ICIC Express Letters
Volume **10**, Number **3**, March **2016**

 ISSN 1881-803X
pp. **649–654**

A STOCHASTIC WORKFLOW PERFORMER-ACTIVITY AFFILIATION NETWORK MODEL

Kwanghoon Pio Kim

Collaboration Technology Research Lab.
Department of Computer Science
Kyonggi University
Suwon-si, Gyeonggi-do 16227, Korea
kwang@kgu.ac.kr

Received August 2015; accepted November 2015

Abstract. *This paper proposes a formal definition of stochastic performer-activity affiliation network model. In dealing with a workflow model, the valid workflow system has to specify its performer-to-activity planned-bindings during buildtime, and fulfill its performer-to-activity observed-bindings on runtime. Applying the probability theory to the performer-to-activity bindings is the very probabilistic performer-to-activity binding formalism. As one of the subsequent works, this paper tries to finish up the conceptual definition of the probabilistic performer-to-activity binding formalism, and the finished formalism is surely the stochastic performer-activity affiliation network model. This paper formalizes the stochastic performer-activity affiliation network model in the information control net of a workflow procedure. Moreover, the graphical representation as well as the mathematical representation of the proposed model is defined and exemplified by an operational example.*

Keywords: Stochastic workflow model, Probabilistic performer-to-activity binding, Stochastic workflow performer-activity affiliation network, Fidelity

1. **Introduction.** In principle, the information control net (which is abbreviated as ICN) [3,11] is a typical open-ended meta-model for understanding, describing, analyzing, and visualizing workflow procedures in organizations. Many organizational information systems have been modeled using ICNs that formally describe several organizational perspectives [3] such as the process, informational, resource (invoked-applications, roles, and performers), and the historical perspectives. So far, there developed several stochastic models of the ICN family in the workflow literature. A typical model is the stochastic information control net [4] that is a formal methodology to specify the probabilistic flow of control [4] in the original information control nets. Another is the stochastic performer-to-activity binding model [2] that is a formal specification to support the probabilistic affiliation bindings between workflow performers and activities in an information control net. Conclusively, the stochastic performer-activity affiliation network model of this paper is a third member of the ICN family, which is based upon the concept of probabilistic affiliation bindings introduced in [2].

We can discover a stochastic performer-activity affiliation network from the probabilistic affiliation binding specification [2] of an extended information control net of workflow procedure. Based upon the stochastic model, we obtain all the possible *planned binding cases*, and then we can compare them with the *observed binding cases* from its enactment event logs. These model-log comparison results used to quantitatively gauge our internal understandings of a workflow-supported organization. In consequence, a new and novel concept of stochastic performer-activity affiliation network as a new member of the ICN family is formally defined in this paper[1], and it eventually allows us to do the qualitative and quantitative analyses (e.g., centrality analyses [5] and prestige analyses [12]), and model-log comparisons [1,13,15] for assessing human-centered resource allocations and distributions as workflow knowledge and intelligence [8,16].

The paper is organized with three main sections excluding introduction and conclusion. In the next section, we summarize the literature survey on the main topics of the paper, and the third section, which is the main part of the paper, starts from introducing the formal representation of the stochastic performer-activity affiliation network model along with its graphical representation of the performer-to-activity bindings in the extended information control net modeling methodology. And the consecutive section describes a mathematical representation of the statistic performer-activity affiliation network model, and exemplifies the proposed model with an operational example of the library book acquisition workflow model, which was firstly introduced in [4]. Finally, we finalize the paper with explaining the implications of the stochastic performer-activity affiliation network model as human-centered organizational knowledge and intelligence.

2. **Related Work.** Recently, technology-supported social networks and organizational behavioral analytics issues have been raised in the IT literature. Naturally, the workflow literature has just started to transit to and focusing on social and collaborative work analysis in workflow-supported organizations, because workflow management systems are "human-centered systems [10]", where workflow procedures must be designed, deployed, and understood within their social and organizational contexts, and they must be conveyed, facilitated, and navigated in large-scaled operational knowledge collections, at the same time. Individuals as employees in workflow-supported organizations have worked intra-organizationally and inter-organizationally under the controls of workflow enactment systems. In particular, it is important to remind that each of the individuals is associated with the essential entity-types of the organizational resources like activity, role, application, and relevant data, when the systems enact the instances of their workflow models. So far, three types of human-centered affiliation networks [7] were introduced in the literature, namely, performer-activity affiliation network [2], performer-role affiliation network [6], and performer-application affiliation network [9].

