Application of Workflow Management System to the Modelling of Processes in Land Administration Systems

Saša VRANIĆ, Hrvoje MATJEVIĆ and Miodrag ROIĆ, Croatia

Key words: Workflow, LADM, Process, Transaction, Correctness

SUMMARY

Cadastral data are maintained through formally defined procedures which need to provide security and consistency. Databases and transactional model enable consistency but lack flexibility in modelling business processes and support for heterogenous IT environments (web services, various programs). Transactional workflow management systems (WFMS) provide flexibility and can provide consistency of data. Land administration domain model (LADM) provides an excellent basis for modelling static component of land administration systems, but doesn’t provide elements to model dynamic component, i.e. processes.

In this paper we define conceptual model of a dynamic component of land administration system. We use the WFMS concept with integrated transactional support. Data model enables storing elements of Petri nets and it is divided to generic and extended. Generic model ensures consistency of processes on object level and is applicable on cadastral data generally. Extended model ensures consistency by spatially defining affected area of the process and it is used to model processes on cadastral parcels spatially represented by polygons. Modelling of processes and transactional support is achieved with Petri nets. Workflow elements enable ensuring consistency of a process in a pessimistic or optimistic manner. Pessimistic approach ensures consistency by locking objects affected by the process and optimistic approach leaves checking of concurrent changes until the very end of the process.

Finally, we demonstrate how the devised model copes with a simple example of two separate processes where each wants to split one of two adjacent parcels, in a pessimistic manner.
1. INTRODUCTION

During the several past decades different levels of digitization of land administration systems have been made all over the world. Especially in the last decade many countries have implemented advanced information systems for land administration, all of which include at least some basic transaction processing mechanisms based on concepts originating from the domain of relational databases. However, sometimes (usually depending on the experience of the implementer) those concepts are not implemented fully or correctly or are not based on clearly defined and explainable concepts.

A significant help for implementers of information systems for land administration is provided by the means of LADM, since it provides an excellent starting point for designing the data models with all the important concepts already included. However, LADM is restricted to the static layout of land administration data and provides little to none help when its dynamic aspect is to be designed. In general, the scientific research of the dynamic aspect of land administration is sparse. The most extensive research was done within the COST Action Model G9 ‘Modelling Real Property Transactions’. One of the outcomes of the project was the terminology and general phases of processes on real estates. Ferlan et al (2007) divide the legal part of real estate traffic (advertising, pre-contract, contract and registration) from creating a new property that represents a technical process (for example, splitting a parcel). Navratil and Frank (2004) analyse cadastral systems (types, organization, documents, stakeholders) and processes in cadastral systems. Other related works on the topic are (Hespanha 2012; Sari 2010).

From the technological standpoint, a notion of transaction comes from the world of relational databases (Gray, 1981). A transaction is a group of database operations that can be considered as a single logical operation over data. The first and most important requirement for transactions is to transform a database from one valid state to another. For that purpose, the ACID (Atomicity, Consistency, Isolation, Durability) concept was devised (Härder and Reuter, 1983). A transaction must be atomic (either all or nothing is executed), must be consistent (integrity/correctness of the data kept) and once executed must always persist as such (duration). Isolation property ensures that transactions executed concurrently (in parallel) bring the database in the same state as if they were executed sequentially. If all transactions over a database comply with ACID, then the database is always valid. The second, optional but sought for concept is concurrency. As much as possible transactions should be executable concurrently whilst still complying with ACID.

Standard database transaction model is limited to the database itself (executing SQL instructions), is meant for short duration of transactions and offers limited to none user management support. This is opposite of the basic characteristics of land transactions which
typically need to access heterogeneous IT environments, are often of long duration and must be processed by various stakeholders. To overcome the limitations of the standard transaction model, advanced transaction models were developed, aiming to remove some limitations of the standard transaction model. Most of the advanced transaction models support long transactions, but do not support spatial data nor provide multiple users support or extended flexibility in terms of operations that can be performed within a transaction. This, along with a fact that there are no industry standard implementations of advanced transactions renders this technology also unsuitable for use in land administration information systems.

