Salvatore Ingrassia • Roberto Rocci
Maurizio Vichi
Editors

New Perspectives in Statistical Modeling and Data Analysis

Proceedings of the 7th Conference
of the Classification
and Data Analysis Group
of the Italian Statistical Society,
Catania, September 9-11, 2009

Springer

Editors
Prof. Salvatore Ingrassia
Università di Catania
Dipartimento Impresa
Culture e Società
Corso Italia 55
95129 Catania
Italy
s.ingrassia@unict.it

Prof. Roberto Rocci
Università di Roma "Tor Vergata"
Dipartimento SEFEMEQ
Via Columbia 2
00133 Roma
Italy
roberto.rocci@uniroma2.it

Prof. Maurizio Vichi
Università di Roma "La Sapienza"
Dipartimento di Statistica
Probabilità e Statistiche Applicate
Piazzale Aldo Moro 5
00185 Roma
Italy
maurizio.vichi@uniroma1.it

ISSN 1431-8814
ISBN 978-3-642-11362-8 e-ISBN 978-3-642-11363-5
DOI 10.1007/978-3-642-11363-5
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011930788

Cover design: deblik, Berlin

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This volume contains revised versions of selected papers presented at the seventh biannual meeting of the Classification and Data Analysis Group (CLADAG) of the Italian Statistical Society, organized by the Faculty of Economics at the University of Catania in September 2009.

The conference encompassed 150 presentations organized in 3 plenary talks and 42 sessions. Moreover, one tutorial on mixture modeling took place before the meeting. With 225 attendees from 11 countries, the conference provided an attractive interdisciplinary international forum for discussion and mutual exchange of knowledge. The topics of all plenary and specialized sessions were chosen to fit the mission of CLADAG which is to promote methodological, computational and applied research within the fields of Classification, Data Analysis and Multivariate Statistics.

The chapters in this volume were selected in a second peer review process after the conference. In addition to the fundamental areas of clustering and discrimination, multidimensional data analysis and data mining, the volume contains some chapters concerning data analysis and statistical modeling in areas like evaluation, economics, finance, environmental and medical sciences, industry and services.

We would like to express our gratitude to the members of the Scientific Program committee for their ability to attract interesting contributions. We wish also thank the session organizers for inviting speakers, the chairpersons and discussants of the sessions for their stimulating comments and suggestions. We are very grateful to the referees for their careful reviews of the submitted papers and for the time spent in this professional activity. We gratefully acknowledge the Faculty of Economics of the University of Catania and the Department of Economics and Quantitative Methods, the Department of Economics and Territory, the Department of Sociology and Social Sciences, and the Faculty of Political Sciences for financial support. We are also indebted to SAS, CEUR Foundation and the Foundation for Subsidiarity for their support. A special thank is due to the Local Organizing Committee for this well-organized conference.

Finally, we would like to thank Dr. Martina Bihn of Springer-Verlag, Heidelberg, for her support and dedication to the production of this volume.

Catania — *Salvatore Ingrassia*
Roma — *Roberto Rocci*
Roma — *Maurizio Vichi*
July 2010

Contents

Contributors

Giada Adelfio Dipartimento di Scienze Statistiche e Matematiche "S. Vianelli", Università di Palermo, Viale delle Scienze, ed. 13, 90128 Palermo (Italy), adelfio@unipa.it

Marco Alfò Dipartimento di Scienze Statistiche, Sapienza Università di Roma, Piazzale Aldo Moro, 5, 00185 Roma, Italy, marco.alfo@uniroma1.it

Angela Alibrandi Department of Economical, Financial, Social, Environmental, Statistical and Territorial Sciences, University of Messina, Via dei Verdi 27, 98122, Messina (Italy), aalibrandi@unime.it

Patrizia Ameli Dipartimento di Istituzioni Economiche e Finanziarie, Università di Macerata, Via Crescimbeni 20–21, 62100 Macerata (Italy), patrizia.ameli@unimc.it

Mouna Akacha Department of Statistics, University of Warwick, Coventry CV4 7AL Warwick (United Kingdom), m.akacha@warwick.ac.uk

Pietro Amenta Department of Social and Economic System Analysis, Università del Sannio, Via delle Puglie 82, 82100, Benevento (Italy), amenta@unisannio.it

Graziano Aretusi D.M.Q.T.E., University of G. d'Annunzio, Viale Pindaro 42, I-65127 Pescara (Italy), graz@inwind.it

Massimo Aria Department of Mathematics and Statistics, University of Naples Federico II, Via Cinthia 26, (Monte Sant'Angelo), 80126 Napoli (Italy), aria@unina.it

Anthony C. Atkinson Department of Statistics, The London School of Economics, London WC2A 2AE (United Kingdom), a.c.atkinson@lse.ac.uk

Antonio Balzanella Dipartimento di Strategie Aziendali e Metodologie Quantitative Facoltà di Economia, Seconda Università degli Studi di Napoli, Via del Setificio, 81100 Caserta (Italy), antonio.balzanella2@gmail.com

Alessandro Barbiero Department of Economics, Business and Statistics, Università degli Studi di Milano, Via Conservatorio 7, 20122 Milano (Italy), barbiero@unimi.it

Francesco Bartolucci Dipartimento di Economia, Finanza e Statistica, Università di Perugia, Via Alessandro Pascoli 20, 06123 Perugia (Italy), bart@stat.unipg.it

Salwa Benammou Faculté de Droit & des Sciences Economiques & Politiques, Université de Sousse, Cité Erriadh - 4023 Sousse (Tunisia), saloua.benammou@fdseps.rnu.tn

Matilde Bini Dipartimento di Economia, Università Europea di Roma, Via degli Aldobrandeschi 190, 00163 Roma (Italy), mbini@unier.it

Giovanna Boccuzzo Dipartimento di Scienze Statistiche, Università di Padova, Via Cesare Battisti 241, 35121 Padova (Italy), boccuzzo@stat.unipd.it

Elvio C. Bonafede Department of Statistics and Applied Economics "L.Lenti", Corso Strada Nuova 65, Pavia (Italy), concetto.bonafede@unipv.it

