Models Integration for the Elaboration of a Domain Multi-view in Accidentology

Walid Ben Ahmed

Mounib Mekhilef

Laboratory of Industrial engineering,

Ecole Centrale Paris, 92295 Châtenay-Malabry,

FRANCE

Michel Bigand

Research Group in Industrial Engineering,

Ecole Centrale de Lille, BP 48, 59651 Villeneuve d'Ascq Cedex,

FRANCE

Yves Page

LAB, Laboratory of Accident research, Biomechanics and human behaviour,

132, rue des Suisses 92000 Nanterre,

FRANCE

Abstract: A mixed models integration approach, both top-down and bottom-up, is presented. Top-down approach is justified by the will to offer a generic and structuring framework; due to the complex nature of the problematic, a systemic approach has been chosen. Bottom-up approach concerns information extraction from database, including knowledge of experts. This mixed approach has been applied to build common representation of accident scenario for accidentologists and designers of new safety systems.

Keywords: models integration, complex system modelling, multi-view model, framework

Walid Ben Ahmed obtained his PhD in 2005, and developed a systemic approach for knowledge extraction from database. This work was performed in the LAB in the accident scenario building and interpretation. He is actually project manager in Renault.

Mounib Mekhilef is a mechanical engineer. He gets his PhD on 1992 at the Ecole Centrale Paris and his Ability to manage Research at the University of Nantes on 2000. Now he is an Associate Professor at the University of Orleans where he is giving courses in the Mechanical Engineering Department. He actually works in the area of Optimisation and Knowledge Management in the Industrial Engineering Department at the Ecole Centrale Engineering School.

Michel Bigand is lecturer in Computer Science and Project Management. He received a M. Sc. (1980) from Ecole Normale Supérieure de Cachan in Mechanical Engineering, and defended a thesis (1988) from Paris 6 University. His research activities in Central Industrial Engineering Laboratory concern the design systems engineering, the product life cycle management, and associated information systems.

Yves Page Page is manager's assistant of the Laboratory of Accident research, of Biomechanics and studies of the human behaviour (LAB) PSA Peugeot Citroën since May 2004. He was previously in charge of primary accidentology and gave knowledge to car manufacturers in road safety in order to design best safety systems. He has coordinated French or European programs of the CEESAR (European Center of Safety Studies and Risks Analysis).

1. Introduction

The aim of this work is to provide in-car safety system designers with accidentology knowledge so as to allow them to understand accident behaviour and therefore to develop new road safety systems. In order to specify such systems, designers and accidentologists have to collaborate. However, designers (specialists in mechanisms and electronics) and accidentologists (specialists in mechanisms, ergonomics, infrastructure and psychology) have different viewpoints, different models for accidents analysis and technical languages and so their communication is difficult.

Accident Scenario (AS) model is a powerful tool for knowledge sharing. A scenario is a prototypical behaviour of a group of subjects or objects (customers, accidents, users, etc.) with similarities; scenario-based approaches are used in different fields [1, 2, 3, 4, 5, 6].

Accidentologists feel that similar accident factors entail similar safety countermeasures [7, 8, 9, 10]. Thus, AS is recognized by the LAB1 as a powerful tool to provide safety system developers with the required knowledge. Figure 1 gives an AS example. It is one of 18 scenarios elaborated using a sample of 750 road accidents; it covers 30 road accidents.

1 Laboratory of Accident research, of Biomechanics and studies of the human behaviour

"The accident happened at a junction of two main roads. The weather was sunny and the road surface was dry. A driver came up to the roundabout at the junction. He did not know which direction he had to take, and was concentrating on the road signs. As he reached the roundabout, he glanced left quickly, and thinking that the road was clear, pulled out. He declared his speed to be about 20 km/h. The crash barrier that runs round the middle of the roundabout reduces the visibility of vehicles coming from the left."

Figure 1: Example of an Accident Scenario (AS)

Within the LAB, one of the key points of primary accidentology is to carry out accidentology studies by using the so-called In-depth Accident Investigation Databases (IAID). The data for the IAID are gathered by the CEESAR2. This Center has two multi-disciplinary teams of experts who deal with the collection of accident-related data. Indeed, in-depth studies are carried out by these teams on the scene of accident, both made of a psychologist, who interviews the persons involved, a vehicle expert, who collects the data related to the vehicle involved, and an infrastructure expert, who collects the data related to the environment of the accident scene.