The named disadvantages of both standard and advanced transaction models are not present in workflow management systems (WFMS) with transactional support. What WFMS provides, and transaction models lack is flexibility. The workflow process consists of a series of related activities. Each activity can be assigned to a different user, and each activity can represent either an SQL command within a database, a web service to an external system, an external or an internal program. WFMS manages execution of activities in order defined by workflow logic.

This paper represents a first attempt to define a conceptual data model for the dynamic component of land administration systems. The data model is based on the WFMS concept with support for the transactional concepts needed for land transactions. We limit the research to the spatial component of the land administration data. Since spatial data is more complex than non-spatial, then it is expected that extending the model to also support non-spatial data should be easier than the other way around. Also, a proposal for integration of the developed data model with LADM is given.

The remainder of the paper is organized as follows. In section two we present WFMS concept and related terminology. Several notations used for modelling processes are presented and decision on using Petri nets is explained. Section three presents the main contribution of the paper, the conceptual model of a dynamic component of land administration system. The model is divided to generic and extended model. Generic model ensures consistency of processes on object level and should be applicable on cadastral data generally. Extended model ensures consistency by spatially defining affected area of the process and it is used to model processes on cadastral parcels spatially represented by polygons. Modelling of processes and transactional support is achieved with Petri nets. At the end of section three we demonstrate model of classical example in land administration systems, splitting a parcel. Section four summarises the outcomes of the papers and offers some possible directions for further research.

2. STATE OF THE ART

In this section we present WFMS concept and related terminology. Also, we describe notations used for modelling processes and explain decision on using Petri nets as notation for modelling processes in this paper.
2.1 Workflow Management Systems

Workflow (WF) is a computer improvement or automation of a business process. A system that fully defines, manages, and executes workflows through performing activities whose order of execution is defined by the workflow logic is called the Workflow Management System (WFMS) (WFMC, 1995). WFMS manages and controls business processes. A process is a set of one or more related activities that contribute to achievement of a goal. Workflow management is case-based, i.e. each process is executed through a specific case. A task represents a set of operations that represent a logical unit within the process. The task that is executed within a certain case is called a work item, and a work item executed by a stakeholder (more general term would be a resource) is called activity. Each task can have preconditions and postconditions. The precondition defines whether the task can be started at all, and a postcondition verifies the correctness of the result returned by the task.

The two disadvantages of standard WFMS related to transaction processing concepts are weak support for enforcing data consistency and recovery support in case of failure. However, by integrating transaction models into WFMS, attempts have been made to reduce or eliminate these disadvantages (Rusinkiewicz and Sheth, 1995; Eder and Liebhart, 1994; Alonso et al, 1996).

Although organizations initially used WFMS to improve their business processes, there is a trend of creating WFMS’ that go beyond individual organizations and allow interaction with non-corporate members (Van der Aalst, 2000; Meng, 2002). Also, there are examples of integration of recent Blockchain technology to the WFMS’ in the financial sector (Fridgen, 2018), which could be applicable to other areas as well as land administration systems (Anand et al, 2016; Anand, 2017; Enemark and McLaren, 2017).

Research in the domain of application of WFMS to spatial data is rather limited in comparison to non-spatial data. Examples of WFMS applications that include some form of spatial manipulation can be found in (Alonso and Hagen, 1997; Weske et al, 1998; Yue et al, 2015; Du and Cheng, 2017). These systems allow the composition of activities in the process to perform complex collaborative operations on spatial data, but without the application of transactional concepts or custom correctness criteria.

Grefen and Vonk (2006) created a taxonomy of transactional support in WFMS’. The authors defined a basic subdivision into integrated models where unique workflow definition model exists and separate models that use two separate specifications, one for the workflow and the other for the transactional aspect. Authors argue that existing commercial WFMSs provide the best support for the WF/Tx model although support for transactional behaviour is limited, since ACID properties can be applied only on task level. A model Tx+WF is a separate model of integration transactional support in WFMS. Both the process specification and transaction specification are on the same abstraction level. Examples of this model are transactional extensions of languages for specification of web services such as Business process execution language (BPEL) or Web services description language (WSDL) (Tai et al, 2001). However, Tx+WF model also provides limited transactional support on a web service level based on two-phase commit protocol. Integrated model WF is also applicable for use in existing commercial management systems because it does not require the definition of new elements.
or the introduction of additional transactional elements. Furthermore, WF model provides a unique formal definition of the process with components of a workflow and provides support for ACID properties on a process level. It supports transactional and as well non-transactional tasks (such as starting a text editor to make an invoice or calling a web service). For this reason, the WF model has been chosen as a basis for this research.