Riccardo Borgoni Dipartimento di Statistica, Università degli Studi di Milano - Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano (Italy), riccardo.borgoni@unimib.it

Giuseppe Bove Dipartimento di Scienze dell'Educazione, Università degli Studi Roma Tre, Via del Castro Pretorio 20, 00185 Roma (Italy), bove@uniroma3.it

Rainer Brüggemann Department Ecohydrology, Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Mueggelseedamm 310, D-12587 Berlin, (Germany), brg@igb-berlin.de

Francesca Bruno Department of Statistics "Paolo Fortunati", University of Bologna, Via delle Belle Arti 41, 40126 Bologna (Italy), francesca.bruno@unibo.it

Furio Camillo Dipartimento di Statistica, Università di Bologna, Via delle Belle Arti 41, 40126 Bologna (Italy), furio.camillo@unibo.it

Maurizio Carpita Department of Quantitative Methods, University of Brescia, Contrada S. Chiara 50, 25122 Brescia (Italy), carpita@eco.unibs.it

Paola Cerchiello Department of Statistics and Applied Economics "L.Lenti", Corso Strada Nuova 65, Pavia (Italy), paola.cerchiello@unipv.it

Andrea Cerioli Dipartimento di Economia, Sezione di Statistica e Informatica, Università di Parma Via Kennedy 6, 43100 Parma Italy, andrea.cerioli@unipr.it

Marcello Chiodi Dipartimento di Scienze Statistiche e Matematiche "S. Vianelli", Università di Palermo, Viale delle Scienze, ed. 13, 90128 Palermo (Italy), chiodi@unipa.it

Daniela Cocchi Department of Statistics "Paolo Fortunati", University of Bologna, Via delle Belle Arti 41, 40126 Bologna (Italy), daniela.cocchi@unibo.it

Claudio Conversano Dipartimento di Economia, Università di Cagliari, Viale Frà Ignazio 17, I-09123 Cagliari (Italy), conversa@unica.it

Pietro Coretto Dipartimento di Scienze Economiche e Statistiche, Università di Salerno, Via Ponte Don Melillo, 84084 Fisciano (Italy), pcoretto@unisa.it

Anthony Cossari Department of Economics and Statistics, University of Calabria, Via Bucci CUBO 0C, 87036 Rende Cosenza (Italy), a.cossari@unical.it

Michele Costa Dipartimento di Scienze Statistiche "P. Fortunati", Università di Bologna, Via delle Belle Arti 41, 40126 Bologna (Italy), michele.costa@unibo.it

Giuseppina D. Costanzo Dipartimento di Economia e Statistica, Università della Calabria, Via P. Bucci, 87036 Arcavacata di Rende, Cosenza (Italy), dm.costanzo@unical.it

Valentina Cozza Dipartimento di Statistica e matematica per la ricerca economica, Università di Napoli "Parthenope", Via Medina 40 I Piano, 80133 Napoli (Italy), valecozza@unina.it

Carles M. Cuadras Department of Statistics, University of Barcelona, Diagonal 645, 08028 Barcelona (Spain), ccuadras@ub.edu

Carmen Cutugno Dipartimento di Impresa Culture e Società, Università di Catania, Corso Italia 55, 95129 Catania (Italy), carmen.cutugno@unict.it

Rosario D'Agata University of Catania, Via Vittorio Emanuele II 8, 95129 Catania (Italy), rodagata@unict.it

Luigi D'Ambra Department of Biological Sciences, Università di Napoli Federico II, Via Mezzocannone 8, 80134 Napoli (Italy), dambra@unina.it

Antonio D'Ambrosio Department of Mathematics and Statistics, University of Naples Federico II, Via Cinthia 26 (Monte S. Angelo), 80125 Napoli (Italy), antdambr@unina.it

Ida D'Attoma Dipartimento di Statistica, Università di Bologna, Via delle Belle Arti 41, 40126 Bologna (Italy), Ida.dattoma2@unibo.it

Cristina Davino Dipartimento di Studi sullo Sviluppo Economico, Università di Macerata, Via Piaggia della Torre 8, 62100 Macerata (Italy), cdavino@unimc.it

Luca De Angelis Dipartimento di Scienze Statistiche "P. Fortunati", Università di Bologna, Via delle Belle Arti 41, 40126 Bologna (Italy), l.deangelis@unibo.it

Angela De Sanctis D.M.Q.T.E., University of Pescara G. D'Annunzio, Viale Pindaro 42, I-65127, Pescara (Italy), a.desanctis@unich.it

Alfredo Del Monte Dipartimento di Economia, Università di Napoli Federico II, Via Cinthia 26 (Monte S. Angelo), 80125 Napoli (Italy), delmonte@unina.it

Laura Deldossi Dipartimento di Scienze Statistiche, Università Cattolica Del Sacro Cuore di Milano, Largo A. Gemelli 1, 20123 Milano (Italy), laura.deldossi@unicatt.it

Francesco Dell'Accio Dipartimento di Matematica, Università della Calabria, Via P. Bucci, 87036 Arcavacata di Rende, Cosenza (Italy), fdellacc@unical.it

Maria R. D'Esposito Dipartimento di Scienze Economiche e Statistiche, Università di Salerno, Via Ponte Don Melillo, 84084 Fisciano (Italy), mdesposi@unisa.it

Tonio Di Battista D.M.Q.T.E., University of Pescara G. d'Annunzio, Viale Pindaro 42, I-65127, Pescara (Italy), dibattis@unich.it

Marina Di Giacomo Dipartimento di Scienze Economiche e Finanziarie "G. Prato", Università di Torino, Corso Unione Sovietica 218 bis, 10134 Torino (Italy)

and

HERMES (Center for Research on Law and Economics of Regulated Services), Collegio Carlo Alberto, Via Real Collegio 30, 10024 Moncalieri (TO), Italy, digiacomo@econ.unito.it

Carlo Drago Dipartimento di Matematica e Statistica, Università di Napoli Federico II, Via Cinthia 26 (Monte S. Angelo), 80125 Napoli (Italy), drago@unina.it

Luigi Fabbris Dipartimento di Scienze Statistiche, Università di Padova, Via Cesare Battisti 241, 35121 Padova (Italy), luigi.fabbris@unipd.it

