

Experimental Verification on Human-Centric Network-Based Resource Allocation Approaches for Process-Aware Information Systems

MOON-SOOK YEON¹, YOUNG-KOO LEE², DINH-LAM PHAM^{3,4},
AND KWANGHOON PIO KIM³, (Member, IEEE)

¹Department of Computer Science and Engineering, Kyunghee University, Yongin-si, Gyeonggi-do 17104, South Korea

²Department of Computer Engineering, College of Electronics and Information, Kyunghee University, Yongin-si, Gyeonggi-do 17104, South Korea

³Data and Process Engineering Research Laboratory, Division of AI Computer, Kyonggi University, Youngtong-gu, Suwon 16227, Republic of Korea

⁴Information Technology Institute, Vietnam National University, Hanoi 100000, Vietnam

Corresponding author: Kwanghoon Pio Kim (kwang@kgu.ac.kr)

ABSTRACT Efficient allocating human resource obtains significant advantages to business operations in enterprises. Manipulating the available human resources of a business optimizes the expense of developing products and services, improves work productivity, and brings more value to the business. In the process of creating products and benefits in the enterprise, a network of working relationships is formed among employees in the organization. In this paper, we propose an approach to support decision-making on human resource allocation in an organization based on the human-centric network discovered from the event logs. To be more precise, the knowledge derived from human-centric networks formed through the operation phases of the information systems is used to make resource allocation strategies in the organizations. Using the appropriateness probabilities and social network metrics for allocating human resources, we clarify efficiently allocating resources based on different metrics. The system architecture and the scenarios are also described for planning human resources allocation based on these networks. Furthermore, we use real-life data set for getting experimental verification and proving the proposal. This approach will be a promising solution to support businesses effectively allocating resources in their organizations.

INDEX TERMS Human-centric network, human resources allocation, business process management, process mining, process-aware information system.

I. INTRODUCTION

Resources allocation activities evaluate employees, facilities, related entities in an organization and then assign them to each process. The evaluation is based on each employee's ability and participation in the organization's works. In process-aware enterprises, the history of enactment operations is often recorded in the IEEE XES format [1], the commonly used standard for logging events. Different organizations maintain different operational processes, but they all require human resources to operate the system in common.

The human resources allocation (HRA) tasks clearly define and plan future processes, then allocate employees to the related entities such as operations and facilities. Resource

allocation is not a significant concern in an organization with few employees hence the small number of activities. Meanwhile, the allocation of human resources is complicated in enterprises with hundreds (or thousands) of employees and procedures. The appropriate allocation of human assets will preserve time and production expenses besides operating expenses. Consequently, to manipulate the system efficiently, improve employee productivity, allocation of human resources is an important task and cannot be lacking in business process management.

In the 1990s of the last century, Business process re-engineering (BPR) has been one of the primary major management systems. In the BPR systems research, scientists investigated numerous studies and surveys on the role of human resource management (HRM). These studies have demonstrated the critical function in human resource management architecture through the BPR systems. Lack of

The associate editor coordinating the review of this manuscript and approving it for publication was Justin Zhang ¹.

consideration for human resources is one of the main reasons for unsuccessfully implementing projects in organizations. Also, the position of managers and teamwork are the key characteristics of the success of businesses as described in studies of Willmott [2], Zucchi [3], Wood [4], and Lepak [5].

With the evolution of information technology, the business process management (BPM) systems grow fastly, providing the infrastructure and context for processing the HRM models. The HRM process model has to support all other activities in the organization for operating these processes. Also, several HRM models were introduced for improving system management according to Cakar [6], Xu [7], Zhao *et al.* [8], Lin [9]. So that, improving the business process enhances human resource allocation, and vice versa.

There are numerous different approaches to allocating human resources in organizations. Each different approach is formed from different studies and perspectives and has its advantages. Nakatumba [10], Liu [13], Zhao *et al.* [16], Arias [24], Drosouli [30] use the process mining approach to analyze the resources through event logs for various types of decision-making. Or using reinforcement learning to learn the appropriate policies for optimizing resource allocation in business process execution, Huang described in [11]. Or use probabilities methods to make resource allocation decisions Cabanillas [14], Wibisono [15]. Alternatively, the other methods for allocating resources such as clustering in the study of Zhao [17], [31]; business process management in the research of Havur [23], Erasmus [27]; human-centric in the investigation of Wibisono [18], Zhang [19], Soltis [20], Bajaj [21], Ballesteros-Perez [22]; task operation model in the analysis of Huang [12]; multi-factor criteria in the study of Arias [26]; fuzzy data mining in the research of Wu [32]; or making the survey about this research topic for summarizing the methodologies described in the results of Arias [25], [28], Zheng [29].

With the multifarious development of business models over time, the strategies to support resource allocation in the organization are continuously changing, especially for enterprises with the participation of hundreds of employees or more. Accordingly, **this paper proposes a new approach to support allocating human assets for organizations based on the discovered human-centric networks.** The principal idea of the approach is that, **at the beginning of distributing human resources phase, we discover the human-centric networks of the organization through mining its operation event log history. Consequently, by analyzing the networks, we get comprehensive views of the operational processes with the association of performers in each work case. Understanding the networks, we get clear on the appropriateness probabilities for allocating employees in each activity, the centrality metrics of employees in the networks in different manners. Based on this knowledge, we effectively conduct allocating resources for the next operation of the enterprises.**

In the remaining sections of this paper, we category the content as follows. The following section discusses the related work compared to our study approach, the similarities

and the differences between our method and these methods. Section 3 describes the fundamental theories and concepts related to our proposal. Section 4 explains the approach in detail, how to allocate human resources based on this information. In section 5, a real-life data set was used for presenting experimental results and verifying our approach. Finally, section 6 gives some conclusions and the future study.

II. RELATED WORKS

In the approach using mining event logs to support resource allocation in workflow systems [13], Tingyu Liu and his colleagues proposed an Apriorilike algorithm to find the common patterns from the event log. After that, a correspondence measure named Lift is used to explain the negatively correlated resource allocation rules for resource reservation. In the proposal, the rules are ranked using the confidence measures as resource allocation rules. Based on this information, they presented a decision-making method to make recommendations to workflow initiators for supporting resources allocation.

To optimize resource allocation based on process mining, Weidong Zhao [16] introduced a resource allocation model by viewing the relationship between process performance and resource allocation. The approach presented a model to minimize process execution time under three constraints: total cost, resource availability, and resource preference. Experiments were also conducted to prove the effectiveness of the proposal.

In [10], Joyce Nakatumba presented the solution for analyzing resource behavior in business processes by using process mining. In the approach, he introduced a new process mining plug-in which is attached to ProM, a framework that supports a general category of process mining techniques [33], to investigate the impact of workload on service times based on historical data and regression analysis. Applying the approach to some real-life data sets from a municipality in the Netherlands, the author's experimentation shows that the relation between workload and performance of workers, described by the "Yerkes-Dodson Law of Arousal" [34], really exists.

In [30], Drosouli introduced a process mining approach for managing resource allocation in a Bike sharing system (BSS) in New York City. The main idea of the study is to analyze the BSS to manage resource allocation by using process mining methods. The author converted the BSS data set to the event log format and then use the ProM tool [33] to apply process mining techniques in event logs for understanding more about the process model, the using-bike time, the using-bike in a station, etc. Through these reports to conduct resource allocation management for the system.

Also using the process mining techniques, in [24], Michael Arias built a framework for recommending resource allocation to the information system. Instead of recommending at the activity level, the framework gave the recommendation at the sub-process level. A resource process cube was introduced in the study and provided an extensible, flexible,

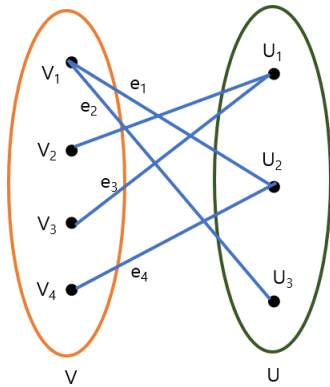


FIGURE 1. Example of a bipartite graph.

and fine mechanism to collect historical information about what happened in past process executions. Based on this information, a final recommendation ranking was obtained through the Best Position Algorithm.

