A Conceptual Approach for Discovering Proportions of Disjunctive Routing Patterns in a Business Process Model

Kyoungsook Kim[1], Moonsuk Yeon[1], Byeongsoo Jeong[1] and Kwanghoon Kim[2*]
[1] Department of Computer Engineering, Kyunghee University
1732, Deogyeong-daero, Giheung-gu, Yongin-si, Gyeonggi-do 17104, Republic of Korea
[e-mail: {khmjmc, msyeon, jeong}@khu.ac.kr]
[2] Department of Computer Science, Kyonggi University
154-42 Gwangkyosan-ro, Youngtong-gu, Gyeonggi-do, 16227, Republic of Korea
[e-mail: kwang@kgu.ac.kr]
*Corresponding author: Kwanghoon Pio Kim

Received October 20, 2016; revised December 13, 2016; accepted January 6, 2017; published February 28, 2017

Abstract

The success of a business process management system stands or falls on the quality of the business processes. Many experiments therefore have been devoting considerable attention to the modeling and analysis of business processes in process-centered organizations. One of those experiments is to apply the probabilistic theories to the analytical evaluations of business process models in order to improve their qualities. In this paper, we excogitate a conceptual way of applying a probability theory of proportions into modeling business processes. There are three types of routing patterns such as sequential, disjunctive, conjunctive and iterative routing patterns in modeling business processes, into which the proportion theory is applicable. This paper focuses on applying the proportion theory to the disjunctive routing patterns, in particular, and formally named proportional information control net that is the formal representation of a corresponding business process model. In this paper, we propose a conceptual approach to discover a proportional information control net from the enactment event histories of the corresponding business process, and describe the details of a series of procedural frameworks and operational mechanisms formally and graphically supporting the proposed approach. We strongly believe that the conceptual approach with the proportional information control net ought to be very useful to improve the quality of business processes by adapting to the reengineering and redesigning the corresponding business processes.

Keywords: Business Process; Workflow Process; Disjunctive, Conjunctive, and Iterative Routing Patterns, Stochastic Information Control Nets, Proportional Information Control Nets, Temporal Workcases, Process Warehouse

A preliminary version of this paper was published with an outstanding paper award at AP-IST 2016.
This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (Grant No. 2012006971).

1. Introduction

According for the business process (or workflow) automation technologies to swiftly grow and be increasingly used by many newly formed and web-based organizations, we need to analyze a new type of requirements and demands concerning about business process intelligence and quality [1][2][4]. Especially, in order to improve the quality of business processes with the high, consistent and predictable efficiency, it is so important for the business process knowledge to be discovered from runtime execution logs of business process models. Note that there exist the various perspectives of the business process knowledge such as control-flow perspective [3][4][10][14][21][23], data-flow perspective, behavioral perspective [7], and organizational resource perspective [8][11][15][16][17] in particular. Out of those perspectives, the main emphasis of this paper is on the control-flow perspective, and from which we are able to derive the answers to the following questions:

- How well is a business process model reflecting the control-flow aspect of all the administrational, managerial and operational activities in the organization?
- How do we quantitatively measure the differences between the planned business process models on build-time and the observed business process models on runtime?

There are four types of routing patterns in the business processes' control-flow perspective such as sequential, disjunctive (alternative), conjunctive (parallel) and iterative (loop) routing patterns. We focus on the disjunctive routing pattern in particular, and try to measure the proportions of control-flow branches by observing their enactment instances of a corresponding business process model. In other words, we are especially interested in the logs (audit trail) of business processes. By discovering the relative frequencies of possible control-flow branches from the enactment event logs, we can figure up the occurrence proportions to all the alternative control-flow branches in a business process model. By observing the proportions of all the control-flow branches on the Exclusive-OR forks, we can calculate the probability of every execution sequence of a corresponding business process model. Then we can compare the observed proportions to the planned proportions, respectively, and use the control-path oriented knowledge to reengineer and redesign the corresponding business process.

In this paper, we propose a conceptual approach for discovering the observed frequency (proportion) of each control-flow branch of the disjunctive routing patterns from the enactment event logs of a business process model. The conceptual approach is based upon the control-path oriented process knowledge analysis algorithm that was firstly introduced in [3]. The most essential issue in the control-path oriented process knowledge is to analyze all the possible execution sequences and their enactment frequencies in a business process model, and the knowledge will be eventually used to redesigning and reengineering with maintaining the higher degree of quality on the corresponding business process model.

2. Related Works and the Scope

Our work is concerned with Proportional Information Control Nets aiming at supporting the model−log comparison of workflow and business process models. The first conceptual initiation of the model−log comparison was ignited by the novel concept of the control-path oriented workflow intelligence [21]. Thereafter, several research outcomes dealing with this issue of rediscovering workflow and business process models have been released in the literature. This paper is one of those trials and outcomes. Recently, there is another conceptual extension of the workflow and business process intelligence issue covering and enhancing the model−log comparisons, which is the stochastic information control net [8][10]. We summarize and describe the relevance and similarity of some of that concept here. In the stochastic information control nets, we have a feature that allows the placement of probabilities, or probability distributions on the out-going arcs of OR-fork nodes. Additionally, it allows meaningful placement of probabilities on AND-fork nodes, too. However, the probabilities placed on OR-fork/AND-fork nodes in the stochastic information control nets are the estimated values (or the probability distributions) for the sake of simulating works, whereas the probabilities placed on OR-fork/AND-fork nodes in the proportional information control nets proposed in this paper are the observed values discovered from the enactment event logs of the corresponding business process models.