Marco Fattore Dipartimento di Metodi Quantitativi per le Scienze Economiche ed Aziendali, Università degli Studi di Milano - Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano (Italy), marco.fattore@unimib.it

Pier A. Ferrari Department of Economics, Business and Statistics, Università degli Studi di Milano, Via Conservatorio 7, 20122 Milano (Italy), pieralda.ferrari@unimi.it

Silvia Figini Department of Statistics and Applied Economics "L. Lenti", University of Pavia, Via Strada Nuova 65, 27100 Pavia (Italy), silvia.figini@unipv.it

Thaìs C.O. Fonseca Department of Statistics, University of Warwick, Coventry CV4 7AL Warwick, (United Kingdom), t.c.o.fonseca@warwick.ac.uk

Lara Fontanella D.M.Q.T.E., University of Pescara G. D'Annunzio, Viale Pindaro 42, I-65127 Pescara (Italy), lfontan@dmqte.unich.it

Antonio Forcina Dipartimento di Economia, Finanza e Statistica, Università di Perugia, Via Pascoli 10, 06100 Perugia (Italy), forcina@stat.unipg.it

Paolo Frumento Scuola Superiore S. Anna di Pisa, Piazza Martiri della Libertà 33, 56127 Pisa (Italy), p.frumento@sssup.it

Giuliano Galimberti Dipartimento di Scienze Statistiche "P. Fortunati", Università di Bologna, Via delle Belle Arti 41, 40126 Bologna (Italy), giuliano.galimberti@unibo.it

Romana Gargano Department of Economical, Financial, Social, Environmental, Statistical and Territorial Sciences, University of Messina, Via dei Verdi 27, 98122, Messina (Italy), rgargano@unime.it

Stefano A. Gattone Department SEFeMEQ, University of Tor Vergata, Rome, Via della Ricerca Scientifica 1, 00133, Roma (Italy), gattone@economia.uniroma2.it

Gianfranco Giordano Department of Mathematics and Statistics, University of Naples Federico II, Via Cinthia 26, (Monte Sant'Angelo), 80126 Napoli (Italy), gianfranco.giordano@unina.it

Giuseppe Giordano Dipartimento di Scienze Economiche e Statistiche, Università di Salerno, Via Ponte Don Melillo, 84084 Fisciano (Italy), ggiordan@unisa.it

Giorgia Giovannetti Dipartimento di Economia, Università di Firenze, Via delle Pandette 6 (FI) and European University Institute, Badia Fiesolana, Via dei Roccettini 9, I-50014 San Domenico di Fiesole (FI), Italy, giorgia.giovannetti@eui.it

Beatriz Goitisolo Department of Applied Economics III, University of Basque Country, Bilbao, Spain, Beatriz.Goitisolo@ehu.es

Jan Graffelman Universitat Politècnica de Catalunya, Department of Statistics and Operations Research, Av. Diagonal, 647, 6th floor, 08028 Barcelona, Spain, jan.graffelman@upc.edu

Luigi Grossi Università di Verona, Via dell'Artigliere 19, 37129 Verona (Italy), luigi.grossi@univr.it

Luigi Ippoliti D.M.Q.T.E., University of Pescara G. D'Annunzio, Viale Pindaro 42, I-65127 Pescara (Italy), ippoliti@unich.it

Ron S. Kenett KPA Ltd., Raanana, Israel and University of Torino, Torino (Italy), ron@kpa.co.il

Deborah Koltai Bar Ilan University, Department of Statistics, Israel, deborah.koltai@yahoo.com

Don Kristt Unit of Molecular Pathology, Laboratory of Histocompatibility and Immunogenetics, Rabin Medical Center, Petach Tikvah, Israel, pdkristt@gmail.com

Michele La Rocca Dipartimento di Scienze Economiche e Statistiche, Università di Salerno, Via Ponte Don Melillo, 84084 Fisciano (Italy), larocca@unisa.it

Fabrizio Laurini Dipartimento di Economia, Università di Parma, Via Kennedy 6, 43100 Parma (Italy), fabrizio.laurini@unipr.it

Yves Lechevallier INRIA, 78153 Le Chesnay cedex, (France), Yves.Lechevallier@inria.fr

Silvia Liverani Department of Mathematics, University of Bristol, (United Kingdom), s.liverani@bris.ac.uk

Marica Manisera Department of Quantitative Methods, University of Brescia, Contrada S. Chiara 50, 25122 Brescia (Italy), manisera@eco.unibs.it

Giancarlo Manzi Department of Economics, Business and Statistics, Università degli Studi di Milano, Via Conservatorio 7, 20122 Milano (Italy), giancarlo.manzi@unimi.it

Donata Marasini Dipartimento di Statistica, Università degli Studi di Milano - Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano (Italy), donata.marasini@unimib.it

Giovanni M. Marchetti Dipartimento di Statistica "G. Parenti", Università di Firenze, Viale Morgagni, 59, 50134 Firenze (Italy), giovanni.marchetti@ds.unifi.it

Paolo Mariani Dipartimento di Statistica, Università di Milano - Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano (Italy), paolo.mariani@unimib.it

Francesca Martella Dipartimento di Scienze Statistiche Sapienza Università di Roma, P.le A. Moro, 5 00185 Roma, Italia, francesca.martella@uniroma1.it

Maria C. Martini Department of Social, Cognitive and Quantitative Sciences, Università degli studi di Modena e Reggio Emilia, Viale Allegri 9, 42121 Reggio Emilia (Italy), mariacristiana.martini@unimore.it

Angelo Mazza Dipartimento di Impresa, Culture e Società, Università di Catania, Corso Italia 55, 95129 Catania (Italy), a.mazza@unict.it

Intissar Mdimagh Institut Superieur de Gestion, Universitè de Sousse, Rue Abedelaziz El Bahi - B.P. 763, 4000 Sousse, (Tunisia), mdimagh@yahoo.fr

Fabrizia Mealli Dipartimento di Statistica "G. Parenti", Università degli Studi di Firenze, Viale G.B. Morgagni 59, 50134 Firenze (Italy), mealli@ds.unifi.it

