.6
W8	0.2	0.0	0.2	0.6
W9	0.2	0.6	0.2	0.0
W10	0.2	0.6	0.2	0.0
W11	0.2	0.6	0.2	0.0
W12	0.2	0.6	0.2	0.0
W13	0.2	0.6	0.2	0.0
W14	0.2	0.6	0.2	0.0
W15	0.2	0.6	0.2	0.0
W16	0.2	0.6	0.2	0.0
W17	0.2	0.6	0.2	0.0
W18	0.2	0.6	0.2	0.0
W19	0.2	0.6	0.2	0.0
W20	0.2	0.6	0.2	0.0

Figure 6. WR: WxR Workflow Roles Matrix

The **WT** Matrix in figure 7 simply lists the time unit for each workflow_instance, which in our case is three for the first eight, and four for the remaining 12 instances.

	Time (hrs)
W1	3.0
W2	3.0
W3	3.0
W4	3.0
W5	3.0
W6	3.0
W7	3.0
W8	3.0
W9	4.0
W10	4.0
W11	4.0
W12	4.0
W13	4.0
W14	4.0
W15	4.0
W16	4.0
W17	4.0
W18	4.0
W19	4.0
W20	4.0

Figure 7. WT: A Column vector of size |W|

Consider the first workflow_instance, **W1.** As shown in figure 6, it requires 20% of the time for a receptionist (which can be satisfied by Andrew, Charlene or Jane), 20% of time by an Office Manager (which can be performed by Andrew or Mary) and 60% of its time by a Claims Inspector (which can be fulfilled by Bob or Mary). The candidate subgroups that can be used to accomplish the first workflow instance are shown in figure 8 along with their cohesiveness for each subgroup. For example, the second candidate subgroup consists of Charlene, Andrew and Bob, and has a cohesiveness of 2/3 = 0.66, since there are 3 possible links between 3 people, but only two members of this group have a social link, as shown in the MM socio-matrix in figure 5.

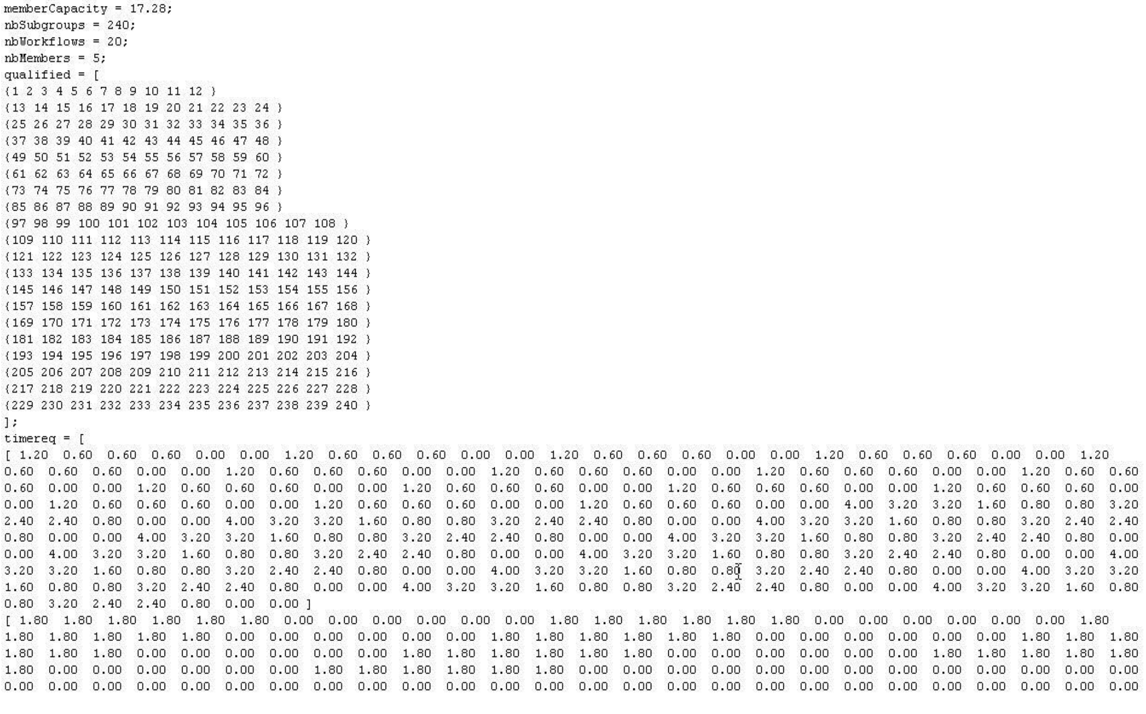
	Receptionist	**Sales Person**	**Office Manager**	**Claims Inspector**	**Cohesiveness**
S1	Andrew	Not Needed	Andrew	Bob	1.00
S2	Charlene	Not Needed	Andrew	Bob	0.66
S3	Jane	Not Needed	Andrew	Bob	0.66
S4	Andrew	Not Needed	Mary	Bob	0.66
S5	Charlene	Not Needed	Mary	Bob	0.33
S6	Jane	Not Needed	Mary	Bob	0.66