The next phase of the human-centered affiliation netowrks' research works is to pose the probability theory [14,15] in information control nets. Ellis et al. [4] were the first trial that posed the concept of probabilistic flow of control in the information control nets, and named the concept as the stochastic information control nets. Another trial was fulfilled in [2] that was concerned with the probabilistic bindings between performers and activities in defining information control nets. A third trial was done by [9] from which this paper is extended. Conclusively, we would say that these pioneering works, which are concerning about the human-centered affiliation networking knowledge until now, are the outputs in the stage of initiative research works of the discovery phase. The next stage ought to be the analysis phase. That is, the current stage of the research works is just a half-finished step on the way of forwarding to the analysis phase shifting. In this paper, we try to complete the step opening the door to the analysis phase by extensively investigating formal approaches of the workflow-supported stochastic performer-activity affiliation networks.

3. **Stochastic Workflow Performer-Activity Affiliation Network Model.** A stochastic workflow performer-activity affiliation network consists of two different types of nodes – a set of performers and a set of activities – and a set of edges connecting the different nodal types with probabilities. The formal representation is defined as follows.

[1]This paper is extended from the conference paper [9] published in the proceedings of the BMFSA2015, Kumamoto, Japan. It is also a subsequent result to the paper [2] published in ICIC Express Letters.

- A stochastic workflow performer-activity affiliation network is formally defined as $\Lambda = (\mathbf{C}, \mathbf{A}, \mathbf{E}^p, \mathbf{E}^a, S^p, S^a)$, and it is a connected bipartite network.
- $\mathbf{C}$ is a set of performers $(= \{\phi_1, \ldots, \phi_n\})$, and $\mathbf{A}$ is a set of activities $(= \{\alpha_1, \ldots, \alpha_m\})$. $|\mathbf{A}|$ and $|\mathbf{C}|$ are the numbers of activities and performers, respectively, that are involved in a specific workflow procedure.
- $\mathbf{E}^a = (\mathbf{A} \times \mathbf{C})$ is a set of edges (pairs of activities and performers) and $pr < \varepsilon_a \in \mathbf{E}^a >$ is a probability of the edge, ε_a; Assume that
  - $\mathbf{E}^a(\alpha)$ returns a group of the edges, $(< \alpha, \phi_1 >, \ldots, < \alpha, \phi_k >, 1 \le k \le |\mathbf{C}|)$, in which all the performers $(\phi_1, \ldots, \phi_k)$ are affiliated with a specific activity, α;
  - $pr < \varepsilon_a \in \mathbf{E}^a(\alpha) >$ returns a probability of the edge, $\varepsilon_a =< \alpha, \phi >$;
  - The summation of the probabilities of all the edges, $< \alpha, \phi_1 >, \ldots, < \alpha, \phi_k >$, must be equal to 1.0.
- $\mathbf{E}^p = (\mathbf{C} \times \mathbf{A})$ is a set of edges (pairs of performers and activities) and $pr < \varepsilon_p \in \mathbf{E}^p >$ is a probability of the edge, ε_p; Assume that
  - $\mathbf{E}^p(\phi)$ returns a group of the edges, $(< \phi, \alpha_1 >, \ldots, < \phi, \alpha_l >, 1 \le l \le |\mathbf{A}|)$, in which all the activities $(\alpha_1, \ldots, \alpha_l)$ are affiliated to a specific performer, ϕ;
  - $pr < \varepsilon_p \in \mathbf{E}^p(\phi) >$ returns a probability of the edge, $\varepsilon_p =< \phi, \alpha >$;
  - The summation of the probabilities of all the edges, $< \phi, \alpha_1 >, \ldots, < \phi, \alpha_l >$, may be either less than or greater than 1.0, or be equal to 1.0.
- In a stochastic performer-activity affiliation network from a corresponding workflow procedure, the activity compartment consists of the only manual activity nodes; The performer compartment consists of a set of workflow-performers;
  - Each manual activity node, α_i, has group-degree, d^{group}, of $k : (d^{group}(\alpha_i) = k \wedge 1 \le k \le |\mathbf{C}| \wedge \alpha_i \in \mathbf{A})$;
  - Each performer node has multifunction-degree, d^{multi}, of $l : \left(d^{multi}(\phi_i) = l \wedge 1 \le l \le |\mathbf{A}| \wedge \phi_i \in \mathbf{C}\right)$;
  - Each manual activity node, α_i, has group-member, m^{group}, of $\{\phi_1, \ldots, \phi_k\} : (m^{group}(\alpha_i) = \{\phi_1, \ldots, \phi_k\} \wedge 1 \le k \le |\mathbf{C}| \wedge \alpha_i \in \mathbf{A})$;
  - Each performer node has multifunction-member, m^{multi}, of
  $$\{\alpha_1, \ldots, \alpha_l\} : \left(m^{multi}(\phi_i) = \{\alpha_1, \ldots, \alpha_l\} \wedge 1 \le l \le |\mathbf{A}| \wedge \phi_i \in \mathbf{C}\right).$$
- In a stochastic performer-activity affiliation network, the activity compartment consists of manual activity nodes and automatic activity nodes. Each manual activity node, α_a^m, has group-degree, d^{group}, of k, and each automatic activity node, α_a^a, has group-degree, d^{group}, of $0(zero) : (d^{group}(\alpha_a^m) = k \wedge 1 \le k \le |\mathbf{C}| \wedge \alpha_a^m \in \mathbf{A})$.
- In a stochastic performer-activity affiliation network, the performer compartment consists of a set of workflow-performers affiliated into a corresponding workflow procedure. Each performer node has multi-degree, d^{multi}, of $l : \left(d^{multi}(\phi_i) = l \wedge 1 \le l \le |\mathbf{A}| \wedge \phi_i \in \mathbf{C}\right)$.
- S^p and S^a are finite sets of work-sharing performers and activities, respectively, in external performer-activity affiliation networks.

