Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : http://oatao.univ-toulouse.fr/
Eprints ID : 17675

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

Modelling complex large scale systems using object oriented Bayesian networks (OOBN)

Q.Liu * **A.Tchangani** ** **F.Pérès** ***

* *Université Fédérale Toulouse Midi-Pyrénée - Laboratoire Génie de Production, Tarbes, 65016 France (e-mail: quan.liu@ enit.fr).*
** *Université Fédérale Toulouse Midi-Pyrénée - Laboratoire Génie de Production, Tarbes, 65016 France (e-mail: ayeley.tchangani@ iut-tabes.fr).*
*** *Université Fédérale Toulouse Midi-Pyrénée, 65016 France (e-mail: francois.peres@ enit.fr).*

Abstract: The aim of this communication is to present a new way of how to structure modelling process of complex and large scale systems by object oriented Bayesian network (OOBN) for risk assessment and management purpose. In the first stage, we extend OOBN by presenting a new definition that introduces some flexibility, in a second stage, dynamic Bayesian networks (DBN) described by OOBN method are presented, that leads to a framework that we refer to as Dynamic Objet Oriented Bayesian Network (DOOBN). A demonstration in the domain of risk assessment of flash floods effect on the infrastructures inoperability is considered to show potential applicability of the extended OOBN.

Keywords: complex systems, modelling, OOBN, DBN, risk management

1. INTRODUCTION

Damage caused to socio-economic and environmental systems by natural disasters such as earthquake, hurricane, floods, forest fire, etc. depends highly on the interactions of many things such as physical phenomena, social organization, geography of concerned zone, industrial and agriculture activities, etc. On the other hand, damage caused by these natural disasters is less and less acceptable by citizens maybe because of the degree of technological advanced era in which humanity is actually living. Decision makers therefore need to dispose of a sound and efficient tools, methods or framework to assists them, analyze, predict, control and manage the effect of these disasters. The challenge for scientists and researchers return then to deriving and supplying such decision support system; and this communication is an attempt toward this goal.

Modeling such systems necessitates to consider modeling relationships (influence, correlation, causality, etc.), modeling uncertainty (no well-known of the nature of relationships, of the intensity of these relationships, or even of the delimitation of the system under consideration), modeling dynamics (the fact that interactions, intensities of these interactions, the structure of the systems, etc. may vary from time to time must be taken into account through time as an important parameter) that characterize these systems. Considering all these phenomena to model leads to very complex and large scale systems on one hand that prevent using classical physical laws to describe them on other hand; appealing therefore for new approaches. In this communication, a framework based on the object oriented Bayesian networks (OOBN) that permit to take into account uncertainty of relationships between different variables or subsystems through Bayesian network (BN) properties and the scale through object oriented (OO) mechanisms in terms of repeatability for instance, is proposed to model natural (with application to floods) disasters. Modeling objectives are multiform and may consist in the purpose of understanding the behavior of the system, of controlling the behavior of the system to optimize for instance some of its performance indices, of measuring and monitoring their relationships, etc. The main purpose of this communication is to model a system in order to facilitate its risk management decisions making. The remainder of this communication is organized as follows: in the second section, main features of complex interdependent systems are briefly presented together with the necessity of disposing of a structured tool for the modelling process; section three presents a state of the art on Bayesian network (BN) family as useful tool of modelling the complex systems; and the forth section is devoted to the main contribution of this communication in terms of extended object oriented Bayesian networks (OOBN) and dynamics Bayesian networks (DBN); and finally section considers applying these approaches for modelling flash floods phenomena in order to assess risk faced by infrastructures in zone where these phenomena take place.