2.2 Notations for modelling of business processes
Various notations can be used to display and model business processes. Business Process Model and Notation (BPMN) (OMG, 2013) enables standard visualisation of business processes in a comprehensible way to various stakeholders. UML provides a wide range of notations for describing different views of computer structures and behaviours. UML Activity Diagrams (UML AD) can be used in various applications, such as modelling business processes, workflows by use cases or workflows in user interfaces (OMG, 2017), and BPMN is used solely for modelling business processes. BPMN and UML AD are the most widely used notations for modelling and presentation of business processes. Although many researchers claim that BPMN is more suitable for modelling business processes than UML AD (White, 2004, Russell et al, 2006), Birkmeier et al (2010) claim that UML AD is at least as convenient as and BPMN. However, BPMN and UML AD are rather comprehensive and complex and lack the ability to test the formal correctness of the process. Petri nets are used in various applications, also for modelling business processes (Petri, 1966). Petri nets have a strong mathematical basis and their graphic nature makes them easy to understand. Moreover, the theoretical basis of BPMN and UML AD notation originates from Petri nets.

Petri nets are directed graphs that consist of three components: places, transitions, and arcs. Places are displayed as circles and represent possible states or conditions of the system. Transitions are shown as rectangles and describe tasks (operations) that can change the system state. Places and transitions are associated via arcs. Arcs between the same type of nodes (place-place or transition-transition) is not allowed. Petri nets used for modelling processes in WFMS are called Workflow nets (or shorter WF nets) introduced by van der Aalst (1998). Since the process modelling in this paper is used to model processes on spatial component of cadastral parcels which requires flexible definition of correctness criteria for different types of processes, it is sufficient to support the basic patterns for modelling processes. For this reason, WF-nets for process modelling have been chosen for modelling processes.

3 PROPOSED MODEL
The basic spatial setup for the research builds on the outcomes of the paper (Vranić et al, 2015). The main concepts, classification of the types of changes, defining the transaction scope and correctness criteria are used from that paper and are slightly extended or modified to be fit for the subject purpose. To reach the goal, the research is executed in three phases.

First a generic conceptual model for modelling of processes on cadastral data is defined. This model implements the mechanisms for fulfilling requirement for long duration of processes and checking correctness on object level. We use concept of the net effect of the process (Widom and Finkelstein, 1990), which defines that only final effect of a set of operations in
the process is checked. Hence, the phases of preparation of the process and checking its correctness is separated. This allows modelling workflows in pessimistic or optimistic manner.

Following, the generic model is extended to specifically support the cadastral parcels spatially represented as polygons, as defined by the LADM. This phase represents connecting the WFMS models with the LADM’s polygonal spatial profile. Within this phase the mechanisms for fulfilling requirement of spatial definition of affected area to ensure serializability of the processes and correctness of spatial component of cadastral parcels are introduced.

Correctness of the process is checked according to the correctness criteria defined in (Vranić et al, 2015) because the testing of correctness on a tuple level is not sufficient to ensure the correctness of the spatial component of cadastral parcels. Due to the possible indirect changes on adjacent objects, if the rules for ensuring topological correctness are defined, affected area of the process must be spatially defined. Authors define the formally affected area of the process over the spatial component of cadastral parcels as a geometric union of formally affected cadastral parcels. Formally affected cadastral parcels are those which are the actual (formal) subject of the process. Authors also define the term technically affected parcels that are added to the process with the purpose of ensuring technical correctness due to the influence of indirect changes at the boundaries of the formally affected area. Geometrical union of technically affected parcels is referred to as the technically affected area. In this paper we define the affected area as a geometrical union of the formally and technically affected area. Since we restrict access to data spatially we also spatially define concurrent process. The concurrent process is every active process whose affected area overlaps with the affected area of another active process. A conflicting concurrent process is every active process in which the formally affected area touch or overlap of another active process.