Applying the human-centric resource allocation method, in [18], Wibisono and his colleagues presented a dynamic approach for managing the execution of business processes. Using the Bayesian inference, they predicted the performances of resources through the data set to organize different types of human resources. The study used the Naive Bayes Model to comprise five nodes; each node implies perspectives participated in business process execution. In particular, there are two types of nodes, i.e., parent node and child nodes. The parent nodes contain performance predictions that influence performance throughout every child node. The evidence of the nodes in the model is used for measuring human performance.

The listed studies indicate that using process mining techniques is one of the efficient methods for allocating resources in information systems. In this paper, we introduce our method using process mining techniques for allocating human resources. Nevertheless, the difference between our proposed method and the studies listed above is that: In our method, we focus on taking people as the center factor to build the human-centric networks. Subsequently, analyzing the networks between performers (employees) and activities that are formed when executing tasks. On the one hand, we obtain exhaustive views of the operating processes with the association of performers in each activity in the organization. On the other hand, we get precise about the appropriateness probabilities for assigning employees in each process, and the centrality metrics of employees in the networks in different manners. Based on this knowledge, we conduct allocating resources for the next operation of the enterprises. By analyzing these information, we have a better understanding of the system and the efficiency of its current operation, thereby making flexible resource allocations. The next section of the paper discusses the relevant basic concepts used

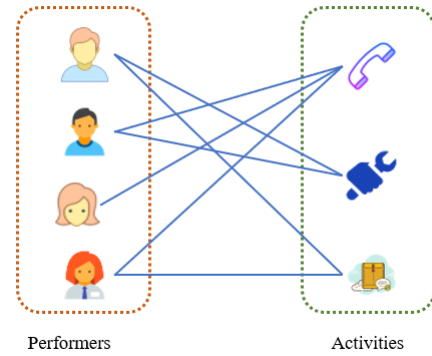


FIGURE 2. Example of process-aware affiliation network.

to discover the human-centric networks to support allocating resources in an organization.

III. BASIC CONCEPTS

A. PROCESS-AWARE AFFILIATION NETWORK (PAN)

1) BIPARTITE GRAPH [35]

In graph theory, a Bipartite graph (or Bigraph) is a graph that its vertices can be split into two independent and disjoint sets U and V such that every edge in the graph connects a vertex in V to one in U . Vertex sets U and V are usually called the parts of the graph.

Figure 1 illustrates an example of a Bipartite graph G . In this graph, $V = \{V_1, V_2, V_3, V_4\}$, $U = \{U_1, U_2, U_3\}$ are two sets of vertices. $E = \{e_1, e_2, e_3, e_4\}$ is the set of edges connecting vertices in V to the others in U .

2) PROCESS-AWARE AFFILIATION NETWORK (PAN)

Discovering event log history from a process-aware information system (PAIS) [36], we get numerous information describing what actually happened in different processes. The information about events and their entities, e.g., activities, roles, performers, relevant data, etc. In this information, activities and performers are related by affiliation with each other in their groups. **Activity is the action/task in the system that runs to archive a specific goal; meanwhile, the performer executes the activities. We can view these relationships as a bipartite graph and call as an affiliation network.** That means we can discover the affiliation network of performers and activities from the event log history. Figure 2 describes an example of an affiliation network obtained from a PAIS. In this example, there are four performers affiliated with three activities in the network. By using the affiliation network visualization, we recognize the list of activities that an individual performer performed and the list of performers involved in an activity through the network.

Definition of PAN: Let ω is the event log recorded from a process-aware information system. $A = \{A_i, i \in \overline{1, n}\}$ is the set of activities in ω . $P = \{P_j, j \in \overline{1, m}\}$ is the set of performers in ω . $T = [T_k, k \in \overline{1, t}]$ is the list of

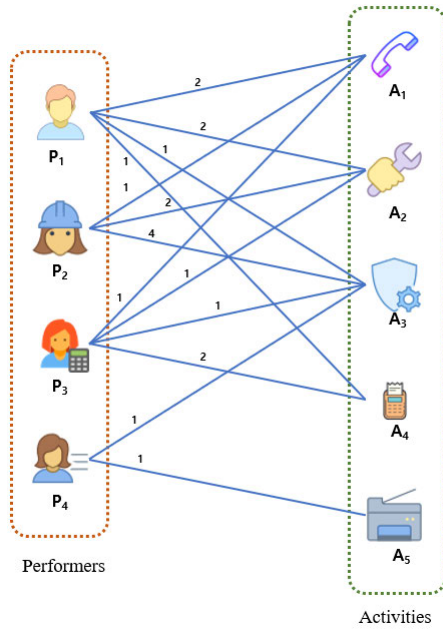


FIGURE 3. The discovered PAN from event log ω .

traces in ω . The process-aware affiliation network $PAN = \{A, P, AN_{P_j, A_l}^i | j \in \overline{1, m}, l \in \overline{1, n}\}$. Where AN is the set of affiliated relationships between performers P_j and activities A_l , “ i ” indicates that how many times the affiliation between P_j and A_l was occurred.

For example (Ex1), we have an event log ω with list of traces T as following:

- $(A_1, P_1) \rightarrow (A_3, P_2) \rightarrow (A_2, P_3) \rightarrow (A_5, P_4)$
- $(A_1, P_2) \rightarrow (A_2, P_1) \rightarrow (A_3, P_2) \rightarrow (A_4, P_3)$
- $(A_1, P_2) \rightarrow (A_3, P_1) \rightarrow (A_3, P_2) \rightarrow (A_4, P_3)$
- $(A_2, P_2) \rightarrow (A_1, P_3) \rightarrow (A_3, P_2) \rightarrow (A_4, P_1)$
- $(A_2, P_1) \rightarrow (A_3, P_2) \rightarrow (A_3, P_1) \rightarrow (A_3, P_3)$
- $(A_1, P_1) \rightarrow (A_2, P_2) \rightarrow (A_4, P_3) \rightarrow (A_3, P_4)$

(A_1, P_1) means that activity A_1 was performed by performer P_1 . Then we have the discovered PAN with the following information and as shown in Figure 3:

- $A = \{A_1, A_2, A_3, A_4, A_5\}$
- $P = \{P_1, P_2, P_3, P_4\}$
- $AN = \{(A_1, P_1)^2, (A_1, P_2)^1, (A_1, P_3)^1, (A_2, P_1)^2, (A_2, P_2)^2, (A_2, P_3)^1, (A_3, P_1)^1, (A_3, P_2)^4, (A_3, P_3)^1, (A_3, P_4)^1, (A_4, P_1)^1, (A_4, P_3)^2, (A_5, P_4)^1\}$

As we can see in the relationship set AN , (A_3, P_2) has the most affiliated value (4), meanwhile activity A_5 is only affiliated with performer P_4 one time. This information will be helpful in supporting allocate human resources in the future.

B. PROCESS-AWARE ENTERPRISE SOCIAL NETWORK (PESN)

When participating in the execution of tasks and activities in the events of the information system, employees (performers)

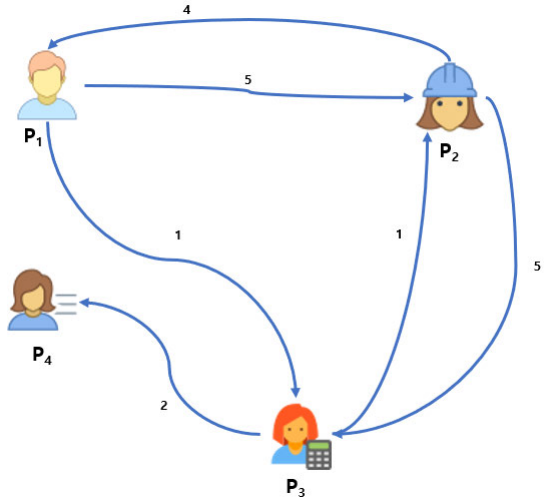


FIGURE 4. The discovered PESN from event log ω .

will interact with each other depending on the job position as well as the role assigned in the organization. This property produces a work-relationships network, shows the working relationship between employees in the same organization, and we call it as *PESN*.