When an accident occurs, the CEESAR is informed by the Emergency Services and a team is sent as fast as possible to the accident scene. The accidentologists look for all the information that might help to explain why the accident happened. The team of experts uses the information collected during the on-site investigation, combined with later investigations, to perform a spatial and temporal reconstitution. Accident-related data are then stored in the IAID. These databases are constituted of data tables describing the driver's behavior, the environment (i.e. the infrastructure, traffic, ambient conditions, and socio-economic context) and the vehicle(s). Each accident is characterized by attributes that are often categorized with several modalities (e.g. the attribute "Road lightning" has the following modalities: No road lightning, Existing one, Lights on, Lights off, and Defective lightning). In addition to the formalized data, the IAID contain pictures of the environment and the vehicle(s), interviews of the person(s) involved, and a dynamic reconstitution of the accident performed by experts on the basis of the information collected.

These IAID are difficult to use for several reasons:
A lot of accidents are already stored (more than 1000);
There is an important number of attributes for each accident (947);
Types of data are various (structured data, interviews, dynamic reconstitution…);
The Driver-Vehicle-Environment (DVE) model is complex.

In order to elaborate an AS, it is possible to use an expert approach: an expert clusters accidents manually from IAID according to their similarity [7, 8, 9, 10]. Next, he elaborates a synthetic description for each cluster. However, this approach has some drawbacks related to the fact that expertise is expensive and scenarios depend not only on the studied objective, but also on the expert viewpoint and discipline. Moreover, different granularity levels and ways of representing accident scenarios exist.

Other methods use data-mining techniques in order to elaborate accident scenarios. In [11, 12, 13, 14] the authors propose classification techniques to elaborate accident configurations. In [15, 16], the authors propose clustering techniques to perform AS. However, the interpretation of the statistical features is a hard task for experts.

So as to identify the different viewpoints and models that are relevant to analyze road accident in order to define new countermeasures, a mixed models integration approach, both top-down and bottom-up, is proposed.

First we propose to use the systemic top-down (also called cybernetic) approach [17, 18, 19] in order to identify the relevant viewpoints and related models (section 2). Secondly, a bottom-up approach is applied for the interpretation of data-mining results and AS representation (section 3).

2. Systemic Top-down Approach for Viewpoints Integration

First, an accident model was built with a systemic approach. More generally, this approach makes it possible to develop a multi-view schema including several granularity levels for a given domain.

2.1. Choice of the Reference Model

Due to the impossibility to predict the DVE system behaviour, road accident can be considered as a complex phenomenon. This unpredictability is notably due to the fact that human actions are strongly involved in accident causation, and that human behaviour is unpredictable. Furthermore, during the road accident, DVE

2 European Center of Safety Studies and Risks Analysis

behaviour may be described through feedbacks and recursive loops. According to Miller's definition of a living system [20], DVE is here defined as an open and living system since each component is constantly interacting with its environment by means of information exchanges. Due to these feedbacks and recursive loops, it is impossible for designers and accidentologists to identify with exhaustiveness and certainty all the failures and dysfunction mechanisms occurring in a road accident. Moreover, a same accident may be seen differently according to the analyst viewpoint. Each designer has an individual perception of the same phenomenon. This assumption is based on constructivist foundations [21], which assume that knowledge depends on how the individual "constructs" meaning from her/his experience.

A system, in a constructivist perspective, is recognized as a representation of reality seen by some people in a given context. The systemic approach assumes that to handle a complex behaviour, it is fundamental to make junction between the ontological, functional, transformational and teleological viewpoints [21] (figure 2).

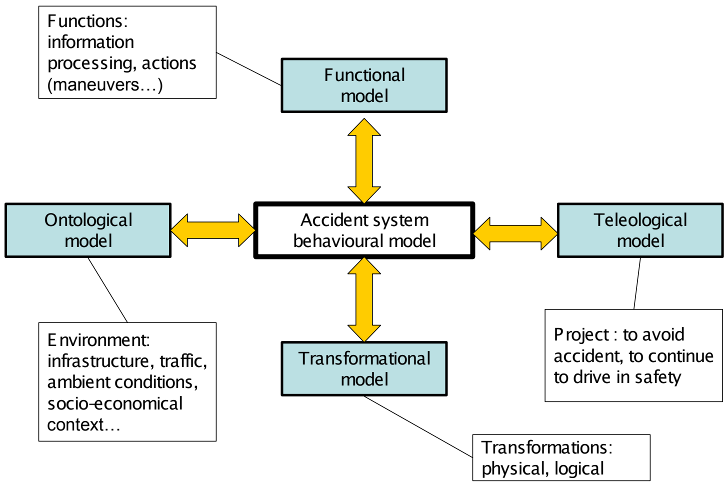

Figure 2: Systemic Model for the Elaboration of the Accident System Behavioural Model

