Generic Evolutionary Framework for Simulating Business Processes

Veng-Ian Chan, Yain-Whar Si

Faculty of Science and Technology, University of Macau
ma76505@umac.mo, fstasp@umac.mo

Abstract. Business process simulation enables detail analysis of resource allocation strategies without actually deploying the processes. Although business process simulation has been widely researched in recent years, less attention has been devoted to automating the simulation of business processes with the help of evolutionary computation. In this research, we aim to implement a generic GA modeling framework which can be used to simulate different kinds of business processes. Specifically, optimum resource allocation scheme for the simulation can be effectively chosen by the evolution process of a genetic algorithm (GA). The proposed generic GA modeling framework is capable of automatically retrieving information regarding available resources, temporal constraints of the tasks, and process models from a given business process and can produce the best resource assignment scheme.

Keywords: Genetic algorithm, resource allocation, business process, Color Petri Nets.

1 Introduction

Business Process Simulation (BPS) [8] artificially implements and assists the management of change in a variety of manufacturing and service settings. BPS enables detail analysis of such performance indicators without actually deploying the processes. BPS is widely used for analyzing business process models in many different domains. For instance, simulation was used to study the effect of stochastic customer shopping traffic for the IBM's Personal Computer Division [9] as well as for analyzing the impact on the user experience and the cost of using the application when a mobile channel presentation (based on GSM, HSCSD, GPRS and UMTS networks) is added [10].

Configuration of process models includes assigning tasks with available resources and inputting temporal constraints. Although BPS has been extensively used in performance analysis, configuration of process models for BPS has been mostly done manually by analysts. As a result, finding the right resource assignment scheme for simulation can be time consuming since it is usually done by trial-and-error approach. In addition, a large number of resource assignment schemes may exist for simulating a complex process model. With the help of evolutionary computation, optimum resource assignment for business process simulation can be effectively chosen by the

evolution process of a Genetic Algorithm (GA) [11]. A Genetic Algorithm can perform a powerful form of hill climbing search by maintaining multiple solutions and eliminating unpromising solutions.

Although GAs can be used to choose optimal resource assignment for BPS, experiments are only limited to modeling process models from a number of application scenarios. In these scenarios [12, 13], different GA designs are required for capturing process dependent parameters. For instance, chromosomes design, crossover and selection operators, and fitness functions are different for each case and they are required to be tailored for each underlying BPS problem. In this research, we propose a generic GA modeling framework which can be used to simulate different kinds of business processes. The proposed generic GA modeling framework is capable of automatically retrieving information regarding available resources, temporal constraints of the tasks, and process models from the Business Process Simulation environment.

The paper is structured as follows. In Section 2, we review related work. In Section 3, the proposed genetic algorithm framework based on CPN is introduced. In Section 4, a case study on resource allocation case from Macau is given. In Section 5, the experimental results are collected and analyzed in detail. In section 6, we conclude the paper with future work.

2 Related Work

Although simulation experiments are widely used in evaluating business processes, to the best of authors' knowledge, no work has been reported on designing generic evolutionary framework for simulating business processes.

Simulation based analysis of business processes were reported in [2], [3]. In [9], a discrete-event simulation model was used to analyze the business process at the Center for Social Work in Slovenia for predicting the effects of the new organizational scheme, the duration of the processes, and potential bottlenecks. Escalation strategies for business processes are evaluated using simulation models in [4]. Escalation actions are used to reduce the deadline violations, or to negotiate an extended deadline with the customer. In [4], Paganos et al. proposed two strategies, which are 1) minimizing the slack time and 2) deadline prediction in order to minimize the number of escalations needed during workflow execution and to mitigate their associated costs. Simulation experiment for an insurance claim process from an Australian Company is also conducted by van der Aalst et al. In [5], they analyze various deadline escalation strategies. In their approach, escalation strategies are evaluated from three perspectives: the process perspective of using alternative path selection, the data perspective of using data degradation, and the resource perspective of using resource redeployment. In [6], four escalation strategies from [4], [5] are evaluated from temporal (workflow time) and cost (execution, resource, compensation) perspectives. In [7], a genetic algorithm is used to find near-optimal resource allocation scheme and the event-driven schedule of a Color Petri Nets.