Definition of PESN: Given that ω is the event log recorded from a process-aware information system. $A = \{A_i, i \in \overline{1, n}\}$ is the set of activities in log ω . $P = \{P_j, j \in \overline{1, m}\}$ is the set of performers who perform an activity A_i . $TRS = \{TR_{A_i, A_k}^a | (i, k) \in \overline{1, n}\}$ is the set of transactions from activity A_i to activity A_k , and “ a ” indicates how many times the transaction occurred. The process-aware enterprise social network $PESN = \{RE_{P_j, P_l}^b | (j, l) \in \overline{1, m}\}$ is the set of relationships between performer P_j and performer P_l , “ b ” indicates how many times the relationship occurred, where P_j performs A_i and P_l performs A_k .

Figure 4 represents the *PESN* discovered from event log ω in the example Ex1. As is evident from Figure 4, The performer P_1 interacted with performers P_2, P_3 , meanwhile, the performer P_4 interacted with only performer P_3 when participating the assigned tasks. As it has been pointed out, through the *PESN*, we can recognize the roles of each performer and their interactions with other performers in the network. This observation will be useful in supporting information and decision-making to allocate human resources for system operation.

IV. METHODS

In most organizations, people are always the main factor that determines the development and success of that organization. Consequently, understanding human resources in an organization, thereby making decisions to allocate resources appropriately, plays a critical role in the organization's development, especially for enterprises. Our human-centric networks-based approach for supporting human resources allocation is based on mining the event log history of

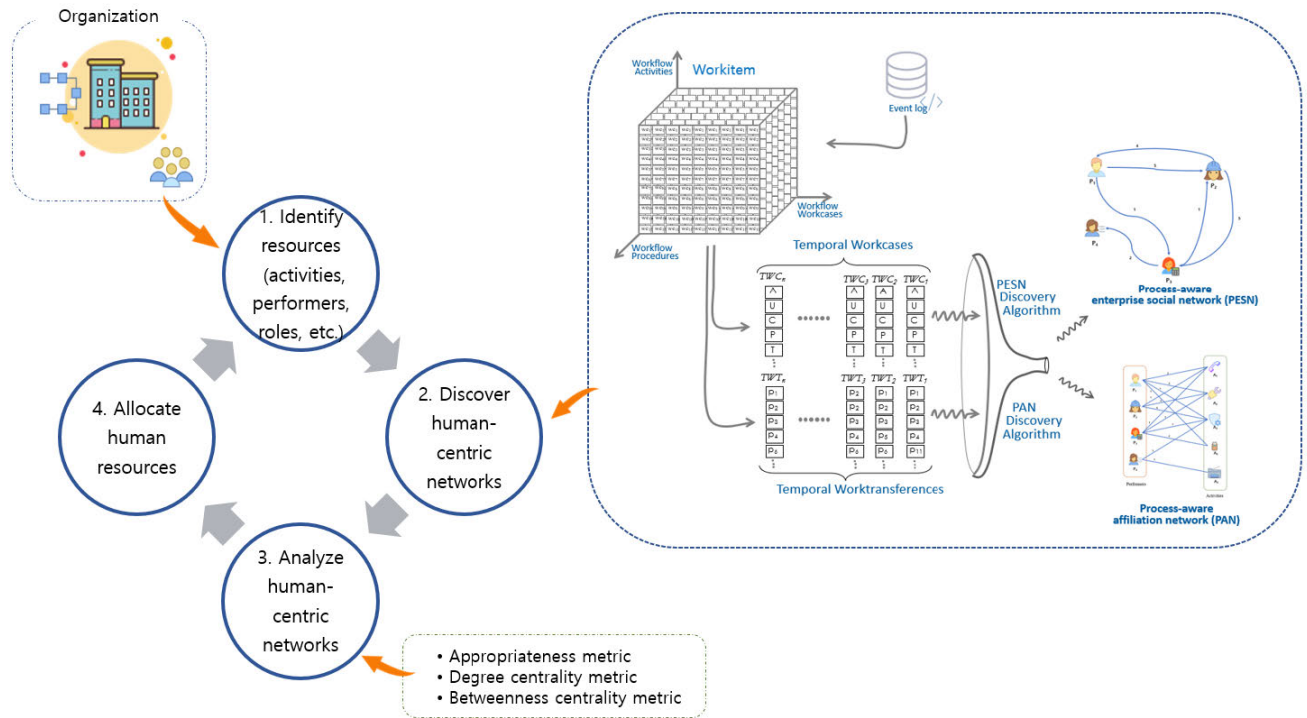


FIGURE 5. The architecture for allocating human resources based human-centric networks.

the information system, thereby building networks where humans play a decisive role in running tasks and activities in the business model. In this procedure, we use PAN and PESN as the basis of analysis to make decisions or recommendations for allocating human resources. Using PAN helps us evaluate the role of an employee when participating in activities in each business process. Meanwhile, using PESN will give us a view of the working relationship between employees in the organization when conducting business activities.

A. SYSTEM ARCHITECTURE

Figure 5 illustrates the architecture of our approach using the human-centric networks for allocating human resources in the process-aware information system. In the architecture, the procedure is divided into four (4) steps:

- **Step 1: Identify resources:** Identify available resources is the first step for preparing allocating resources. At the beginning, we need to identify the organization's available resources, including activities, roles, performers assigned to activities, relevant data, and materials for executing the activities in each business process. To ascertain the available resources, we can base on the operational plan, current business operating state. Furthermore, we can investigate the event log history to get information about the activities and relevant data that have been involved in the operation of the business process. Based on the different sources, we can synthesize and determine the available resources of the

organization/department to prepare for the operational business process in the next phase.

- **Step 2: Discover human-centric networks:** This step conducts discovering the human-centric networks from the event log history. In this step, the work-items in event log information are extracted into three (3) dimensions, i.e., workflow procedures, workflow workcases, and workflow activities.
- **Step 3: Analyze human-centric networks:** After obtaining information about PAN and PESN, we proceed to analyze the relationship between the components and the role of each performer in creating these networks by evaluating the following metrics (the calculation of these values will be described in the next section):
 - Appropriateness metric.
 - Degree centrality metric.
 - Betweenness centrality metric.

These values provide information that represents the performer's role in running the tasks and activities and the relationship between the performers. By evaluating these values, combined with the business plan and human resources capabilities, we can make efficient decisions for allocating human resources to operate the system's business process in the next phase.

- **Step 4: Allocate human resources:** After gathering all information collected in the previous steps, we understand more about what actually happened inside the system. Based on the analysis of human-centric networks and the business process plan, we conduct allocating

human resources to each activity by answering these questions:

- Who is the most appropriate performer for executing a specific activity?
- Which activity is most appropriate for each performer?
- Which performers play an essential role in the system's operation and are difficult to replace by the others?
- Which performers can be transferred to perform other tasks?
- Which performer acts as a working bridge in the process to execute the task?

Each different organization and business model have additional questions for allocating human resources. With the obtained metrics and the visualization of networks, our method will help answer these questions.

B. HUMAN-CENTRIC NETWORKS METRICS SUPPORTING RESOURCES ALLOCATION

1) DEGREE CENTRALITY

In graph theory, Degree centrality indicates the number of links between a vertex and other vertices [37]. Using this metric in PESN, the Degree centrality shows the working relationship between employees in the organization. The Degree centrality index in PESN indicates activity transfer/receipt information among employees. Lets consider ($C_D(P_k)$) is the Degree centrality value of performer P_k , To calculate ($C_D(P_k)$), we use the following equation [37]:

$$C_D(P_k) = \frac{\sum_{i=1}^m (P_i, P_k)}{m-1} \quad (i \neq k)$$

where:

- $(P_i, P_k) = 1$ if $(P_i \rightarrow P_k)$ or $(P_k \rightarrow P_i)$.
- $(P_i, P_k) = 0$, otherwise.
- $P_i \rightarrow P_k$: There is a work transfer from P_i to P_k .
- m : the total performers in the PESN.

