



Société Francophone de Classification

Bulletin numéro 20

Novembre 2006

Président : Y. Lechevallier

Vice-présidents : R. Verde et D. Zighed

Secrétaire : A. Hardy

Trésorier : B. Leclerc

Éditeur : F.-X. Jollois

Société Francophone de Classification, 61 rue de Bruxelles, 5000 Namur, Belgique

Correspondance : A. Hardy, Facultés Universitaires Notre-Dame de la Paix, 8 rempart de la vierge, 5000 Namur, Belgique

Sommaire

Mot du président	1
Assemblée générale	2
Parution	5
Annonces	6

Mot du président

Cher(e)s collègues et ami(e)s de la Classification,

Voici notre bulletin annuel qui a été réalisé par notre nouvel éditeur du bulletin, François-Xavier Jollois, qui a été élu cette année à Metz lors de l'Assemblée Générale. J'en profite pour féliciter nos deux élus : François-Xavier Jollois et André Hardy.

Notre nouveau éditeur du bulletin, François-Xavier Jollois, est un jeune Maître de Conférence à Paris 5. Il a soutenu sa thèse dans le domaine de la classification automatique et il a aussi obtenu le prix EGC en 2004. François-Xavier Jollois remplace François Brucker. J'en profite pour remercier François de son travail d'éditeur qui est l'élément clé de la vie de notre association et pour souligner sa ténacité dans ses nombreux rappels de nos textes afin de pouvoir éditer notre bulletin dans les temps impartis. André Hardy a été, à nouveau, élu au poste de secrétaire, nous pouvons le remercier notamment pour le travail qu'il a réalisé lors du renouvellement des statuts de

notre association et le féliciter de son élection comme *Publication Officer* au Conseil de l'IFCS.

Nos journées 2006 se sont déroulées à Metz et nous remercions Mohammed Nadif, président du comité d'organisation, et son équipe pour l'organisation impeccable et l'accueil qu'ils nous ont réservés et aussi souligner le travail du comité scientifique, présidé par Madame Le Thi Hoai An. Il est prévu qu'une sélection de papiers sera publiée dans la Revue RNTI (Revue des Nouvelles Technologies de l'Information). Sur les six conférenciers invités il y avait deux conférenciers européens et deux conférenciers appartenant à des pays francophones. Durant ces journées nous avons eu droit à un temps magnifique ce qui nous a permis d'apprécier les charmes de la ville de Metz.

Durant cette assemblée générale nous avons retenu Caserte (Italie) comme lieu des journées de la SFC en 2008. Aussi il faut se préparer à repartir à l'extérieur de l'hexagone dans la région de Naples au mois de juin 2008. Je me permets de vous rappeler que nos journées de la SFC se dérouleront à Paris à l'ENST au mois de septembre 2007.

La journée commune avec la SFdS, EGC, AFIA et INFORSID a été cette année sur le thème de la classification et elle s'est déroulée sur deux jours, les 20-21 mars 2006. L'intervenant de la SFC a été Hans-Hermann Bock et nous avons eu de nombreuses in-

terventions de doctorants sur le thème de la classification et, je l'espère, futurs membres de la SFC.

Grâce à notre trésorier, Bruno Leclerc, nous pouvons maintenant devenir membre affilié de la CSNA (Classification Society of North America) et ainsi recevoir le *Journal of Classification*. Actuellement la SFC est en discussion avec Springer-Verlag afin que les membres de la SFC puissent bénéficier d'une réduction importante sur l'abonnement au nouveau journal ADAC (Advances in Data Analysis and Classification).

Je voudrais terminer mon propos en vous rappelant que cette année il y a les élections des deux vice-présidents de la SFC, l'appel à candidature est maintenant lancé.

A tous et à toutes, à bientôt.

Yves Lechevallier

Assemblée Générale

Compte-rendu de l'Assemblée Générale de la S.F.C. Metz - 7 septembre 2006. L'Assemblée Générale est ouverte sous la présidence de Yves Lechevallier.

1. Mot du Président de la SFC (Yves Lechevallier)

Yves Lechevallier remercie, au nom de la SFC, toute l'équipe de Metz pour l'organisation des Rencontres SFC'2006.