In the final part of the paper a demonstration on fulfilling the posed requirements are given.

### 3.1 Generic conceptual model

The generic conceptual model (Figure 1 below), besides the concepts defined by standard WFMS requirements should enable:

- describing of the process with WF nets;
- transactional concepts of atomicity and isolation;
- defining the correctness criteria on a transition level, i.e. task-level;
- limit access to data on an object level;
- process recovery with definition of compensating tasks (backward-recovery);
- storing geometric information on errors.

Class `WF_Workflow` can have multiple places (`WF_Place`) and transitions (`WF_Transition`) associated with arcs (`WF_Arc`). Conditions (`WF_Condition`) can be associated to any arc which allows modelling of preconditions and postconditions of an individual transition. Temporal attributes `activeBegin` and `activeEnd` enable archiving of process definitions even after cases are instantiated. The `WF_Place` class contains a type attribute (`WF_PlaceType`) that specifies the type of a place: initial, final, or middle. Class `WF_Transition` contains the attributes `trigger` and `transitionTimeout`. The trigger determines how the task defined by the transition is started. Another important attribute of the transition is the attribute `vital` that
indicates whether the transition is vital or non-vital according to the concept defined in (Schick et al, 2011). Failure of a vital task causes rejection of the task and rollback of the process to last consistent state which means that process can be completely cancelled. Non-vital tasks can fail to execute, but this will not affect the outcome of the process. Workflow can have one or more transitions. Transition may have a reflexive association to one or more compensating transitions which are used to discard changes in case of an error or failure to preserve atomicity of a process. The attribute type of a WF_Arc class allows different types of routing workflows (for instance, sequential, parallel, or choice) which increases flexibility of a workflow. The WF_Condition class contains a definition of conditions that can be associated with a place as a precondition or a postcondition.

**Figure 1. WFMS conceptual model**

WF_Case is an instance of a workflow and represents a process that needs to be solved. Multiple cases can be instantiated according to the same workflow. Class WF_Workitem represents a single task (work item) within a process that can be assigned to a user. When it becomes enabled, or when it is executed by the resource, then it is also called activity. The activity contains temporal attributes which determine the status of the activity. enabledDate is written when the transition preconditions have been fulfilled. The finishedDate attribute is the moment when the transition is completed. The attribute canceledDate is written if the user cancels the activity or if the compensating activity is started. Class WF_Token keeps information on current state of the process. Along with the identifier and status, it contains temporal attributes that determine the status of the token. The attribute enabledDate is written when postconditions of the previous transition are fulfilled. The consumedDate attribute is written when the next activity is triggered, i.e. when the preconditions of a following transition are fulfilled. One case may have one or more tags (in case of various splits).
The association between WFMS and LADM is a WF_Case and all objects involved in a process should have association to it (Figure 2). The abstract LA_Source class with its descendants LA_SpatialSource and LA_AdministrativeSource currently enables association between the data and sources based on which a change occurred. Process itself is not modelled in LADM. Therefore, WF_Case is set as a descendant of the LA_Source class, and the LA_SpatialSource and LA_AdministrativeSource classes are associated with the WF_Case. WF_Case is further associated with a VersionedObject that is a parent class for most of the LADM classes. Thus, all descendants of the VersionedObject inherit the two associations caseCreated and caseDestroyed to the processes in which the object is created or destroyed. The VersionedObject is extended with an attribute locked that allows the object to be locked. Also, reflexive association checkedOutVersion is added to enable the tracking on which version of the object a process is making changes, i.e. which version a process checked-out.