So far, many publications were released in the area of the model−log comparisons, which is named as workflow mining or business process mining, and they are mainly concerning about the quality and reengineering of business process models through the comparisons of workflow models [20] to event logs [3][5][6][7][13] as well as the comparisons of models to other models by using the probability theory [8][11][14][17][19]. Some of those works were especially issuing the managerial concerns with stochastic definitions of human-centric behavioral knowledge [4][12][15][16][18] of business process models. Summarily speaking, however, all of these works mainly focused on the conformance of models by either specifically comparing a model to a log or probabilistically estimating the behavioral patterns of a model. Our work addressed in this paper tries to answer to the question, "How well (or how much) is a business process model conformed to the planned activities and their behavioral patterns?," by directly ransacking its enactment event logs. Conclusively, through the outcome of our work like the proposed conceptual approach and its related formalisms, we are able to judge of over-fitting or under-fitting on a business process model, and eventually to calculate the level of fidelity of the corresponding model. We note that none of these previous works have used proportions on branches of OR-forks, AND-forks and LOOP-forks in the way that this paper proposes to help quantify their analysis results. In this paper, we confine the scope to the OR-forks' branches.

3. The Conceptual Initiation

The conceptual initiation of the idea comes from a novel concept of the control-path oriented workflow intelligence [21], which was firstly announced by the authors' research group. As illustrated in **Fig. 1**, we introduce the conecpt of the control-path oriented business process intelligence. The middle part of the figure depicts a series of observed routing patterns (*e.g.* 5 types of control-paths) and their frequencies that are discovered from the running time traces and execution logs of a corresponding busienss prpcess model, and the left-hand side of the figure represents a proportional business process model constructed from the observed routing patterns. The examplified business process model in the figure consists of three mutually Exclusive-OR gateways, through which we generate five different types of control-paths based on the corresponding business process model of the build-time and also through which we are able to possibly discover 5 difference types of firing sequences based upon the run-time's. The proportional business process model in the figure conveys a very valuable knowledge, which is called the concept of control-path oriented intelligence, that represents the indicator of control-paths' significance after being elapsed a certain amount of time. Based upon the indicator, we can evaluate the corresponding business process model, and qualitatively redesign the model by separating the control-paths having the smaller frequencies from the control-paths having the higher frequencies in the next phase of the quality improvement and refinement of the model.

More precisely speaking, the conceptual approach proposed in this paper is deployed by a series of procedural transformations for analyzing control-paths and their enactment frequencies from the enactment event histories of a specific business process model. In this paper, we try to extensively refurbish the concept and apply it into discovering a proportional business process model from the enactment event logs. The first transformation is to rediscover a set of observed routing patterns, each of which can be modeled into a workcase model, from the enactment event histories logged by executing a corresponding business process model. At the same time, it is necessary to count the frequencies of the rediscovered workcase models and measure the proportions. The second transformation is to discover a proportional business process model, which is formally and graphically modeled by a proportional information control net, from the rediscovered temporal models (observed routing patterns) and their frequencies. In terms of defining the proportional business process model, we need to enhance the graph model so as to add the proportions on each edge. In other words, we are unable to discover the proportions of all the associated edges with the Exclusive-ORs through only the frequencies of the observed routing patterns. So, we need to develop an additional scheme for discovering the proportions of the observed routing patterns from the enactment event logs of a corresponding business process model. In this paper, we conceive a conceptual approach to theoretically support this additional scheme.

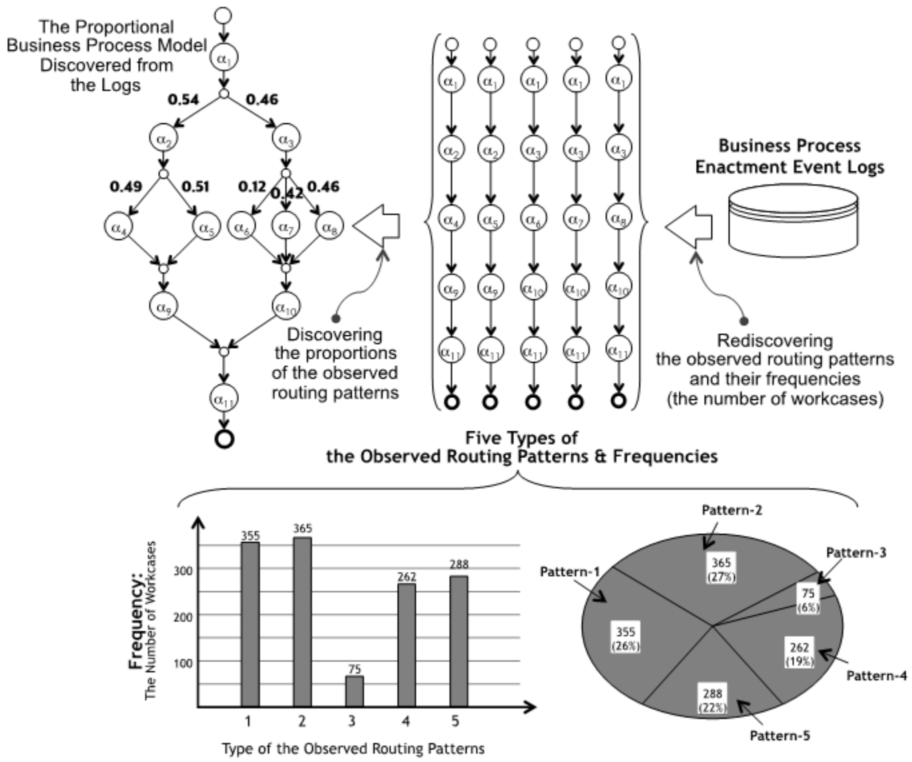

Fig. 1. The Conceptual Initiation for Discovering Proportions in a Business Process Model