The graphical representation of a stochastic performer-activity affiliation network is depicted by an undirected bipartite graph that consists of two types of graphical nodes – a set of performers (shaped in the identifiers of performers) and a set of workflow activities (shaped in circle) – and a set of undirected edges between the two nodal types. Figure 1 is a graphical representation of the stochastic workflow performer-activity affiliation network discovered from the extended information control net of the library book acquisition workflow model [4,17]. Also, a stochastic performer-activity affiliation network with the g number of performers and the h number of workflow activities can be mathematically represented by an adjacent matrix with 2-dimension of $g \times h$, and the entries of the adjacent matrix are the probabilities of the corresponding edges.

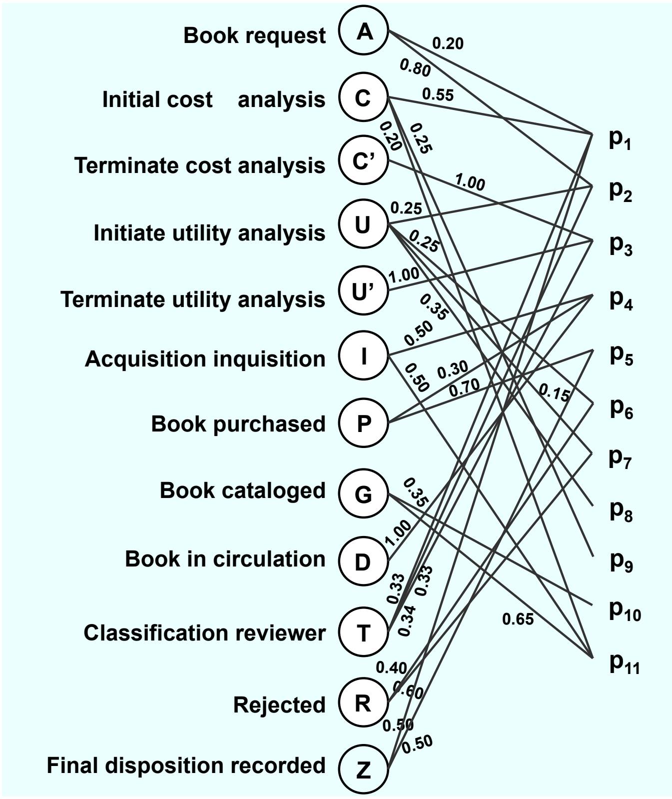

FIGURE 1. A stochastic workflow performer-activity affiliation network from the library book acquisition workflow model [17]

