Business Process Analysis Method based on Petri nets

Jianglong Qin\textsuperscript{1,2}, Na Zhao\textsuperscript{1,2}, Zhongwen Xie\textsuperscript{1,2} and Qi Mo\textsuperscript{1,2}
\textsuperscript{1}(School of Software, Yunnan University
Kunming, China)
\textsuperscript{2}(Key Laboratory in Software Engineering of Yunnan Province
Kunming, China)
e-mail:king8414@163.com, zhaonayx@126.com, xiez56@126.com, moqiyueyang@163.com

Abstract—To better reorganize and optimize business process, Petri nets is a good formal tool. However, the prototype Petri nets can’t fully describe the related properties of the business process. In this paper, the time and cost attributes are added to the Petri nets and it is called the extended generalized stochastic Petri nets. On this basis, the performance evaluation model is proposed. Execution efficiency and cost of Business process can be got by the evaluation model. They provide a direct basis for business process reengineering and optimization. Finally, the effectiveness of the model is validated through case study.

Keywords-Quantitative Analysis; Business Process Optimization; Petri nets; Performance Evaluation Model

I. INTRODUCTION

Business process analysis is the foundation of business process management and processing capabilities has very important significance to improve work efficiency for the government and to improve business performance for the enterprise. With the development of information technology, it creates the technical foundation for a substantial increase for business process management capabilities and processing power\textsuperscript{[1]}. Business process analysis is the foundation of business process reengineering, reorganization and optimization. To cognitive current business process of the enterprise, quantitative analysis of business process performance is done, which identifies bottleneck through the analysis. A relatively unified research platform is provided for the business process analysis, design and implementation. This will extend the Generalized Stochastic Petri nets applied to business process analysis phase.

II. DEFINITION OF BUSINESS PROCESS

Definition 2.1\textsuperscript{[2]} A triple $N=(C, A, F)$ is called a net if

1. $C \cup A \neq \emptyset$;
2. $C \cap A = \emptyset$;
3. $F \subseteq (C \times A) \cup (A \times C)$ is a binary relation, the flow relation of $N$;
4. $\text{dom}(F) \cup \text{cod}(F) = C \cup A$.

$C$ which is called the condition set of $N$, $A$ is called the activity set, $F$ is called arc set. $X = C \cup A$ is called the $N$ element set. Elements in $C$ are called the condition or element $C$, Elements in $A$ are called activity or element $A$. The $\emptyset$ is empty, $\times$ is the two-set (space) of the Cartesian product operation, so $F$ is the set of ordered pair composed of a $C$ element and a $A$ element. $\text{dom} (F)$ is the order dual of $F$ contains the first element (starting) of the set, $\text{cod} (F)$ is the second element (the end) of the collection: $\text{dom} (F) = \{x \mid y \in F \}
(x, y) \in F$; $\text{cod} (F) = \{x \mid y \in F \}$. Net standard graphics is that circle represents the condition, with the box that transition $s$ from $x$ to $y$ by a directed arc (arrow) that ordered dual $(x, y)$, (hence $(x, y) \in F$), it is also known as directed arcs.

Definition 2.2 Satisfy the following six conditions set $\Sigma = (C, A, F, M, \lambda, g)$ is called a business process system:

1. $(C, A, F)$ is a net;
2. Where $T$ is divided into time transition Set $A_t$ and instantaneous transition set $A_i$ 2 sub-sets. The implementation of the time transition will take some time, the instantaneous transition implementation time is negligible, that is, $A = A_t \cap A_i$, $A_t \cap A_i \neq \emptyset$;
3. $F$ in the permitted inhibitor arcs, inhibitor arcs exist only from the position to transition the arc. Inhibitor arcs can be implemented by the conditions of the original link becomes unenforceable (disable) the conditions, the original non-implementation of the conditions can be implemented into the conditions, and transition in the linked implementation, not marked out from the connected position;
4. $\lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_m\}$, $\lambda_i$ is the average implementation rate of the transition $a_i (a_i \in A)$, which represents average implementation times in the case of the implementation in unit time. The unit is times/hour;
5. $g = \{g_1, g_2, \ldots, g_m\}$, $g$ is the average implementation costs of the transition $a_i (a_i \in A)$, which represents implemented cost in the case of implementation per unit time, The unit is the yuan/hour.