S7	Andrew	Not Needed	Andrew	Mary	0.33
S8	Charlene	Not Needed	Andrew	Mary	0.33
S9	Jane	Not Needed	Andrew	Mary	0.00
S10	Andrew	Not Needed	Mary	Mary	0.33
S11	Charlene	Not Needed	Mary	Mary	0.33
S12	Jane	Not Needed	Mary	Mary	0.33

Figure 8. W1-1: Possible Subgroups for the first Workflow Instance

The **WI** matrix for this illustrative example was generated based on the algorithm in figure 3, and lists 240 possible member sets (subgroups) as the solution space. There are 96 possible subgroups for the first eight workflow instances (12 subgroups each), and 144 possible subgroups for the second set of 12 workflow_instances (also 12 each). The first 12 rows in the **WI** matrix are shown in figure 8. It is important to note that **only one** of these subgroups will be selected in the solution set, and will represent the subgroup allocated to work on **W1**.

As part of this research, we implemented the ideas described in section 3 in Java ™. The program takes as **input** the MR, MM, WR and WT matrices, and **outputs** a file that lists all the possible subgroups for each workflow instance in the workload. This output file is then further input to a CPLEX solver that selects the subgroups that will globally optimize the cohesiveness. Figure 9(a) is a partial snapshot of the data file generated by our program for the insurance workload problem.