4. **Mathematical Representation of the Proposed Stochastic Model.** Eventually, the stochastic workflow performer-activity affiliation network model has to be transformed into a form of mathematical representation in order to be analyzed to obtain quantitative organizational knowledge. The stochastic model is mathematically represented by a bipartite affiliation matrix as though it is graphically represented by a bipartite affiliation graph. The bipartite affiliation matrix can be realized by either an involvement adjacent matrix or a participation adjacent matrix. That is, a stochastic performer-activity affiliation network is mathematically transformed into a bipartite affiliation matrix that records the presence-probabilities and the absences of g performers at h workflow activities; thus its dimensions are g rows and h columns, respectively. If a certain performer ϕ_i attends on a workflow activity α_j within a probability $pr << \phi_i, \alpha_j >>$, then the entry in the i^{th} and j^{th} cell in the matrix equals $pr << \phi_i, \alpha_j >>$; otherwise the entry is 0. Denoting a stochastic performer-activity affiliation matrix as $\mathbf{Z}$, its $x_{i,j}$ values meet the conditions of Equation (1) as follows:

$$x_{i,j} = \begin{cases} pr << \phi_i, \alpha_j >> & \text{if } \phi_i \text{ is affiliated to } \alpha_j \text{ on probability, } pr << \phi_i, \alpha_j >> \\ \mathbf{0} & \text{otherwise} \end{cases} \tag{1}$$

Assuming that a stochastic performer-activity affiliation network is made up of g workflow-performers and h workflow-activities, then its stochastic performer-activity affiliation adjacent matrix has the elements of $(g+h)\times(g+h)$ in two dimensions. Consequently, by using the involvement adjacent matrix $(\mathbf{Z}_p \Leftrightarrow \mathbf{E}^p)$ and the participation adjacent matrix $(\mathbf{Z}_a \Leftrightarrow \mathbf{E}^a)$, we obtain a stochastic performer-activity affiliation adjacent matrix, $\mathbf{Z}$, as shown in Equation (2). For example, we apply to the stochastic workflow performer-activity affiliation network of Figure 1; its eventual stochastic performer-activity affiliation adjacent matrix and its involvement and participation relationships are shown in Table 1. Based on the stochastic affiliation adjacent matrix, we are able to possibly analyze

TABLE 1. Stochastic performer-activity affiliation matrix of Figure 1

Z	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	A	C	C'	U	U'	I	P	G	D	T	R	Z
p_1	–	0	0	0	0	0	0	0	0	0	0	.20	.55	0	0	0	0	0	0	0	.33	0	.50
p_2	0	–	0	0	0	0	0	0	0	0	0	.80	0	0	.25	0	0	0	0	0	.33	0	0
p_3	0	0	–	0	0	0	0	0	0	0	0	0	0	1.	0	1.	0	0	0	0	.34	0	0
p_4	0	0	0	–	0	0	0	0	0	0	0	0	0	0	0	0	.50	.30	0	1.	0	0	0
p_5	0	0	0	0	–	0	0	0	0	0	0	0	0	0	0	0	0	.70	0	0	0	0	.50
p_6	0	0	0	0	0	–	0	0	0	0	0	0	0	0	.25	0	0	0	0	0	0	.40	0
p_7	0	0	0	0	0	0	–	0	0	0	0	0	0	0	.15	0	0	0	0	0	0	.60	0
p_8	0	0	0	0	0	0	0	–	0	0	0	0	0	0	.35	0	0	0	0	0	0	0	0
p_9	0	0	0	0	0	0	0	0	–	0	0	0	.25	0	0	0	0	0	0	0	0	0	0
p_{10}	0	0	0	0	0	0	0	0	0	–	0	0	0	0	0	0	0	0	.35	0	0	0	0
p_{11}	0	0	0	0	0	0	0	0	0	0	–	0	.20	0	0	0	.50	0	.65	0	0	0	0
A	.20	.80	0	0	0	0	0	0	0	0	0	–	0	0	0	0	0	0	0	0	0	0	0
C	.55	0	0	0	0	0	0	0	.25	0	.20	0	–	0	0	0	0	0	0	0	0	0	0
C'	0	0	1.	0	0	0	0	0	0	0	0	0	0	–	0	0	0	0	0	0	0	0	0
U	0	.25	0	0	0	.25	.15	.35	0	0	0	0	0	0	–	0	0	0	0	0	0	0	0
U'	0	0	1.	0	0	0	0	0	0	0	0	0	0	0	0	–	0	0	0	0	0	0	0
I	0	0	0	.50	0	0	0	0	0	0	.50	0	0	0	0	0	–	0	0	0	0	0	0
P	0	0	0	.30	.70	0	0	0	0	0	0	0	0	0	0	0	0	–	0	0	0	0	0
G	0	0	0	0	0	0	0	0	0	.35	.65	0	0	0	0	0	0	0	–	0	0	0	0
D	0	0	0	1.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	–	0	0	0
T	.33	.33	.34	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	–	0	0
R	0	0	0	0	0	.40	.60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	–	0
Z	.5	0	0	0	.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	–

the stochastic affiliation network model, proposed in this paper, and eventually to obtain several types of human-centered knowledge [8], such as degree, closeness, and betweenness centralities, in workflow-supported organizations.