3 Process Simulation Framework Based on CPN

In this research, we propose a generic GA modeling framework which is capable of automatically retrieving information regarding available resources, temporal constraints of the tasks, and the process models. Based on the extracted information, the GA modeling framework constructs chromosomes and fitness functions required for the evolution process. The overall scheme of the generic GA modeling framework for BPS is depicted in Figure 1. First, the user imports the process model to the GA framework for analysis. Next, the GA framework uses the resource information of the model to form the members of the population for the 1st generation. Each chromosome represents a potential resource allocation scheme. For each chromosome, one round of simulation is performed to calculate the degree of fitness. Chromosome fitness is calculated based on the workflow completion time and the total cost. These chromosomes are then ordered according to the fitness and genetic operators such as selection, crossover, and mutation are applied to the chromosomes to form new members for the next generation. The overall process iterates until the change in the fitness of the best members in the population for several consecutive generations is less than a predefined threshold. The resource allocation scheme encoded in the fittest chromosome from the last generation is then returned as the best resource allocation scheme for the process model. The GA modeling framework and input process models are both implemented using Color Petri Nets and simulation of process models are conducted in CPN Tools [14].

Algorithm 1 shows the pseudo-code of our general genetic algorithm framework. For different resources types, a predefined chromosome structure is used to capture their upper and lower bounds. In the proposed model, there is no limit on the number of resource types for the simulation. The predefined chromosome structure is in the form of: [(upper bound of resource n, resource n), (upper bound of resource n-1, resource n-1), …,(upper bound of resource 1, resource 1)].

Population size depends on the nature of the problem and chromosomes are randomly generated to cover the entire range of possible solutions. In our proposed genetic algorithm framework, we define a fitness function *f(c,g)* for calculating the fitness of chromosome *c* from generation *g* based on the total workflow completion time and total cost of all tasks *n* in the workflow.

$$f(c,g) = \frac{1}{\sum_{1}^{n} time_i^{c,g}} + \frac{1}{\sum_{1}^{n} cost_i^{c,g}} \tag{1}$$

Evolution process of genetic algorithm stops when a termination condition has been reached. In our case, the simulation will stop when the change in the fitness of several generations is less than a predefined threshold. The genetic algorithm then returns the fittest chromosome from the final generation. The resulted chromosome represents the best possible resource allocation scheme for the process model.

Although we have used some default settings for the simulation, they can be easily changed by the user to meet their requirements. In our prototype, the population size is set to 10 in each generation and there is no limit on the number of resource types for the process model. The upper bound of a resource can be defined in the chromosome structure and the selection and elimination rate is set to 90% (40% for

crossover and 50% for mutation) and 10% respectively. The crossover point is randomly generated by the GA CPN model and the mutation rate is 1/L where L is the length of the chromosome. All these settings are stored in the *places* in the Color Petri Nets and can be altered according to the user's requirement.