In terns of defining the concept of proportions in a business process model, we would adopt a way of formalism for the stochastic information control nets [10], and semantically extend the meaning of stochastics to the concept of proportions. There are a number of AND/OR graph models [7][10][18][19][22] in the business process modeling literature. In particular, we take into account an Exclusive-OR (disjunctive routing pattern) construct with holding proportions. Through such construct, we are able to model decision-making activities in a business process model, and the conceptual approach to be proposed in this paper is able to measure probability (proportion) of each selection of the decision-making choices by rediscovering the observed Exclusive-OR routing patterns and their frequencies from the enactment event logs of a corresponding business process model. For the sake of the formal representation of a proportional business process model, we extensively revise the original information control net methodology [9][19] by supporting the concept of proportions. That is, we conceive the idea of proportional information control nets (abbreviated as p-ICNs) by extending the classic ICN definitions with appending proportions to outgoing arcs of an Exclusive-OR construct. Each arc in the p-ICN graph so has a proportional value after all. In any given Exclusivel-OR node, the sum of all the proportions of the arcs must be one as shown in **Fig 2**. Also, all the activity nodes in a single p-ICN have to have only one outgoing arc (`out-degree = 1`), the proportion of the activity node's arc so must be one. For convenience, such an activity node's arcs are unlabeled in our graphical representation of p-ICNs.

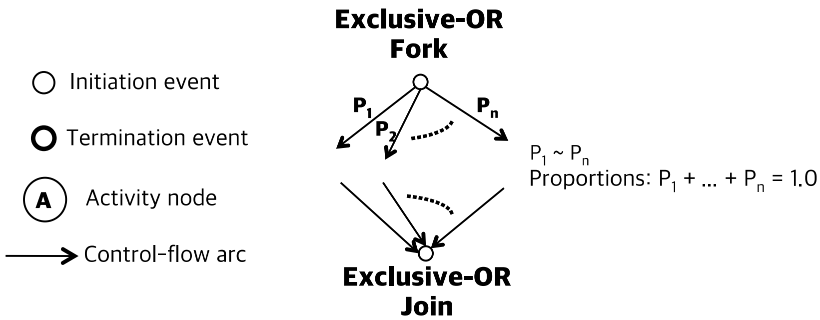

Fig. 2. The Proportional Properties on the p-ICN's Disjunctive (Exclusive-OR) Routing Patterns

Fig. 2 illustrates the proportional properties of an Exclusive-OR routing pattern of p-ICN. Note that the Exclusive-OR fork has the proportions of $P_1 \sim P_n$ that are attached to the outgoing arcs (with $\sum p_i = 1$), but all the incoming arcs of the Exclusive-OR join have no specified proportions, because the proportion of each arc is `1.0`. The number of outgoing arcs of the Exclusive-OR join node must be one. Also, all of Exclusive-OR fork nodes and Exclusivel-OR join nodes in a p-ICN have to keep the properties of matched-pair and proper-nesting. The formal definition of p-ICNs will be presented in the next section.

4. The Conceptual Approach for Discovering the Proportions

In this paper, our emphasis is placed on the quality and reengineering of business process models that are especially built by a combination of disjunctive routing patterns. That is, we used to model a business process model by a combination of four types of routing patterns, such as sequential, disjunctive, conjunctive, and iterative routing patterns, in defining its control-flow perspective. We are particularly interested in only the disjunctive routing pattern, because this pattern produces a series of alternative control-flow paths, and its enactment history gives us very meaningful knowledge to be used not only for evaluating the quality of a corresponding business process model, but also for redesigning and reengineering the model, itself. The eventual goal of the conceptual approach is to discover a proportional business process model from its enactment event histories logged by executing all the four types of routing patterns. However, we confine the conceptual approach to the matter of the sequential and disjunctive routing patterns in this paper. We so assume that the enactment event log is only containing the execution histories of those business process models built by a combination of sequential and/or disjunctive routing patterns. In this section, we deploy the detailed formalisms of the conceptual approach for rediscovering a proportional business process model from its enactment event histories.