![Figure 2. Association between LADM classes and WFMS](image)

### 3.2 Extending generic model to support spatial component of cadastral parcels

In this section, the generic model is extended to enable modelling of the process over the cadastral parcels spatially represented as polygons with the following requirements:

- spatial definition of the affected area of the process;
- ensuring serializability of the process by locking the parcels in the affected area;
- ensuring of geometric and topological correctness of the spatial component.
To enable definition the affected area of the process, the attributes *formalAffectedArea* and *technicalAffectedArea* have been added to the *WF_Case*. Additionally, the association classes *CP_AffectedParcel* and *CP_ResultingParcel* have been used from (Vranić et al, 2015), enabling the definition of parcels that are affected by the process and parcels resulting from the process (Figure 3). In a complex processes where the identity of an object changes or disappears (such as splitting, merging or re-allocation) these classes enable tracking of history of objects.

![Diagram](image)

**Figure 3.** Association between cadastral parcels, points and the case

The developed model supports pessimistic approach by locking objects within affected area or optimistic (non-locking) approach. In either approach affected area needs to be defined to limit required set of spatial data which needs to be checked within a process. In this paper only pessimistic approach will be explained.

### 3.3 Modelling processes on spatial component of cadastral parcels

Now the developed model is applied to modelling two simple concurrent processes that simultaneously try to split adjacent parcels (Figure 4 below). Although it may seem that the result of the simultaneous execution of the process is correct this is not so because of the processes’ isolation i.e. one process is not aware of changes that a concurrent process prepares on its versions of the same objects. Since processes must comply with isolation property to preserve correct temporal order of versions of parcels, if execution is not handled correctly, the problem of duplicated objects may occur. In *Process 1* the affected parcel $A$ is split into resulting parcels $A1$ and $A2$. Indirect result is that a new node $v_1$ is created on a common border with cadastral parcel $B$. To maintain the topological consistency of the spatial component of cadastral parcels, node $v_1$ must be added to the boundary of the cadastral parcel
so the parcel B needs to be technically included in that process and a new resulting version of the parcel \(B_0\) must also be created there. In the Process 2, parcel B is split to parcels B1 and B2 and a node \(v_1\) is created at the same position as in Process 1. Although in this example the node \(v_1\) is on the same position in both processes it could be on a different position. Although it seems that concurrent execution of processes preserves topological correctness of the spatial component, this is not the case because in the Process 1 two new parcels A1 and A2 are created while in the Process 2 creates a new version of the parcel A00 on the same place. The same situation is with the parcel B. Therefore, it is necessary to limit access to parcels.

![Concurrent processes (splitting a parcel)](image)

**Figure 4. Concurrent processes (splitting a parcel)**

Figure 5 shows how previous example can be modelled by using pessimistic workflow which preserves consistency by locking formally and technically affected parcels of a process as described in (Vranić et al, 2015). Locking of parcels within affected area ensures serializability of processes. Preparation of a change is separated and followed by the checking of correctness of the spatial component of the cadastral parcels. If all correctness criteria are fulfilled, the change which process causes is marked as correct and the process can be executed.

![Pessimistic workflow](image)

**Figure 5. Pessimistic workflow**
Figure 6 shows two possible time schedules of execution of concurrent processes that ensure data consistency. First option (Process 2a), would be to discard changes and wait for the concurrent Process 1 to finish. Second possibility (Process 2b) would be to wait until concurrent Process 1 is completed and then proceed to prepare the change. The first phase of acquiring of locks on objects ends when the token comes to the place $P_2$. From that moment on, the process will not acquire additional locks, and the concurrent processes cannot lock any of the locked objects. Object locks are released when the token comes to the final position of the workflow $o$. From that moment on, other processes can lock and modify those objects.

![Diagram of concurrent processes](image)

**Figure 6. Time schedule of concurrent processes according to pessimistic workflow**

4 CONCLUSION

In this paper we made a first attempt to define a conceptual data model for the dynamic component of land administration systems based on the WFMS concept. Introduction of WFMS concept enables flexibility in modelling processes which increases usability of a system, while integration of transactional concepts into WFMS can ensure consistency of data which is a crucial property of land transactions. Research is limited to the spatial component of land administration data, although it is expected that the model could be easily extended to support non-spatial data.