2. Elections (Yves Lechevallier)

Après dépouillement, on constate les résultats suivants :

- François-Xavier JOLLOIS est élu *Editeur du Bulletin de la SFC*.
- Patrice BERTRAND est élu *Représentant de la SFC au Conseil de l'IFCS*.
- André HARDY est réélu *Secrétaire de la SFC*.

Le Président remercie François BRUCKER et Alain GUENOCHÉ pour l'excellent travail qu'ils ont fourni dans le cadre de leurs mandats.

3. Bilan de SFC'2006

Mohamed NADIF fait le point sur l'organisation de SFC'2006. 70 personnes sont présentes au congrès. 40 exposés sont programmés. Certains papiers seront publiés dans un numéro spécial de la revue RNTI. Mohamed NADIF remercie les sponsors pour leur aide financière. Le package *hébergement-repas* a bien fonctionné ; il permet de maintenir les frais d'inscription à un niveau le plus bas possible. Le Thi Hoai An remercie les conférenciers, ainsi que les membres du comité scientifique qui ont sélectionné les exposés pour le congrès et les papiers

pour le numéro spécial de la revue RNTI. Elle remercie plus particulièrement Mohamed NADIF et François-Xavier JOLLOIS pour l'organisation de ce congrès. Elle suggère qu'une publicité plus importante soit faite par la SFC auprès des chercheurs qui travaillent dans le domaine de l'optimisation et des applications de la classification. Le congrès de la SFC a traditionnellement lieu début septembre. D'autre part les dates pour la rentrée universitaire en France ont été avancées. On suggère que l'on réfléchisse sur les meilleures dates possibles pour l'organisation des futures Rencontres de la SFC.

Le Président remercie une fois encore très chaleureusement le comité d'organisation et le comité scientifique pour l'excellent travail fourni pour l'organisation de SFC'2006.

4. Rapport moral (André Hardy, Secrétaire de la SFC)

(a) Membres de la SFC

Voici l'évolution des membres de la SFC depuis 1994.

1994	1995	1996	1997
63	82	63	123

1998	1999	2000	2001
116	119	110	113

2002	2003	2004	2005	2006
118	118	125	100	81+...

Tous les membres en ordre de cotisation à la SFC reçoivent les Bulletins de la SFC et de l'IFCS.

(b) Bulletin de la SFC

L'éditeur est François BRUCKER. Le numéro 19 est paru en octobre 2005. Le prochain numéro paraîtra en octobre 2006.

(c) Rencontres de la SFC

Voici le nombre de participants aux différentes Rencontres de la SFC.

Lieu	Année	Nombres
Brest	1992	50
Tours	1994	80
Namur	1995	55
Vannes	1996	103
Lyon	1997	107
Montpellier	1998	105
Nancy	1999	82
Pointe-à-Pitre	2001	113
Toulouse	2002	120
Neuchâtel	2003	80
Bordeaux	2004	120
Montréal	2005	86
Metz	2006	70

(d) *SFC'2006*

Les 13^{èmes} Rencontres de la SFC ont lieu à Metz du 6 au 8 septembre 2006.

(e) *SFC'2007*

Les 14^{èmes} Rencontres de la SFC auront lieu à Paris (France) du 5 au 7 septembre 2007. Elles sont organisées par l'Ecole Nationale Supérieure des Télécommunications de Paris. Le Président du comité scientifique est Olivier HUDRY. Le comité local d'organisation comprend Irène CHARON, Georges HEBRAIL et Olivier HUDRY. Le site du congrès se trouve à l'adresse <http://www.enst.fr/SFC2007>

(f) *Contacts avec d'autres sociétés*

- La SFC est membre fondateur de l'IFCS ; des informations sur la SFC sont reprises dans chaque bulletin de l'IFCS ; trois membres de la SFC participent au Conseil de l'IFCS (Alain Guenoche, Jean-Paul Rasson et André Hardy). André Hardy a été élu *Publication Officer* de l'IFCS. Il fait donc aussi partie de l'Exécutif de l'IFCS.
- La SFC a des accords privilégiés avec la *Gesellschaft für Klassifikation* et avec l'*Associação Portuguesa de Classificação e Análise de Dados (CLAD)*.
- Une convention a été signée entre la SFC et la SFdS. Un membre de la SFC fait partie du Conseil de la SFdS (Yves Lechevallier). Un membre de la SFdS fait partie du Bureau élargi de la SFC (Christian Derquenne). Une cotisation conjointe est prévue depuis 1999.
- La SFC a parrainé, pour la sixième année consécutive, les *Journées francophones d'Extraction et de Gestion des Connaissances* (Lille, France, 17-20 janvier 2006).