4.1 Formalisms for the Conceptual Framework

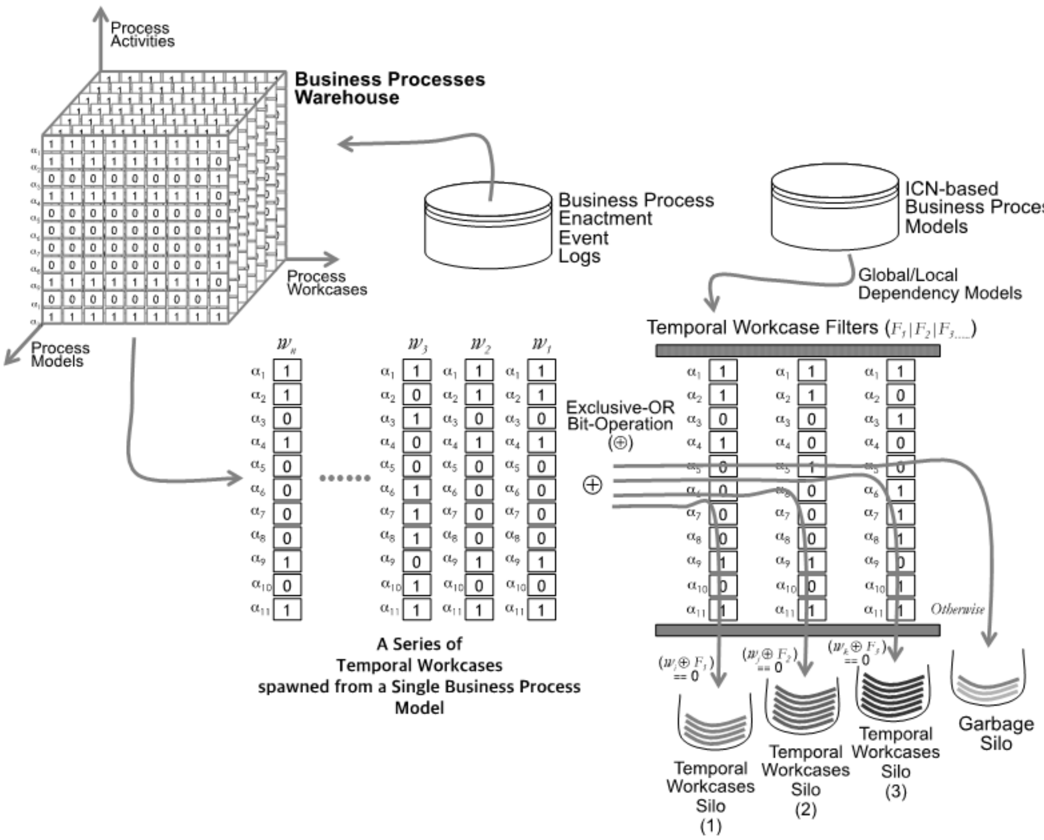

Fig. 3. A Procedural Framework to Discover the p-ICN's Proportions

The conceptual approach is concretized by a conceptual framework that is composed of a series of procedural concepts and their algorithmic formalisms such as business process enactment event logs, business process warehouses, temporal workcases, and the cascading filters. **Fig. 3** is to illustrate these procedural concepts and the names of their algorithms in the conceptual framework to be used for realizing the conceptual approach proposed in this paper. Especially, our focus is on a concept of the cascading filters, and its algorithmic details, further.

4.2 p-ICNs with Disjunctive Routing Patterns

Proportional Information Control Nets (which is abbreviated as p-ICN) using the terminology of the graph theory for representing proportional business process models can be formally defined as follows:

> *A p-ICN is a directed graph where every node (or vertex) has a unique ID number, a (not necessarily unique) label, and a type; every arc (or edge) is defined as an ordered pair $<n_1, n_2>$ of nodes, and has a proportion (in the (0,1] range) associated with it denoted $pr<n_1, n_2>$. Directed arcs connect nodes, and semantically denote precedence. There are 5 node types: initiation (initial-node) event, termination (final-node) event, activity, OR-fork, and OR-join. Note that p-ICNs must be properly nested, and non-empty. The properly nested means that all the control-flow constructs must be entirely contained within other control-flow constructs, and that they may be neither interleaved nor structured with "GO TO" arcs.*

- Each activity node, n_a, has in-degree, din, of one, and out-degree, dout, of one. The exception is the one unique initial event-node, n_I such that $d_{in}(n_I) = 0$; and the one unique final event-node, n_F, such that $d_{out}(n_F) = 0$. Semantically, activity nodes denote activities, events or actions; and have at most a single predecessor node, n_{pred}, and at most a single successor node, n_{next}. Proportionally, $pr<n_a, n_{next}> = 1.0$.
- Each Exclusive-OR fork node, n_{EOF}, has in-degree, $d_{in}(n_{EOF}) = 1$, and out-degree, $d_{out}(n_{EOF}) = m$ where m is an integer, $m > 0$. Semantically, Exclusive-OR forks denote exclusive choice or conditional branch control nodes. They have a single predecessor node, and more than one successor nodes, $n_1, n_2, \ldots, n_m$, of which only one can be chosen to execute next during any execution sequence (Exclusive-OR semantics.) Proportionally, $pr<n_{EOF}, n_{next}> > 0.0$ for all out-arcs, and the node is normalized if $pr<n_{EOF}, n_{next}> = 1.0$, where the sum is from 1 to m.
- Each Exclusive-OR join node, n_{EOJ}, has out-degree, $d_{out}(n_{EOJ}) = 1$, and in-degree, $d_{in}(n_{EOJ}) = m$ where m is an integer, $m > 0$. Proportionally, Exclusive-OR joins de- note control-flow alternatives coming together. They have a single successor node, and more than one predecessor nodes, $n_1, n_2, ..., n_m$. The completion of any one predecessor can trigger the execution of n_{EOJ} during any execution sequence (Exclusive-OR semantics.) Proportionally, $pr<n_{EOJ}, n_{next}> = 1.0$.

4.3 Business Process Enactment Event Logs

On the procedural framework of **Fig. 3**, the starting point is the business process enactment event logs that were stored as the enactment histories of all the business process models (or packages) managed by a underlying business process management system. According as an instance of a business process model executes, its temporal execution sequence is produced and logged into the backup databases or files; this temporal execution sequence is called business process event trace (or temporal workcase). The business process event trace is made up of a set of activities' event logs.

In general, we consider storing the business process event traces in an XML format. The WfMC (Workflow management coalition) has leased a standardized specification for the XML-based business process event traces, which is named as BPAF (Business Process Audit Format). Also, prior to the BPAF specification, an XML-based workflow (or business process) event log language was studied and proposed in [13] as one of the business process mining mechanisms. In this paper, we formally define the format of business process enactment event logs through the following definitions:

[Definition 1] ***Business Process Enactment Event Log***. *Let bpe=(α,pc,wf,c,ac,ε,p*,t,s) be an event, where α is a workitem (activity instance) number, pc is a package number, wf is a business process number, c is a business process instance (workcase) number, ac is an activity number, ε is an event type, which is one of {Scheduled, Started, Changed, Completed}, p is a participant or performer, t is a timestamp, and s is an activity state, which is one of {Inactive, Active, Suspended, Completed, Terminated, Aborted}. Note that * indicates multiplicity.*