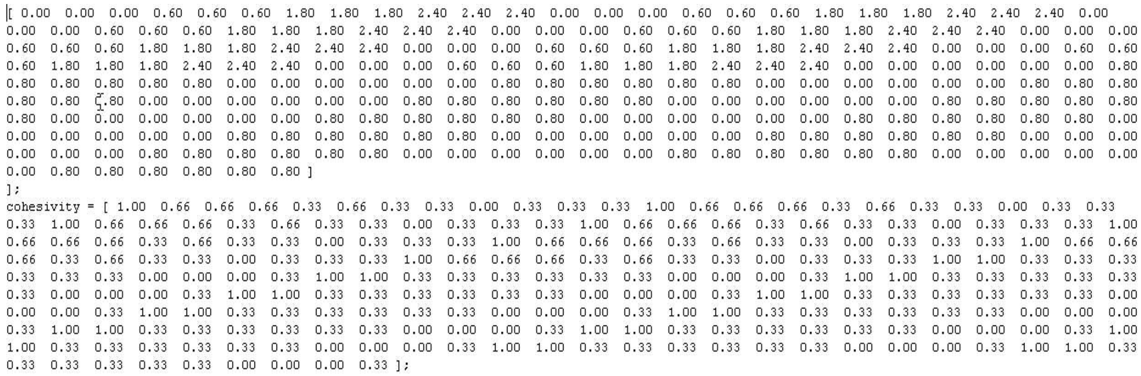
```
memberCapacity = 17.28;
nbSubgroups = 240;
nbWorkflows = 20;
nbMembers = 5;
qualified = [
{1 2 3 4 5 6 7 8 9 10 11 12 }
{13 14 15 16 17 18 19 20 21 22 23 24 }
{25 26 27 28 29 30 31 32 33 34 35 36 }
{37 38 39 40 41 42 43 44 45 46 47 48 }
{49 50 51 52 53 54 55 56 57 58 59 60 }
{61 62 63 64 65 66 67 68 69 70 71 72 }
{73 74 75 76 77 78 79 80 81 82 83 84 }
{85 86 87 88 89 90 91 92 93 94 95 96 }
{97 98 99 100 101 102 103 104 105 106 107 108 }
{109 110 111 112 113 114 115 116 117 118 119 120 }
{121 122 123 124 125 126 127 128 129 130 131 132 }
{133 134 135 136 137 138 139 140 141 142 143 144 }
{145 146 147 148 149 150 151 152 153 154 155 156 }
{157 158 159 160 161 162 163 164 165 166 167 168 }
{169 170 171 172 173 174 175 176 177 178 179 180 }
{181 182 183 184 185 186 187 188 189 190 191 192 }
{193 194 195 196 197 198 199 200 201 202 203 204 }
{205 206 207 208 209 210 211 212 213 214 215 216 }
{217 218 219 220 221 222 223 224 225 226 227 228 }
{229 230 231 232 233 234 235 236 237 238 239 240 }
];
timereq = [
[ 1.20  0.60  0.60  0.60  0.00  0.00  1.20  0.60  0.60  0.60  0.00  0.00  1.20  0.60  0.60  0.60  0.00  0.00  1.20  0.60  0.60  0.60  0.00  0.00  1.20
0.60  0.60  0.60  0.00  0.00  1.20  0.60  0.60  0.60  0.00  0.00  1.20  0.60  0.60  0.60  0.00  0.00  1.20  0.60  0.60  0.60  0.00  0.00  1.20  0.60  0.60
0.60  0.00  0.00  1.20  0.60  0.60  0.60  0.00  0.00  1.20  0.60  0.60  0.60  0.00  0.00  1.20  0.60  0.60  0.60  0.00  0.00  1.20  0.60  0.60  0.60  0.00
0.00  1.20  0.60  0.60  0.60  0.00  0.00  1.20  0.60  0.60  0.60  0.00  0.00  1.20  0.60  0.60  0.60  0.00  0.00  4.00  3.20  3.20  1.60  0.80  0.80  3.20
2.40  2.40  0.80  0.00  0.00  4.00  3.20  3.20  1.60  0.80  0.80  3.20  2.40  2.40  0.80  0.00  0.00  4.00  3.20  3.20  1.60  0.80  0.80  3.20  2.40  2.40
0.80  0.00  0.00  4.00  3.20  3.20  1.60  0.80  0.80  3.20  2.40  2.40  0.80  0.00  0.00  4.00  3.20  3.20  1.60  0.80  0.80  3.20  2.40  2.40  0.80  0.00
0.00  4.00  3.20  3.20  1.60  0.80  0.80  3.20  2.40  2.40  0.80  0.00  0.00  4.00  3.20  3.20  1.60  0.80  0.80  3.20  2.40  2.40  0.80  0.00  0.00  4.00
3.20  3.20  1.60  0.80  0.80  3.20  2.40  2.40  0.80  0.00  0.00  4.00  3.20  3.20  1.60  0.80  0.80  3.20  2.40  2.40  0.80  0.00  0.00  4.00  3.20  3.20
1.60  0.80  0.80  3.20  2.40  2.40  0.80  0.00  0.00  4.00  3.20  3.20  1.60  0.80  0.80  3.20  2.40  2.40  0.80  0.00  0.00  4.00  3.20  3.20  1.60  0.80
0.80  3.20  2.40  2.40  0.80  0.00  0.00 ]
[ 1.80  1.80  1.80  1.80  1.80  1.80  0.00  0.00  0.00  0.00  0.00  0.00  1.80  1.80  1.80  1.80  1.80  1.80  0.00  0.00  0.00  0.00  0.00  0.00  1.80
1.80  1.80  1.80  1.80  1.80  0.00  0.00  0.00  0.00  0.00  0.00  1.80  1.80  1.80  1.80  1.80  1.80  0.00  0.00  0.00  0.00  0.00  0.00  1.80  1.80  1.80
1.80  1.80  1.80  0.00  0.00  0.00  0.00  0.00  0.00  1.80  1.80  1.80  1.80  1.80  1.80  0.00  0.00  0.00  0.00  0.00  0.00  1.80  1.80  1.80  1.80  1.80
1.80  0.00  0.00  0.00  0.00  0.00  0.00  1.80  1.80  1.80  1.80  1.80  1.80  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00  0.00
```

Figure 9 (a). Output File for Insurance Problem fed to CPLEX Solver

For the workload in this problem, the time "capacity" of each member in the group was computed to be 72/5 with a 20% slack allowed for overtime (the slack can be set to any value in the program), which equals 17.28 hours. The number of possible subgroups for the workload, the number of members and number of workflow instances are listed next in the file. The "qualified" section lists the subgroups qualified to perform each workflow instance; for example, subgroups 1-12 are listed as potential subgroups for workflow_instance 1.