- Une collaboration scientifique a lieu entre des associations françaises et francophones (AFIA, BDA, INFORSID, SFdS, SFC, ?).
 - La représentante de la SFC est Pascale KUNTZ.
 - Les deuxièmes Rencontres Inter-Associations ont eu lieu au Laboratoire ERIC, Université Lumière Lyon 2, les 20 et 21 mars 2006 sur le thème *La classification et ses applications*. Le conférencier SFC était Hans BOCK.
 - *CSNA (Classification Society of North America)* : une possibilité est donnée aux membres de la SFC de rejoindre la CSNA comme *membre affilié*, avec une réduction (10 dollars) sur la cotisation CSNA, et quelques avantages substantiels :
 - Le CD *Classification Literature Automated Search Service* qui inclus un certain nombre d'ouvrages numérisés.
 - L'abonnement au *Journal of Classification*, et la disponibilité en ligne de la totalité des anciens numéros de ce Journal.
 - La possibilité d'acheter des livres Springer avec un rabais de 20 %.
- 6 membres de la SFC se sont inscrits comme *affiliate* à la CSNA.

(g) *Site Web et envois électroniques*

La SFC est sur Internet. Le site est pris en charge par François Brucker (ENST Bretagne - Brest). Des liens ont été créés avec d'autres sites pouvant intéresser la SFC (IFCS, SFdS, ...). Cette année, plusieurs annonces ont été faites aux membres de la SFC par courrier électronique (congrès, annonces diverses, ...). L'adresse du site est <http://sfc.enst-bretagne.fr>

(h) *Divers*

- Habituellement, l'Assemblée Générale de la SFC se tient lors des rencontres de la SFC.
- Prochaines élections au Conseil d'Administration de la SFC : les mandats des Vice-Présidents arrivent à échéance.

Le rapport moral est approuvé à l'unanimité.

Deux suggestions sont faites

- Inclure dans le fascicule des résumés des communications, lors d'un congrès de la SFC, quelques informations sur la SFC (noms des membres du Conseil d'Administration de la SFC, adresse du site web, ...).
- Mettre les résumés des communications sur le site web de la SFC

5. Rapport financier (Bruno Leclerc)

Rapport Financier pour 2005-2006 (au 1er septembre 2006)

Aux journées de la S. F. C. 2005 à Montréal, le compte de la Société se montait à	10822,30
Il se monte au 1er septembre 2006 à	10170,54
Recettes :	
Cotisations 2004 perçues durant ou après les journées de Montréal :	16,00
Cotisations 2005 perçues durant ou après les journées de Montréal :	604,00
Cotisations perçues à ce jour au titre de l'année 2006 :	833,00
Cotisations perçues à ce jour au titre de l'année 2007 :	16,00
TOTAL	1469,00
Dépenses :	
Prix Simon Régner 2005 et bourses de voyage	1400,00
Dépenses pour SFC 2004 (remboursements et clôture) :	519,85
Frais de tenue de compte :	22,00
Contribution à l'IFCS pour les années 2005 et 2006	178,91
TOTAL	2120,76
Prévisions sur 2006	
Recettes :	
Cotisations en cours d'encaissement (factures) :	170,00
Régularisation de comptes avec la SFdS au titre de 2005 :	600,00

Le rapport financier est approuvé à l'unanimité.

6. Bulletin de la SFC et site internet

La rédaction du Bulletin de la SFC et la tenue du site web incombent maintenant à François-Xavier JOLLOIS. Deux suggestions sont faites :

- Mettre dans le bulletin des informations sur les thèses qui sont défendues.
- Publier deux numéros du Bulletin par an.

7. Cotisation SFC et cotisation conjointe SFC/SFds

On décide de ne pas modifier les montants des cotisations à la SFC pour 2007.

