Modeling and Performance Evaluation of Workflow Systems

Juliane Dehnert
Kommunikations- und Softwaretechnik
Technische Universität Berlin, Sekr. E-N 7
Einsteinufer 17, 10587 Berlin, Germany

Jörn Freiheit, Armin Zimmermann
Prozeßdatenverarbeitung und Robotik
Technische Universität Berlin, Sekr. FR 2-2
Franklinstr. 28/29, 10587 Berlin, Germany

ABSTRACT

There is a need for modeling and performance evaluation techniques and tools for a fast and reliable design of workflow systems. This paper introduces a modeling methodology based on colored stochastic Petri nets. It allows the integration of functional, organizational, and timing aspects in one modeling framework. The processing delays include stochastic distributions in addition to deterministic times. Several workflows and the effects of constrained shared resources needed for different tasks can easily be described and analyzed together. Functional and organizational aspects are modeled separately in resource and workflow models. These models are automatically compiled into one model, which can then be used for qualitative analysis or performance evaluation. The proposed modeling and evaluation method is supported by the software tool TimeNET. An application example shows the use of the method.

Keywords: Modelling, Performance Evaluation, Workflow, Petri Nets

1. INTRODUCTION

Because of the increasing demand for business process (re-)engineering, tools and techniques supporting their planning and evaluation are needed. A model of a business process allows to verify the correct document processing steps and cooperation between the responsible departments. Additionally the performance evaluation is of special importance. It helps to predict and optimize the degree of service to the customer, the number of staff and resources needed, and nevertheless the business process efficiency. Some requirements have therefore to be met by the modeling formalism (e.g. see [5]). The modeling technique should combine a graphical representation with possibilities of composition and decomposition. This makes the identification and modeling part especially for processes of real-life size easier. Syntax and semantic of the model have to be formally defined and there has to be a mathematical background of analysis techniques. This is required to guarantee a common understanding and to prove the correctness of the specification. The executability of the model ensures that it can be simulated or numerically analyzed, aiming at an evaluation of the performance.

To avoid a more complex description, many approaches focus mainly on one of the views of a business process (e.g. functional or organizational), thus hiding information related to the others. Mixing model elements related to different views in one model, on the other hand, leads to descriptions that are not easy to understand. Moreover, if resources are specified in a functional model, the relations between the resources cannot be expressed properly, and it is difficult to understand their interdependencies. Different business processes carried out by one organization and the effects of restricted shared resources between several processing steps have to be incorporated.

Business (re)engineering projects typically try to reduce turnaround times and improve execution durations in order to improve competitiveness and customer contentedness. These questions can be answered by an evaluation of the business process performance, necessitating the exact description of all timing aspects of the process. Although this seems to be a crucial point for business process engineering, it has not gained much interest. There are only few approaches (e.g. [3, 4, 6]) considering time, which offer analysis or simulation as a means to evaluate the behavior of the workflow model. Often the notion of time is limited to the functional view and to deterministic delays and time constraints. Stochastically distributed durations are often neglected. Durations that correspond only to employees and resources are not considered.

In the field of workflow performance modeling a mixture of both stochastic and fixed (deterministic) durations are necessary. Fixed times are needed to model durations of automated processes, deadlines, and routine activities with previously known delay. Stochastic durations are necessary to model human activities, stochastic events like employee absence due to sickness or holidays, and other events with unknown duration. Consequently, deterministic and stochastic durations should be integrated into the model.

In the following we will introduce a methodology for the modeling and performance evaluation of workflow processes, which tries to overcome the above mentioned limitations. The approach is not limited to the functional aspects, but includes a resource description as well. Both aspects are described using the same modeling technique, namely colored stochastic Petri nets. Both aspects are modeled separately to improve clarity.

Petri nets allow an appropriate description of business processes. The suitability of Petri nets for this field of application has been examined and discussed extensively in the literature [2, 6, 7]. Hierarchical refinement makes it possible to construct large systems easier. In contrast to the *critical path technique* or *event process chains* (EPK's) they can be used to describe the dynamic behavior of systems, because they are directly executable.