2. COMPLEX SYSTEM

A complex system is defined as a set of elements interacting with each other and with the outside. However the complex systems have four characteristics —large, interconnected, dynamic and uncertain. The behavior of the system is highly unpredictable without a sound model. The

purpose of this communication is to develop an approach that can be used for the modelling of complex systems to evaluate the risk when one or some of the components are destabilized by an external event or internal cause. Such kind of model can be used in the many domains to assist the evaluation of indicators or the decision making, such as economy, medicine, production and so on. Because of the characteristics of this kind of systems, the model must be built by using the logical tool to have a more reliable model. Next paragraph reviews existing approaches and proposes those that can be used or adapted to fulfill our needs.

Classical modelling methods suppose many simplifying hypothesis which are not suitable to deal with all the aspects of an interdependent, large, dynamic systems. The simplifying hypothesis sometimes make the problem even worse see Le Moigne (1990). Finding another approach to simulate without too much hypothesis is a problem. However today it is possible to collect data from a complex processes, using these data to simulate the behavior of the system becomes an approach which can keep the system's original characteristics as much as possible. Learning approach offers the possibility to build a model with these data that can reproduce the behavior of the system. Learning is an artificial intelligence method that permits to relate input data (causes) of a system to the observations (consequences or output). The learning process may consist in determining the internal structure of the system that is identifying relationships (or interactions) between its different components referred to as structure learning; or in determining the strength of interactions between components of a system with known structure known in the literature as parameter learning. Within this framework, Bayesian Networks (BN) are very efficient for modelling uncertainties. Dynamic Bayesian Network (DBN) may be used when temporal dimension in the behavior of the system is to be taken into see Murphy (2002). In DBN, each sample instant t of time horizon is constituted by a BN. For BN, there exist many learning (mainly in what concern parameters learning) sound and powerful algorithms in the literature ; this is not the case for DBN for which existing learning algorithms are so complicated that their deployment in real world applications is not easy mainly in the case of systems with a huge number of components. A possibility to reduce this complexity is to use the so called Object Oriented BN (OOBN) in order to exploit possibilities offered by this modelling technique. The idea of modelling repeatable systems by object oriented techniques has been already considered in a certain number of studies such as works undertaken in references see Jaeger (2000),Weber and Jouffe (2006),Xiang et al. (2005) to mention just a few.

3. STATE OF THE ART

3.1 Bayesian Network

Bayesian Networks are used to formalize knowledge in the form of a causal graph associated with a probability space Nielsen and Jensen (2009), Pearl (1988). They are directed acyclic graphs (DAGs) where knowledge is represented by variables. Each node of the graph corresponds to a variable and arcs represent the probabilistic dependencies between these variables. Formally, a Bayesian network is defined by:

- a graph-oriented without circuit, noted $\mathcal{G} = (\mathcal{V}, \varepsilon)$, with $\mathcal{V}$, the set of nodes of $\mathcal{G}$, and ε, the set of arcs of $\mathcal{G}$,
- a finite probability space $(\Omega, \mathcal{A}, \mathcal{P})$, where Ω is the universe, i.e. the set of all the elements considered in the problem, $\mathcal{A}$ is a $\sigma - algebra$ on Ω and $\mathcal{P}$ is a measure on Ω such that $\mathcal{P}(\Omega) = 1$; $\mathcal{P}(\emptyset) = 0$; $\mathcal{P}(A) \leq \mathcal{P}(B)$ if A included in B,
- a set of random variables defined on $(\Omega, \mathcal{A}, \mathcal{P})$, corresponding to each node of the graph, such that the set of probabilities associated with these variables defines the distribution of probabilities attached to the network: $\mathcal{P}(\mathcal{V}_1, \mathcal{V}_2, ..., \mathcal{V}_n) = \Pi_{i=1}^n \mathcal{P}(\mathcal{V}_i | pa(\mathcal{V}_i))$