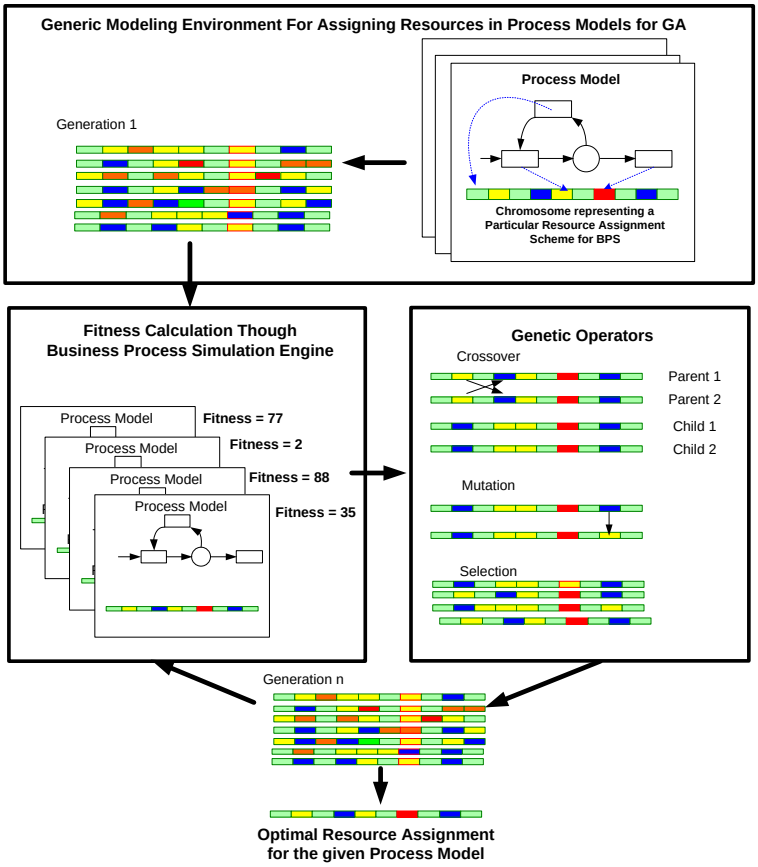

Fig. 1: The overall scheme of the generic Genetic Algorithm modeling framework for resource assignment in Business Process Simulation

The main contribution of the proposed framework is that, the process models depicted in the upper right hand corner of Figure 1 can be substituted with any Color Petri Nets process models. Since the remaining modules of Genetic Algorithm modeling framework in Figure 1 are also implemented in Color Petri Nets, these modules can be seamlessly integrated with any process model for simulation (see Figure 7 from Appendix).

To enable this integration, we define an interface based on *places* in the Petri Nets model for importing and exporting of chromosomes from the business process model to the GA algorithm. These chromosomes are programmed as *tokens* in the Petri Nets model and stored in the places (depicted as ellipses). The high level processes "Pass in" and "GA" which are depicted as rectangles with double border lines in the

highlighted area of Figure 7 are implemented as separate Petri Nets models and encompass all the modules of the proposed Generic Modeling Framework. These modules communicate with the business process model which is programmed as transitions and places on the left side of Figure 7 via the places called “Pool” and “Results”. In Figure 7, the place “Pool” which belongs to the business process model contains tokens for the simulation. These tokens are referenced during the simulation by the respective tasks in the business process model. After the simulation, the result is returned to the GA modules from the process “Measurement system” via the place “Results”. In our prototype, all transitions (rectangles with double border lines) in Figure 7 are implemented as separate Petri Nets models. Due to the limited space, only the highest level of Petri Nets model is shown in the Appendix.

```
procedure genetic algorithm
begin
        set generation g:=1;
        set i = 1;
        initialize the population P(g) = pop;
    Repeat
        While i ≤ pop then Do
            run the simulation based on chromosome_i;

            evaluate fitness of chromosome_i;

            sort the chromosome_i by fitness;
            i:=i+1;
        end
            select 40% of parents from the population P(g);
            crossover to produce offspring from these pairs;
            mutate the remaining 50% of the candidates of P(g);
            randomly generate 10% candidates for the new population;
            replace the weakest candidate of P(g) with these new
            offspring in P(g+1);
            set generation g := g+1;
            set i:=1;
     Until change in degree of fitness <= predefined threshold
     return the chromosome with best fitness result;
   end
```

Algorithm 1. Pseudo-code of genetic algorithm CPN model

4 Case Study

In this section, we demonstrate the application of the proposed framework based on an archival management workflow. Archive [1] is a collection of records and documents which need to be kept and conserved as an invaluable asset. We conceptualize the workflow model of Macau Historical Archives in Figure 2. Color Petri Nets representation of the workflow model is depicted in Figure 7 of Appendix. First, records for archiving are evaluated in the “Appraisal” (T1) process. Appraisal is the process of assessing whether these records have sufficient value to warrant acquisition by an archival institution. After the appraisal task, these records are formally accepted by the archive institute in “Receive record transfer” (T2) process. Next, in “Initial conservation” (T3) process, basic cleaning of the records is

performed before they are grouped in certain order in “Arrangement” (T4) process. Then “Description” (T5) process is carried out to analyze, organize and record details of the archive based on international description standards. After descriptions are added, “Paper surrogating” (T6) process is carried out to create digital archives from the physical format. The next steps in the process involve creating backups and rebinding archives. In “Backup” (T8) process, the original files are copied into storage media so that it can be restored if the original data is deleted or damaged. “Rebinding” (T9) process is performed for repackaging as well as for associating the related meta-information with the specific archival record for permanent storage. “Microfilm surrogating” (T7) process captures and stores images of the archives in microfilm formats. Finally these records are stored at the permanent storage. Depending of the nature of the archived material, periodic maintenance tasks (i.e. preservation) are also scheduled at the end of the process.