The model introduced here is based on a concept of independent models for the functional aspects and the resources, both using a dedicated class of colored stochastic Petri nets. The functional aspects (possibly of different workflows) are specified in the *workflow models* and the organizational aspects in the *resource models*. A workflow model describes the possible executions using the

For the evaluation both parts are merged together. In the approach presented here, evaluation of the performance is facilitated by associating stochastic, deterministic, or zero firing delays with transitions. Basic quantitative measures like the throughput, utilization, queue length, processing time, and others can be computed either by direct numerical analysis or discrete event simulation.

To support the proposed methodology, a software tool is necessary. Throughout this paper the tool TimeNET is used. A recent enhancement of this tool is a modeling and evaluation environment originally intended for manufacturing systems [9]. We found that due to the similarities to business processes it is possible to adapt the tool for the modeling and evaluation of workflows as well.

2. AN APPLICATION EXAMPLE

In a damage event handling system there are many persons and institutions involved. Planning and managing the corresponding business process and its resources is very complex. Many concurrent activities and their synchronization are necessary. Figure 1 gives a rough overview of the involved offices and organizations as well as possible information flows of the transport company. It should be noted that this is only intended as an informal way of describing the organization.

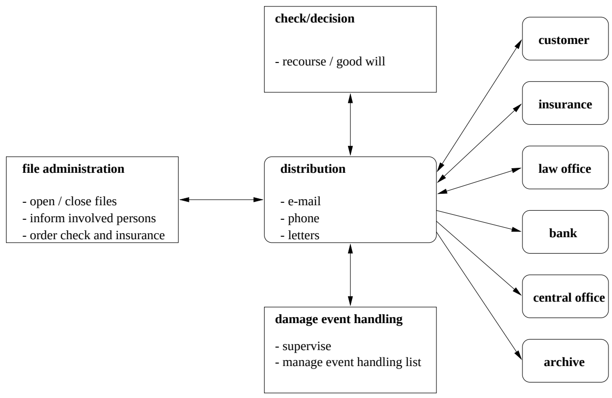

Figure 1: Coarse structure of the damage event handling system example

In the example three offices are considered in detail, namely file administration, check/decision, and the main damage event handling office. In Figure 1 they are depicted by rectangles. The damage event handling office supervises the business process that is started for each incoming message of a damage event. The file administration office processes the associated files, informs the involved persons, orders checks, and passes the necessary information to the involved insurance companies. The check/decision office checks the case and decides whether the damage event is treated as a recourse or a fair dealing. All necessary communication between the involved offices and the exchange of documents between them is depicted as distribution. Arcs in the figure describe possibilities of communication.

In addition to the organizational structure the workflow processes of the example are described in the following. Figure 2 shows a part of the example workflow.

The workflow is initiated when a customer message arrives, informing about a new damage event. The file administration office creates a new file, in-

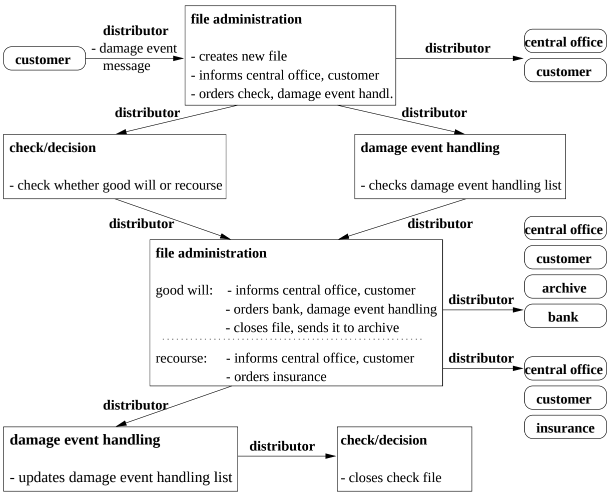

Figure 2: Part of workflow in the system example

forms the central office and the customer via distributor, and orders an analysis of the case by the check/decision office. The damage event handling office is informed in parallel, checks the damage event handling list and sends it to the file administration office. The check/decision office decides whether the damage event is treated as a recourse or a good will. If it is a recourse, the insurance has to pay for the damage, and is therefore informed. In the case of a fair dealing the company orders its bank to pay the damage itself. In Figure 2 only a selected part of the whole workflow is depicted, which is later also shown as a model.