with $pa(\mathcal{V}_i)$, the parent set (also called predecessors or causes) of $\mathcal{V}_i$ in graph $\mathcal{G}$. There are two types of probability tables in Bayesian Networks Godichaud et al. (2012). Tables of prior probabilities characterizes the chances that the variable $\mathcal{V}_a$ without any parent is in state a_i. Tables of conditional probabilities establish the chances that a variable $\mathcal{V}_b$ is in state b_j based on the state of its parents Matthieu et al. (2012), Godichaud et al. (2012). Inference in a Bayesian network consists in propagating information in the network Ben Hassen et al. (2013b), Ben Hassen et al. (2015). Indeed, a model using this formalism is generally not intended to be a static representation of knowledge. Beyond the a priori reasoning, evidences may be introduced to update the observed situation and to insert into the model the changes enabling the refinement of the results Ben Hassen et al. (2013a). This new knowledge, takes the form of a so-called elementary information, denoted $\mathcal{J}$, relative to a particular node. There are two types of basic information. The deterministic information allows instantiating a variable, that is affecting it a precise value, (eg $\mathcal{P}(\mathcal{V}_a = a_1 | \mathcal{J}) = 1$). The imprecise information modifies the distribution of probability of the variable, either by excluding a value of the universe of the variable ($\mathcal{P}(\mathcal{V}_a = a_1 | \mathcal{J}) = 0$) or, more usually, by changing the law ($\mathcal{P}(\mathcal{V}_a = a_1 | \mathcal{J}) \neq \mathcal{P}(\mathcal{V}_a = a_1)$).

3.2 Object Oriented Bayesian Network

Systems to model for risk assessment become more and more large. The classic BN has some limits for modelling large scale systems. That is why the structured representation offered by the object oriented techniques enables to improve the performance of the Bayesian Networks in terms of complexity of specification and inference of large systems. An object-oriented Bayesian network, is a direct application of the object paradigm see Bangsø and Wuillemin (2000), Koller and Pfeffer (1997). The basic element is the class, fragment of a Bayesian network whose nodes are broken down into three sets: input and output interfaces together with internal nodes. The object oriented Bayesian network takes advantage of classic Bayesian networks but introduce the concept of instance nodes. An instance node is an abstraction of a part of a network into a single unit. Consequently, instance nodes can be used to represent different network classes within other networks. The notion of encapsulation allows the transmission of all properties of the net fragment. An object oriented network

can be viewed as a hierarchical description/model of a problem domain. This makes the modelling easier since the OOBN-fragments at different levels of abstraction are more readable.
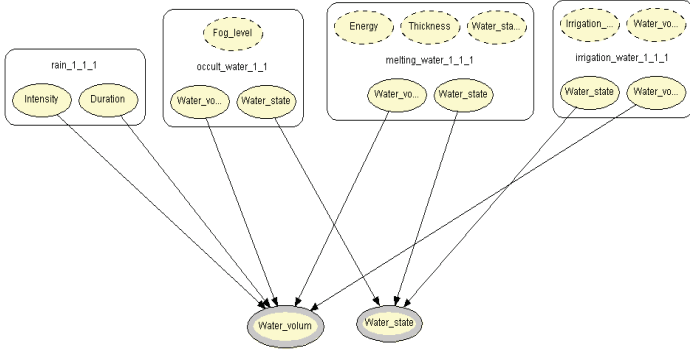

Fig. 1. OOBN model of water supply

An example of an OOBN is presented in figure 1. Based on Ambroise (1999) the model introduces four classes, namely: rain, melting water, irrigation water and occult water. Inputs are represented by dotted line such as energy or thickness, outputs are characterized by solid lines like for instance water state and water volume. Interstructure associated with internal nodes is encapsulated in each class.

Once the structure and relationships between the nodes has been established, the main work consists in characterizing the Conditional Probability Tables (CPT). Building OOBN model may be somehow difficult when it concerns a complex system with many variables, states or relationships. The use of learning techniques might bring some help for the identification of the relationship between nodes as well as the CPTs values. InLangseth and Nielsen (2003), Wuillemin and Torti (2012) the authors give some insight over OOBN structure learning. In Langseth and Bangsø (2001) the author extends the parameter learning based on OO assumption (the parameters of the objects who have the same structure are identical) and propose a parameter learning method based on the reducing of the parameter number during the learning phase. Using OOBN to modelling a complex system has not only reducing the design work, but also facility the update of structure work. However, the limit of the OO assumption and other researches of OOBN is the parameter part, they suppose that the same structure has the identical parameters. But actuality it not always likes this.