Arcangelo Merla Clinical Sciences and Bioimaging Department; Institute of Advanced Biomedical Technologies, Foundation University of Pescara G. D'Annunzio, a.merla@itab.unich.it

Simona C. Minotti Dipartimento di Statistica, Università degli Studi di Milano - Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano (Italy), simona.minotti@unimib.it

Isabella Morlini Department of Economics, University of Modena, Viale Berengario 51, 41121 Modena (Italy), imorlini@unimore.it

Mauro Mussini Dipartimento di Statistica, Università di Milano - Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano (Italy), mauro.mussini1@unimib.it

Miki Nakai Department of Social Sciences, College of Social Sciences, Ritsumeikan University, 56-I Toji-in kitamachi, Kyoto 603-8577, (Japan), mnakai@ss.ritsumei.ac.jp

Federico Niccolini Dipartimento di Istituzioni Economiche e Finanziarie, Università di Macerata, Via Crescimbeni 20 – 1, 62100 Macerata (Italy), fniccolini@unimc.it

Luciano Nieddu LUSPIO, Via C. Colombo, 200, 00145 Roma, Italy, l.nieddu@luspio.it

Akinori Okada Graduate School of Management and Information Sciences, Tama University, 4-1-1 Hijirigaoka Tama city Tokyo 206-0022, okada@tama.ac.jp

Jan Owsiński Department of Systems Methods Applications, Systems Research Institute, Polish Academy of Sciences, Jan.Owsinski@ibspan.waw.pl

Barbara Pacini Dipartimento di Statistica e Matematica, Università di Pisa, Via C. Ridolfi 10, 56124 Pisa (Italy), barbara.pacini@sp.unipi.it

Francesco Palumbo Dipartimento di Scienze Relazionali "Gustavo Iacono", Università di Napoli Federico II, Via Porta di Massa 1, 80133 Napoli (Italy), fpalumbo@unina.it

Fulvia Pennoni Dipartimento di Statistica, Università degli Studi di Milano-Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano (Italy), fulvia.pennoni@unimib.it

Domenico Perrotta European Commission, Joint Research Centre, Institute for the Protection and Security of the Citizen, Via E. Fermi 1, 21020, Ispra (Italy), domenico.perrotta@ec.europa.eu

Alessandra Petrucci Dipartimento di Statistica "G. Parenti", Università di Firenze, Viale Moragni 59, 50134 Firenze (Italy), alex@ds.unifi.it

Massimiliano Piacenza Dipartimento di Scienze Economiche e Finanziarie "G. Prato", università di Torino, Corso Unione Sovietica 218 bis, 10134 Torino (Italy)

and

HERMES (Center for Research on Law and Economics of Regulated Services), Collegio Carlo Alberto, Via Real Collegio 30, 10024 Moncalieri (TO), Italy, piacenza@econ.unito.it

Marilena Pillati Dipartimento di Scienze Statistiche "P. Fortunati", Università di Bologna, Via delle Belle Arti 41, 40126 Bologna (Italy), marilena.pillati@unibo.it

Antonella Plaia Department of Statistical and Mathematical Sciences "S. Vianelli", University of Palermo, Viale delle Scienze 13, 90128, Palermo (Italy), plaia@unipa.it

Mariano Porcu Dipartimento di Ricerche Economiche e Sociali, Università di Cagliari, Viale S. Ignazio 78, 09123 Cagliari (Italy), mrporcu@unica.it

Antonio Punzo Dipartimento di Impresa, Culture e Società, Università di Catania, Corso Italia 55, 95129 Catania (Italy), antonio.punzo@unict.it

Piero Quatto Dipartimento di Statistica, Università degli Studi di Milano - Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano (Italy), piero.quatto@unimib.it

Marco Riani Dipartimento di Economia, Sezione di Statistica e Informatica, Università di Parma Via Kennedy 6, 43100 Parma Italy, mriani@unipr.it

Emilia Rocco Dipartimento di Statistica "G. Parenti", Università di Firenze, Viale Morgagni 59, 50134 Firenze (Italy), rocco@ds.unifi.it

Giulia Roli Dipartimento di Scienze Statistiche "P. Fortunati", Università di Bologna, Via Zamboni 33, 40126 Bologna (Italy), g.roli@unibo.it

Antonio A. Romano Dipartimento di Statistica e Matematica per la Ricerca economica, Università di Napoli "Parthenope", Via Medina 40 I Piano, 80133 Napoli (Italy), antonio.romano@uniparthenope.it

Elvira Romano Dipartimento di Studi Europei e Mediterranei, Seconda Università degli studi di Napoli, Via del Setificio 15, 81100 Caserta (Italy), elvira.romano@unina2.it

Mariantonietta Ruggieri Department of Statistical and Mathematical Sciences "S. Vianelli", University of Palermo, Viale delle Scienze 13, 90128, Palermo (Italy), mariantonietta.ruggieri@unipa.it

Silvia Salini Dipartimento di Scienze Economiche Aziendali e Statistiche, Università degli studi di Milano, Via Conservatorio 7, 20122 Milano (Italy), silvia.salini@unimi.it

Giuseppe Scandurra Dipartimento di Statistica e Matematica per la Ricerca Economica, Università di Napoli "Parthenope", Via Medina 40 I Piano, 80133 Napoli (Italy), giuseppe.scandurra@uniparthenope.it

Germana Scepi Dipartimento di Matematica e Statistica, Università di Napoli Federico II, Via Cinthia 26 (Monte S. Angelo), 80125 Napoli (Italy), scepi@unina.it

Manuela Scioni Dipartimento di Scienze Statistiche, Università di Padova, Via Cesare Battisti 241, 35121 Padova (Italy), scioni@stat.unipd.it

Luca Scrucca Dipartimento di Economia, Finanza e Statistica, Università degli Studi di Perugia, Via Alessandro Pascoli 20, 06123 Perugia (Italy), luca@stat.unipg.it

Biagio Simonetti Department of Social and Economic System Analysis, Università del Sannio, Via delle Puglie 82, 82100, Benevento (Italy), simonetti@unisannio.it