Where each delay of the transition $1/\lambda = \{1/\lambda_1, 1/\lambda_2, \ldots, 1/\lambda_m\}$ are exponentially distributed, the cost of each transition $g = \{g_1, g_2, \ldots, g_m\}$ are exponentially distributed.

Definition 2.3 Let $\Sigma = (C, A, F, M, \overrightarrow{A}, G)$ represents a business process system. Let $M_0 \in M (M_0 \subseteq C)$, $p = (C, A,$
\( F, M_0, \lambda, g \) \). \( M_0 \) is called the initial marking of \( p \) and \( p \) is called a business process. If confusing, then, \( C, A, F, M_0, \lambda, g \) respectively \( p.C, p.A, p.F, p.M_0, p.\lambda, p.g \) to represent.

**Definition 2.4** Let \( p = (C, A, F, M_0, \lambda, g) \) represents a business process. For \( p \) in any one activity \( a \in A \), Inflow \( (a) = \{(x, a) | (x, a) \in F, x \in C\} \), Inflow \( (a) \) is called the input stream for the activity of \( a \), Outflow \( (a) = \{(a, y) | (a, y) \in F, y \in C\} \), Outflow \( (a) \) is called the output stream for the activity of \( a \) \[3\].

**Definition 2.5** \[2\] \( M_0 = \{(c, 1), (c, 2)\} \) is called the initial marking of \( p \) and \( p \) is called business process. For \( p \) in any one activity \( a \), Input stream for the activity of \( a \), Outflow \( (a) = \{(a, y) | (a, y) \in F, y \in C\} \), Outflow \( (a) \) is called the output stream for the activity of \( a \) \[3\].

**Definition 2.6** Let \( N = \{C, A, F\} \) as a net. For \( x \in C \cup A \), \( X = \{y \mid y \cup A \wedge (y, x) \in F\} \), \( x = \{y \mid y \cup A \wedge (x, y) \in F\} \). \( X \) is called the front set or the input set of \( x \). \( x \) is called the back set or the input set of \( x \).

**Definition 2.7** Let \( N = \{C, A, F\} \) as a net, \( \text{StartPlace}(N) = \{c \in C \mid a \in A : (a, c) \notin F\} \), \( \text{StartPlace}(N) \) is called the initial condition of \( N \); \( \text{EndPlace}(N) = \{c \in C \mid a \in A : (a, c) \notin F\} \), \( \text{EndPlace}(N) \) is called termination condition of \( N \); \( \text{SpecialPlace}(N) = \text{StartPlace}(N) \cup \text{EndPlace}(N) \), \( \text{SpecialPlace}(N) \) is called special condition of the net.

**Definition 2.8** \[3\] Sequence block, concurrency block, selection block and iteration block are called basic block. A basic block can be described as a 5-tuple \( b = (C, A, F, A_e, A_i) \) where

1. \( C, A \) and \( F \) are called the condition set, the activity set and the arc set respectively.
2. \( A_e, A_i \subseteq A \) are called the entrance and the exit of \( b \) respectively.

Basic blocks, which are enclosed by dotted lines in figures, are described as follows

Where each \( e_i \) \((i = 1, 2, \ldots) \) denotes an activity.

**Sequence block:** It describes activities \( e_i \) and \( e_j \) are executed sequentially, shown in Figure 2.1. Formally, \( C = \{c\}, A = \{e_i, e_j\}, F = \{(e_i, c), (e_j, c)\}, A_e = \{e_i\}, A_i = \{e_j\} \)

**Concurrency block:** It describes activities \( e_i \) and \( e_j \) are executed concurrently, shown in Figure 2.2. Formally, \( C = \{c_1, c_2, c_3, c_4\}, A = \{e_0, e_1, e_2, e_3\}, A_e = \{e_0, e_1, e_2, e_3\}, F = \{(e_0, c_1), (e_0, c_2), (c_1, e_1), (c_2, e_2), (e_1, c_3), (e_2, c_4), (c_1, e_3), (c_2, e_4)\}, A_i = \{e_0, e_1\}, A_e = \{e_2, e_3\} \)

**Selection block:** It describes activities \( e_i \) and \( e_j \) are executed selectively, shown in Figure 2.3. Formally, \( C = \{\}, A = \{e_0, e_j\}, F = \{\}, A_e = \{e_0, e_j\}, A_i = \{e_0, e_j\} \)