3.3 Dynamic Bayesian Network

A dynamic Bayesian network is a BN representing a temporal probabilistic model. A DBN is an extension of a BN, it adds a variable evaluates over time. The DBN is a series of time-slice, each slice contains a set of random variables, some observable and some not. The basic of DBN can be present by the Hidden Markov Model (HMM) Rabiner and Juang (1986), learning techniques in a DBN

extends directly form the classical BN see Murphy (2002). The extending parts are:

- *the transition model:* $A = P(X_{t_i}|X_{t_{i-1}})$
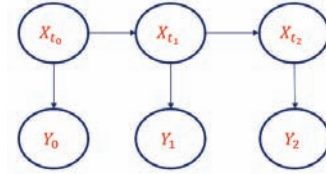- *the intial state:* $\pi = P(X_{t_0})$

Fig. 2. Dynamic Bayesian Network

Figure 2 presents a basic DBN with 3 time-slices, where the dynamic state is X_{t_i}, the observable state is Y_i. The communication between the time uses the transition model. In Bangsø and Wuillemin (2000), it uses an OO approach to present the dynamic part of BN.

4. EXTENDED OBJECT ORIENTED BAYESIAN NETWORK

In this section we present an extended OOBN (EOOBN) approach. After having described the different parameters associated with the objects, we will propose a definition, will show the construction steps and illustrate the concepts through an example. The extension of OOBN to time dynamic dimensions will be helpful in particular to characterize the failure of the system with time.

4.1 Extended OOBN

Due to the limits of the classical OOBN which do not consider the possibility of changing their parameters between the different instantiations Bangsø and Wuillemin (2000) and the difficulty to take into account dynamic interactions Koller and Pfeffer (1997), we propose here an extended structure of OOBN much more flexible in terms of time dependent process characterization.

Class and Object

In Bangsø and Wuillemin (2000), the authors define together the structure and the CPT with respect to a given object. Once the CPT of an object is defined, all the others will inherit its properties including the CPT parameters based on OO assumption. In this section we will consider that the definition of a class is made at structure level (that is the nodes and their connexions in the object) but the object itself will be instantiated both through the input values and with respect to its CPT parameters which are likely to evolve with context or time.

Class: A class (C) is the structure part S in a BN independently of the CPT parameters values. It has three kind of nodes namely: input nodes, output nodes and internal nodes. Only the input and output nodes are visible from outside the class.

Object: An object $(O\{S, P\})$ in the OOBN is an instantiation of the corresponding class. There are two parts in

an object, the structure (S) which inherits from the class and the parameters (P) which will be defined by experts or learning processes.

Since only the input and output nodes in the class/object can be seen from outer place, they are called and considered as *communication channel* for the class/object entity.

- Input node cannot have parents inside the class/object
- Input node is a reference node who is the projection of an output or a normal node from outside
- Input nodes can be split into two parts: the local input does not receive any information from other objects and the external output which requires exchanging information with outer nodes
- Internal node can have neither parents nor children outside the class/object
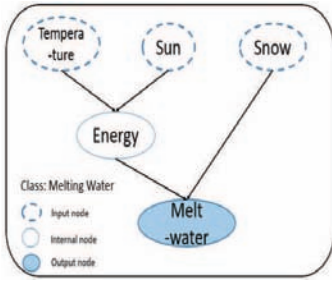- Output node cannot have children outside the class /object

Fig. 3. Class: Melting water

In Fig.3 an example of class is proposed dealing with the Melting Water variable. Temperature, Sun, Snow stand for the inputs while energy is an internal node and Melting Water represents the output in terms of liquid water volume. The class characterizes only the structure of the network. If a class has to be used, it must be converted into an object.