The ontological viewpoint describes what the system is. It represents the sub-systems (driver, infrastructure, traffic, ambient conditions, vehicle, etc.), their taxonomic groups, their contexts (the driver's professional status, family status, etc.), their structures, as well as the various interactions between these sub-systems and their components.

The functional viewpoint describes what the system makes. It represents the global process of the DVE functioning during the road accident, which combines several procedures (perception, diagnostic, prognostic, decision and action) [22].

The transformational (or evolutionary) viewpoint describes how the system evolves. The DVE system behaviour can be described as an evolution that goes through several states. The transformational viewpoint integrates the accident's sequential and causal models developed by the INRETS3 [23, 24];

Finally, the teleological (or intentional) viewpoint describes what the goal of the system is. It assumes that each of the DVE system components or functions has to serve a purpose in an active context in order to ensure the safety of the DVE system.

2.2. Multi-view Elaboration of Accident Scenario

Using the systemic viewpoints, a framework was developed that ensures to represent the same scenario according to different models specific to different fields (safety system design field and accidentology fields). Each scenario user has the possibility to represent the scenario according to his/her own model.

The used approach is described through the following steps:

1. Find and/or construct accident representation models according to each systemic viewpoint. For example, the DVE model is assigned to the ontological view, the sequential model is assigned to the transformational view, the information processing model is assigned to the functional view.
2. Each model is composed of one or more concepts. For example, "Normal driving step", "Failure step", "Emergency step" and "Crash step" are the concepts composing the sequential model. "Perception",

3 French National Institute on Transports and Safety

"Diagnosis", "Prognosis", "Decision" and "Action" are the concepts composing the information processing model.

3. Each concept is characterized by one or more attributes. Each attribute may characterize many concepts in different models. For example, the attribute "steering angle" characterizes, at the same time, the concept *Driver/Vehicle interaction* in the DVE model, the concept *Emergency* in the sequential model and the concept *Action* in the information processing model. The attributes classification according to the model concepts can be perceived as the construction of metadata since it is a "data about data". Figure 3 shows how XML is used to represent these metadata and how an attribute (here *Steering angle*) can be assigned to various concepts. This assignation of all the attributes present in the IAID was made by in parallel by two teams of three experts, with a convergence method. A level of pertinence and a level of reliability were assigned to each attribute.

```
<?xml version="1.0"?>
<Accident_Metadata>
<Viewpoint>
  <ViewpointName> Ontological_View </ ViewpointName >
  <Model>
  <ModelName> DEV_Model </ModelName>
  ...
  <Concept>
     <ConceptName> Driver/Vehicle interaction </
ConceptName >
     <Attributs>
        Steering angle
     </Attributs>
  </Concept>
  ...
  </Model>
  </Viewpoint>
<Viewpoint>
  <ViewpointName> Functional_View </ ViewpointName >
  <Model>
  <ModelName> Information_Processing_Model
</ModelName>
  ...
  <Concept>
     <ConceptName> Action </ ConceptName >
     <Attributs>
        Steering angle
     </Attributs>
  </Concept>
  ...
  </Model>
  </Viewpoint>
...
</ Accident_Metadata >
```

Figure 3: Example of XML Representation of the Metadata: each Attribute is Assigned to One or Several Concepts According to the Various Models

The accident clusters are characterized by attributes and these attributes are classified according to the concepts in the different models; so it is possible to perform a multi-view projection of an AS according to the different models.

Figure 4 presents the UML class diagram from which the software allowing the multi-view projection of the models was developed.