**Iteration block:** It describes activities \( e_i \) and \( e_j \) are executed repeatedly, shown in Figure 2.4. Formally, \( C = \{c_1, c_2\}, A = \{e_i, e_j, e_k\}, A_e = \{e_0, e_i, e_j, e_k\}, F = \{(e_0, c_1), (c_1, e_i), (e_i, c_2), (c_2, e_j), (e_j, c_3), (c_3, e_k)\}, A_i = \{e_0\}, A_e = \{e_1\} \)

III. ANALYSIS OF BUSINESS PROCESS

**Theorem 3.1** \[4\] positions of any kind with the limited, a transition of finite continuous time SPN (Stochastic Petri Nets) is isomorphic to a one-dimensional continuous time MC.

It can be seen from the definition 2.3, business process system is essentially an extension of the generalized stochastic Petri nets (Extended Generalized Stochastic Petri Nets). Generalized stochastic Petri nets extended a number of signs that contains \( M \), the change process of mark is a random process. If the only time some signs of change is enabled, then we call non-instant flag, which indicates a change in the activation, the system should stay in the state before sending a non-zero time; the rest of the tag as an instant symbol. It shows that the system state change occurs in the moment (i.e., after the change occurs, the system state before sending the residence time is zero). Generalized
stochastic Petri net is a symbol of the process of semi-discrete state space Markov process[5], generalized stochastic Petri extended net costs incurred to transition exponentially distributed, so it remains a symbol of the semi-Markov process with discrete state space. The semi-Markov process, hidden Markov chain state and EGSPN symbol correspondence. In the state diagram, a directed arc between states connected by the arc, on the graph with the trigger up to change the state flag change in the average activation rate of $\lambda$ and the transition firing probability $p$ mark, forms of expression such as $\lambda (p)$, for example: $20(0.1), 20$ of which expressed the activation rate of transition (ie, changes in unit time the average number of the state), $0.1$ represents the probability of transition firing, for the choice transition, which identifies the M in a certain state, in which both N-transition can occur, but can only choose one to trigger, transition in this N-transition can be triggered in the probability of their implementation, and is equal to 1, in which the implementation of any changes in the probability of less than 1. For the extended workflow net sign of generalized random process, the retention time of the moment marks is zero, so from the perspective of performance evaluation, simply visit non-transient signs on the line. Assume that expansion of generalized random net has k non-transient signs: $M_1, M_2, ..., M_k$, $\pi_i$ said to mark the course of the steady-state probability of $M_i$ (the signs on the course of time spent on the ratio $M_i$), clearly $\sum_{i=1}^{k} \pi_i = 1$ to $m_i$ that marks the average residence time, there are formulas (3.1): 

$$m_i = 1/ \sum_{a_i \in M_i} F(M_i, a_i) \tag{3.1}$$

$A_i$ in the formula 3.1 represents that $M_i$ enable transition in the set, $F(M_i, a_i)$ is used to marked sending time of the transition $a_i$ of $M_i$. Let $Y = [y_1, y_2, ..., y_k]$, where $y_i$ marked $M_i$ access flag as a symbol rate of the process, which consists of the following limits of discrete time stochastic process the probability formula (3.2) calculated[5]: 

$$YQ = Y$$

$$\sum_{j=1}^{k} y_i = 1 \quad Q \text{ is transition probability matrix}; \tag{3.2}$$

According to the definition of steady-state probability (ie the signs on the course of time spent on Mi ratio), steady-state probability of $i$ derived the formula (3.3) 

$$\pi_i = y_im_i / \sum_{j=1}^{k} y_jm_j \tag{3.3}$$

According to Theorem 3.1 and the down, we can draw extended generalized stochastic Petri nets (EGSPN) model for performance evaluation as follows: Step 1: business process modeling, business process is described by extended generalized stochastic Petri (EGSPN). Step 2: all reachable marking graphs of extended generalized stochastic Petri nets are got. Step 3: According to the reachable marking graphs, constructed Markov chain of being isomorphic with EGSPN, the reachable graphs of EGSPN and the odd time continuous Markov chain is isomorphic. Based on the principle that the trigger of instantaneous transition in the EGSPN is prior than time transition, the corresponding reachable graphs are got. The average implementation rate and trigger probability of each arc are marked, thus getting the Markov chain. Step 4: seeking Markov chain transition probability matrix after reduction (the stage of business processes to ensure the identification process is a continuous random process of being non-absorbing state, resulting in steady-state solution of extended generalized stochastic Petri nets, One arc that transition probability is 1 between the terminate state and the initial state is added, assuming that $M_{start}$ represents initial state of the net, $M_{end}$ represents termination state of the net, State transition exists from $M_{start}$ to $M_{end}$, and the probability is 1.) Step 5: Find the probability that the marking process to access $M_i$. Step 6: Find the steady state probability of identification process. Step 7: Analyse utilization of transition. Utilization of transition in the solution as follows: 