The higher this metric, the greater the degree of an employee involved in completing the trace in a business process case. Even though the Degree centrality metric does not indicate the importance of a performer in connecting others, it is a good measure of the total number of connections a performer has. From the manager's perspective, we perceive the level of work engagement and collaboration among employees in the organization.

2) BETWEENNESS CENTRALITY

In graph theory, the Betweenness centrality of a vertex measures whether or not the vertex is a bridge on the shortest path of two other vertices in a graph [38]. Using this index in PESN, the Betweenness centrality represents the important role each performer plays when performing the activities in each trace from the event log. Due to connecting performers in the network as a bridge, the effectiveness of the performer with a higher index value will collectively affect the overall

TABLE 1. The list of activities and performers in the Help desk data set [39].

Name of activity	Name of Performer
Assign seriousness	Value 1
Create SW anomaly	Value 2
DUPLICATE	Value 3
Insert ticket	Value 4
INVALID	Value 5
Require upgrade	Value 6
Resolve SW anomaly	Value 7
Resolve ticket	Value 8
RESOLVED	Value 9
Schedule intervention	Value 10
Take in charge ticket	Value 11
VERIFIED	Value 12
Wait	Value 13
Closed	Value 14
	Value 15
	Value 16
	Value 17
	Value 18
	Value 19
	Value 20
	Value 21
	Value 22

outcome of each trace. Let's consider $C_B(P_k)$ is the Betweenness centrality value of performer P_k . To calculate ($C_B(P_k)$), we use the following equation [38]:

$$C_B(P_k) = \frac{\sum_{i=1}^m \sum_{j=1}^m b_{ij}(P_k)}{m-1} \quad (i \neq j \neq k)$$

where:

- $b_{ij}(P_k) = \frac{g_{ij}(P_k)}{g_{ij}}$.
- g_{ij} : The total number of shortest path from $i \rightarrow j$.
- $g_{ij}(P_k)$: The total number of shortest path from $i \rightarrow j$ which pass through P_k .
- m : the total performers in the PESN.

Using the Betweenness centrality metric, the manager perceives the crucial performers in the network (with high Betweenness centrality values). They occupy the significant tasks for completing the entire trace in the event log.

3) APPROPRIATENESS PROBABILITIES

As shown in Figure 2, the PAN gives information of operating activities in the system and the participating performers who executed these activities in the event log. Now we will examine the appropriate probabilities for the performers to perform activities through the discovered PAN. Let's consider:

- PAN is a discovered process-aware affiliation network from an event log ω .
- n : The total activities in PAN .
- m : The total performers in PAN .
- $A^i P^j$: The total times activity A^i was performed by performer P^j .
- $\sum_{i=1}^n P^j(A^i)$: The total times performer P^j perform all activities A^i in ω .
- $\sum_{i=1}^m A^i(P^j)$: The total times activity A^i was performed by all performers P^j in ω .

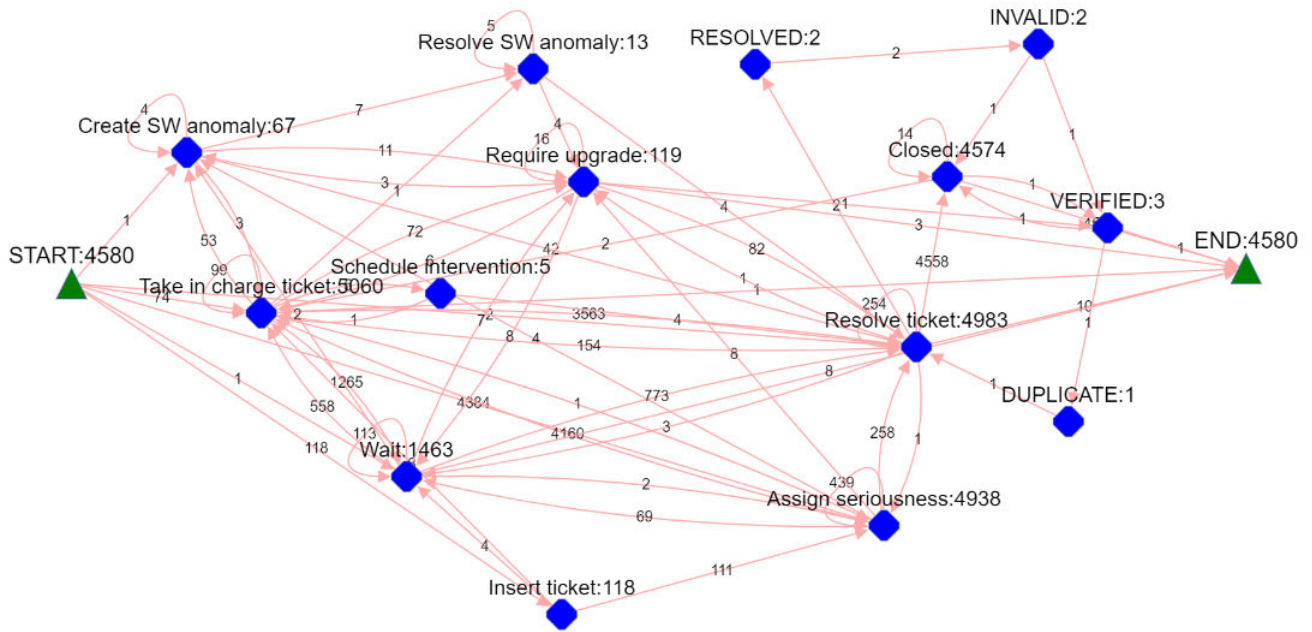


FIGURE 6. Processes discovered from the help desk data set.

TABLE 2. The degree and betweenness centrality values of performers in help desk data set.

Name of Performer	Degree centrality	Betweenness centrality
Value 1	0.76744186	0.496492901
Value 2	1	0.782260164
Value 3	0.534883721	0.188922693
Value 4	0.604651163	0.114824569
Value 5	0.674418605	0.420356537
Value 6	0.744186047	0.141098548
Value 7	0.651162791	0.053892194
Value 8	0.511627907	0.068304171
Value 9	0.906976744	0.486805619
Value 10	0.604651163	0.104032584
Value 11	0.558139535	0.079424994
Value 12	0.790697674	0.309532891
Value 13	0.697674419	0.111876129
Value 14	0.674418605	0.27040901
Value 15	0.534883721	0.057286022
Value 16	0.581395349	0.065486408
Value 17	0.348837209	0.042585445
Value 18	0.348837209	0.027738928
Value 19	0.348837209	0.007284382
Value 20	0.23255814	0
Value 21	0.23255814	0.014359504
Value 22	0.162790698	0

- The appropriateness probability AP_{A^i} of activity A^i is calculated by equation:

$$AP_{A^i} = \{\xi_{pj} | \xi_{pj} = [\xi_{p1}, \xi_{p2}, \dots, \xi_{pm}]; 1 \leq j \leq m\}$$

where:

$$\xi_{pj} = \frac{A^i P_j}{\sum_{k=1}^m A^i(P^k)}$$

- The appropriateness probability AP_{P_j} of performer P_j is calculated by equation:

$$AP_{P_j} = \{\xi_{Ai} | \xi_{Ai} = [\xi_{A1}, \xi_{A2}, \dots, \xi_{An}]; 1 \leq i \leq n\}$$

where:

$$\xi_{Ai} = \frac{A^i P_j}{\sum_{k=1}^n P_j(A^k)} \quad (1)$$

By calculating the Appropriateness values, we receive the probability of performers completed an activity and the reasonableness value of activities that a performer executed. These assessments help modelers to plan and allocate human resources in the re-engineering phase of the process life-cycle in the organization.