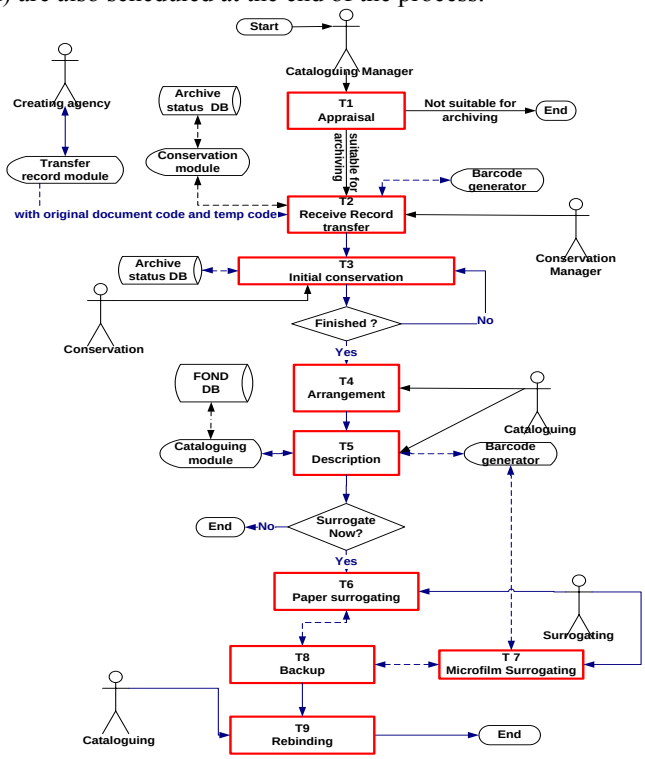

Fig. 2: Archival management workflow (Level 0)

In Macau Historical Archives, at the beginning of each year, transfer lists are sent to the Macau Historical Archives by various organizations. A transfer list contains approximately 225 records for archiving. In normal condition, approximately two transfer lists and a box of microfilm are received per year. Each box of microfilm contains approximately 2000 images. According to the recent statistics in Macau Historical Archives, approximately 14% microfilm pages are delayed and the bottleneck was located at T4, T5 and T6 since relatively high number of assigned

records cannot be processed on time. We also find that delays in these tasks have a ripple effect on the whole workflow process. In the GA CPN models, each task is assigned with an "estimated average completion time" for processing job (see Table 1). These values are calculated from the recent statistical data from Macau Historical Archives. We can see that the total completion time of the workflow is approximately 1580 minutes.

Table 1: Average completion time of each task in the workflow

Task	Task description	Estimated average task cost (per record)	Estimated average completion time (per record)
T1	Appraisal	$5.28	2 minutes
T2	Received record transfer	$9.5	4 minutes 25 seconds
T3	Initial conservation	$182.83	85 minutes
T4	Arrangement	$4.3	2 minutes
T5	Description	$3810.15	24 Hours
T6	Paper surrogating	$37.23	19 Minutes 30 seconds
T7	Microfilm surrogating	$12.84	6 Minutes
T8	Backup	$17.83	8 Minutes 20 seconds
T9	Rebinding	$18.43	10 Minutes
Total (per record)		**$4098.39**	**1577 Minutes 15 seconds**

For better accuracy, each simulation experiments are run for 4 years period and average values are calculated for comparison. Also in this experiment, there are 14 types of resources and constraints on these resources are defined in the chromosomes. The chromosome [(2,14), (2,13), (1,12), (1,11), (1,10), (1,9), (7,8), (4,7), (6,6), (8,5), (1,4), (1,3), (1,2), (1,1)] is used to encode the resource assignment scheme which is currently used in Macau Historical Archives. For illustration, we denote this scheme as "original as-is model".