Many questions however remain open after this. The research has shown how one of two structured spatial data structures, the polygonal can be handled. Polygons are easy to manipulate since the entire object is always assembled, however when planar partitions are to be handled they also introduce a lot of redundancy. Subsequent researches need to analyse whether and how the described approach would handle topological data structures. Also, the research only tackled with pessimistic approach to handling concurrency. Especially with polygons, strict locking can lead to rather restrictive situations (e.g. when large elongated parcels are locked causing many other parcels to be inaccessible). Potentially, optimistic concurrency control could improve the concurrency in such situations or boundary points need somehow to be introduced into the overall data model to enable finer grained locking. Furthermore, within this research the WF approach to modelling transactional WFMS has been chosen based on a rather limited set of selection criteria. Possibly some of the other concepts could prove more appropriate for the purpose. Finally, one of the characteristics of transactional WFMS is the ability to introduce a level of relaxation of atomicity of processes. This has not been researched, but it is the authors vision that complex transactions, involving multiple data sets (e.g. parcels and building legal areas) could benefit from introducing relaxation of atomicity.
ACKNOWLEDGEMENT

The authors would like to thank Dr. Mario Mader for the proof reading the text and for some constructive critics that helped us to improve it. The authors would also like to thank the reviewers for their constructive comments and suggestions that helped us to improve the quality of the text even further. This work has been fully supported by Croatian Science Foundation under the project HRZZ-IP-2013-11-7714.

REFERENCES


**BIOGRAPHICAL NOTES**

Saša Vranić graduated at University of Zagreb, Faculty of Geodesy in 2009th with diploma thesis Interface of cadastral database. After graduation he worked in Croatian geoinformatics company Geofoto LLC for several years as GIS Consultant on projects of implementing spatial information systems. Since 2012, he is employed at the University of Zagreb, Faculty of Geodesy as University Assistant on Chair for land surveying. He is also PhD student and his main research interests are land administration systems, spatial databases, workflow management systems and land survey data management.

Hrvoje Matijević is currently working as Project Manager at the Croatian geoinformatics company IGEA L.L.C. and is a guest lecturer at the Faculty of Geodesy in Zagreb. Prior to this Hrvoje has worked in another Croatian geoinformatics company, Geofoto L.L.C. as Senior Consultant, and at the Faculty of Geodesy in Zagreb as University Assistant. He
received master’s degree in 2004 with thesis "Cadastral Data Modeling" and defended in 2006 his Ph.D. thesis "Modeling Changes in Cadastre" at same university. His main research interests were and still are GIS and Spatial DBMS technology in service of cadastral data management.

**Miodrag Roić** graduated in Geodesy from the University of Zagreb, Faculty of Geodesy. In 1994, he received a PhD from the Technical University Vienna. Since 1996, he is a professor at the University of Zagreb, Faculty of Geodesy. He was Dean of the Faculty 2011-2015. The topics that he specializes in are Cadastre, Land Administration Systems, Engineering Geodesy and Geoinformatics. He was an editor-in-chief of Geodetski list. He is a corresponding member of the German Geodetic Commission (DGK) and many other national and international scientific and professional institutions.

**CONTACTS**

**Saša Vranić**  
University of Zagreb  
Faculty of Geodesy  
Kačićeva 26  
HR-10000 Zagreb  
CROATIA  
Phone: +385 1 4639 210  
Fax: +385 1 4828 081  
E-mail: svranic@geof.hr  
Website: https://www.researchgate.net/profile/Sasa_Vranic2

**Hrvoje Matijević**  
IGEA LLC  
Frana Supila 27/B  
HR-42000 Varaždin  
CROATIA  
E-mail: hmatijev@geof.hr  
Website: https://www.researchgate.net/profile/Hrvoje_Matijevic

**Miodrag Roić**  
University of Zagreb  
Faculty of Geodesy  
Kačićeva 26  
HR-10000 Zagreb  
CROATIA  
Phone: +385 1 4639 229  
Fax: +385 1 4828 081  
E-mail: mroic@geof.hr  
Website: http://www.geof.unizg.hr/~mroic
Saša Vranić, Hrvoje Matijević and Miodrag Roić
Application of Workflow Management System to the Modelling of Processes in Land Administration Systems

7th International FIG Workshop on the Land Administration Domain Model
11-13 April 2018, Zagreb, Croatia