V. EXPERIMENTAL RESULTS

In this section, we present the experimental results by applying the human-centric network-based for allocating resources in organizations with a real-life data sets. The main objectives of the analysis in this section are as follows:

- Assess the feasibility of this approach to support in informing decision-making on human resource allocation within the organization.
- Evaluate the effectiveness of human resource allocation through human-centric networks.
- Answer the questions raised in the previous section (System architecture, Step 4).

This data set is about the process in an Italian software company [39] (Help desk data set). The log describes the process of the help desk about ticketing management and consists of 14 activities, 21,348 events, 22 performers performed in 4580 process instances. In these processes, performers are assigned to different activities to manage tickets of software problems. The list of activities and performers are represented in Table 1.

Using the process mining framework described in [40], we get the discovered process model of this data set as shown

TABLE 3. The appropriateness probabilities value of the most executed activities in help desk data set.

Activity name	Performer name	AP value	Activity name	Performer name	AP value
Resolve ticket	Value 2	0.267509532	Assign seriousness	Value 1	0.31490482
	Value 9	0.144089906		Value 2	0.180234913
	Value 1	0.138069436		Value 8	0.139935196
	Value 13	0.10295003		Value 9	0.07654921
	Value 4	0.07525587		Value 13	0.07654921
	Value 6	0.049568533		Value 4	0.068651276
	Value 12	0.035119406		Value 14	0.054880518
	Value 16	0.033313265		Value 6	0.014580802
	Value 14	0.032711218		Value 7	0.013770757
	Value 8	0.023479831		Value 15	0.013770757
	Value 11	0.01986755		Value 18	0.010935601
	Value 10	0.019466185		Value 19	0.010935601
	Value 15	0.018262091		Value 12	0.009518023
	Value 7	0.016656633		Value 16	0.005872823
	Value 17	0.011639575		Value 11	0.005062778
	Value 19	0.004615693		Value 10	0.003037667
	Value 18	0.003010235		Value 20	0.000405022
	Value 21	0.003010235		Value 21	0.000202511
	Value 20	0.001404776		Value 5	0.000202511

(AP value: Appropriateness probabilities value)

Activity name	Performer name	AP value	Activity name	Performer name	AP value
Take in charge ticket	Value 2	0.301383399	Closed	Value 5	0.813073896
	Value 9	0.11916996		Value 3	0.183209445
	Value 1	0.112450593		Value 21	0.002404897
	Value 13	0.08201581		Value 2	0.000874508
	Value 8	0.081225296		Value 10	0.000218627
	Value 6	0.060474308		Value 16	0.000218627
	Value 4	0.056719368			
	Value 16	0.031027668			
	Value 12	0.029051383			
	Value 14	0.028853755			
	Value 11	0.020948617			
	Value 10	0.020750988			
	Value 7	0.017588933			
	Value 15	0.014031621			
	Value 17	0.005533597			
	Value 19	0.005533597			
	Value 3	0.004545455			
	Value 18	0.004545455			
	Value 5	0.001581028			
	Value 20	0.001383399			
	Value 21	0.001185771			

in Figure 6. For convenience in understanding the discovered processes, we added two virtual activities named “START” and “END” to this model. These activities aid in recognizing the activities that started and ended a process in the system.

Now, we conduct resource analysis from the event log history to contribute information to support future human resource allocation according to the architecture presented in the previous section (Figure 5):

- Step 1. Identify resources: By reading the event log, we get the list of activities and performers of the organization as described in Table 1. Through data in table 1, we perceive information on the tasks to be performed and the available human resources to allocate to these tasks. As we can see in the table, there are 14 activities and 22 performers in this company. These activities will be executed in the future operation with the list of performers. We will evaluate the potential performers for assigning in each task.
- Step 2: Discover the human-centric networks: After gathering information about activities and performers, we discover the PESN and PAN networks depicted in Figure 7, Figure 10, respectively. In the discovered PESN, we also added two virtual performers named “START” and “END” to help us realize which performers started and ended a process. For representing these networks, we use the Graphviz [41] and Cytoscape library [42].
- Step 3: Analyze the human-centric networks of the Help desk data set: In this step, we examine the Appropriateness probabilities, Degree centrality, and Betweenness centrality metric of performers for the networks. The values of each performer are described in Table 2 and Table 3. Table 3 describes the Appropriateness probabilities of the most executed activities discovered through the data set, i.e., *Resolve ticket*, *Assign seriousness*, *Take in charge ticket*, and *Closed*. For seeing the values of

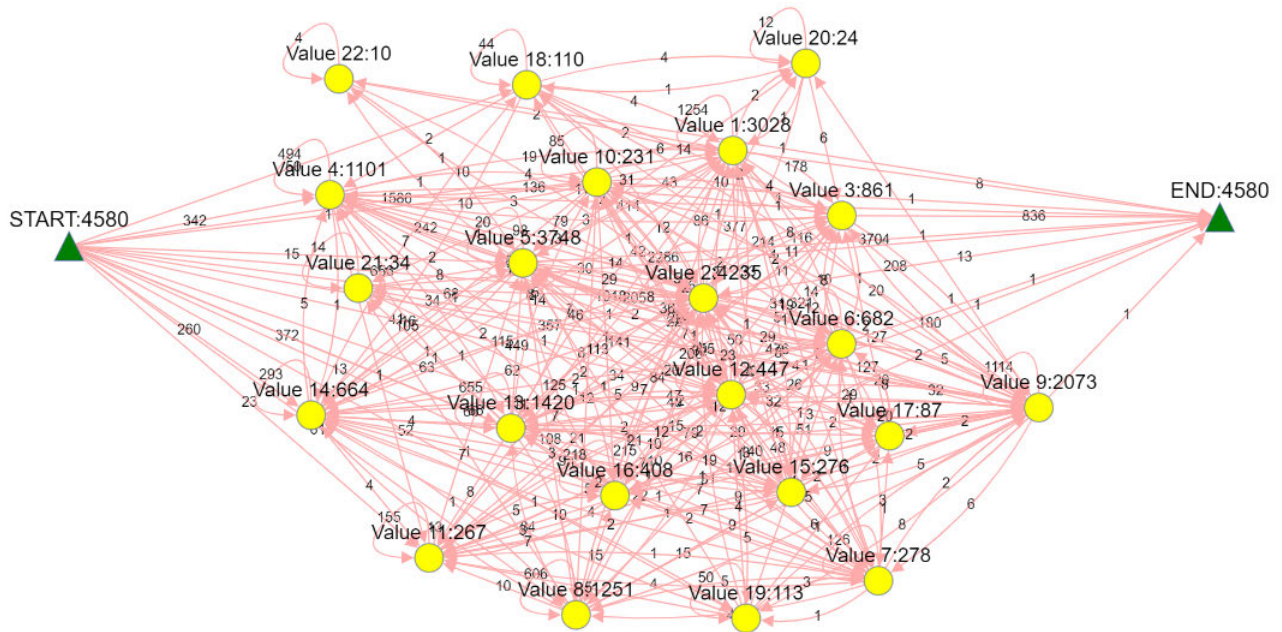


FIGURE 7. The process-aware enterprise social network (PESN) discovered from the Help desk data set.