3. COLORED PETRI NET WORKFLOW MODELING

In this paper a special class of colored Petri nets is used for the modeling of the workflow management system example described above. Two kinds of models are distinguished. The *resource model* describes the abilities and workflow independent properties of the workflow management system resources, such as communication connections, staff, technical resources etc. The *workflow models* specify the actual business processes that take place for every case, and describe the workflow dependent features of the system. Each workflow model can be thought of as a path through the resource model. Later on, the different model parts are automatically merged resulting in a complete model, which then includes both the resource constraints of the system and the synchronization of the workflow steps. The models are hierarchically structured, which is necessary to handle complex systems.

As it is usual for Petri net models, transitions model activities of the system, while places are passive elements and contain tokens that model moving and/or state changing entities. Transitions with thick bars are called substitution transitions, acting as place holders for submodels describing their behaviour in more detail on a lower level of hierarchy. Places shown as dotted circles connect the submodel with its environment. Transitions depicted as a bar fire immediately without delay. Transitions drawn as empty rectangles have an exponentially distributed firing time, while transitions with deterministic delay are depicted as filled rectangles.

Two colour types are predefined in the model class: *Object tokens* model files, orders, letters etc. inside the workflow management system, and consist of a name and the current state. *Elementary tokens* cannot be distinguished. They are used to model states of the resources, for instance whether an employee is busy or not. Places can contain only tokens of one type. Therefore it is possible to graphically separate object places and arcs (drawn thick) as opposed to thinly drawn elements that correspond to elementary tokens. Textual descriptions needed in colored Petri nets for the definition of variables and colour types can thus be omitted, and the specification of the types of places and arcs are implicitly given.

THE RESOURCE MODEL

The resource model contains all used resources, like employees, computers, data bases, etc. and abstract possible actions of the resources, even if they are not used for the processing. Figure 3 shows the highest level of the hierarchical colored model.

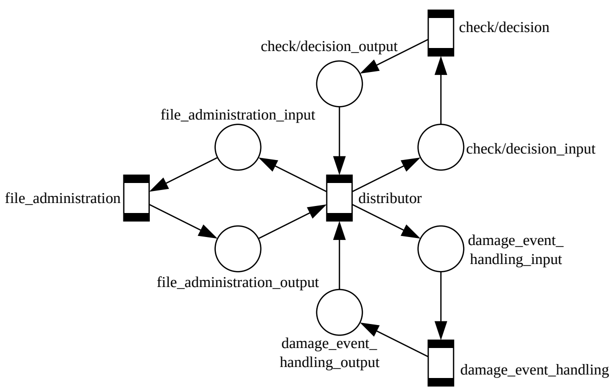

Figure 3: Structure model of the damage event handling system example

Places model possible locations of objects, like files, orders, letters etc. All offices exchange documents via `distributor`. The `input` and `output` places model in- and outgoing matters of the corresponding offices. In principal, there are two different operations that can be performed: transport and processing of objects. The former corresponds to moving a token to another place, while the latter is modeled by a change in the color of the token that corresponds to the object.