5 Experimental Result

From the experiment result (Figure 3, 4, 5) we can see the original as-is model has relatively long average waiting time for each task. It also has the longest workflow time and largest total workflow cost among all experimental models. The resource allocation schemes found by GA from last five generations are depicted as G26, G27, G28, G29, and G30. The fittest chromosome found in 30th generation is: [(3,14), (2,13), (3,12), (2,11), (3,10), (3,9), (4,8), (13,7), (1,6), (9,5), (2,4), (1,3), (1,2), (1,1)]. In Figure 3, we can see that T4, T5, T6 and T8 have relatively high average waiting time.

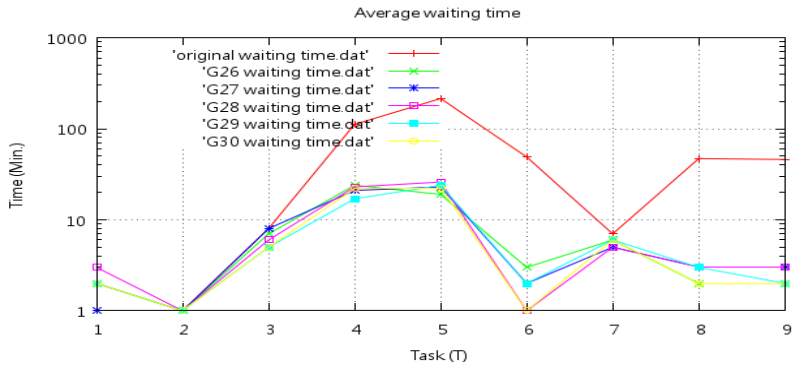

Fig. 3: Average waiting time

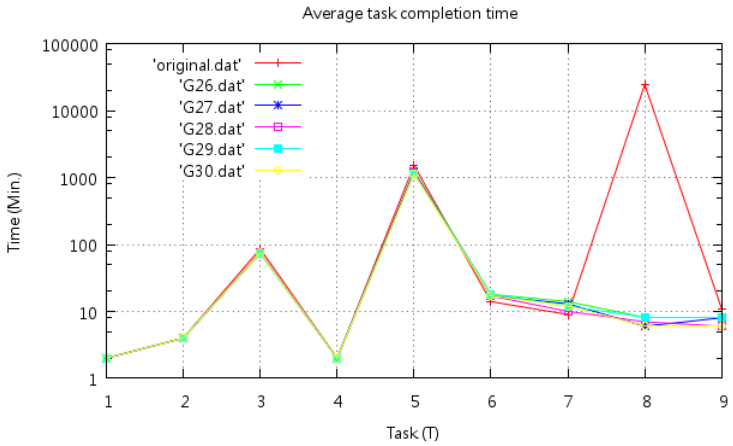

Fig. 4: Average task completion time

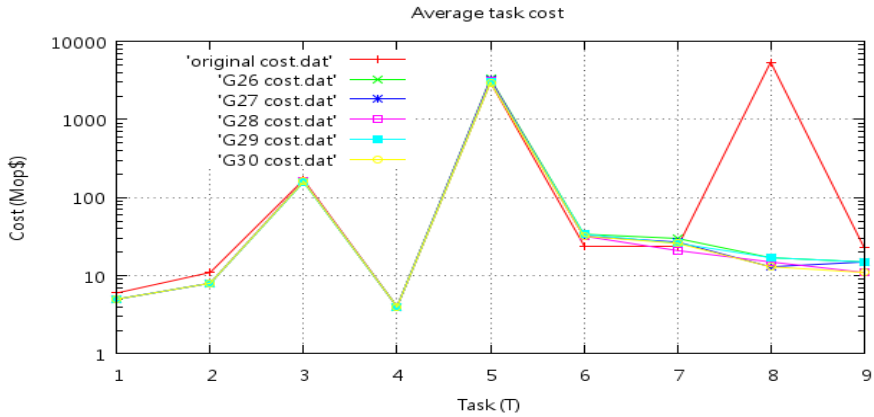

Fig. 5: Average task cost

In figure 6, experiment results show that the estimated average total workflow time and average total workflow cost gradually decrease in later generations.