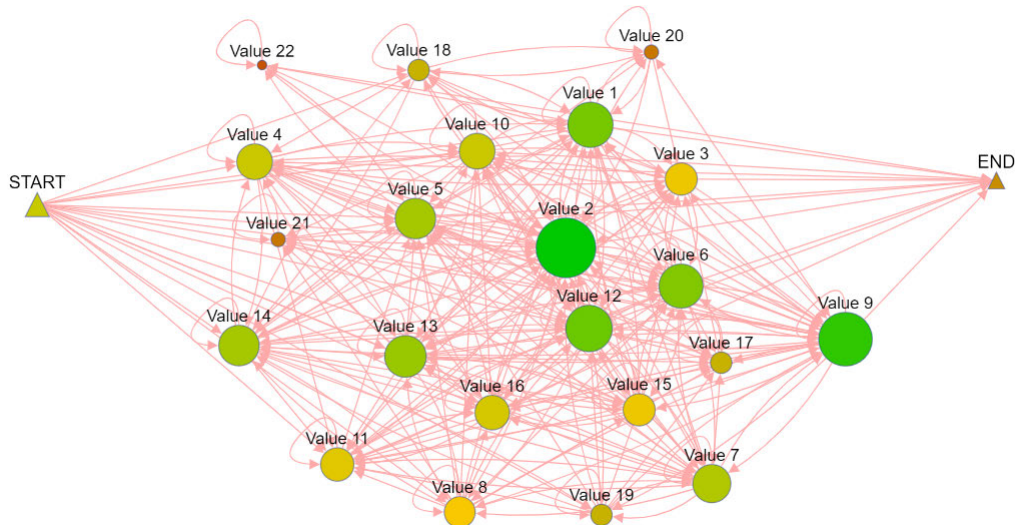


FIGURE 8. Visualization of the discovered PESN in degree centrality metric.

other activities, the reader can refer to the *Reproducibility* section at the end of this paper.

- Step 4. Allocate human resources for activities in the Help desk data set: Gathering collected information in previous steps, we identify the role of each performer in participating in the system's tasks and the relationship between activities and performers. Now we proceed to allocate human resources for some situations as below. Each different organization or each different operating stage will have other measures to conduct resource allocation. In this example, we make a scenario-based allocation that answers some of the questions raised in the previous section.

- Who is the most appropriate performer for executing a specific activity: Table 4 describes which performer is most appropriate to execute an activity. The meaning of the first row data in Table 4 is that: The most appropriate performer for executing activity named “Assign seriousness” is Value 1 and its AP Value is 0.31490482.
- Which activity is most appropriate for each performer: See Table 5 for getting answers of this question. The meaning of the second row data in Table 5 is that: The Activity named *Take in charge ticket* is the most appropriate activity for executing by performer Value 2 with the AP Value 0.360094451.

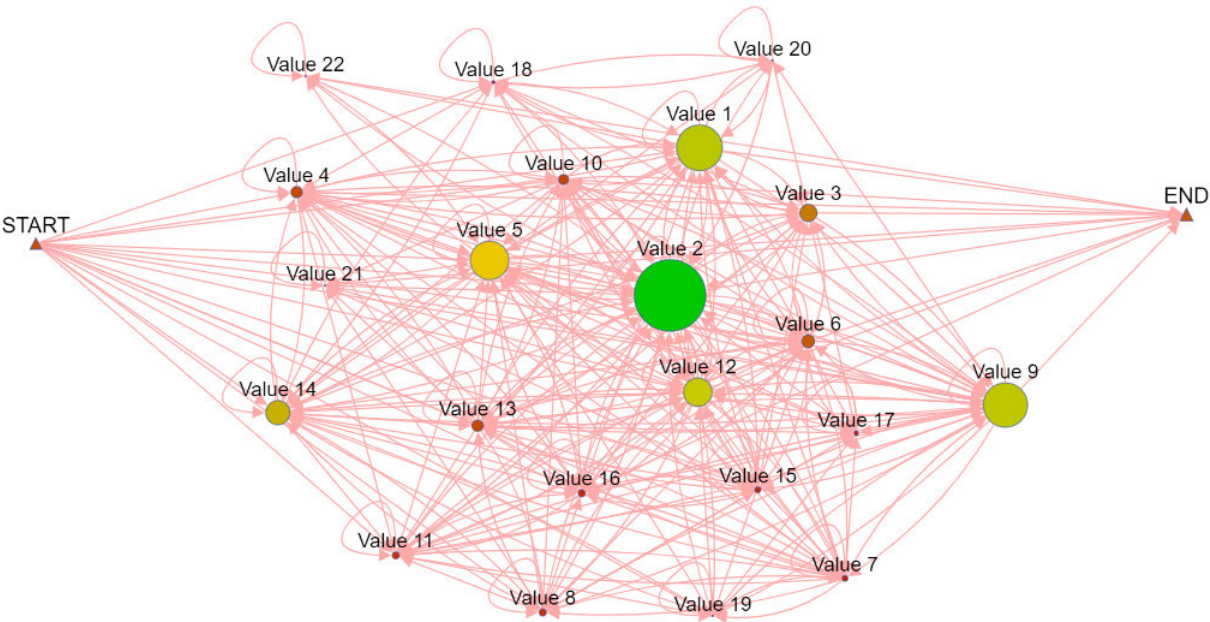


FIGURE 9. Visualization of the discovered PESN in Betweenness centrality metric.

TABLE 4. The list of appropriate performers for executing a specific activity.

Name of Activity	Name of Performer	AP Value
Assign seriousness	Value 1	0.31490482
Closed	Value 5	0.813073896
Create SW anomaly	Value 2	0.641791045
DUPLICATE	Value 2	1
Insert ticket	Value 15	0.245762712
INVALID	Value 2/ Value 1	0.5
Require upgrade	Value 2	0.386554622
Resolve SW anomaly	Value 22	0.615384615
Resolve ticket	Value 2	0.267509532
RESOLVED	Value 2/ Value 1	0.5
Schedule intervention	Value 12	0.6
Take in charge ticket	Value 2	0.301383399
VERIFIED	Value 22	0.666666667
Wait	Value 2	0.263157895

TABLE 5. The list of appropriate performers for executing a specific activity.

Name of Performer	Name of Activity	AP Value
Value 1	Assign seriousness	0.513540291
Value 2	Take in charge ticket	0.360094451
Value 3	Closed	0.973286876
Value 4	Resolve ticket	0.340599455
Value 5	Closed	0.99226254
Value 6	Take in charge ticket	0.448680352
Value 7	Take in charge ticket	0.320143885
Value 8	Assign seriousness	0.552358114
Value 9	Resolve ticket	0.346357935
Value 10	Take in charge ticket	0.454545455
Value 11	Take in charge ticket	0.397003745
Value 12	Resolve ticket	0.391498881
Value 13	Resolve ticket	0.361267606
Value 14	Assign seriousness	0.40813253
Value 15	Resolve ticket	0.329710145
Value 16	Resolve ticket	0.406862745
Value 17	Resolve ticket	0.666666667
Value 18	Assign seriousness	0.490909091
Value 19	Assign seriousness	0.477876106
Value 20	Wait	0.333333333
Value 21	Resolve ticket	0.441176471
Value 22	Resolve SW anomaly	0.8

- Which performers play an essential role in the system’s operation and are difficult to replace by the others: To answer this question, we visualize the PESN with the view of the Degree centrality value of each performer (see Figure 8). The larger size of the circle implies the essential role a performer executed activities in the system. As shown in Figure 8, Value 2, Value 9, Value 12, Value 1, and Value 6 are the performers who play essential roles in the system.
- Which performers can be transferred to perform other tasks? Based on the discovered PAN (see Figure 10), we get the list of performers doing the same activity. Furthermore, we get the list of activities that a single performer performed. With this information, we can transfer a performer to perform

other activities in the same cases. E.g., The activity named *Solve SW anomaly* was performed by Value 22, Value 2. Meanwhile, both Value 22 and Value 2 performed VERIFIED and Solve SW anomaly. So that these performers can be transferred to do each other’s tasks.