[Definition 2] ***Business Process Warehouse***. *Let* $I_i = \{c_1^i, ..., c_m^i\}$ *be a set of completed business process instances (m is the number of the business process instances) that have been instantiated from a business process model,* I_i*. A business process warehouse consists of a set of business process logs,* $WL(I_1), ..., WL(I_n)$*, where* $WL(I_i) = \forall WT(c^i \in I_i)$*, and n is the number of business process models managed in a system.*

[Definition 3] ***Business Process Trace*** *(Temporal Workcase). Let BPT(c) be a trace of business process instance, c, where* $BPT(c) = (bpe_1, . . . , bpe_n)$*. Especially, the business process trace is called a temporal workcase, TW(c), if all activities of its underlying business process instance are successfully completed. There are three types of temporal workcases according to the events type—Scheduled, Started, Completed:*

- ScheduledTime Temporal Workcase
  $\{bpe_i \,|\, bpe_i.c = c \wedge bpe_i.e = \text{'Scheduled'} \wedge bpe_i.t \leq bpe_j.t \wedge i < j \wedge 1 \leq i, j \leq n\}$, which is a temporally ordered business process event sequence based upon the scheduled time stamp.
- StartedTime Temporal Workcase
  $\{bpe_i \,|\, bpe_i.c = c \wedge bpe_i.e = \text{'Started'} \wedge bpe_i.t \leq bpe_j.t \wedge i < j \wedge 1 \leq i, j \leq n\}$, which is a temporally ordered business process event sequence based upon the stated time stamp.
- CompletedTime Temporal Workcase
  $\{bpe_i \,|\, bpe_i.c = c \wedge bpe_i.e = \text{'Completed'} \wedge bpe_i.t \leq bpe_j.t \wedge i < j \wedge 1 \leq i, j \leq n\}$, which is a temporally ordered business process event sequence based upon the completed time stamp.

Based on these formal definitions, we are able to extract a full set of the temporal workcases spawned from a specific business process model, which become the input of the temporal workcase filters proposed in this conceptual approach. Consequently, according to the timestamp types, we can build three different types of business process enactment event logs and their business process warehouses, such as Scheduled Time-based Warehouse, StartedTime-based Warehouse, and CompletedTime-based Warehouse.

4.4 A Formal Model of Temporal Workcases: The Workcase Model

The formal concept of the workcase model was firstly defined in [14], in where the authors proposed the concept and its amalgamating algorithm as a novel algorithmic approach for rediscovering business process models from the corresponding enactment event histories. Each of the temporal workcases is formally represented by a workcase model, and at the same time it is mechanically modeled by a string of bits, where each bit of the string implies the presence (or absence) of activity in a corresponding temporal workcase, as shown in **Fig. 3**. In the following **[Definition 4]**, we formally define the workcase model, and also it can be graphically represented by a shape of bit-string. The primary reason that we use a bit-string representation of a temporal workcase is on enhancing the operational efficiency in implementing the cascading filter mechanism that is able to classfy all the temporal workcases into one of the silos, each of which archives all such temporal workcases having the same bit-string.

[Definition 4] ***Workcase Model (WCM)***. *A workcase model is formally defined through 3-tuple* $W = (\omega, P, S)$ *over an activity set,* **A**, *where*

- *P is a predecessor activity of some external workcase model, which is connected into the current workcase model;*
- *S is a successor activity of some external workcase model, which is connected from the current workcase model;*
- $\omega=\omega_i \cup \omega_o$, *where,* $\omega_o : \mathbf{A} \rightarrow \wp(\alpha \in \mathbf{A})$ *is a single-valued mapping function of an activity to its immediate successor in a temporal workcase, and* $\omega_i : \mathbf{A} \rightarrow \wp(\alpha \in \mathbf{A})$ *is a single-valued mapping function of an activity to its immediate predecessor in a temporal workcase.*

In the procedural framework of **Fig. 3**, there are, for example, three patterns of temporal workcases associated with those Exclusive-OR nodes in a specific business process model, which are differentiated from each other according to such temporal information (the events' timestamps) logged at the time when the corresponding activities' workitems was performed. Especially, each of the temporal workcase patterns is used to build a temporal workcase filter as shown in the righthand-side of **Fig. 3**. The enactment engine of a business process management system maintains three sorts of timestamps such as scheduled, started, and completed timestamps when it stores the enactment event timestamps of activities; there are three sorts of temporal workcases in the business process enactment event logs, and from which we can build three sorts of busienss process warehouses, too. Accordingly, the workcase model subdivides into three types, such as scheduled, started, and completed workcase models. In this paper, we do not cope with these specific types of the workcase models in detail.

4.5 The Temporal Workcase Filters

This section describes a procedural approach for generating a series of temporal workcase filters from a business process model. We will use the approach to measure the proportions of the disjunctive routing patterns, and eventually to discover a proportional business process model from its enactment event logs. As illustrated in **Fig. 4**, the temporal workcase filters, which can be represented by bit-sequences, are generated from a specific business process model by analyzing the possible firing sequences of activities along with the Exclusive-OR split-nodes. In order to analyze a possible set of firing sequences in a business process model, it is necessary to develop a sort of control-path analysis algorithm. Fortunately, [21] proposed a novel algorithm for analyzing control-paths of an information control net, and we would use the algorithm to generate the firing sequences from business process models. Each of the activity firing sequences is formally represented by the workcase model defined in the previous subsection. The approach also needs another algorithm that is able to automatically transform from each of the temporal workcases to a filter being realized in a bit-sequence to be applied into the filtering function *via* a series of exclusive-OR bit-operations.

Fig. 4 is to illustrate an operational example of the procedural framework, which shows constructing the temporal workcase filter with 5 filters from an information control net of the examplified business process model that consists of 3 Exclusive-OR nodes with 5 firing sequences (control-paths). To obtain a possible set of the firing sequences, we applied the examplified information control net to the activity firing sequences analysis algorithm [21], and graphically represented the result of the 5 firing sequences in the workcase models. Also, the bit-sequences generation algorithm produced the 5 temporal workcase filters from the input of the 5 workcase models, each of which is graphically represented by a box-sequence in **Fig. 4**.