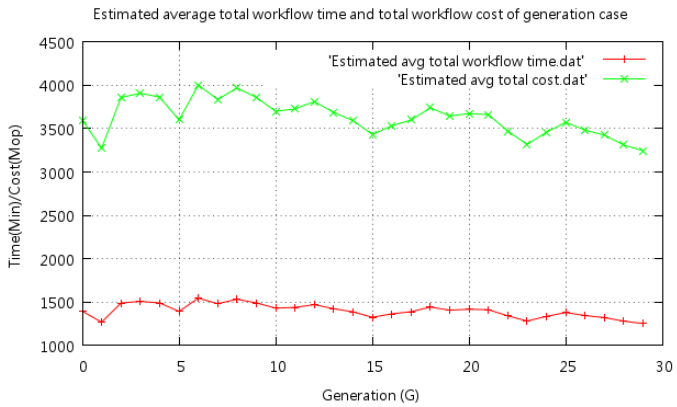

Fig. 6: Estimated average total workflow time and average total workflow cost for each generation

From the above experiments, we find that GA-based approach significantly reduces workflow time and the total cost compared to original resource allocation scheme. We can see that original as-is model has the longest workflow time among all experimental models. The performance of GA-based resource allocation in generation 30 achieves 20% reduction in total workflow time and 21% reduction in total workflow cost.

6 Conclusion

In this paper, we detail a generic evolutionary framework which can be used to simulate different kinds of business processes. The proposed generic GA modeling framework is capable of automatically retrieving information regarding available resources, temporal constraints of the tasks, and process models from a given business process and can produce the best resource assignment scheme. The framework was developed in Color Petri Nets and can be used to simulate various workflows for identifying the best resource allocation scheme. The use of the framework was evaluated against a case study on Macau Archival Management Workflow.

So far, the proposed generic evolutionary framework can be used to simulate any process models defined in Color Petri Nets. As for the future work, we are planning to extend our framework for simulating workflows which are defined in other process modeling notations.

Acknowledgments. This research is funded by the University of Macau.

References

1. Pearce-Moses, R.: SAA: A Glossary of Archival and Records Terminology, The Society of American Archivists, http://www.archivists.org/glossary/index.asp
2. Kovacic, A., Pecek, B.: Use of Simulation in a Public Administration Process. Simulation 83(12), 85-861, (2007)
3. Greasley, A.: Using Business-Process Simulation within a Business-Process Reengineering Approach. Business Process Management Journal. 9(4), 408-420, (2003)
4. Panagos, E. & Rabinovich, M.: Reducing Escalation-Related Costs in WFMSs. In: Proceedings of NATO Advanced Study Institute on Workflow Management Systems and Interoperability, pp. 107-27, Springer, (1998)
5. van der Aalst, W.M.P., Rosemann, M., & Dumas, M.: Deadline-based Escalation in Process-Aware Information Systems. Decision Support Systems 43(2):492-511, (2007)
6. Chan, K.L., Si, Y.-W., & Dumas, M.: Simulation-based Evaluation of Workflow Escalation Strategies. In: Proceedings of the 2009 IEEE International Conference on e-Business Engineering (ICEBE 2009), pp. 75--82, IEEE Press, (2009)
7. Saitou K., Malpathak S., and Qvam H.: Robust Design of FMSs using CPN and Genetic Algorithm, Journal of Intelligent Manufacturing, no.13, pp. 339-351, (2002)
8. An, L. & Jeng, J.-J.: On developing system dynamics model for business process simulation, in Proceedings of the 37th conference on Winter Simulation (WSC), pp. 2068 - 2077, (2005)
9. Lee, Y.M.: Simulating Availability Outlook for E-Commerce Business for Personal Computer Sales, in Proceeding of the 36th conference on Winter Simulation (WSC), pp. 1201-1204, (2004)
10. Book, M., Gruhn, V., Hülder, M., Köhler, A., & Kriegel, A.: Cost and Response Time Simulation for Web-based Applications on Mobile Channels, in Proceedings of the 5th International Conference on Quality Software, pp. 83-90, (2005)
11. Holland, John H.: Adaptation in Natural and Artificial Systems, University of Michigan Press, Ann Arbor, (1975)
12. Lee, H. & Kim, S.-S.: Integration of Process Planning and Scheduling Using Simulation Based Genetic Algorithms, Journal The International Journal of Advanced Manufacturing Technology, Volume 18, Number 8/October, pp. 586-590, Springer London, (2001)
13. Hegazy, T. & Kassab, M.: Resource Optimization Using Combined Simulation and Genetic Algorithms, Journal of Construction Engineering and Management, Volume 129, Issue 6, pp. 698-705, (2003)
14. CPN Group, University of Aarhus, Denmark. CPN Tools Home Page. http://wiki.daimi.au.dk/cpntools/.