Gabriele Soffritti Dipartimento di Scienze Statistiche "P. Fortunati", Università di Bologna, Via delle Belle Arti 41, 40126 Bologna (Italy), gabriele.soffritti@unibo.it

Nadia Solaro Dipartimento di statistica, Università di Milano - Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milano (Italy), nadia.solaro@unimib.it

Giuseppe Storti Dipartimento di Scienze Economiche e Statistiche, Università di Salerno, Via Ponte Don Melillo, 84084 Fisciano (Italy), storti@unisa.it

Isabella Sulis Dipartimento di Ricerche Economiche e Sociali, Università di Cagliari, Viale S. Ignazio 78, 09123 Cagliari (Italy), isulis@unica.it

Nicola Tedesco Dipartimento di Ricerche Economiche e Sociali, Università di Cagliari, Viale S. Ignazio 78, 09123 Cagliari (Italy), tedesco@unica.it

Venera Tomaselli University of Catania, Via Vittorio Emanuele II 8, 95129 Catania (Italy), tomavene@unict.it

Francesca Torti University of Parma, Faculty of Economics, Via J.F. Kennedy 6, 43100 Parma (Italy), francesca.torti@nemo.unipr.it

Giulio Trombetta Dipartimento di Matematica, UNICAL, Via P. Bucci, 87036 Arcavacata di Rende, Cosenza (Italy), trombetta@unical.it

Gilberto Turati Dipartimento di Scienze Economiche e Finanziarie "G. Prato", Università di Torino, Corso Unione Sovietica 218 bis, 10134 Torino (Italy)

and

HERMES (Center for Research on Law and Economics of Regulated Services), Collegio Carlo Alberto, Via Real Collegio 30, 10024 Moncalieri (TO), Italy turati@econ.unito.it

Valerio A. Tutore Department of Mathematics and Statistics, University of Naples Federico II, v.tutore@unina.it

Margherita Velucchi Dipartimento di Statistica "G. Parenti", Università di Firenze, Viale G.B. Morgagni, 59 (FI), Italy, velucchi@ds.unifi.it

Rosanna Verde Dipartimento di Studi Europei e Mediterranei, Seconda Università degli studi di Napoli, Via del Setificio 15, 81100 Caserta (Italy), rosanna.verde@unina2.it

Marika Vezzoli Department of Quantitative Methods, University of Brescia, Contrada Santa Chiara 50, 25122 Brescia (Italy), vezzoli@eco.unibs.it

Donatella Vicari Dipartimento di Scienze Statistiche, Sapienza Università di Roma, Piazzale Aldo Moro, 5, 00185 Roma, Italy, donatella.vicari@uniroma1.it

Domenico Vistocco Dipartimento di Scienze Economiche, Università di Cassino, Via Marconi 10, 03043 Cassino (Italy), vistocco@unicas.it

Maria P. Vitale Dipartimento di Scienze Economiche e Statistiche, Università di Salerno, Via Ponte Don Melillo, 84084 Fisciano (Italy), mvitale@unisa.it

Sergio Zani Department of Economics, University of Parma, Via Kennedy 6, 43100 Parma (Italy), sergio.zani@unipr.it

Diego Zappa Dipartimento di Scienze Statistiche, Università Cattolica Del Sacro Cuore di Milano, Largo A. Gemelli 1, 20123 Milano (Italy), diego.zappa@unicatt.it

Amaya Zárraga Department of Applied Economics III. University of Basque Country. Bilbao. Spain. Amaya.Zarraga@ehu.es

Emma Zavarrone IULM. Institute of Consumer, Behaviour and Corporate Communication. Via Carlo Bo 1. 20143 Milano (Italy). emma.zavarrone@iulm.it

Latent Ties Identification in Inter-Firms Social Networks

Patrizia Ameli, Federico Niccolini, and Francesco Palumbo

Abstract Social networks are usually analyzed through manifest variables. However there are social latent aspects that strongly qualify such networks. This paper aims to propose a statistical methodology to identify latent variables in inter-firm social networks. A multidimensional scaling technique is proposed to measure a latent variable as a combination of an appropriate set of two or more manifest relational aspects. This method, tested on an inter-firm social network in the Marche region (Italy), it is a new way to grasp social aspect with quantitative tools that could be implemented under several different conditions, using also other variables.

1 Introduction

In the last few years, the synergic integration between the organizational and statistical disciplines has been producing relevant scientific outputs. However this cross-disciplinary synergy can bring to new important evolutions, especially in the Social Network Analysis (SNA).

SNA studies the social resources exchange between actors and their relationships within a social system (Wasserman and Faust 1994). Ties measurement represents a relevant concern of SNA. Some authors postulate the existence of a latent 'social space' within which the presence of a tie between two actors is determined as function of some measures of [dis]similarity between the latent space positions of these actors. In this direction Leydesdorff et al. (2008) and Okada (2010) propose to use scaling models to determine the actors latent space positions and consequently the ties between actors.

In this framework, the paper aims to propose a statistical procedure to study ties in the latent space positions where actors are firms and ties are function of two or more measures of their relationships (Hoff et al. 2002).

As matter of fact, in the organizational science, when focusing on inter-firm social network many authors remark the complex nature of the social relationships. Grandori and Soda (1995) define the inter-firm network as 'a mode of regulating interdependence between firms, which is different from the aggregation of these

S. Ingrassia et al. (eds.), *New Perspectives in Statistical Modeling and Data Analysis*, Studies in Classification, Data Analysis, and Knowledge Organization,
DOI 10.1007/978-3-642-11363-5_21,

units within a single firm and from coordination through market signals, and which is based on a co-operative game with partner-specific communication'. Particularly, social inter-firm networks can be defined as long term informal relationships among two or more organizations.

Nowadays in the emerging inter-organizational architectures, such as learning and visionary networks, relationships are built and planned more focusing on social basis instead than on formal agreements. In this context, partnerships' social latent aspects become the main ground for knowledge sharing and for long term performance.

Members of inter-firms social networks aim at sharing core values, knowledge, and, in some cases, they share also a vision. Topics like social capital, cohesiveness, and the embeddedness have a strong social component, that is not directly measurable but that can be identified through the observation of their manifest variables. Consequently the key research question concern the possibility to better study and quantify deep latent aspects of a social network, using manifest variables.