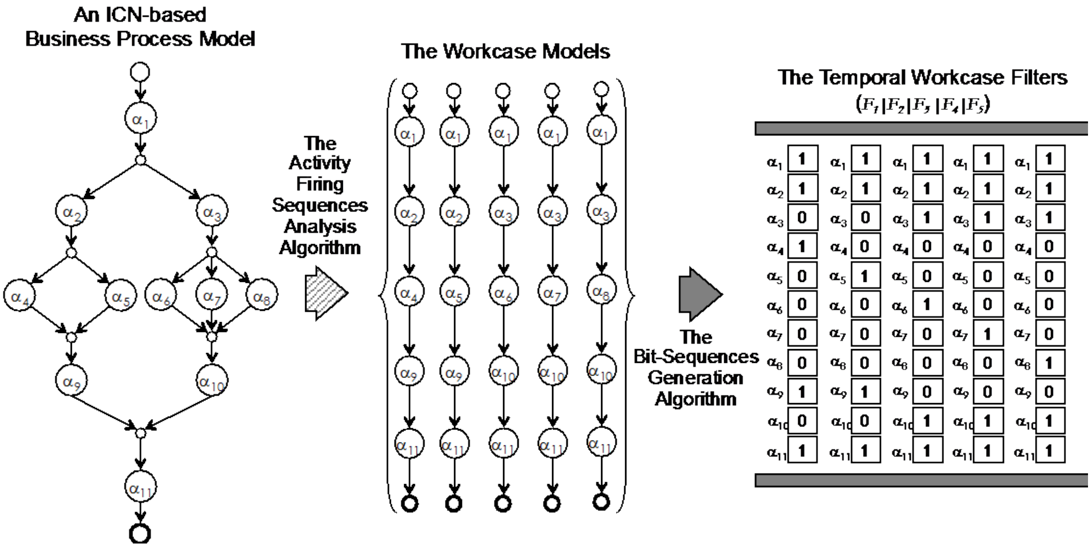

Fig. 4. A Procedural Framework for Constructing the Temporal Workcase Filters

4.6 The Proportion Decision-Making Model

The eventual goal of the conceptual approach proposed in this paper is to discover a proportional business process model from its enactment event logs. Until now, we have been excogitating a series of conceptual and procedural formalisms that is able to discover the number of workcases belonging to each of the disjunctive routing patterns (the firing sequences of activities) in a corresponding business process model. However, the discovered numbers are definitely not enough to complete this goal. We need an additional step that is to build a proportional business process model by applying the discovered numbers of the temporal workcases to measure the proportions of all the branches on the Exclusive-OR routings. In order to complete the final step of the proportions discovery works, we propose a novel concept of the proportion decision-making model that is formally defined as follows:

[Definition 5] ***Proportion Decision-Making Model***. *Let* $\mathbf{M}$ *be a proportion decision-making model, that is formally defined as* $\mathbf{M} = (\chi, \theta)$ *over a set of activities,* $\mathbf{A}$, *where*

– $\chi = \chi_i \cup \chi_o$

where, $\chi_o : \mathbf{A} \rightarrow \wp(\mathbf{A})$ *is a multi-valued mapping of an activity to an another set of activities, each member of which has the disjunctive-type dependency, and*

$\chi_i : \mathbf{A} \rightarrow \wp(\mathbf{A})$ *is a single-valued mapping of an activity to an another activity that is one of the members in* $\{\alpha_I$, *Exclusive-OR–split*$\}$;

– $\theta = \theta_i \cup \theta_o$

where, θ_i *: the number of discovered temporal workcases on each arc,* $(\chi_i(\alpha), \alpha)$*; and*

θ_o : the number of discovered temporal workcases on each arc, $(\alpha, \chi_o(\alpha))$, where $\alpha \in \boldsymbol{A}$;

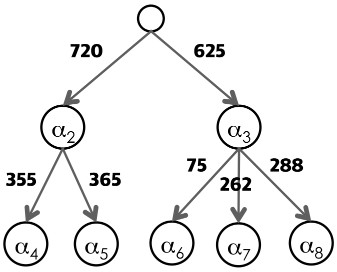

Fig. 5. A Proportion Decision-Making Model

In order to derive a proportion decision-making model from a business process model, we need to devise an algorithm that is able to construct a proportion decision-making model from an information control net of the correponding business process model. Fortunately, we can adopt the concept of activity control dependencies that was proposed in [3], and we revise the algorithm [3] so as to extract a proportion decision-making model from an activity dependent net that is the formal represention of the control dependency relationship among the activities in a specific information control net. In this paper, we won't provide the details of the revised algorithm. The graphical representation of the proportion decision-making model is given in **Fig. 5**; the numerical values on the edges imply the numbers of temporal workcases on the disjunctive routing patterns of the business process model examplified in the previous sections.