Below this, a fragment of the "timereq" section is shown, that lists the number of hours used in each of the 240 possible subgroups. Thus, Andrew would work for 1.20 hours in subgroup 1, 0.6 hours in subgroups 2-4, 0.0 hours in subgroups 5-6 and so on. Similarly, looking at the next open bracket, towards the bottom of the fragment, we find that Bob would work 1.8 hours in the subgroups 1-6. Again, we note that from subgroups 1-12 only one subgroup will be selected to implement the first workflow instance of the first workflow type, which is handling a customer insurance claim. Similarly, from subgroups 13-24, one will be selected to implement the second workflow instance, and so on. In all 20 subgroups from the 240 potential subgroups will be selected, one for each of the 20 workflow_instances that make up the workload in this example.

Figure 9 (b) shows another fragment of the data file for this problem. The top of the fragment indicates the times that would be used by the last member (Mary) for each of the 240 potential subgroups. The lower half of the fragment shows the cohesiveness values for each of the 240 subgroups. For example, the cohesiveness value for subgroup 2 is 0.66. The cohesiveness is a group level measure, dependent on the structure of the composition of each subgroup, and the overall MM socio-matrix.



```
[ 0.00  0.00  0.00  0.60  0.60  0.60  1.80  1.80  1.80  2.40  2.40  2.40  0.00  0.00  0.00  0.60  0.60  0.60  1.80  1.80  1.80  2.40  2.40  2.40  0.00
0.00  0.00  0.60  0.60  0.60  1.80  1.80  1.80  2.40  2.40  2.40  0.00  0.00  0.00  0.60  0.60  0.60  1.80  1.80  1.80  2.40  2.40  2.40  0.00  0.00  0.00
0.60  0.60  0.60  1.80  1.80  1.80  2.40  2.40  2.40  0.00  0.00  0.00  0.60  0.60  0.60  1.80  1.80  1.80  2.40  2.40  2.40  0.00  0.00  0.00  0.60  0.60
0.60  1.80  1.80  1.80  2.40  2.40  2.40  0.00  0.00  0.00  0.60  0.60  0.60  1.80  1.80  1.80  2.40  2.40  2.40  0.00  0.00  0.00  0.00  0.00  0.00  0.80
0.80  0.80  0.80  0.80  0.80  0.00  0.00  0.00  0.00  0.00  0.00  0.80  0.80  0.80  0.80  0.80  0.80  0.00  0.00  0.00  0.00  0.00  0.00  0.80  0.80  0.80
0.80  0.80  0.80  0.00  0.00  0.00  0.00  0.00  0.00  0.80  0.80  0.80  0.80  0.80  0.80  0.00  0.00  0.00  0.00  0.00  0.00  0.80  0.80  0.80  0.80  0.80
0.80  0.00  0.00  0.00  0.00  0.00  0.00  0.80  0.80  0.80  0.80  0.80  0.80  0.00  0.00  0.00  0.00  0.00  0.00  0.80  0.80  0.80  0.80  0.80  0.80  0.00
0.00  0.00  0.00  0.00  0.00  0.80  0.80  0.80  0.80  0.80  0.80  0.00  0.00  0.00  0.00  0.00  0.00  0.80  0.80  0.80  0.80  0.80  0.80  0.00  0.00  0.00
0.00  0.00  0.00  0.80  0.80  0.80  0.80  0.80  0.80  0.00  0.00  0.00  0.00  0.00  0.00  0.80  0.80  0.80  0.80  0.80  0.80  0.00  0.00  0.00  0.00  0.00
0.00  0.80  0.80  0.80  0.80  0.80  0.80 ]
];
cohesivity = [ 1.00  0.66  0.66  0.66  0.33  0.66  0.33  0.33  0.00  0.33  0.33  0.33  1.00  0.66  0.66  0.66  0.33  0.66  0.33  0.33  0.00  0.33  0.33
0.33  1.00  0.66  0.66  0.66  0.33  0.66  0.33  0.33  0.00  0.33  0.33  0.33  1.00  0.66  0.66  0.66  0.33  0.66  0.33  0.33  0.00  0.33  0.33  0.33  1.00
0.66  0.66  0.66  0.33  0.66  0.33  0.33  0.00  0.33  0.33  0.33  1.00  0.66  0.66  0.66  0.33  0.66  0.33  0.33  0.00  0.33  0.33  0.33  1.00  0.66  0.66
0.66  0.33  0.66  0.33  0.33  0.00  0.33  0.33  0.33  1.00  0.66  0.66  0.66  0.33  0.66  0.33  0.33  0.00  0.33  0.33  0.33  1.00  1.00  0.33  0.33  0.33
0.33  0.33  0.33  0.00  0.00  0.00  0.33  1.00  1.00  0.33  0.33  0.33  0.33  0.33  0.33  0.00  0.00  0.00  0.33  1.00  1.00  0.33  0.33  0.33  0.33  0.33
0.33  0.00  0.00  0.00  0.33  1.00  1.00  0.33  0.33  0.33  0.33  0.33  0.33  0.00  0.00  0.00  0.33  1.00  1.00  0.33  0.33  0.33  0.33  0.33  0.33  0.00
0.00  0.00  0.33  1.00  1.00  0.33  0.33  0.33  0.33  0.33  0.33  0.00  0.00  0.00  0.33  1.00  1.00  0.33  0.33  0.33  0.33  0.33  0.33  0.00  0.00  0.00
0.33  1.00  1.00  0.33  0.33  0.33  0.33  0.33  0.33  0.00  0.00  0.00  0.33  1.00  1.00  0.33  0.33  0.33  0.33  0.33  0.33  0.00  0.00  0.00  0.33  1.00
1.00  0.33  0.33  0.33  0.33  0.33  0.33  0.00  0.00  0.00  0.33  1.00  1.00  0.33  0.33  0.33  0.33  0.33  0.33  0.00  0.00  0.00  0.33  1.00  1.00  0.33
0.33  0.33  0.33  0.33  0.33  0.00  0.00  0.00  0.33 ];
```