In social inter-firms networks typical examples of manifest measures are *trust*, *frequency of exchanges* among parties, and *reciprocity* (Jones et al. 1997). Many scholars indicate *trust* as fundamental ingredient for every typology of network or alliance and as facilitator of knowledge transmission (Nonaka 1991) and vision sharing. To have trust relationship is necessary to built it in a daily behavior (Lomi 1997, p. 214).

The *frequency* concerns how often specific parties exchange with one another (Jones et al. 1997, p. 917). Frequency is important for three reasons. First, frequency facilitates transferring tacit knowledge, second, frequent interactions establish the conditions for relational and structural embeddedness, which provide the foundation for social mechanisms to adapt, coordinate, and safeguard exchanges effectively and third, frequent interactions provide cost efficiency in using specialized governance structures (Williamson 1985, p. 60; Jones et al. 1997, p. 917).

Reciprocity 'transforms a unilateral supply relationship into a bilateral one' (Williamson 1985, p. 191; Jones et al. 1997, p. 922) and creates the perception of a similar 'destiny' with greater 'mutual interest' (Williamson 1985, p. 155; Jones et al. 1997, p. 922).

This paper proposes that the latent space can be defined as 'basic social relational embeddedness', and it is obtained as a function of the three chosen manifest variables: trust, frequency of exchanges among parties, and reciprocity. The basic social embeddedness is a component of the strength of the tie (Granovetter 1973), that is especially important for knowledge and vision sharing (Uzzi 1999).

The basic social embeddedness is identified in this work as the core part of the 'relational embeddedness': a component of the strength of the tie (Granovetter 1973), that is especially important for knowledge and vision sharing (Uzzi 1999). Relational embeddedness is an indicant of the motivational aspect of tie strength (Rindfleisch and Moorman 2001), essentially refers to the quality and depth of a single dyadic tie, it is fundamental mechanism of social governance, and captures the quality of dyadic exchanges and the behaviours exchange parties exhibit, such as trust, confiding, and information sharing (Jones et al. 1997, p. 924).

2 The Unique Social-Relational Variable

Scholars are familiar with the idea of *latent variable* in both organisational and in behavioral sciences. Latent variables refer to concepts that cannot be directly measured and are opposed to the observable variables. However, latent variables may be defined on the basis of properly identified sets of observable variables, called also manifest variables (Bartholomew 1987).

In the SNA framework, the network graphical visualization implies the identification of a *metric* space where actors and ties are represented. Generally the strength of the ties are represented in terms of distance between two actors or/and by the ties weight. In some cases suitable strength thresholds are defined: when strengths are lower than the threshold, the corresponding tie is omitted. However, the approach presented in the present paper assumes that ties are function of two or more measures and they only exist in latent space and not otherwise (Hoff et al. 2002).

The proposed procedure can be summarized in the following steps: (a) according to a set of two or more measures of relationships, actors are displayed in a latent metric space; (b) ties between actors depend on the distances between them in the latent (unobserved) space.

Formally we define the following data structure. Given a set of n statistical units and K manifest measures (variables), the notation δ_{ijk} indicates the general proximity measure between the statistical units i and j for the measure, which are arranged in K asymmetric $n \times n$ matrices $\{\Delta_1, \Delta_2, \ldots, \Delta_k, \ldots, \Delta_K\}$, where $(i, j) = 1, \ldots, n$. We remember that $\delta_{ijk} = \max, \forall k$ and if $i = j$ (diagonal elements), and that $\delta_{ijk} \geq 0$ and the value 0 indicates absence of any relationship, by definition.

The final aim of the paper is to identify and visualize ties in the latent space of the inter-firm social network; from a technical point of view, the problem consists in finding a good approximation of the 3-way proximity matrix by a $n \times n$ distance matrix.

2.1 *Multidimensional Scaling for Ties Identification*

This section shortly presents the MultiDimensional Scaling (MDS) basic principles and then it motivates the choice of the PROXSCAL model with respect to other scaling models. For sake of space, we do not go to deepen the PROXSCAL; interested readers can refer more specifically to Borg and Groenen (2005) for the MDS foundations and to Meulman and Heiser (2001) for the PROXSCAL model, more specifically.

The MDS aims at finding a configuration of n points into a p dimensional space. Generally, p is set equal to 2 in order to get graphic representation of the n points (Takane 2007). The scaling transformation of the proximity matrix has two advantages: (1) to summarize different proximity measures into one single distance; (2) to permit distances graphical representations into 2 and 3 dimensional spaces.

Moreover, it is worth noticing that the transformed relationship measures into metric measures can be displayed according to different approaches to visualize the network.

In the simplest two dimensional case, where $K = 1$, the MDS aims at defining a function $\varphi(\cdot)$ such that:

$$\varphi\left(\delta_{ij}\right) = d(x_i, x_j) + \epsilon_{ij} \tag{1}$$

In other words, $\varphi(\cdot)$ indicates a function that maps the dissimilarities δ_{ij} into a metric space, where the distance $d(\cdot)$ (generally the Euclidean distance) is defined. The quantity ϵ_{ij} indicates the residual.

Some alternative models have been proposed to deal with 3-Way data structures. In this paper we refer to the INDividual SCALing (INDSCAL) model, and more specifically to the PROXSCAL algorithm for 3Way dis(similarities) data structures. INDSCAL model is also referred to as weighted Euclidean model. Starting from K proximity matrices $\Delta_1, \Delta_2, \ldots, \Delta_k, \ldots, \Delta_K$, INDSCAL minimizes the Stress (squared Euclidean distances) defined by the following equation:

$$\sigma(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_k, \ldots, \mathbf{X}_K) = \sum_k \sum_{i<j} \left(\delta_{ijk} - d_{ij}(\mathbf{G}\mathbf{W}_k)\right)^2, \tag{2}$$

where $\mathbf{G}$ indicates the common space $n \times n$ matrix and $\mathbf{W}_k$ represents the generic weighting matrix. The point coordinates are then defined as $\mathbf{X}_k = \mathbf{G}\mathbf{W}_k$ and the scaled distances as $d_{ij}(\mathbf{X}_k)$.