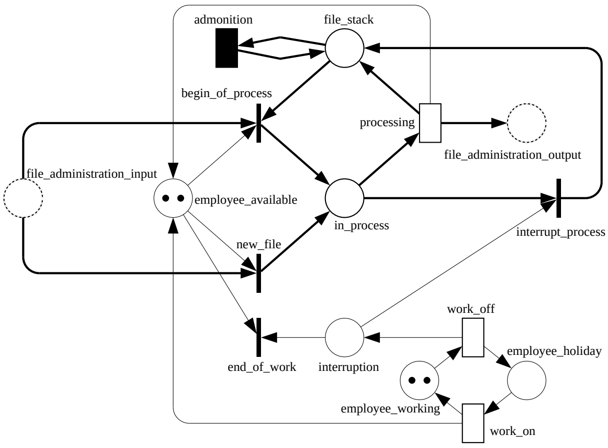

Figure 4: Structure submodel of the file administration office

Figure 4 shows the resource submodel of the file administration office in more detail. The model has two parts. The transitions and places connected by thin arcs describe the behaviour and states of the employees working in this office. The place `employee_working` models the working employees. In this example there are two employees associated to this office. The `work_off` transition models the case that an employee goes to holidays or becomes ill. This interrupts the current work of the employee (transition `interrupt_process` fires), or only decreases the number of available idle employees (transition `end_of_work` fires). If the employee comes back to work (`work_on`), he is available again (`employee_available`).

The places and transitions connected by thick arcs describe the behaviour of objects like documents or memos. The places `file_administration_input` and `file_administration_output` connect the submodel with its environment. The place `file_stack` contains all files which are currently not being processed, while place `in_process` contains the active ones. The processing of an existing file starts with the firing of the immediate transition `begin_of_process`. For each new damage event message the transition `new_file` creates a new file. The transition `processing` describes the processing of a file. Due to the modeling of human actions in this case the transition has an exponentially distributed delay. The deterministic transition `admonition` fires after a fixed time and thus models deadlines for expected answers from other offices.

THE WORKFLOW MODELS

In the resource model the structure of the system and possible ways of communication is summarized. The workflow model now describes the actual processes using the resources described in the resource model. Figure 5 shows the top level workflow model of the example.

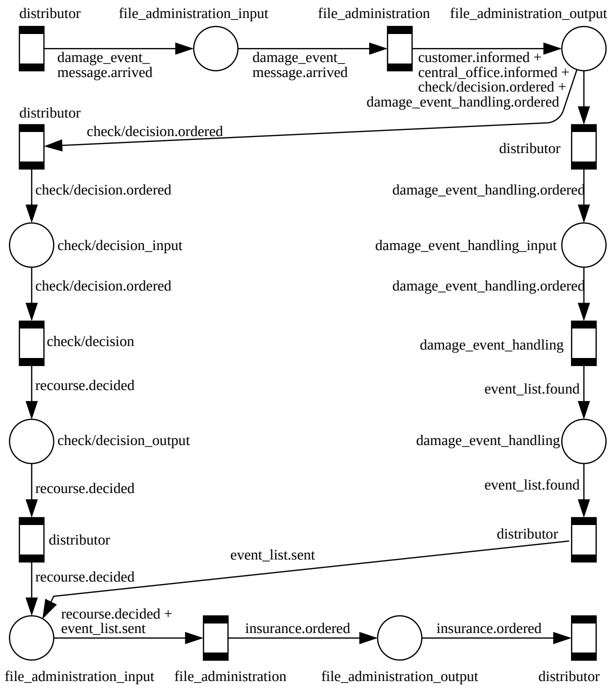

Figure 5: Part of the top level workflow model of damage event handling

The workflow models describe paths through the resource model (see Figure 3). In the workflow model the arcs are inscribed by the tokens which flow through the arcs. After a damage event message has arrived via `distributor`, the `file administration` creates a new file and informs the customer and the central office. Additionally it sends the necessary information to the check/decision office and the damage event handling office. The check result `recourse.decided` is send to the `file administration` office. The `damage event handling` office sends the damage event handling list to the `file administration` office. For simplicity, Figure 5 only shows the model until the point where the insurance is informed about the recourse case.

4. EVALUATION OF THE WORKFLOW MODEL

The previous sections described how to construct models of the workflows and used resources. Subsequently, both model parts are merged automatically to create a complete model that can be analyzed. The information contained in the workflow models is added to the structure model during this process. Invisible for the modeller, the transitions are enriched with their *firing possibilities*.