Figure 9 (b). Second Fragment of Output File fed to CPLEX Solver

The task for the CPLEX optimizer is to create a set of subgroups (one subgroup for each workflow instance) so that global cohesiveness (across all selected subgroups) is maximized. It needs to do this given the time capacity permitted to each member. Consider the 12 candidate subgroups for workflow instance **W1**, shown in figure 8. If each member was allowed to work for unlimited time, the optimizer would obviously select subgroup 1, since it has the highest cohesiveness (1.00) within the first 12 candidate subgroups. However, if time capacity is a constraint, then subgroup 1 may actually be **suboptimal** for the overall global cohesiveness for the workload, since it may prevent the selection of other subgroups in other workflow instances that contribute more to the global cohesiveness. The second constraint in our model is that one subgroup from each set needs to be selected, so that the entire workload consisting of 20 workflow instances is satisfied. Given the file with the constraints specified, the optimal solution set that was generated by the CPLEX optimizer (< 1 sec for the problem) is shown in figure 10.

	Receptionist	**Sales Person**	**Office Manager**	**Claims Inspector**	**Subgroup Selected**	**Cohesiveness**
W1	Charlene	Not needed	Mary	Bob	5 (from 1-12)	0.33
W2	Charlene	Not needed	Mary	Bob	17 (from 13-24)	0.33
W3	Charlene	Not needed	Mary	Bob	29 (from 25-36)	0.33
W4	Andrew	Not needed	Andrew	Bob	37 (from 37-48)	1.00
W5	Charlene	Not needed	Mary	Bob	53 (from 49-60)	0.33
W6	Charlene	Not needed	Mary	Bob	65 (from 61-72)	0.33
W7	Charlene	Not needed	Mary	Bob	77 (from 73-84)	0.33
W8	Charlene	Not needed	Mary	Bob	89 (from 85-96)	0.33
W9	Charlene	Jane	Mary	Not needed	107 (from 97-108)	0.00