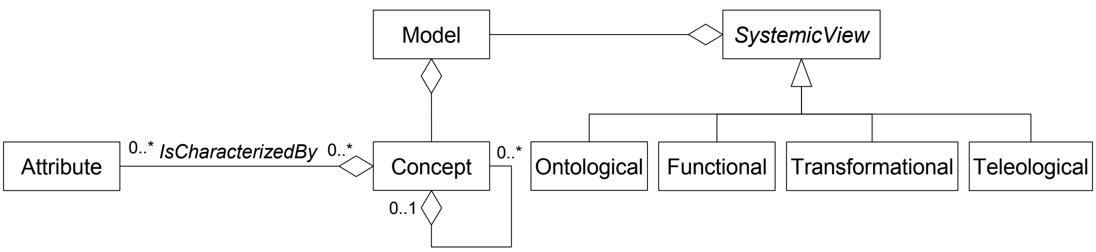

Figure 4: UML Class Diagram for the Developed Software: a view (Ontological, Functional, Transformational or Teleological) is Composed of Several Models and a Model is Composed of

Concepts; a Concept can be Decomposed in Concepts and can be Characterized by aAttributes

3. Bottom-up Approach for Accident Scenario Representation

The second proposal, complementary to previous modelling work, consists in incorporate experts' knowledge in a Knowledge Discovery in Database (KDD) process. It is a bottom-up approach because the start point is a database (the IAID) and the result consists in AS, of high conceptual level.

Experts' knowledge intervenes in two of the three steps of a KDD process [25, 26]: data preparation and results interpretation (figure 5).

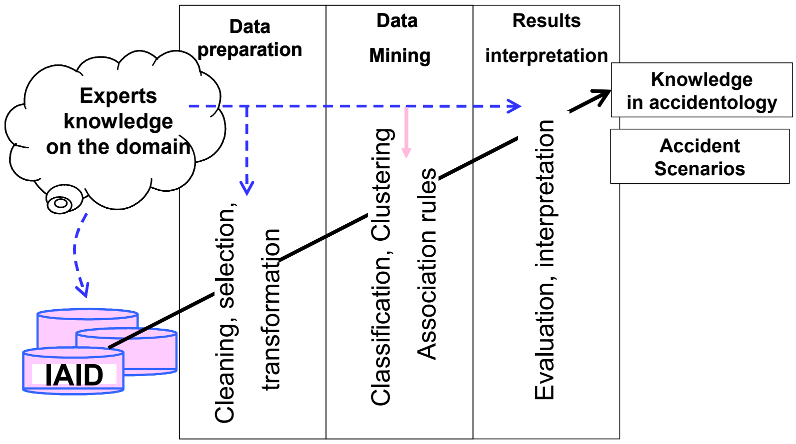

Figure 5: Overview of the KDD Process and Expert's Knowledge Integration: Starting by Databases (here IAID), the First Step is the Data Preparation, the Second is Data Mining and the Third is Results Interpretation; Experts' Knowledge Intervenes in the First and the Third step to Produce Accident Scenarios and Increase Knowledge in Accidentology for the User

3.1. Data Preparation

Each work of a user concerns a particular study for a given order, like for example ABS braking system impact.

In order to make a clustering of accidents, the user begins to select the pertinent attributes related to the objectives of the study. The attributes of the model corresponds to the columns of the database tables.

An interface was developed allowing the user to express an attributes selection request like figure 6. The user checks boxes to select the systemic viewpoints, the models and concepts which are appropriate to his/her study.

```
(Ontological or Functional or Transformational) and (Driver/Vehicle interaction or Driver/Environment interaction) and (Failure step or Emergency step) and (Perception or Diagnosis or Prognostic)
```

Figure 6: Example of Attributes Selection Request

Consequently, it is possible to translate study objective in pertinent attributes using metadata formalized by the experts of the domain. By comparison with manual expert clustering, this approach presents several advantages. Firstly, it systematizes the attributes selection from different viewpoints, and the risk to miss a pertinent viewpoint is reduced. Secondly, the reduction of the attributes number is made automatically, but by the use of experts' team knowledge, and no *a priori* by an isolated expert (however, a manual attributes selection is always possible). Lastly, not only the selection quality is improved but the time is markedly reduced (some minutes comparing to several hours previously).

The second step uses clustering techniques. We used k-means algorithm [27]; this step is not detailed in this paper.

3.2. Results Interpretation

When the clustering is achieved, we obtain a table like figure 7 for each cluster; the SPAD tool was used to obtain this table. The designers and accidentologists have to interpret these tables using statistical features.