$$\mathbf{Z} = \begin{bmatrix} \mathbf{0} & \mathbf{Z}_p \\ \mathbf{Z}_a & \mathbf{0} \end{bmatrix} \tag{2}$$

5. **Conclusion.** This paper is aimed to deliver a new and novel concept of the stochastic workflow performer-activity affiliation network model through defining its formal representation as well as its mathematical presentation. That is, we extended the original information control nets so as to support the new concept of stochastic performer-to-activity bindings. From the extended information control net, we can discover the proposed stochastic affiliation network model that is eventually used to analyze the *fidelity* [4] issue of a workflow model in terms of measuring the degree of efficiency in conducting workflow-supported resources allocations and managements without referring the event logs of workflow enactment histories. At this moment, it is important to emphasize that we did not differentiate the single-performer binding activity type from the group-performer binding activity (realtime groupware activity) type that implies that a group of performers is simultaneously assigned to cooperatively perform a single manual activity. Therefore, as future work, we need to consider the concept of stochastic group-performer bindings in formalizing the proposed concept.

Acknowledgements. This research was supported by the sabbatical year research grant (2014) of Kyonggi University. Additionally, this work was backed from the GRRC research program of the Gyeonggi Province, Republic of Korea.

REFERENCES

[1] W. P. M. Aalst, A. K. Alves de Medeiros and A. J. M. M. Weijters, Process equivalence: Comparing two process models based on observed behavior, *Lecture Notes in Computer Science*, vol.4102, pp.129-144, 2006.

[2] H. Ahn and K. P. Kim, A stochastic activity-to-performer affiliation binding formalism in ICN-based workflow models, *ICIC Express Letters*, vol.9, no.12, pp.3213-3219, 2015.

[3] C. A. Ellis, Information control nets: A mathematical model of office information flow, *Proc. of the 9th Annual ACM Conference on Simulation, Measurement, and Modeling of Computer Systems*, 1979.

[4] C. A. Ellis, K. Kim, A. Rembert and J. Wainer, Investigations on stochastic information control nets, *Information Sciences*, vol.194, pp.120-137, 2012.

[5] K. Faust, Centrality in affiliation networks, *Journal of Social Networks*, vol.19, pp.157-191, 1997.

[6] H. Kim, H. Ahn and K. Kim, Modeling, discovering, and visualizing workflow performer-role affiliation networking knowledge, *KSII Trans. Internet and Information Systems*, vol.8, no.2, pp.691-708, 2014.

[7] K. P. Kim, Discovering activity-performer affiliation knowledge on ICN-based workflow models, *Journal of Information Science and Engineering*, vol.29, no.1, pp.79-97, 2013.

[8] K. P. Kim, Human-centered workflow intelligence, *International Conference on Advanced Communications Technology*, Phoenix Park, Pyeongchang, South Korea, 2014.

[9] K. P. Kim, A stochastic workflow performer-to-activity affiliation network model, *Proc. of the 28th Annual Conference of Biomedical Fuzzy System Association*, Kumamoto, Japan, 2015.

[10] K. Kim, S. Paik and C. Ellis, Actor-oriented workflow model, *Proc. of the 2nd International Symposium on Cooperative Database Systems for Advanced Applications*, Wollongong, Australia, 1999.

[11] K. Kim and C. A. Ellis, ICN-based workflow model and its advances, *Handbook of Research on Business Process Modeling*, IGI Global, ISR, 2009.

[12] D. Knoke and S. Yang, Quantitative applications in the social sciences, *Social Network Analysis*, 2nd Edition, SAGE Publications, 2008.

[13] Y. Lei and M. P. Singh, A comparison of workflow metamodels, *Proc. of the ER-97 Workshop on Behavioral Modeling and Design Transformations*, Los Angeles, USA, 1997.

[14] T. A. B. Snijders, G. G. van de Bunt and C. E. G. Steglich, Introduction to stochastic actor-based models for network dynamics, *Social Networks*, vol.32, pp.44-60, 2010.

[15] D. Varacca and M. Nielsen, Probabilistic petri nets and Mazurkiewicz equivalence, *Proc. of the Conference on Decision and Control*, vol.5, pp.4104-4109, 2001.

[16] The Gartner Group, *Gartner Reveals Five Business Intelligence Predictions for 2009 and Beyond*, http://www.gartner.com/it/page.jsp?id=856.

[17] *University of Colorado Library Renovation*, http://ucblibraries.colorado.edu/learningcommons/index.htm.