The EOOBN inherits all the advantages of the classical OOBN such as hierarchy or encapsulation. Moreover the object in EOOBN are distributed the different parameters.

Construction method

Because the EOOBN is hierarchical, one can still use the top-down method to make easier the network building. This can be done by carrying out the following steps:

(1) Formalize the hierarchical structure S of a system
(2) Design the structure of each class (C) with respect to S
(3) Instantiate the class by introduction the parameters corresponding to the object
(4) Connect the objects through their communication channels (because in a large scale system, there are a lot of objects)

4.2 Extended DBN

Although in Bangsø and Wuillemin (2000) a DBN simulation approach is given based on a self-reference node in an object, a confusion might appear when trying to add the dynamic part within a large OOBN. To overcome this issue

we introduce a virtual input/output node in the extended OOBN to simulate a dynamic network.

Virtual node: The virtual node is either an input or an output in the class/object. It stands for the temporal node in the class/object as a communication channel allowing exchanging with other time-slices. The connexion between the objects at time slices $t-1$ and t is as follows. The link between the temporal node (at $t-1$) and the virtual input (at t) is characterized by the CPT while the relationship between the virtual output (at t) and the temporal node (at t) is certain (conditional probability equal to 1).
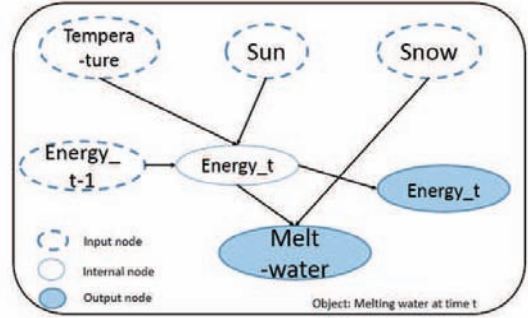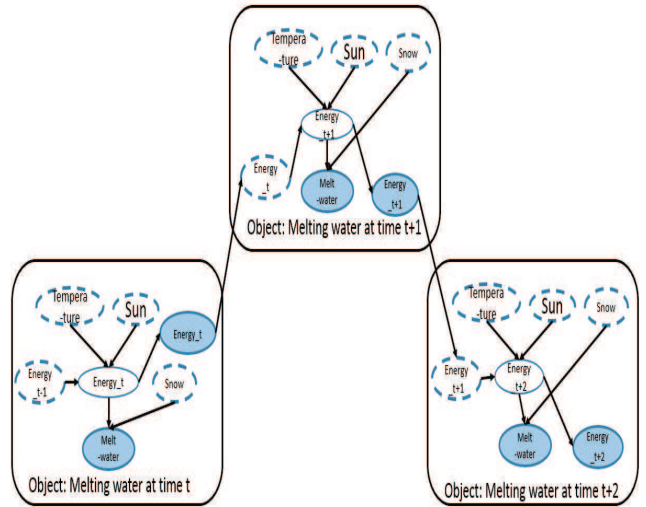

Fig. 4. Object: Melting water at time t

Fig. 5. Dynamic Bayesian Network in 3 time-slices

In Fig.4, the temporal node is the energy who changes through the influence of the sunshine, temperature and energy cumulated at $t-1$. Within this set of assumptions the virtual nodes associated with the variation of energy are represented at time t by the input node receiving the information from the $t-1$ time slice and the output node transmitting the information to the t+1 time-slice. The transition model characterizing the relationship between the object at time t-1 and time t is given by:

$$A_{t,t-1} = P(energy_t|energy_{t-1})$$

The transition model characterizing the relationship between the object at time t and for the input of time t+1 is given by:

$$A_{t,t} = P(energy_t|energy_t) = 1.$$

$A_{t,t}$ just stores the information to the outside communication channel and can be considered from the next time-slice as an input node (see Figure 5).