W10	Charlene	Jane	Mary	Not needed	119 (from 109-120)	0.00
W11	Charlene	Jane	Mary	Not needed	131 (from 121-132)	0.00
W12	Charlene	Andrew	Andrew	Not needed	134 (from 133-144)	1.00
W13	Charlene	Andrew	Andrew	Not needed	146 (from 145-156)	1.00
W14	Charlene	Andrew	Andrew	Not needed	158 (from 157-168)	1.00
W15	Charlene	Andrew	Andrew	Not needed	170 (from 169-180)	1.00
W16	Charlene	Andrew	Andrew	Not needed	182 (from 181-192)	1.00
W17	Charlene	Jane	Mary	Not needed	203 (from 193-204)	0.00
W18	Charlene	Jane	Mary	Not needed	215 (from 205-216)	0.00
W19	Charlene	Jane	Mary	Not needed	227 (from 217-228)	0.00
W20	Charlene	Jane	Mary	Not needed	239 (from 229-240)	0.00

Figure 10. Solution for the insurance problem to maximize the overall cohesiveness

Note that the solution in Fig. 10 maximizes the **global** cohesiveness for the workload, while still taking into account the capacity constraint that no member can work more than a certain amount of time (we set 17.28 hours in this case). In our model, the greater the slack in how much "overtime" a member is theoretically allowed, the larger the global cohesiveness that can be achieved, since subgroups with higher cohesiveness can be selected from each group. In many real world situations, overtime may need to be allocated, if some people are critical to several different workflows and perform roles that others cannot. In some of the empirical test cases we ran (described ahead in section 3.4), setting the "overtime" to 0% led to no feasible solution, because it became impossible to find a candidate subgroup for each workflow instance and still allow each person to work with no overtime.

CONCLUSION

In this work, we presented a formal method of modeling workflow teams so as to optimize group cohesiveness for the workload. We have conducted extensive empirical studies that are not shown in this work in progress, for lack of space. The experiments used sets of up to 3000 workflows allocated amongst groups of a maximum of 60 members with 30 roles within the group. These empirical studies demonstrate that our methodology can solve most real world size workloads in less than 6 minutes. This data will also be presented at the conference.

From a theoretical standpoint, our work here opens up a new opportunity to research which group level variables actually drive team performance for certain types of tasks. As an example, consider a real organization with predefined workflows. Two different group level variables can be compared by first creating two sets of optimal subgroups, based on each variable, and then asking managers and other stakeholders to compare these subgroups. Similar, our methodology can also facilitate the study of which criteria drive team performance in which types of workflows (for example transactional versus creative types of workflows).

From a practical perspective, the methodology can be immediately applied in diverse settings, to create workgroup teams based on a selected optimizing criterion, hopefully resulting in an immediate positive impact on organizational effectiveness.

REFERENCES

1. Alba, R.D. "A graph-theoretic definition of a sociometric clique," *Journal of Mathematical Sociology* (3) 1973, pp 113-126.
2. Austin, J.T., and Bobko, P. "Goal Setting Theory: Unexplored Areas and Future Research Needs," *Journal of Occupational Psychology* (58) 1985, pp 289-308.

3. Bock, R.D., and Husain, S.Z. "An Adaptation of Holzinger's B-coefficients for the Analysis of Sociometric Data," *Sociometry* (13) 1950, pp 146-153.

4. Casati, F., Castano, S., and Fugin, M. "Management Workflow Authorization Constraints through Active Database Technology," *Information Systems Frontiers* (3:3) 2001, pp 319-338.

5. Collins, R. *Theoretical Sociology* Harcourt Brace Jovanovich, New York, 1988.

6. Corts, K. "Teams vs. Individual Accountability: Solving Multi-Task Problems through Job Design," *Rand Journal of Economics*) forthcoming.

7. Evans, M.G. "Organizational Behavior: The Central Role of Motivation," *Journal of Management* (12) 1986, pp 203-222.

8. Festinger, L. "Informal Social Communication," *Psychological Review* (57) 1950, pp 271-282.

9. Freeman, L.C. "Cenrtality in Social Networks," *Social Networks* (1) 1979, pp 215-239.

10. Friedkin, N.E. "Structural Cohesion and Equivalence Explanations of Scoial Homogeneity," *Sociological Methods and Research* (12) 1984, pp 235-261.