Differently from INDISCAL, PROXSCAL minimizes Normalized Stress σ_n:

$$\sigma_n(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_k, \ldots, \mathbf{X}_K) = \frac{\sum_k \sum_{i<j} \left(\delta_{ijk} - d_{ij}(\mathbf{X}_k)\right)^2}{\sum_k \sum_{i<j} \delta_{ijk}^2}. \tag{3}$$

The main point in favor of PROXSCAL is that it works on the Euclidean distance and not on the squared distances. Avoiding the square transformation, it prevents putting more emphasis on large dissimilarities. The second, but not less important point, is the possibility to consider asymmetric (dis)similarity matrices. However, it is worth noticing that dissimilarity matrices are transformed in symmetric ones. Alternative scaling models could be taken into account, and this surely represents a research direction for future works. Last, the normalized stress is a relative measure that allows us to appreciate the overall quality of the solution and to make comparisons among several models.

3 Empirical Evidence

This section presents the output obtained on a small real dataset. The case study is an Italian inter-firms network in the Marche region (Ancona administrative district).

Nexus can be defined as an 'heritage network' as well as their firms are all located in a 'heritage area': a delimited territory (Vallesina) with a specific cultural identity. Network membership needs to share some values (mainly coming from the traditional farmers' culture of that area). The whole network consists of 25 firms, operating in different and potentially integrated field of activities (automotive and energy, consultancy, software house, clothing...), with a low level of reciprocal competition among actors. Firms to be officially included in Nexus need to share a collaborative mission, consequently relationships are mainly collaborative. Even if some official activities are periodically organized (such as meetings), the network is aimed to stimulate informal relationships and knowledge exchange among firms.

The latent space considered refers to the *basic social embeddedness*; the manifest measures are: *trust*, *frequency of exchanges among parties*, and *reciprocity*. The same questionnaire was given to a sample of eleven firms belonging to the network; using a scale $\{0, 1, \ldots, 5\}$, respondents have evaluated their partners on *trust*, *frequency of exchanges*, and *reciprocity*. For these three variables, each respondent was asked to indicate her/his feeling with respect to other actors in the net: scores indicate the measure of relationship measured on a non-linear similarity scale from 1 to 5. Higher values indicates high higher involvement degrees with other actors in the net. By definition we assumed that the self evaluation is equal to 5.

Data have been arranged into a three-way data matrix. Each slice has dimension 11×11, and the first row/column refer to Gruppo Loccioni, which has been the promoter of Nexus. Other firms are labelled with the capital letters 'B, C, ...'.

The matrix has been analyzed with the PROXSCAL procedure, and choosing a three dimensional solution. Generally researcher prefer the two-dimensional solution because it permits to get good graphical representations. However, the aim of the present paper is to visualize the scaled similarities using a network display.

In the following part of this section is summarized the PROXSCAL output of the SPSS package. To be consistent with the paper general aim, the *weighted Euclidean model* has been selected. This model defines the points co-ordinates in the common space as a weighted sum of the single space co-ordinates. Proximity transformations have been imposed on interval scale and applied across all sources simultaneously. The algorithm, starting from the Torgeson initial solution, reached the minimum value at $\sigma_n = 0.0277$, which represents a very good result. Table 1 displays the weighting matrix $\mathbf{W}$ having order 3×3 (three sources and three dimensions). It is worth noticing that all terms in $\mathbf{W}$ are positive. However, coefficients are very similar: in particular, *reciprocity* and *trust* have almost the same values. This coefficient signs concordance is consistent with the choice of an additivity model for defining the latent tie. It highlights that *reciprocity* and *trust* were perceived as equivalent concepts by the interviewed entrepreneur.

Figure 1 displays the statistical units in the 3D scaled space. All actors are connected with Gruppo Loccioni indicating the central role played by this firm inside Nexus.

Ties are function of the scaled distances among statistical units in the common space: $\sum_k d_{ij}(\mathbf{X}_k)$. The common space display (Fig. 1) allows to appreciate the reciprocal positions of the statistical units, and the network display (Fig. 2) shows

Table 1 Dimension weights

Source	Dimension Weights		
	Dim1	Dim2	Dim3
TRUST	0.397	0.334	0.347
REC	0.426	0.392	0.362
CONN	0.425	0.392	0.369

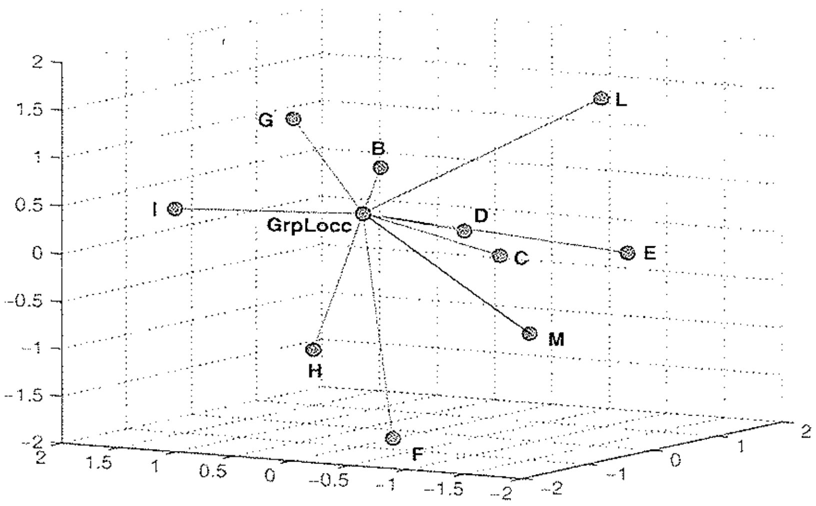

Fig. 1 Points in the 3D common space

the basic social embeddedness. Furthermore, vertices of the network represent the actors and the arcs thickness is function of the embeddedness.

Network latent ties display is represented in Fig. 2 it has been obtained using the Pajek© package. Scaled distances in the latent space among vertices (actors) represent the strength of the ties, it can be interpreted in terms of reciprocal 'social closeness'; at the same time the arcs thickness represents the basic social embeddedness as synthesis of the three considered variables. According to the most largely used techniques in SNA, when displaying the network proper thresholds can be defined to omit the visualisation of the weak ties.