Adding the virtual node in the class/object protects the encapsulation structure of an object. The communication between the time-slice through the virtual node keeps the independence of each time-slice. In the next section we try to make compatible the dynamic behaviour with the object oriented approach through the use of Dynamic Object Oriented Bayesian Networks (DOOBN).

4.3 DOOBN

Today systems are made complex not only by their large structure but also by their propensity to evolve with time. To deal with this issue a suitable tool is required. Given that OOBN and DBN seem appropriate to model complex systems the idea is to join their specific capabilities to generate a new class of tools called DOOBN.

This matter has already started to be tackled in Weber and Jouffe (2006) where the authors tried to define a DOOBN for the modelling and simulating complex systems. Nevertheless they did not formalize neither the definition nor the construction of such a model. In our approach we propose to extend the previously described Extended Dynamic Bayesian Network by using the Object Oriented concept. In order to reproduce the dynamic behaviour related to a large and complex OOBN one just need to identify the temporal nodes (that is the ones corresponding to the time-dependent variables) and introduce the virtual nodes to each one. The network shown figure 4 corresponds indeed to a DOOBN since it makes clearly appear an encapsulated object combined with virtual nodes as input and/or output of the time-dependent nodes. By using the extended OOBN approach we can set up the dynamic part of the OOBN which solve the problem of setting up a dynamic part to an OOBN from Koller and Pfeffer (1997).

5. APPLY TO THE FLASH FLOOD

Nowadays due to industrial development and urban progress, the climate is changing all over the planet. Scientist predict that climate changes will increase the frequency of heavy rains, putting many communities at risk of ooding. Especially in the mountainous areas, this risk is increased by the relief and its consequences in terms of water ow kinematics. Flood directly threatens human life and material issues due to the intensity and suddenness of the events. Flood risk analyses are necessary for protecting the population and infrastructure. Facing a potential, announced or proven crisis, they lead to a better design of insurance policies and actions of anticipation or remediation implemented by companies, municipalities or even citizens. The most common approach to dene ood risk requires to combine a hazard characterized by statistical aspects (frequency of occurrence) and physical aspects (ow intensity) and an impact expressed in terms of vulnerability; i.e. exposure and sensitivity of persons and goods to potential damages. Torrential oods are rapid gravity phenomena which include a share of irreducible uncertainty related to randomness events (rain, snow...) and to the knowledge of the involved processes. Risk

management decisions must compose and integrate this stochastic dimension. Bayesian networks seem a suitable tool for the implantation of such a model see Villeneuve et al. (2011). The modelling work which is in progress has been divided into 2 parts. It consists first in identifying the inuential parameters in the generation of the ood phenomenon by using extended OOBN (class) approach which can build a main structure of the phenomenon for the problem. The second part aims at initialise the class to have the objects and distribute the CPT for them. As a first step, an elementary time-independent extended OOBN class will be established to characterize the inuence of variables on a small geographical area homogeneous in terms of topology. In order to characterize the spatiality of the phenomenon, the idea is then to exploit this basic model as a generic block of modelling, characteristic of the evolution of all of the variables that may be brought into play, and then to associate the dierent bricks instantiated with respect to the considered area and sequenced temporally according to the phenomenon timeline. The Fig.6 is a class of the zone. There we don't make any different between internal and external input. In this class we add the dynamic part to the temporal variable by introducing the virtual node like $snow\_thichness\_t\_1$ (virtual input) and $snow\_thickness\_t$ (virtual output). The instantiation of this class is the object which simulates the river level at time t. Because the objective is to simulate the whole area, all the objects will have to be placed together and exchange the information, for example the node $river\_zone\_i$. Through input/output node in the objects such as the node $river\_zone\_i\_1\_t$ (the river level of zone $i-1$) which gives the information for the next zone i as the input (at time t). And modelling the dynamic of this place we can link all the corresponding virtual nodes together.