Appendix: CPN models

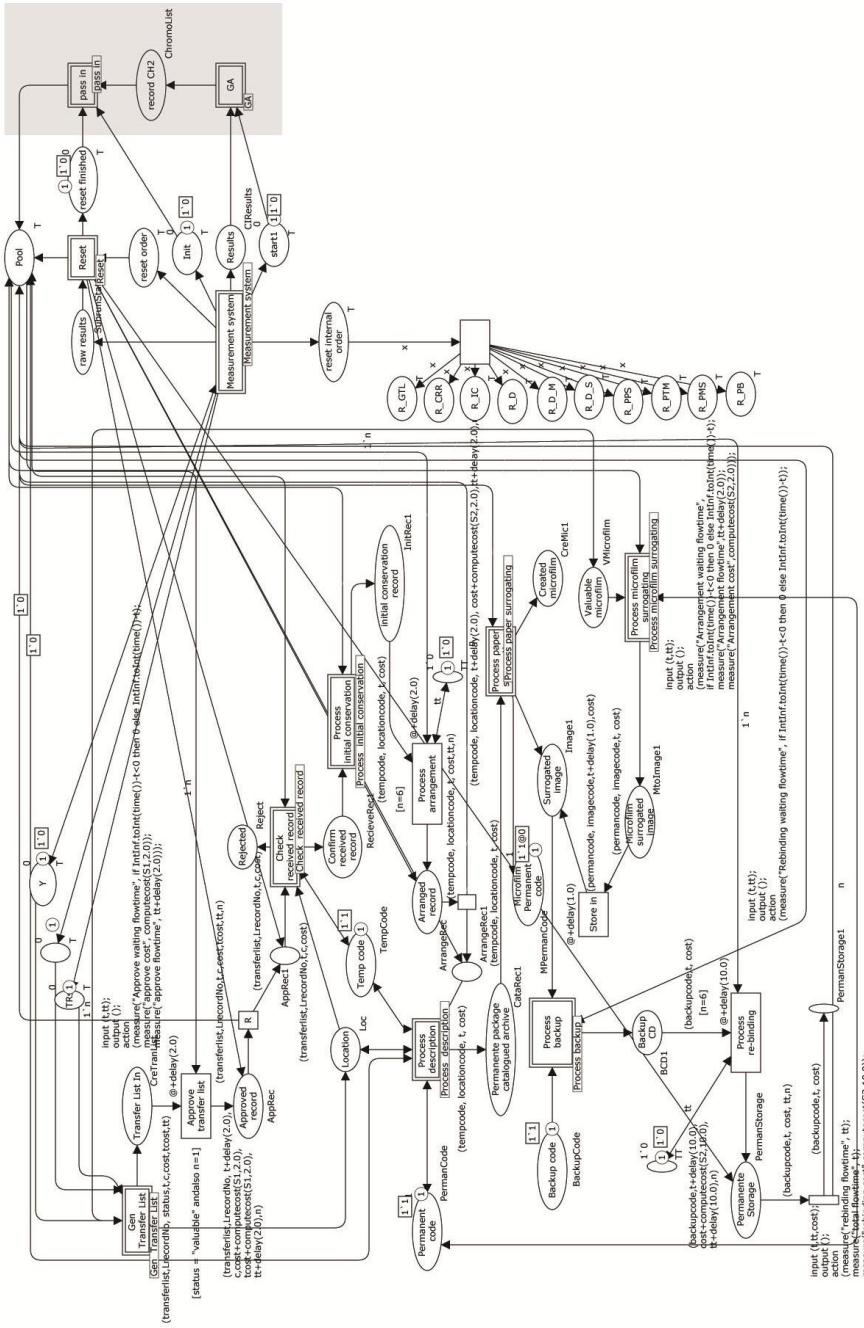

Fig. 7: Level – 0 Color Petri Nets model of Archival Management Workflow