Clustering Attributes	Attribute modality	% of the modality in the study sample	% of the modality in the cluster

Crash position	Offroad	26.64	96.72
Crash Type	Rollover	21.76	78.69
Obstacle	Obstacle=ground	18.97	68.85
…			
failure mecanism	Panic	5.72	14.75

Figure 7: Example of accident cluster

In this example, *Crash_position=Offroad* is a relevant attribute (it provides a good characterization of the cluster) because its percentage in this cluster is 96.72%, which is significantly higher than its percentage in the studied sample (26.64%). However, *Failure_Mecanism=Panic* does not characterize this cluster because its percentage (14.75%) is too small.

Some expert rules were identified and formalized allowing an automatic use of statistical feature in order to identify relevant attributes. Hence, the time that users require to perform an interpretation task is sharply reduced. Each attribute characterizing a cluster has been classified according to the several concepts in the multi-view model. The attribute classification is expressed in XML language. The use of XML metadata combined with the use of XSL (eXtensible Stylesheet Language) allows performing a multi-view representation of each accident cluster (figure 8).

CVE Model	Driver	Vehicle	Environment	Driv/Veh	Driv/Env	Env/Veh
	failed_task Lane_change_manoeuv external_perturbation typacc=pilotabilite driver_distraction Drug failure_mec=Panic failure_typ=Action	rollover Obstacl=ground	atmosph=clear/normal Dry_road_surf Secondary_road No_Mask driver_distraction straight_line	Control_Probl Dry_road_surf loss_lateral_control	Guidance_infrastr Dry_road_surf Lane_change_manoeuv	Obstacl=ground Dry_road_surf Single_Vehicle

Sequential Model	Permanent State	Normal Driving step	Failure step	Emergency step	Crash step
		atmosph=clear/normal Dry_road_surf Secondary_road straight_line	Control_Probl Guidance_infrastr No_Mask failed_task Single_Vehicle Lane_change_manoeuv driver_distraction	failure_mec=Panic	Obstacl=ground Dry_road_surf rollover Obstacl=ground Offroad

Inform. process. Model	Perception	Diagnostic	Prognostic	Decision	Action	Global
	No_Mask			failure_mec=Panic	fondef=Action failed_task loss_lateral_control	driver_distraction Drug/font>

Task Model	Navigation	Guidage latéral	Guidance longit.	Control latéral	longit. Control
	No_Mask	Guidance_infrastr straight_line	Guidance_infrastr	loss_lateral_control rollover	

Figure 8: Example of Cluster Multi-view Projection

The same table can be represented according to the different identified models. The same attributes are represented according to the concepts they characterize in each model. For instance, the attribute *Failure_Mecanism=Panic* characterizes at the same time the concept *Driver* in the *CVE model* and the concept *Emergency step* in the S*equential Model*. The cluster presentation in figure 8 is semantically more relevant and useful than in figure 7. In addition, the user access three textual descriptions of the nearest accidents of the centre of the cluster. It allows scenarios interpretation according to the several models stemming from designers and accidentologists and thereby allows a valuable communication between the different users.

4. Conclusion

Nowadays large databases are created, due to the facilities to capture data. To mine adequate information from these databases, KDD process can be used. The complexity of the data, the domain and the researched knowledge need the expert knowledge integration.

For this purpose, a new approach is proposed, mixing top-down and bottom-up models of knowledge integration. The top-down method consists in the proposition of a framework based on a systemic approach allowing a multi-view model elaboration with a repartition of the attributes of the database in an architecture based on three abstraction levels (concepts, models and views).

The bottom-up approach is the integration of experts' knowledge for an appropriate attribute selection in the first step of a KDD process (data preparation). The results of the KDD process are then easier to interpret because an automatic projection on the different models is realised.

This approach is generic; it has been applied in the context of accidentology in order to facilitate Accident Scenarios elaboration, but may be used in other domain as marketing, maintenance, etc. For example, it is possible to elaborate clusters of failures from incident database.

Acknowledgements

This research is carried out in the LAB (Laboratory of Accident research, of Biomechanics and studies of the human behaviour), which is a shared laboratory between the two French car manufacturers, PSA (Peugeot-Citroën) and Renault.