6. CONCLUSION

This communication considered the issue of modelling complex and large scale systems for risk management purpose. The necessity to dispose of a structure and flexible tool to tackle modelling of such systems, lead as to extends the classical object oriented Bayesian networks (OOBN) that we refer as (EOOBN) by defining the structure in the class and distributing the parameters in the instantiation level of object. This approach not only makes the OOBN much more flexible, but also keeps all the benefits of classical OOBN, such as encapsulation, hierarchy, top-down design. The extended OOBN is more suitable for modelling large complex situations encountered in risk management framework. To take into account dynamics of the systems and the possibility for parameters to vary over time, we adapt the EOOBN to obtain dynamic OOBN referred to as (DOOBN) where each time-slice is occupied by an object and parameters are allowed to change from slices through introduced virtual nodes. A multi-dimension OOBN (MOOBN) can be built by using the virual nodes which makes the model much more flexible. The construction of MOOBN will be developed in the next paper. An application of the developed approach for flash floods effect modelling show its usability at structural level; nevertheless, inference and learning processes have to be adapted and this issue is one of the future work to be carried up.
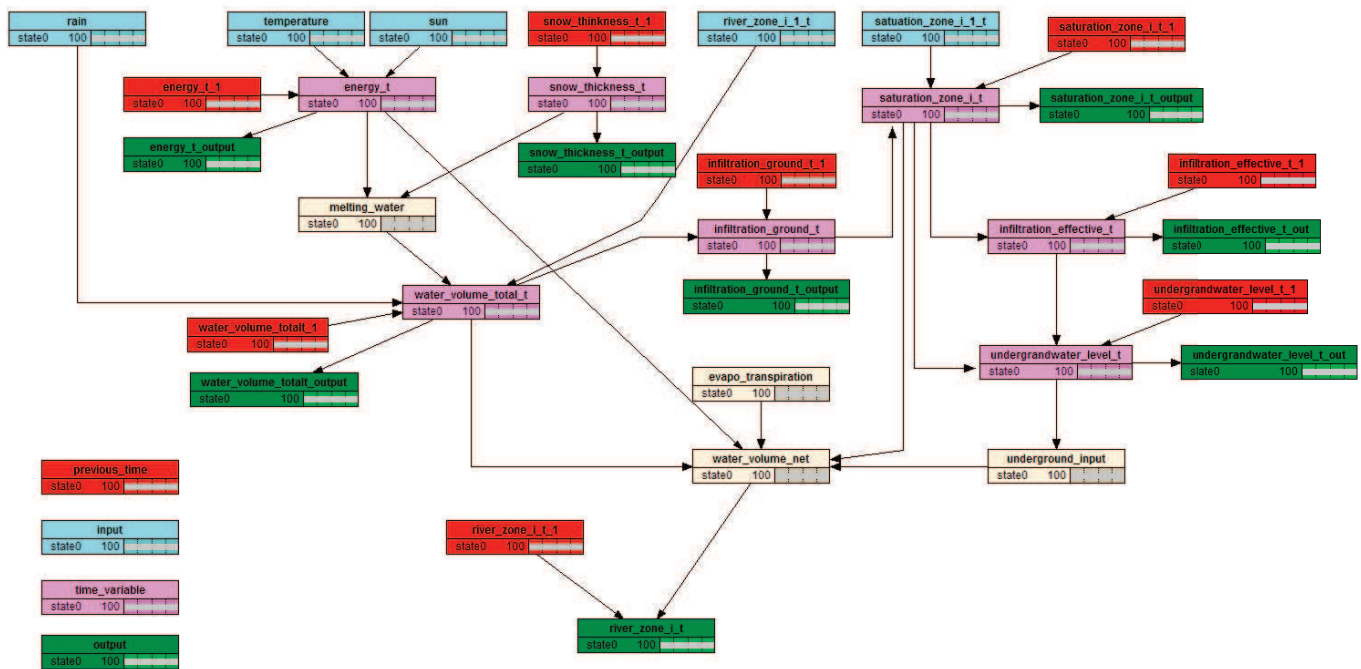
Fig. 6. Flash flood structure in a class level

REFERENCES

Ambroise, B. (1999). *La dynamique du cycle de l'eau dans un bassin versant: processus, facteurs, modèles.* HGA.