- Which performer acts as a working bridge in the process to execute the task? A bridge connects two sides across a river. In a working case (or trace), there are some performers involved in doing the

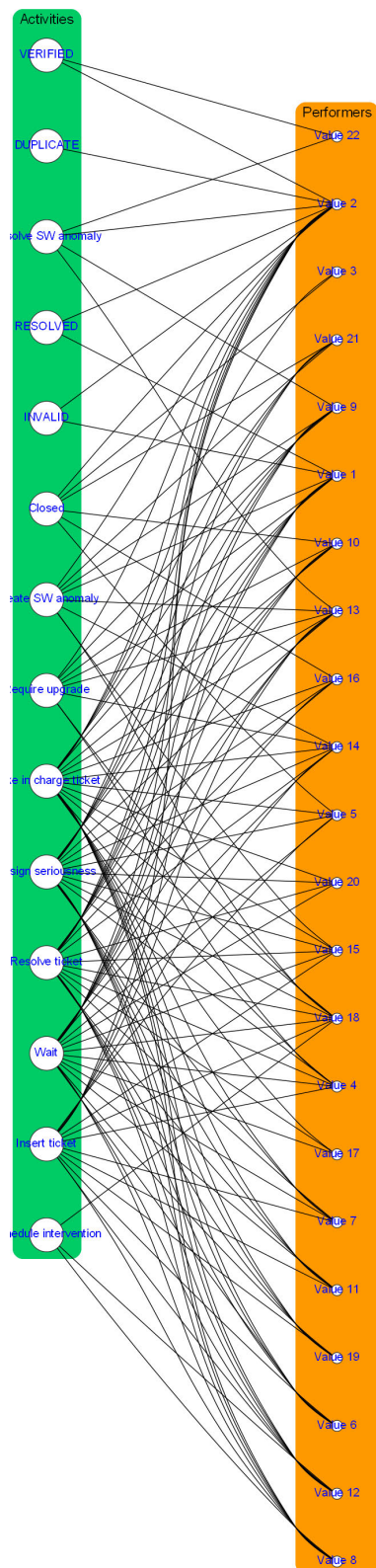


FIGURE 10. The Process-aware affiliation network (PAN) discovered from the Help desk data set.

activities. A performer in this group acted on some tasks and made a connection with the others in the

process-aware information system. We call them a working bridge person. Usually, this person's role plays an essential part in completing the entire workflow. In Figure 9, we visualize the discovered PESN in view of the Betweenness centrality metric. The bigger circle, the higher performer acted as a working bridge in the organization.

VI. CONCLUSION

Comprehending the resources to operate, utilize and serve business activities is the solution to the success and growth of the organization. Therefore, effectively allocating resources to the organization is one of the main concerns in businesses. In order to efficiently allocate resources, managers oblige to understand how employees have associated in operating the activities in the organization, effectively manipulating the available resources of employees in the system, and making appropriate adjustments to the situation of businesses. In this paper, we have introduced our approach for allocating human resources in enterprises based on the human-centric networks metrics. We:

- Summarized some different approaches to allocating resources in the organization. With various methods, researchers represented other efficient solutions to support allocating resources in organizations.
- Introduced our approach using Human-centric networks to analyze the relationships between employees and tasks and employees with the others in the organization while operating the system. Our solution focuses on taking employees as the core factor to build Human-centric networks. Based on the metric values of different networks, we explained how to allocate human resources to the organization with different scenarios.
- Verified the approach with a real-life data set to analyze the discovered networks and obtain values for allocating human resources in the organization. By investigating these discovered networks, we get a comprehension view of the system and thereby give flexible resource allocations for the next operational phase of the system.

In addition to the benefits indicated, our approach still has some limitations:

- With a huge number of employees (thousands of employees), analyzing the relationship between organization members with Human-centric networks will be complicated and require more time and effort.
- Our approach has not considered the performance evaluation index of each employee in the organization.
- The solution is also not connected to the time efficiency of employees.

In future works, we will take these concerns into account and provide solutions that help predict the work efficiency of employees

REPRODUCIBILITY

The full information of analyzed networks, source code, and all related resources to produce the experimental results in this paper can be found and downloaded at <https://github.com/MSYEONKHU/human-centric-network-based-resources-allocation>.