REFERENCES

1. LEITE, J., DOORN, J., and GN, K., **A scenario construction process**, Requirements Engineering, 5: p. 38-61, 2000.
2. SCHERINGER, M. et al., **Scenario-based risk assessment of multi-use chemicals: application to solvents**, Risk Analysis, 21(3), p. 481-497, 2001.
3. CAROLL, J.M., **Scenario-based design: envisioning work and technology in system development**, New York, JOHN WILEY AND SONS Eds., 1995.
4. CAROLL, J.M., HELANDER, M., LANDAUER and T.K., PRABHU, P., **Scenario-Based Design**, in Handbook of Human-Computer Interaction, 2nd edition, Ch. 17, North-Holland: Amsterdam, 1998.
5. GANDON, F. and DIENG R., **Ontologie pour un système multi-agents dédié à une mémoire d'entreprise**, in Ingénierie des Connaissances, IC'2001, Grenoble, France, 2001.
6. JARKE, M., T. BUI, and CAROLL, J.M., **Scenario management: an interdisciplinary approach**, Requirements Engineering 3(4), p. 155–173, 1998.
7. BRENAC, T. and FLEURY, D., **Le concept de scénario type d'accident de la circulation et ses applications**, Recherche Transport Sécurité, 63: p. 63-77, 1999.
8. VAN ELSLANDE, P. and ALBERTON, L., **Scénarios-types de production de l'erreur humaine dans l'accident de la route, problématique et analyse qualitative**, INRETS, 1997.
9. BRENAC, T. and MEGHERBI, B., **Diagnostic de sécurité routière sur une ville : intérêt de l'analyse fine de procédures d'accidents tirées aléatoirement,** Recherche Transport Sécurité, 52: p. 59-71, 1996.
10. FLEURY, D. and BRENAC, T., **Accident prototypical scenarios, a tool for road safety research and diagnosis studies**, Accident Analysis & Prevention, 33, p. 267-276, 2001.
11. CHOVAN, J.D. et al., **Examination of intersection, left turn across path crashes and potential ivhs countermeasures,** National Highway Traffic Safety Administration, 1994.
12. NAJM, W.G., SMITH, J.D., and SMITH, D.L., **Analysis of Crossing Path Crashes**, National Highway Traffic Safety Administration: Washington D. C., 2001.
13. SOHN, S.Y. and SHIN, H.W., **Pattern recognition for road traffic accident severity in korea**. Ergonomics, 44(1), p. 107-117, 2001.
14. SOHN, S.Y. and LEE, S.H., **Data fusion, ensemble and clustering to improve the classification accuracy for the severity of road traffic accident in korea**, Safety Science, 41(1), p. 1-14, 2003.
15. PAGE, Y., **Elaboration de scénarios types d'accident pour le développement des systèmes de sécurité active embarqués dans les véhicules**, LAB PSA Peugeot-Citroën Renault, Nanterre, 2002.

16. PAGE, Y. et al. **Combination of statistical and case-by-case approach for accident situations classification**, in FISITA, Spain, 2004.
17. ASHBY, W.R., **An introduction to cybernetics**, C. HALL, London, 1965.
18. LE MOIGNE, J.L., **La théorie du système général**, Paris,. P. U. F. Ed., 1974.
19. VON FOERSTER, H., **The cybernetics of cybernetics** (2nd edition), MINNEAPOLIS: FUTURE SYSTEMS INC., 1995.
20. MILLER, J.G., **Living Systems**, UNIVERSITY PRESS OF COLORADO Ed., 1995.
21. LE MOIGNE, J.-L**., La modélisation des systèmes complexes,** DUNOD Ed., 1999.
22. VAN ELSLANDE, P. and ALBERTON, L., **Scénarios-types de production de l'erreur humaine dans l'accident de la route, problématique et analyse qualitative**, INRETS, 1997.
23. BRENAC, T., **L'analyse séquentielle de l'accident de la route: comment la mettre en pratique dans les diagnostics de sécurité routière, Outil et méthode**, INRETS, 1997.
24. FLEURY, D. and PAGE, Y., **Etudes détaillées d'accidents de la circulation routière**, PREDIT, 2002.
25. FAYYAD, U., PIATETSKY-SHAPIRO, G., and SMYTH P., **The KDD process for extracting useful knowledge from volumes of data**, Communications of the ACM, vol. 39, no. 11, 1996.
26. HAN, J. and KAMBER, M., **Data Mining: Concepts and Techniques**, MORGAN KAUFMANN PUBLISHERS Eds., ISBN 1-55860-489-8, 2000.
27. MACQUEEN, J.B., **Some methods for classification and analysis of multivariate observations**, in Proc. Symp. Math. Stat. and probability (5th), Univ. of California, Berkeley, 1967.