Bangsø, O. and Wuillemin, P.H. (2000). Object oriented bayesian networks a framework for topdown specification of large bayesian networks and repetitive structures.

Ben Hassen, W., Auzanneau, F., Incarbone, L., Péres, F., and Tchangani, A.P. (2013a). Omtdr using ber estimation for ambiguities cancellation in ramified networks diagnosis. In *Intelligent Sensors, Sensor Networks and Information Processing, 2013 IEEE Eighth International Conference on*, 414–419. IEEE.

Ben Hassen, W., Auzanneau, F., Péres, F., and Tchangani, A.P. (2013b). Diagnosis sensor fusion for wire fault location in can bus systems. In *SENSORS, 2013 IEEE*, 1–4. IEEE.

Ben Hassen, W., Auzanneau, F., Incarbone, L., Pérès, F., and Tchangani, A.P. (2015). Distributed sensor fusion for wire fault location using sensor clustering strategy. *International Journal of Distributed Sensor Networks*, 2015.

Godichaud, M., Tchangani, A., Pérès, F., and Iung, B. (2012). Sustainable management of end-of-life systems. *Production Planning & Control*, 23(2-3), 216–236.

Jaeger, M. (2000). On the complexity of inference about probabilistic relational models. *Artificial Intelligence*, 117(2), 297–308.

Koller, D. and Pfeffer, A. (1997). Object-oriented bayesian networks. In *Proceedings of the Thirteenth conference on Uncertainty in artificial intelligence*, 302–313. Morgan Kaufmann Publishers Inc.

Langseth, H. and Bangsø, O. (2001). Parameter learning in object-oriented bayesian networks. *Annals of Mathematics and Artificial Intelligence*, 32(1-4), 221–243.

Langseth, H. and Nielsen, T.D. (2003). Fusion of domain knowledge with data for structural learning in object oriented domains. *The Journal of Machine Learning Research*, 4, 339–368.

Le Moigne, J.L. (1990). La modélisation des systèmes complexes. *Paris: Bordas, Dunot, 1990*, 1.

Matthieu, G., François, P., and Tchangani, A. (2012). Optimising end-of-life system dismantling strategy. *International journal of production research*, 50(14), 3738–3754.

Murphy, K.P. (2002). *Dynamic bayesian networks: representation, inference and learning.* Ph.D. thesis, University of California, Berkeley.

Nielsen, T.D. and Jensen, F.V. (2009). *Bayesian networks and decision graphs.* Springer Science & Business Media.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: networks of plausible inference.* Morgan Kaufmann.

Rabiner, L. and Juang, B.H. (1986). An introduction to hidden markov models. *ASSP Magazine, IEEE*, 3(1), 4–16.

Villeneuve, E., Béler, C., Pérès, F., and Geneste, L. (2011). Hybridization of bayesian networks and belief functions to assess risk. application to aircraft deconstruction.

Weber, P. and Jouffe, L. (2006). Complex system reliability modelling with dynamic object oriented bayesian networks (doobn). *Reliability Engineering & System Safety*, 91(2), 149–162.

Wuillemin, P.H. and Torti, L. (2012). Structured probabilistic inference. *International Journal of Approximate Reasoning*, 53(7), 946–968.

Xiang, Y., Jensen, F.V., and Chen, X. (2005). Inference in multiply sectioned bayesian networks: Methods and performance comparison. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 36(3), 546–558.