$$\forall a \in A \quad \text{utility } U(a) = \sum_{M_i \in E} P[M_i]$$

Where E is a set of all reachable marking that can be implemented.

**Theorem 3.2[5]** System B is comprised of n transitions that is in series. Assume the delay time of n transitions that is in series are n-independent random variables, which are subject to the exponential distribution function that parameters are $\lambda_1, \lambda_2, ..., \lambda_n$, the average delay time of n transition are $1/\lambda_1, 1/\lambda_2, ..., 1/\lambda_n$, the total delay time of n-transition is: 

$$1/\lambda = \sum_{i=1}^{n} 1/\lambda_i$$

**Theorem 3.3** Assummar the delay time of n parallel transitions are n-independent order statistics $x_1, x_2, ..., x_n$ and are subject to the exponential distribution function that parameters are $\lambda_1, \lambda_2, ..., \lambda_n$, then the total average time of n parallel transitions is: 

$$1/\lambda = \sum_{i=1}^{n} 1/\lambda_i - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 1/\lambda_i + 1/\lambda_j + \sum_{i=1}^{n-2} \sum_{j=i+1}^{n} \sum_{k=j+1}^{n} (1/\lambda_i + 1/\lambda_j + 1/\lambda_k) + \cdot \cdot \cdot + (1)$$

**Theorem 3.4** Assume the delay time of n optional transitions $e_1, e_2, ..., e_n$ are n-independent random variables
\( t_1, t_2, \ldots, t_n \), which are subject to the index distribution function that parameters are \( \lambda_1, \lambda_2, \ldots, \lambda_n \).
Assume the implementation probability of transition \( t_i \) is \( \alpha_i \), \( \sum_{i=1}^{n} \alpha_i = 1 \), the total average delay time of \( n \) transitions is:

\[
1/\hat{\lambda} = \sum_{i=1}^{n} \alpha_i / \hat{\lambda}_i
\]

Theorem 3.5 [4] Assume the delay time of two circular transitions \( e_i, e_j \) are two independent random variables \( t_1, t_2 \), which are subject to the exponential distribution function that parameter are \( \lambda_1, \lambda_2 \), assuming after executing transition \( e_j \) the probability of loop transition \( e_j \) is \( \alpha \), the total average delay time of the two circular transitions is:

\[
1/\hat{\lambda} = (\hat{\lambda}_2 + \alpha \cdot \hat{\lambda}_1) / (1-\alpha) \cdot \hat{\lambda}_1 \cdot \hat{\lambda}_2
\]

Corollary 3.6 System B is composed of two transitions in series. the delay time of the two transitions in series are two independent random variables, which are subject to exponential distribution function that parameters are \( \hat{\lambda}_1, \hat{\lambda}_2 \), the average delay time of two transitions are \( \frac{1}{\hat{\lambda}_1}, \frac{1}{\hat{\lambda}_2} \), the total delay time of the two transitions is:

\[
1/\hat{\lambda} = \frac{1}{\hat{\lambda}_1} + \frac{1}{\hat{\lambda}_2}
\]

Corollary 3.7 Assume the delay time of two parallel transitions are the independent order statistics \( x_1, x_2 \), which are subject to the exponential distribution function that parameters are \( \hat{\lambda}_1, \hat{\lambda}_2 \), the total average time of the two parallel transitions is:

\[
1/\hat{\lambda} = (\hat{\lambda}_1^2 + \hat{\lambda}_2^2) / (\hat{\lambda}_1 + \hat{\lambda}_2) \cdot \hat{\lambda}_1 \cdot \hat{\lambda}_2
\]