8. Rencontres Inter-Associations (Djamel ZIGHED)

Georges Hébrail a été l'initiateur de ces Rencontres. L'objectif est de promouvoir les échanges scientifiques entre les différentes associations proches dans les domaines de l'informatique et des mathématiques appliquées, notamment la statistique. Ces associations sont AFIA, ARIA, EGC, INFORSID, LMO, SFC, SFdS.

Les 2èmes Rencontres ont eu lieu à Lyon les 20 et 21 mars 2006. Elles ont regroupé une centaine de participants. Le site web est toujours disponible : <http://eric.univ-lyon2.fr/rias2006/>

Les prochaines rencontres auront lieu à Toulouse en 2007.

9. SFC'2007

Olivier HUDRY, ne pouvant pas participer à l'AG, a rencontré le Conseil d'Administration de la SFC ce midi, afin de communiquer les dernières informations sur les Rencontres SFC'2007.

Ces Rencontres se dérouleront à Paris du 5 au 7 septembre 2007. Elles sont organisées par Telecom Paris (Ecole Nationale Supérieure des Télécommunications de Paris). Olivier HUDRY présidera le comité scientifique. Le comité d'organisation est composé de Irène CHARON, Georges HEBRAIL et Olivier HUDRY. Il sera bientôt élargi. Des articles longs seront sélectionnés et paraîtront dans un numéro de la revue *Mathématiques et Sciences Humaines*. Le site du congrès se trouve à l'adresse <http://www.enst.fr/SFC2007>

10. SFC'2008

Rosanna VERDE présente la candidature de CASERTE, site de la deuxième Université de Naples. Ce congrès sera organisé en collaboration avec le Groupe de Classification et d'Analyse des Données (CLADAG) de la Société Italienne de Statistique. Les dates proposées sont les 11, 12, 13 juin ou les 16, 17, 18 juin 2008. Le comité d'organisation sera entre autres composé de Carlo LAURO, Maurizio VICHI et Rosanna VERDE.

Rosanna signale qu'il y a 3 ou 4 vols par jour de Paris vers Naples, et 2 vols entre Nice et Naples.

Un vote est organisé. La candidature de CASERTE est acceptée.

11. SFC'2009, 2010, ...

Deux candidatures ont été proposées pour 2009

- L'Ile de la Réunion, présentée par Henri RALAMBONDRAIN et Jean DIATTA
- Grenoble, présentée par Ahlame CHOUAKRIA

Pour les années suivantes on enregistre dès à présent les manifestations d'intérêt de Brest et de Nantes.

12. Divers

Maurizio VICHI signale qu'il souhaite que les membres de la SFC soumettent des papiers pour le nouveau Journal *Advances in Data Analysis and Classification*. Ce Journal pourrait d'ailleurs servir entre autres à publier les meilleurs articles de congrès (IFCS, ?). Il signale par ailleurs qu'un prix spécial pour ce nouveau Journal pourrait être proposé aux membres des Sociétés de Classification de l'IFCS.

André HARDY,
Secrétaire de la SFC

Bruno LECLERC,
Trésorier de la SFC

Parution

Matériaux pour l'histoire de notre temps

L'historien face à l'ordre informatique : classification et histoire (revue),
Nanterre, BDIC, Avril-Juin 2006
www.bdic.fr/librairie

Fuzzy Clustering and Its Applications

Sous la direction de Jose Valente De Oliveira et Witold Pedrycz, Wiley, 16 Fév 2007, 416 pages.

Divided into four sections the first section, *Fuzzy clustering-the primer*, covers the essentials of fuzzy clustering including motivation, basic algorithms, computing aspects, realisations, cluster validity assessment and ensuing interpretation of the results along with several representative areas of applications. Chapters presenting the underlying fundamentals i.e. the *Foundations* of fuzzy clustering follow this. The authors elaborate on the role of fuzzy sets in data analysis, discuss the principles of data organization, present fundamental algorithms and their augmentations, identify pitfalls and embark on the interpretation issues. Different paradigms of unsupervised learning along with so-called knowledge-based clustering and data organisation are also addressed. The third section *Algorithms and Computational Aspects* focuses on the algorithmic and computational augmentations of fuzzy clustering and demonstrates its effectiveness in highly dimensional problems, distributed problem solving and uncertainty handling. To conclude the book, *Applications and*

Case Studies is devoted to a series of applications in which fuzzy clustering plays a pivotal role. The primary intent is to discuss its role in the overall design process in various tasks of prediction, classification, control and modelling. Here it becomes highly instructive to highlight at which phase of the design clustering is of relevance, what role it plays and how the results facilitate further detailed development of models or enhance interpretation aspects.