11. Governatori, G., Rotolo, A., and Sadiq, S. "A model of dynamic resource allocation in workflow systems," Fifteenth Australasian database conference Dunedin, New Zealand, 2004, pp. 197-206.

12. Hemmer, T. "On the interrelation between production technology, job design, and incentives," *Journal of Accounting and Economics* (19) 1995, pp 209-245.

13. Holmstrom, B.R., and Migrom, P. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization* (7) 1991, pp 24-52.

14. Itoh, H. "Job design, delegation, and cooperation: a principal-agent analysis," *European Economic Review* (38) 1994, pp 691-700.

15. Jordan, M.H., Feild, H.S., and Armenakis, A.A. "The Relationship of Group Process Variables and Team Performance:
A Team-Level Analysis in a Field Setting " *Small Group Research* (33:1) 2002, pp 121-150.

16. Kautz, H., Selman, B., and Shah, M. "ReferralWeb: Combining Social Networks and Collaborative Filtering," *Communications of the ACM* (30:3) 1997.

17. Kumar, A., Aalst, W.M.v.d., and verbeek, E.M. "Dynamic Work Distribution in Workflow Management systems: How to Balance Quality and Performance?," *Journal of Management Information Systems* (18:3) 2001, pp 129-157.

18. Liben-Nowell, D., and Kleinberg, J. "The Link Prediction Problem for Social Networks," 12th Annual ACM INternational Conference on Information & Knowledge Management (CIKM), 2003, pp. 556-559.

19. Momotko, M., and Subieta, K. "Dynamic Changes in Workflow Participant Assignment," East European Conference on Advances in Databases and Information Systems, Bratislava, 2002.

20. Moreno, J.l. *Who shall survive? A new approach to the problems of human interrelations.* Nervous and Mental Disease Publishing Co. , Washington, DC, 1934.

21. Mudrack, P.E. "Group Cohesiveness and Productivity: A Closer Look," *Human Relations* (42:9) 1989, pp 771-785.

22. Raghavan, P. "Social Networks: From the Web to the Enterprise," *IEEE Internet Computing* (6:1) 2002, pp 91-94.

23. Russell, N., Aalst, W.v.d., Hofstede, A.H.M.t., and Edmond, D. "Workflow Resource Patterns: Identification, Representation and Tool Support," Conference on Advanced Information Systems Engineering (CAiSE), Springer Verlag, Porto, Portugal, 2005.

24. Sailer, L.D., and Gaulin, S.J.C. "Proximity, sociality and observation: The definition of social groups," *American Anthropologist* (86) 1984, pp 91-98.

25. Sanders, K., and Nauta, A. "Social Cohesiveness and Absenteeism:
The Relationship Between Characteristics of Employees and Short-Term Absenteeism Within an Organization " *Small Group Research* (35:6) 2004, pp 724-741.

26. Seers, A., Petty, M.M., and Cashman, J.F. "Team-Member Exchange Under Team and Traditional Management
A Naturally Occurring Quasi-Experiment " *Group & Organization Management* (20:1) 1995, pp 18-38.

27. Simmel, G. "The Web of Group Affiliations," in: *Conflict and the Web of Group Affiliations,* G. Simmel (ed.), Free Press, Glencoe, IL, 1955.

28. Steinmark, R. "Group Cohesiveness and Extrinsic Motivation in Virtual Groups: Lessons from an Action Case Study of Electronic Brainstorming," 35th Hawaii International Conference on System Sciences (HICSS), IEEE, Hawaii, 2002, p. 16b.

29. Thomas, J., and Griffin, R. "The Social Information Processing Model of Task Design: A review of the Literature," *Academy of Management Review* (8) 1983, pp 672-682.

30. Valsecchi, I. "Policing team production through job design.," *Journal of Law,Economics, and Organization* (12:2) 1996, pp 361-375.

31. Wasserman, S., and Faust, K. *Social Network Analysis: Methods and Applications* Cambridge University Press, Cambridge, 1999.