The transition firing delays adopted here can either be zero (immediate), exponentially distributed, deterministic, or may belong to a class of general distributions called *expolynomial*. If no more than one transition with non-exponentially distributed firing delay is enabled in each marking, the underlying stochastic process is *semi-regenerative* and can be directly numerically analyzed [1].

The evaluation of the model can be used to answers questions like

- How many documents can be processed per week with the modeled organization?
- What is the mean time for a case to be finished?
- How big is the utilization of the resources?
- What are the bottlenecks?
- How much time does a document spend during processing, waiting, or being transported?
- How will the above numbers change if the available staff decreases e.g. due to holidays?

Now the application example is evaluated. We assume that there are funds available for a total staff of seven employees. The following question is considered: how many employees should be associated with each one of the three offices (file administration, check/decision, damage event handling)?

For the realistic description of times for the process steps, exponentially distributed firing delays have been associated with the normal office tasks, while transitions modeling deadlines and in-house communication fire after a fixed (deterministic) delay. The simulation component of the tool TimeNET is used for the performance evaluation of the application example. For all evaluations, a confidence interval of 99% and a relative error probability of 3% is chosen, to enforce a comparably high accuracy of the simulation. The simulations have been carried out on a cluster of ten UltraSparc workstations, and each run took typically 49 seconds of total CPU time to finish.

In the cases where the distribution of employees is more or less balanced for the three offices, the utilization results clearly point out that the damage

	utilization			
employee distr.	file adm.	check	dam. ev. handling	thr. put
1 / 1 / 5	99.7 %	94.1 %	99.9 %	2.28
1 / 2 / 4	100 %	74.3 %	99.9 %	2.90
1 / 3 / 3	99.8 %	71.5 %	99.8 %	2.40
1 / 4 / 2	99.3 %	67.8 %	99.7 %	1.98
1 / 5 / 1	98.7 %	64.9 %	100 %	1.87
2 / 1 / 4	91.2 %	99.0 %	99.6 %	0.68
2 / 2 / 3	91.3 %	96.6 %	99.7 %	0.70
2 / 3 / 2	91.1 %	87.4 %	99.9 %	0.67
2 / 4 / 1	91.0 %	80.4 %	100 %	0.64
3 / 1 / 3	78.6 %	99.6 %	99.8 %	0.56
3 / 2 / 2	77.7 %	98.1 %	99.9 %	0.53
3 / 3 / 1	77.3 %	90.4 %	100 %	0.51
4 / 1 / 2	72.2 %	100 %	99.7 %	0.39
4 / 2 / 1	71.9 %	95.1 %	99.9 %	0.38
5 / 1 / 1	68.2 %	100 %	99.7 %	0.35

Figure 6: Simulation results of the performance evaluation

event handling office is the main bottleneck of the system. It is therefore not surprising that a considerable gain in the throughout can be achieved by associating more people to this office. After this has been done, it is advisable to have more employees working in the check/decision office than in the file administration, because this seems to be the second bottleneck. The optimal configuration is 1/2/4. In addition to the shown evaluations one possible configuration (1/1/4) for only six employees has been evaluated, resulting in a throughput of 1.77, which is better than many of the evaluated combinations of seven employees.

5. TOOL SUPPORT

Modeling and evaluation of complex systems is only feasible with the support of appropriate software tools. For the work presented here the software tool TimeNET is used, a tool for the modeling and performability evaluation using stochastic Petri nets (shown in Figure 7). A general overview of the software package can be found in [8].