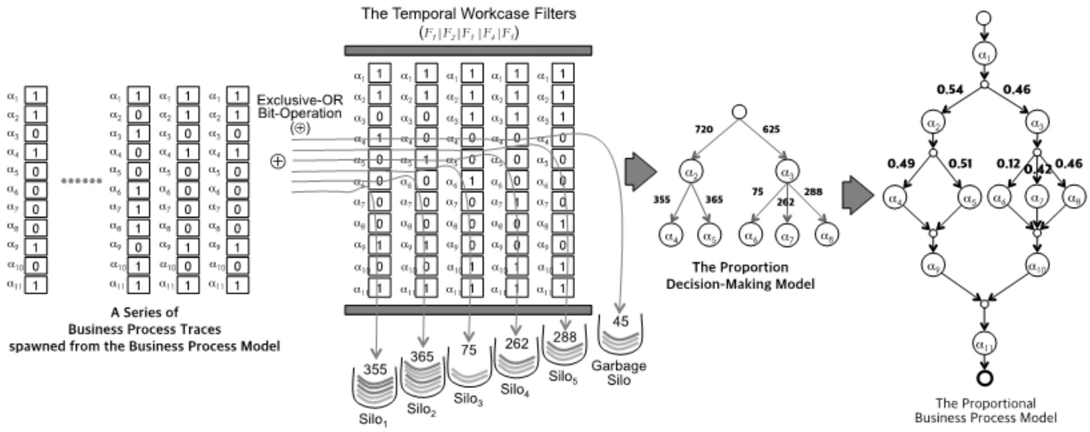

Fig. 6. An Operational Example Discovering the Proportional Business Process Model

Conclusively, the proportion decision-making model of **Fig. 5** and its numerical values are calculated from the counts in the silos as illustrated in **Fig. 6**. Through the temporal workcase filters, we can collectively sort out all the business process traces into the proper silos, and the numerical values in the fronts of the silos, such as 355, 365, 75, 262, and 282, are the counts of temporal workcases filtered into the corresponding silos. Based upon these counts and the dependency information (χ) of the proportion decision-making model, we can calculate the counts of the edges (the disjuctive routing patterns), such as (720, 625) and ((355, 365), (75, 262, 282)). After all, we can calculate the proportions of the edges, such as (0.54, 0.46), and ((0.49, 0.51), (0.12, 0.42, 0.46)), respectively, and produce a proportional information control net that is a formal represention of the corresponding business process model.

5. Conclusion

In this paper, we have introduced the conceptual approach of discovering the proportional business process model from its enactment event histories and logs, and we have also described the formal and graphical representations of all the conceptual ideas for the proposed conceptual approach. Particularly, in this paper we newly conceived a concept of the proportional business process model, and its rediscovery framework with a serial of feasible and conceptual solutions. The proposed approach eventually gives us a higher-level of efficiency in terms of improving the quality of business process models. In recent, the literature needs various, advanced, and specialized process mining techniques and analysis approaches to redesign and reengineer the existing workflow and business process models. This situation is why the fidelity issue of business process models has been raised in the literature. Through the outcome of our work, we strongly believe that we are able to implement a very reasonable business process intelligence system based on the novelties of the proposed conceptual approach and the fully supported functions to judge of over-fitting or under-fitting on a business process model and eventually to measure the fidelity of the corresponding business process model, as well.

Acknowledgement

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (Grant No. 2012006971). We would like to extend our appreciation to our collaborative researchers in the collaboration technology research laboratory at Kyonggi University and the database research laboratory at Kyunghee University.

References

[1] Daniela Grigori, Fabio Casati, Malu Castellanos, Umeshwar Dayal, Mehmet Sayal and Ming-Chien Shan, "Business Process Intelligence," *JORNAL OF COMPUTERS IN INDUSTRY*, Vol. 53, Issue 3, 2004. Article (CrossRef Link)

[2] Fabio Casati, et al, "Business Process Intelligence," Technical Report, HPL-2002-119, HP Laboratories Palo Alto, 2002.

[3] Kwang-Hoon Kim and Clarence A. Ellis, "Workflow Reduction for Reachable-path Rediscovery in Workflow Mining," *Series of Studies in Computational Intelligence: the Foundations and Novel Approaches in Data Mining*, Vol. 9, pp.289-310, Springer, 2006. Article (CrossRef Link)

[4] Kwang-Hoon Kim, "Control-path Oriented Workflow Intelligence Analysis on Enterprise Workflow Grids," in *Proc. of the International Conference on Semantics, Knowledge, and Grid*, Beijing, China, 2005. Article (CrossRef Link)

[5] W. M. P. van der Aalst, B. F. van Dongena; J. Herbst, L. Marustera, G. Schimm and A. J. M. M. Weijters, "Workflow mining: A survey of issues and approaches," *JOURNAL OF DATA AND KNOWLEDGE ENGINEERING*, Vol. 47, Issue 2, pp. 237-267, 2003. Article (CrossRef Link)

[6] W. M. P. van der Aalst and A. J. M. M. Weijters, "Process mining: a research agenda," *Journal of Computers in Industry*, Vol. 53, Issue 3, 2004. Article (CrossRef Link)

[7] Aalst, W.P.M., Alves de Medeiros, A.K., and Weijters, A.J.M.M., "Process Equivalence: Comparing Two Process Models Based on Observed Behavior," *BPM2006, Lecture Notes in Computer Science*, Vol. 4102, pp. 129-144, 2006. Article (CrossRef Link)

[8] Ahn, H. and Kim, K., "A Stochastic Activity-to-Performer Affiliation Binding Formalism in ICN-based Workflow Models," *ICICI Express Letters*, Vol. 9, No. 12, 2015.