REFERENCES

- [1] *The IEEE XES Standard*. Accessed: Aug. 2021. [Online]. Available: <https://xes-standard.org/>
- [2] H. Willmott, "Business process re-engineering and human resource management," *Personnel Rev.*, vol. 23, no. 3, pp. 34–46, May 1994.
- [3] F. Zucchi and J. S. Edwards, "Human resource management aspects of business process reengineering: A survey," *Bus. Process Manage. J.*, vol. 5, no. 4, pp. 325–344, Dec. 1999.
- [4] S. Wood, "Human resource management and performance," *Int. J. Manage. Rev.*, vol. 1, no. 4, pp. 367–413, 1999.
- [5] D. P. Lepak and S. A. Snell, "The human resource architecture: Toward a theory of human capital allocation and development," *Acad. Manage. Rev.*, vol. 24, no. 1, pp. 31–48, Jan. 1999.
- [6] F. Cakar, U. S. Bititci, and J. MacBryde, "A business process approach to human resource management," *Bus. Process Manage. J.*, vol. 9, no. 2, pp. 190–207, Apr. 2003.
- [7] J. Xu, C. Liu, and X. Zhao, "Resource allocation vs. business process improvement: How they impact on each other," in *Proc. Int. Conf. Bus. Process Manage.*, 2008, pp. 228–243.
- [8] W. Tan, L. Zhao, N. Xie, A. Tang, X. Hu, and S. Tang, "Methods for optimal resource allocation on cooperative task scheduling in cross-organizational business process," in *Proc. Int. Conf. Comput. Social Netw.*, 2018, pp. 126–138.
- [9] H. Y. Chiang and B. M. T. Lin, "A decision model for human resource allocation in project management of software development," *IEEE Access*, vol. 8, pp. 38073–38081, 2020.
- [10] J. Nakatumba and W. M. P. van der Aalst, "Analyzing resource behavior using process mining," in *Proc. Int. Conf. Bus. Process Manage.*, 2009, pp. 69–80.
- [11] Z. Huang, W. M. P. van der Aalst, X. Lu, and H. Duan, "Reinforcement learning based resource allocation in business process management," *Data Knowl. Eng.*, vol. 70, no. 1, pp. 127–145, Jan. 2011.
- [12] Z. Huang, X. Lu, and H. Duan, "A task operation model for resource allocation optimization in business process management," *IEEE Trans. Syst., Man, Cybern. - Part A: Syst. Humans*, vol. 42, no. 5, pp. 1256–1270, Sep. 2012.
- [13] T. Liu, Y. Cheng, and Z. Ni, "Mining event logs to support workflow resource allocation," *Knowl.-Based Syst.*, vol. 35, pp. 320–331, Nov. 2012.
- [14] C. Cabanillas, J. M. García, M. Resinas, D. Ruiz, J. Mendling, and A. Ruiz-Cortés, "Priority-based human resource allocation in business processes," in *Proc. 11th Int. Conf. Service-Oriented Comput. (ICSOC)*, Berlin, Germany, 2013, pp. 374–388.
- [15] A. Wibisono, A. Nisafani, H. Y. Bae, and Y.-J. Park, "On-the-fly performance-aware human resource allocation in the business process management systems environment using Naïve Bayes," in *Proc. Asia-Pacific Conf. Bus. Process Manage.*, 2015, pp. 70–80.
- [16] W. Zhao, L. Yang, H. Liu, and R. Wu, "The optimization of resource allocation based on process mining," in *Proc. Int. Conf. Intell. Comput.*, 2015, pp. 341–353.
- [17] W. Zhao, H. Liu, W. Dai, and J. Ma, "An entropy-based clustering ensemble method to support resource allocation in business process management," *Knowl. Inf. Syst.*, vol. 48, no. 2, pp. 305–330, Aug. 2016.
- [18] A. Wibisono, A. Nisafani, H. Y. H. Bae, and Y.-J. Park, "A dynamic and human-centric resource allocation for managing business process execution," *Int. J. Ind. Eng.: Theory, Appl. Pract.*, vol. 23, no. 4, pp. 1–13, 2016. [Online]. Available: https://repository.its.ac.id/3062/1/198303062012121001_dynamic_wibisono.pdf
- [19] L. Zhang, X. Guo, Z. Lei, and M. K. Lim, "Social network analysis of sustainable human resource management from the employee training's perspective," *Sustainability*, vol. 11, no. 2, p. 380, 2019.
- [20] S. M. Soltis, D. J. Brass, and D. P. Lepak, "Social resource management: Integrating social network theory and human resource management," *Acad. Manage. Ann.*, vol. 12, no. 2, pp. 537–573, Jun. 2018.
- [21] A. Bajaj and R. Russell, "AWSM: Allocation of workflows utilizing social network metrics," *Decis. Support Syst.*, vol. 50, no. 1, pp. 191–202, Dec. 2010.
- [22] P. Ballesteros-Pérez, M. C. González-Cruz, and M. Fernández-Diego, "Human resource allocation management in multiple projects using socio-metric techniques," *Int. J. Project Manage.*, vol. 30, no. 8, pp. 901–913, Nov. 2012.
- [23] G. Havur, C. Cabanillas, J. Mendling, and A. Polleres, "Resource allocation with dependencies in business process management systems," in *Proc. Int. Conf. Bus. Process Manage.*, 2016, pp. 3–19.
- [24] M. Arias, E. Rojas, J. Muñoz-Gama, and M. Sepúlveda, "A framework for recommending resource allocation based on process mining," in *Proc. Int. Conf. Bus. Process Manage.*, 2016, pp. 458–470.
- [25] M. Arias, J. Muñoz-Gama, and M. Sepúlveda, "Towards a taxonomy of human resource allocation criteria," in *Proc. Int. Conf. Bus. Process Manage.*, 2017, pp. 475–483.
- [26] M. Arias, J. Muñoz-Gama, M. Sepúlveda, and J. C. Miranda, "Human resource allocation or recommendation based on multi-factor criteria in on-demand and batch scenarios," *Eur. J. Ind. Eng.*, vol. 12, pp. 364–404, May 2018.
- [27] J. Erasmus, I. Vanderfeesten, K. Traganos, X. Jie-A-Looi, A. Kleingeld, and P. Grefen, "A method to enable ability-based human resource allocation in business process management systems," in *Proc. IFIP Work. Conf. Pract. Enterprise Modeling*, 2018, pp. 37–52.
- [28] M. Arias, R. Saavedra, M. R. Marques, J. Muñoz-Gama, and M. Sepúlveda, "Human resource allocation in business process management and process mining: A systematic mapping study," *Manage. Decis.*, vol. 56, no. 2, pp. 376–405, Feb. 2018.
- [29] X. Zheng, Y. Le, A. P. C. Chan, Y. Hu, and Y. Li, "Review of the application of social network analysis (SNA) in construction project management research," *Int. J. Project Manage.*, vol. 34, no. 7, pp. 1214–1225, Oct. 2016.
- [30] I. Drosouli, G. Theodoropoulou, G. Miaoulis, and A. Voulodimos, "A process mining approach for resource allocation management in a bike sharing system," in *Proc. 24th Pan-Hellenic Conf. Informat.*, Nov. 2020, pp. 327–333.
- [31] W. Zhao, S. Pu, and D. Jiang, "A human resource allocation method for business processes using team faultlines," *Int. J. Speech Technol.*, vol. 50, no. 9, pp. 2887–2900, Sep. 2020.
- [32] Y. Wu, Z. Wang, and S. Wang, "Human resource allocation based on fuzzy data mining algorithm," *Complexity*, vol. 2021, pp. 1–11, Jun. 2021.
- [33] *The ProM Tool*. Accessed: Sep. 2021. [Online]. Available: <https://www.promtools.org/>
- [34] C. Wickens, *Engineering Psychology and Human Performance*. New York, NY, USA: HarperCollins, 1992.
- [35] A. S. Asratian, T. M. J. Denley, and R. Häggkvist, *Bipartite graphs and Their Applications*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
- [36] M. Dumas, W. M. Van der Aalst, and A. H. Ter Hofstede, *Process-Aware Information Systems: Bridging People and Software Through Process Technology*. Hoboken, NJ, USA: Wiley, 2005.
- [37] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social Netw.*, vol. 1, no. 3, pp. 215–239, Jan. 1978.
- [38] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, pp. 35–41, 1977.
- [39] Verenich, Ilya. (2016). *Helpdesk Data Set*. Mendeley Data, V1. Accessed: Oct. 18, 2021. [Online]. Available: <http://dx.doi.org/10.17632/39bp3vv62t.1>
- [40] K. S. Kim, D. L. Pham, and K. P. Kim, " ρ -algorithm: A SICN-oriented process mining framework," *IEEE Access*, vol. 9, pp. 139852–139875, 2021.
- [41] *Graphviz Library*. Accessed: Oct. 18, 2021. [Online]. Available: <https://graphviz.org/>
- [42] *Cytoscape Library*. Accessed: Oct. 18, 2021. [Online]. Available: <https://jts.cytoscape.org/>



respectively. Her research interests include data and knowledge engineering, business process intelligence, process mining, and database systems.

MOON-SOOK YEON received the B.S. degree in computer engineering from Woosong University, South Korea, in 2000, and the M.S. degree in computer engineering from Kyung Hee University, South Korea, in 2004, where she is currently pursuing the Ph.D. degree with the Department of Computer Science and Engineering. She worked as a Researcher and the Director at Ahaetel Company Ltd., Global Networks Company Ltd., and International Company Ltd., from 2008 to 2020,



BPM, workflow-supported social and affiliation networks discovery and analysis, the process-aware Internet of Things architecture and services, process mining, and deep-learning-based process predicting.

DINH-LAM PHAM received the B.S. and M.S. degrees in computer science from Thai Nguyen University, in 2008 and 2010, respectively, and the Ph.D. degree in computer science from Kyonggi University, South Korea, in 2021. He is currently a Research Professor with the Contents Convergence Software Research Institute as well as a Research Member of the Data and Process Engineering Research Laboratory, Kyonggi University. His research interests include workflow systems,



been a Professor with the Department of Computer Engineering, College of Electronics and Information, Kyung Hee University, South Korea. His research interests include ubiquitous data management, data mining, activity recognition, bioinformatics, on-line analytical processing, data warehousing, database systems, spatial databases, and access methods.

YOUNG-KOO LEE received the B.S., M.S., and Ph.D. degrees in computer science from the Korea Advanced Institute of Science and Technology (KAIST), South Korea, in 1988, 1994, and 2002, respectively. From 2002 to 2004, he was a Post-doctoral Fellow with the Advanced Information Technology Research Center, KAIST, and a Post-doctoral Research Associate with the Department of Computer Science, University of Illinois at Urbana-Champaign, USA. Since 2004, he has



at the Electronics and Telecommunications Research Institute (ETRI), South Korea. He is currently a Full Professor with the Computer Science Department and the Founder and a Supervisor of the Collaboration Technology Research Laboratory, Kyonggi University. At Kyonggi University, he is the Dean of the Computerization and Informatics Institute, and the Director of the Contents Convergence Software Research Center, as well. His research interests include groupware, workflow systems, BPM, CSCW, collaboration theory, Grid/P2P distributed systems, process warehousing and mining, workflow-supported social networks discovery and analysis, process-aware information systems, data intensive workflows, and the process-aware Internet of Things. He is also the Vice-Chair of the BPM Korea Forum. He has been in charge of a country-chair (South Korea) and the ERC Vice-Chair of the Workflow Management Coalition. He has also been on the Editorial Board of the journal of *KSII*, and the committee member of the several conferences and workshops.

KWANGHOON PIO KIM (Member, IEEE) received the B.S. degree in computer science from Kyonggi University, South Korea, in 1984, the M.S. degree in computer science from Chungang University, in 1986, and the M.S. and Ph.D. degrees from the Computer Science Department, University of Colorado Boulder, in 1994 and 1998, respectively. He worked as a Researcher and a Developer at Aztek Engineering, American Educational Products Inc., and IBM, USA, as well as

...