Corollary 3.8 Assume the delay time of two optional transitions \( e_1, e_2 \) are two independent random variables \( e_1, e_2 \), which are subject to the exponential distribution function that parameters are \( \hat{\lambda}_1, \hat{\lambda}_2 \), assuming the implementation probability of transition is \( \alpha_i \), \( \sum_{i=1}^{n} \alpha_i = 1 \), the total average delay time of the two transitions is:

\[
1/\hat{\lambda} = \alpha_1 / \hat{\lambda}_1 + \alpha_2 / \hat{\lambda}_2
\]

Theorem 3.4, Corollary 3.5, Corollary 3.6 and Corollary 3.7 are corresponding to a equivalent model of performance (sequence, choice, concurrency, iteration), equivalent model of performance of sequence block shown in figure 3.1, equivalent model of performance of choice block shown in figure 3.2, equivalent model of performance of concurrency block shown in figure 3.3, equivalent model of performance of iteration block shown in figure 3.4.
The principle that is simplified and calculated: from the inside out, firstly simplifying the innermost layer, which are four basic structures and then calculated from the inside out gradually, through continuous performance equivalent simplification, from complex to simple, and ultimately obtained performance parameters of the entire system.

Based on the equivalent time performance of business processes, the cost that the business process is executed one time can be further calculated. The cost formula of the business process sees formula (3.4)

\[
\text{Cost} = \left( \sum_{i=1}^{n} \frac{g_i}{\lambda_i} \right) \cdot T_{\text{flow}} = \left( \sum_{i=1}^{n} \frac{1}{\lambda_i} \right)
\]  

(3.4)

\(\lambda_i\) represents the implementation rate of transition \(a_i\), \(g_i\) presents the implementation cost of unit time, \(T_{\text{flow}}\) presents execution time of business process.

IV. CASE STUDY

An example is given to illustrate the business process analysis based on Petri net. The logistics company receives the customer’s delivery notice. According to customer requirements, the appropriate mode of transport is selected. After checking the goods and warehousing, storage is completed. The customer will receive the receipt and the cargo insurance is handled at the same time. If the customer chooses a road transport, the logistics company delivers to the company’s cargo terminal and carries the goods. When the problem is found, the existing problems will be resolved and a new round of assembly. If the customer chooses the rail transport, the logistics company makes the corresponding shipping document and sends the delivery certificate to the customer, then sends the goods to the train station. If the customer chooses air transport, the logistics company makes the corresponding documents and sends the goods to the airport, and then informs the customer of the carrier. After the goods are sent, the document is checked. The operation flow of a logistics company is shown in Figure 4.1.

According to the steps of the performance evaluation model, the representation of the semi-Markov chain of business process can be got. It is shown in Figure 4.2.

After that, state transition probability matrix can be got. It is shown in Figure 4.3. Assume the implementation rates
of changes a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8,a_9,a_{10},a_{11},a_{12},a_{13},a_{14},a_{15},a_{16}, a_{17},a_{18},a_{19},a_{20},a_{21},a_{22},a_{23} are respectively \( \lambda_i = 60 \), \( \lambda_i = 100 \), \( \lambda_i = 120 \), \( \lambda_i = 80 \), \( \lambda_i = 30 \), \( \lambda_i = \infty \), \( \lambda_i = 60 \), \( \lambda_i = \infty \), \( \lambda_i = 40 \), \( \lambda_i = \infty \), \( \lambda_i = 12 \), \( \lambda_i = 10 \), \( \lambda_i = 6 \), \( \lambda_i = 40 \), \( \lambda_i = 20 \), \( \lambda_i = 30 \), \( \lambda_i = 50 \), \( \lambda_i = 25 \).

Through the analysis of the utilization rate of the change, it is found that the changes a_5,a_{12},a_{16},a_{18} and a_2 have higher utilization rate during the execution of the process. To improve this part of the change can improve the efficiency of the implementation of the process.

V. CONCLUSIONS

In this paper, the concept of cost is introduced to the generalized stochastic Petri nets. It is called it extended generalized stochastic Petri nets. At the same time, basic block is introduced to business process modeling. Business process is got described with the extended stochastic Petri net. In the analysis phase, because generalized stochastic Petri nets and semi Markov chain is homogeneous. Based on it, steady-state probability can be got, which can calculate the utilization rate of transitions, analyses the bottleneck of business process and identifies problems in business processes. At the same time, equivalent performance models of the four basic blocks in the workflow are introduced into the business process. The equivalent time performance of the business processes can be solved.

ACKNOWLEDGMENT

This work has been supported by the National Science Foundation of China under Grant No. 61379032.

REFERENCES