[9] Ellis, C.A. "Information Control Nets: A Mathematical Model of Office Information Flow," *The Proceedings of the Ninth Annual ACM Conference on Simulation, Measurement, and Modeling of Computer Systems*, 1979.
[10] Ellis, Clarence A., Kim, K., Rembert, A., Wainer, J., "Investigations on Stochastic Information Control Nets," *INFORMATION SCIENCES*, Vol. 194, pp. 120-137, 2012. Article (CrossRef Link)
[11] Kim, H., Ahn, H. and Kim, P. K, "Modeling, Discovering, and Visualizing Workflow Performer-Role Affiliation Networking Knowledge," *KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS*, vol. 8, no. 2, pp. 689-706, 2014. Article (CrossRef Link)
[12] Kim, J., Ahn, H., Park, M., Kim, S., and Kim, K. P., "An Estimated Closeness Centrality Ranking Algorithm and Its Performance Analysis in Large-Scale Workflow-supported Social Networks," *KSII Transactions on Internet and Information Systems*, 10, 3, 1454-1466, 2016. Article (CrossRef Link)
[13] Kim, Kwanghoon, "A XML-Based Workflow Event Logging Mechanism for Workflow Mining," in *Proc. of the International Workshop on APWeb*, pages 132-136, 2006. Article (CrossRef Link)
[14] Kim, Kwanghoon and Ellis, Clarence A., "σ-Algorithm: Structured Workflow Process Mining Through Amalgamating Temporal Workcases," in *Proc. of PAKDD2007, Advances in Knowledge Discovery and Data Mining, Lecture Notes in Artificial Intelligence*, Vol. 4426, pp. 119-130, 2007. Article (CrossRef Link)
[15] Kim, Kwanghoon Pio, "Discovering Activity-Performer Affiliation Knowledge on ICN-based Workflow Models," *JOURNAL OF INFORMATION SCIENCE AND ENGINEERING*, Vol. 29, No. 1, pp. 79-97, 2013. Article (CrossRef Link)
[16] Kim, Kwanghoon Pio, "Human-Centered Workflow Intelligence," in *Proc. of Tutorial at the International Conference on Advanced Communications Technology*, Phoenix Park, Pyeongchang, South Korea, Feb. 2014.
[17] Kim, Kwanghoon Pio, "A Stochastic Workflow Performer−to−Activity Affiliation Network Model," in *Proc. of the 28th Annual Conference of Biomedical Fuzzy System Association (BMFSA2015)*, Kumamoto, Japan, 21-22 Nov. 2015.
[18] Kim, K., Paik, S, and Ellis, C., "Actor-Oriented Workflow Model," in *Proc. of the Second International Symposium on Cooperative Database Systems for Advanced Applications (CODAS'99)*, Wollongong, Australia, March 27-28, pp. 150-164, Springer, Singapore, ISBN 9814021644, 1999.
[19] Kim, Kwanghoon and Ellis, Clarence A., "Section II / Chapter VII. An ICN-based Work- flow Model and Its Advances," *HANDBOOK OF RESEARCH ON BUSINESS PROCESS MODELING*, pp. 142-172, IGI Global, ISR, 2009.
[20] Lei, Y., and Singh, M.P., "A Comparison of Workflow Metamodels," in *Proc. of the ER-97 Workshop on Behavioral Modeling and Design Transformations*, Los Angeles, USA, November 1997.
[21] Park, Minjae and Kim, Kwanghoon, "Control-path Oriented Workflow Intelligence Analyses," *JOURNAL OF INFORMATION SCIENCE AND ENGINEERING*, Vol. 24, No. 2, pp. 343-359, 2008. Article (CrossRef Link)
[22] Snijders, Tom A.B. , Bunt, Gerhard G. van de, Steglich, Christian E.G., "Introduction to Stochastic Actor-Based Models for Network Dynamics," *SOCIAL NETWORKS*, Vol. 32, pp. 44-60, 2010. Article (CrossRef Link)
[23] Varacca, D. and Nielsen, M., "Probabilistic Petri Nets and Mazurkiewicz Equivalence," in *Proc. of the Conference on Decision and Control*, vol. 5, pp. 4104-4109, 2001.

Kyoungsook Kim is a Ph. D. candidate in computer engineering as well as a graduate member of the database research laboratory at Kyunghee University, South Korea. She received B.S. degree in computer science from Kyonggi University in 1984. She also received M.S. degree in computer engineering from Kyunghee University in 2001. She had worked as a researcher and developer at FAST Systems Inc. and OSI Corp. in South Korea. Her research interests include large-scale database systems and applications, enterprise information systems, real-time databases, distributed databases, temporal databases, federate databases, object-oriented database management systems, and the Internet of Things collaboration architectures and services.

Moonsuk Yeon is a Ph. D. candidate in computer engineering as well as a graduate member of the database research laboratory at Kyunghee University, South Korea. She received B.S. degree in computer information engineering from Woosong University in 2000. She also received M.S. degree in computer engineering from Kyunghee University in 2004. Her research interests include large-scale database systems and applications, enterprise information systems, multimedia databases, real-time databases, distributed databases, data warehousing and mining, and big data analysis.

Byeongsoo Jeong is a full professor in computer engineering department and the founder of the database research laboratory at Kyunghee University, South Korea. He received B.S. degree in computer science from Seoul National University in 1983. He received M.S. in computer engineering from KAIST in 1985, and also received Ph. D. degree in computer science and engineering from Georgia Institute of Technology in 1995. His research interests include large-scale database systems and applications, enterprise information systems, distributed databases, real-time databases, data warehousing and mining, and object-oriented database management systems.

Kwanghoon Pio Kim is a full professor of computer science department and the founder and supervisor of the collaboration technology research laboratory at Kyonggi University, South Korea. He received B.S. degree in computer science from Kyonggi University in 1984. And he received M.S. degree in computer science from Chungang University in 1986. He also received his M.S. and Ph. D. degrees from the computer science department at University of Colorado Boulder, in 1994 and 1998, respectively. At Kyonggi University, he is Dean of the Computerization and Informatics Institute, and the director of the contents convergence software research center, as well. He had worked as researcher and developer at Aztek Engineering, American Educational Products Inc., and IBM in USA, as well as at Electronics and Telecommunications Research Institute (ETRI) in South Korea. In present, he is a vice-chair of the BPM Korea Forum. He has been in charge of a country-chair (Korea) and 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. 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 process-aware Internet of Things.