Results show that the Nexus form is similar to a 'constellation of firms' where there is a focal enterprise, with a central role in the network. Regarding the basic social embeddedness, it is possible to observe that is deeper among the founders of the network, included the focal company. In fact, among the founders there is a high level of information exchange and knowledge creation. The firms that have been joined Nexus later, probably have some difficulties to integrate themselves with the founders. This means that the basic social embeddedness is stronger in the founders group than in the group of entrepreneurs that have been joined Nexus later.

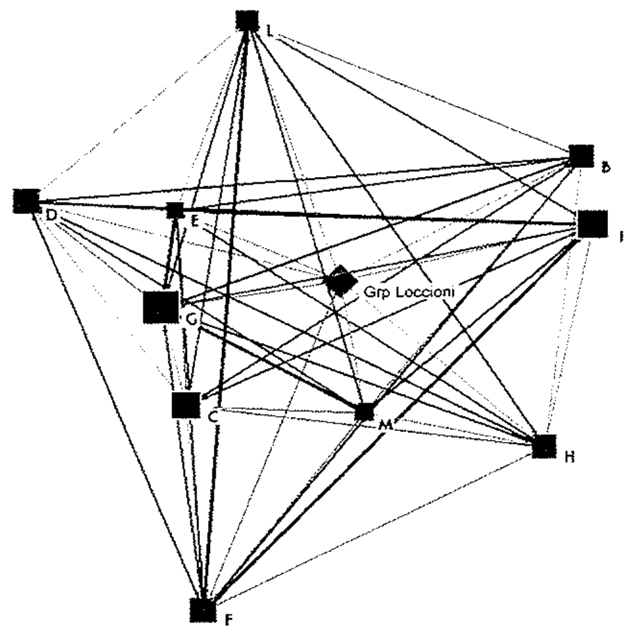

Fig. 2 Network display

This phenomenon has also a positive aspect as well as weak ties could be able to connect the network to others 'world' and to create new contacts with news firms (Granovetter 1983).

4 Conclusive Remarks: Possible Applications of the Proposed Method

Continuing to work at the network level, the proposed methodology can be used as a valid tool to study different kinds of social latent ties among organizations. *Trust*, *frequency*, *reciprocity* and *basic social embeddedness* are only an example of a set of manifest and latent variables: choosing sets of different manifest variables it will be possible to study other latent relational inter–organizational elements, such as network centrality and network density, that are important elements in a network analysis (Lomi 1997).

Measuring that kind of variables, it can help the management in the decisions concerning the investment in organizational structures that can facilitate the development of social relations. The method can also be viewed in an evolutionary perspective, comparing different measures of the latent tie over the time.

For this reason, this methodology is suitable to identify emerging significant networks' latent characteristics and consequently can offer new opportunities to study emerging inter-organizational settings that are becoming more relevant in the actual complex co-operative scenario, such as learning and visionary networks.

Researcher can also imagine applications of this method to identify latent ties, working at different organizational levels.

The research method can be used at the team level to analyze the main characteristics of the collaborative behavior between workers, such as the cohesiveness of the team, that are becoming more and more important especially in the perspective of the organizational learning studies (Senge 2006). At the organizational level, this method can be experimented to measure the latent relationship between the different units, providing organizational charters where the lines thickness connecting two units represents the measure of the latent tie studied.

Moreover, is it possible to imagine applications of the method at the macro-systemic level, especially for the authors that need to better understand the latent ties among organizations and their more relevant stakeholder groups.

References

Bartholomew, D. J. (1987). *Latent variable models and factors analysis*. New York, NY, USA: Oxford University Press, Inc.

Borg, I., & Groenen, P. J. F. (2005). *Modern multidimensional scaling: Theory and applications* (2nd ed.). New York: Springer.

Grandori, A., & Soda, G. (1995). Inter-firm networks: Antecedents, mechanism and forms. *Organization Studies, 16*(2), 183–214.

Granovetter, M. (1973). The strength of weak ties. *American Journal of the Sociology, 78*(6), 1360–1380.

Granovetter, M. (1983). The strength of weak ties: A network theory revisited. *Sociological Theory, 1*, 201–233.

Hoff, P. D., Raftery, A. E., & Handcock, M. S. (2002). Latent space approaches to social network analysis. *JASA, 97*(460), 1090–1098.

Jones, C., Hesterly, W. S., & Borgatti, S. P. (1997). A general theory of network governance: Exchange conditions and social mechanism. *Academy of Management Review, 22*(4), 911–945.

Leydesdorff, L., Schank, T., Scharnhorst, A., & de Nooy, W. (2008). Animating the development of Social networks over time using a dynamic extension of multidimensional scaling. *El profesional de la información, 17*(6), 611–646.

Lomi, A., (1997). *Reti organizzative, teorie tecniche e applicazioni*. Bologna: Il Mulino.

Meulman, J. J., & Heiser, W. J. (2001). SPSS Inc., *SPSS Categories©*. Chicago: SPSS Inc.

Nonaka, I. (1991). The knowledge creating company. *Harvard Business Review, 69*(6), 96–104.

Okada, A. (2010). Two-Dimensional centrality of asymmetric social network. In F. Palumbo, N. C. Lauro, & M. J. Greenacre (Eds.), *Data analysis and classification proceedings of the 6th conference of the ClaDAG* (pp. 93–100). Berlin Heidelberg: Springer.

Rindfleisch, A., & Moorman, C. (2001). The acquisition and utilization of information in new product alliances: A strength of ties perspective. *Journal of Marketing, 65*, 1–18.

Senge, P. (2006). *La quinta disciplina*. Milano: Sperling & Kupfer Editori.

Takane, Y. (2007). Applications of multidimensional scaling in psychometrics. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics, Vol. 26, Psychometrics*. USA: Elsevier B.V.

Uzzi, B. (1999). Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking financing. *American Sociological Review, 64*, 481–505.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, UK: Cambridge Univ. Press.

Williamson, O. E. (1985). *The economic institutions of capitalism – firms, markets, relational contracting*. New York: The Free Press.