Clustering Large And High Dimensional Data

de Jacob Kogan, Cambridge University Press, 30 Nov 2006, 224 pages.

There is a growing need for a more automated system of partitioning data sets into groups, or clusters. For example, digital libraries and the World Wide Web continue to grow exponentially, the ability to find useful information increasingly depends on the indexing infrastructure or search engine. Clustering techniques can be used to discover natural groups in data sets and to identify abstract structures that might reside there, without having any background knowledge of the characteristics of the data. Clustering has been used in a variety of areas, including computer vision, VLSI design, data mining, bioinformatics (gene expression analysis), and information retrieval, to name just a few. This book focuses on a few of the most important clustering algorithms, providing a detailed account of these major models in an information retrieval context. The beginning chapters introduce the classic algorithms in detail, while the later chapters describe clustering through divergences and show recent research for more advanced audiences.

Pattern Classification

de Richard O. Duda, Peter E. Hart et David G. Stork, Wiley-Interscience, 2nd edition, 21 Jui 2007, 168 pages.

The first edition, published in 1973, has become a classic reference in the field. Now with the second edition, readers will find information on key new topics such as neural networks and statistical pattern recognition, the theory of machine learning, and the theory of invariances. Also included are worked examples, comparisons between different methods, extensive graphics, expanded exercises and computer project topics.

Cluster And Classification Techniques for the Biosciences

Recent advances in experimental methods have resulted in the generation of enormous volumes of data across the life sciences. Hence clustering and classification techniques that were once predominantly the domain of ecologists are now being used more widely. This book provides an overview of these important data analysis methods, from long-established statistical methods to more recent machine learning techniques. It aims to provide a framework that will enable the reader to recognise the assumptions and constraints that are implicit in all such techniques. Important generic issues are discussed first and then the major families of algorithms are described. Throughout the focus is on explanation and understanding and readers are directed to other resources that provide additional mathematical rigour when it is required. Examples taken from across the whole of biology, including bioinformatics, are provided throughout the book to illustrate the key concepts and each technique's potential.

Annonces

Risques en Assurance-Finance et Modèles Statistiques

Cette conférence aura lieu à l'*Institut de Mathématiques Appliquées (IMA)*, Université Catholique de l'Ouest, les 15 et 16 novembre 2006, Angers, France
www.ima.uco.fr/assurance2006

Ce colloque rassemble des cadres des sociétés d'assurance ainsi que de réassurance et des universitaires intéressés par l'analyse et la prédiction du risque .

L'enjeu est de montrer que les méthodes statistiques sont des techniques d'aide à la décision utiles aux responsables de l'entreprise et contribuent donc à la définition de stratégies adéquates.

Plus particulièrement, les différentes conférences prévues sur le thème Data Mining montreront comment exploiter les richesses des grandes et très grandes bases de données.

Un autre thème abordé portera sur l'analyse des durées de vie et présentera l'utilisation des modèles de survie dans la construction de tables de mortalité et dans l'étude des durées de vie des contrats.

Enfin autour du thème de la dépendance stochastique, l'application des copules en assurance vie et dommage sera développée ainsi que la mesure de dépendance entre risques financiers, risques d'actifs et risques passifs.

— —

La 7ème édition de la conférence francophone **Extraction et Gestion des Connaissances** aura lieu à l'*Institut d'Informatique, Facultés Universitaires Notre Dame de la Paix*, Namur, Belgique, du 23 au 26 janvier 2007.

www.info.fundp.ac.be/egc2007

L'objectif de ces journées est de rassembler, d'une part les chercheurs des disciplines connexes (apprentissage, statistique et analyse de données, systèmes d'information et bases de données, ingénierie des connaissances, etc.), et d'autre part les spécialistes d'entreprises qui déploient des méthodes d'extraction et de gestion des connaissances, afin de contribuer à la formation d'une communauté scientifique dans le monde francophone autour de cette double thématique de l'extraction et de la gestion de connaissances.