6. CONCLUSION

In this paper a method for the modeling and performance evaluation of workflow systems based on colored stochastic Petri nets is introduced. The approach combines functional and organizational parts, which are described in clearly separated models. There is no restriction on the number of different business processes of the organization that can be

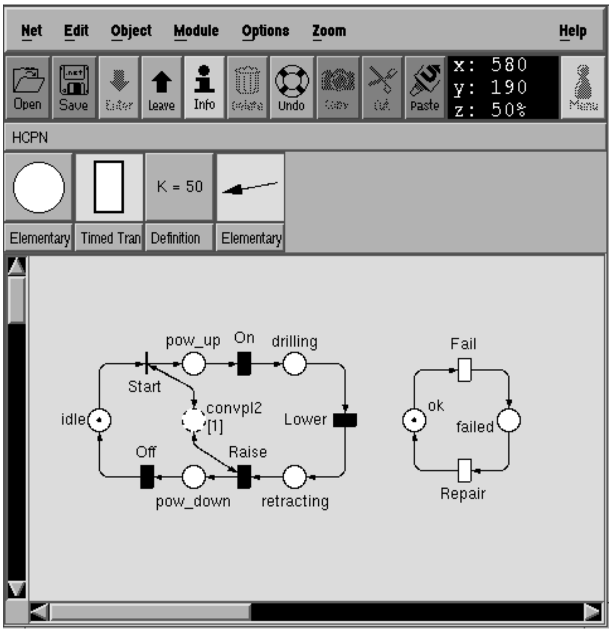

Figure 7: Sample screenshot of the graphical user interface

modeled and analyzed together, making the evaluation of their combined effects possible. Deterministic as well as stochastic delays of activities are allowed in the approach and facilitate a more realistic modeling of the processes. The performance of a modeled system can be evaluated by direct numerical analysis or discrete-event simulation. For the application of the described techniques to an application example of medium size the software tool TimeNET has been used.

The paper concentrates on the modeling and simulation of business processes. In the future it should be possible to augment the tool such that run time support including resource monitoring and deadline checking is possible.

REFERENCES

[1] G. Ciardo, R. German, and C. Lindemann, A Characterization of the Stochastic Process Underlying a Stochastic Petri Net, *IEEE Transactions on Software Engineering* **20** (1994) 506–515.

[2] W. Deiters, V. Gruhn, and R. Striemer, Der FUNSOFT-Ansatz zum integrierten Geschäftsprozessmanagement, *Wirtschaftsinformatik* **5** (37) (1995) 459–466.

[3] J. Eder, E. Panagos, H. Pozewaunig, and M. Rabinovich, Time Management in Workflow Systems, in: W. Abramowicz and M. Orlowska, eds., *BIS'99 3rd International Conference on Business Information Systems* (London Berlin Heidelberg, 1999, Springer Verlag) 265–280.

[4] T. Erwin, Leistungsbewertung von Geschäftsprozessen durch Auswertung halbgeordneter Petrinetz-Abläufe, Master's thesis, Universität Karlsruhe, 1998.

[5] P. Kawalek and P. Kueng, The Usefulness of Process Models: A Lifecycle Description of how Process Models are used in Modern Organisations, in: K. Sian, Y. Wand, and J. Parsons, eds., *International Workshop on Evaluation of Modeling Methods in System Analysis and Design (EMMSAD'97)*, Proceedings of the 2nd CAISE/IFIP of the EMMSAD'97 (Barcelona, Spain, 1997) I:1–12.

[6] A. Oberweis, *Modellierung und Ausführung von Workflows mit Petri-Netzen*, Teubner-Reihe Wirtschaftsinformatik (B. G. Teubner Verlagsgesellschaft, Stuttgart, Leipzig, 1996).

[7] W. van der Aalst, The Application of Petri Nets to Workflow Management, *The Journal of Circuits, Systems and Computers* **8** (1) (1998) 21–66.

[8] A. Zimmermann, J. Freiheit, R. German, and G. Hommel, Petri Net Modelling and Performability Evaluation with TimeNET 3.0, in: B. H. et.al., ed., *11th Int. Conf. on Modelling Techniques and Tools for Computer Performance Evaluation, Lecture Notes in Computer Science*, Vol. 1786 (Schaumburg, Illinois, USA, 2000) 188–202.

[9] A. Zimmermann and G. Hommel, Modelling and Evaluation of Manufacturing Systems Using Dedicated Petri Nets, *Int. Journal of Advanced Manufacturing Technology* **15** (1999) 132–137.