Cette manifestation vise à stimuler le développement de la communauté EGC en offrant un lieu privilégié d'échanges entre chercheurs et spécialistes industriels, d'où devraient émerger les besoins et enjeux futurs de la discipline.

— —

GfKI 2007

La 31ème conférence de la société allemande de classification **Gesellschaft für Klassifikation** aura lieu à l'*Universität Albert-Ludwigs*, Freiburg, Allemagne, du 7 au 9 mars 2007.

www.informatik.uni-freiburg.de/gfki

The 31st Annual Conference of the German Classification Society - Gesellschaft für Klassifikation (GfKI), a member of the International Federation of Classification Societies (IFCS), focuses on data analysis, machine learning and their various applications.

The scientific program will include sessions with invited talks and contributed presentations from a broad range of topics. Special emphasis will be laid on interdisciplinary research and the interaction between theory and practice.

The program will be complemented by a doctoral workshop for PhD students working on the conference topics.

The conference language is English (and German).

Participants interested in presenting a talk are invited to submit an abstract. All abstracts will undergo a review process. Accepted abstracts will be distributed to the conference participants.

Full papers can be submitted shortly after the conference. They will be reviewed by at least two members of

the program committee and successful contributions will be published in the post-conference proceedings in the Springer series Studies in Classification, Data Analysis, and Knowledge Organization.

Find here the Call for Papers as a pdf file.

The German Classification Society grants an award for the best paper.

WDSA

Le premier atelier européen sur l'analyse de flux de données **European Workshop on Data Stream Analysis** se déroulera au *Belvedere di San Leucio*, Caserta, Italie, du 14 au 16 mars 2007.

www.jsda.unina2.it/wdsa/home.html

A growing number of applications in areas like networking, retail industry or sensor networks are dealing with a challenging type of data : data is produced over time in an unpredictable fashion, representing streams of network traffic, retail transactions or sensor-measured values. A key requirement of such applications is to continuously monitor and react to interesting phenomena occurring in the input streams. Streaming applications are usually characterized by transient relations, continuous queries, approximate answers and one-pass evaluation. These characteristics make them incompatible with several assumptions usually made in traditional databases as well as in statistical techniques. Indeed, simply storing the arriving data into a traditional database management system and manipulating the stored data is impossible.

This special European Workshop seeks to bring together researchers working on knowledge discovery problems in data streams. While the primary objective of this Workshop is to provide an avenue for dissemination of research results and works in progress, we also seek to discuss the future development of knowledge discovery in data streams, and how the research will benefit applications. We invite authors to submit their work that demonstrates current research and novel applications in this area.

SFdS

Les 39 èmes Journées de Statistique de la **Société Française de Statistique** auront lieu à l'*Université Catholique de l'Ouest*, Angers, France du 11 au 15 juin 2007.

www.uco.fr/jds2007

Les Journées de Statistique, organisées annuellement par la Société Française de Statistique (SFdS) depuis 1970, constituent le plus important congrès annuel des statisticiens français. Les JdS rassemblent ainsi tous les ans plus de 500 participants, chercheurs, enseignants et praticiens, français, francophones ou non.

Répondant à l'invitation de collègues francophones, les Journées de Statistique se sont aussi tenues à Bruxelles (2002, 1992, 1982), Fès (2000), Québec (1996), Neuchâtel (1994), et Lausanne (1987).

En 2007, et pour la première fois, cette manifestation, en sa 39ème édition, a lieu à Angers (Département du Maine-et-Loire, Région des Pays-de-Loire, France).

SFC 2007

Les XIVème Rencontres de la **Société Francophone de Classification** auront lieu à l'*École Nationale Supérieure des Télécommunications*, Paris, France, du 5 au 7 septembre 2007.

www.enst.fr/SFC2007

RIA 2007

Les 3èmes **Rencontres Inter-Associations** sur le thème de la **Recherche d'Information** auront lieu à Toulouse.

site à venir