

SIMONE GARATTI: CURRICULUM VITAE

DATI ANAGRAFICI

Nome: **Simone Garatti**
Nato a: **Brescia, il 21 Agosto 1976**
Cittadinanza: **Italiana**
Lingua: **Italiano (madrelingua), Inglese (parlato e scritto)**

Contatti: **Politecnico di Milano – Dipartimento di Elettronica ed Informazione
Piazza L. da Vinci 32, 20133, Milano**
Telefono: **+39.02.2399.3650 (ufficio)**
Email: sgaratti@elet.polimi.it, simone.garatti@polimi.it

ATTUALE POSIZIONE

Ricercatore nel settore scientifico disciplinare ING-INF/04 presso il Dipartimento di Elettronica ed Informazione del Politecnico di Milano.

POSIZIONI PRECEDENTEMENTE RICOPERTE

08/06 – 09/06: visiting scholar presso il “Department of Mechanical and Aerospace Engineering – University of California San Diego”, San Diego, CA, USA.
03/04 – 12/04: assegnista di ricerca presso il Dipartimento di Elettronica e Informazione del Politecnico di Milano.
Agosto 2003: visiting student presso il “Lund Institute of Technology”, Lund, Svezia.
03/01 – 02/04: allievo del corso di Dottorato di Ricerca in Ingegneria dell’Informazione (XVI ciclo) presso il Dipartimento di Elettronica e Informazione del Politecnico di Milano.

FORMAZIONE

Dottorato in Ricerca in Ingegneria dell’Informazione, presso il Politecnico di Milano.

Esame finale sostenuto il 5 Maggio 2004. Valutazione: **Eccellente (Lode)**.

Titolo della tesi di dottorato: “Stima della qualità dei modelli identificati: una rivisitazione della teoria asintotica con applicazioni al controllo iterativo”.

Relatore: Prof. Sergio Bittanti; *Correlatore:* Prof. Marco Campi.

Controrelatore: Prof. Giorgio Picci.

Laurea in Ingegneria Informatica (indirizzo Automatica) presso il Politecnico di Milano.

Esame finale sostenuto il 21 Dicembre 2000. Valutazione: **100/100 Lode**.

Titolo della tesi: “Una tecnica iterativa per la sintesi del regolatore basata su un approccio robusto in media”.

Relatore: Prof. Sergio Bittanti; *Correlatore:* Prof. Marco Campi.

Controrelatore: Prof. Patrizio Colaneri.

Durante il Dottorato di Ricerca ha frequentato i seguenti corsi:

- Corso “Complementi di analisi e controllo di sistemi lineari multivariabili” organizzato da Patrizio Colaneri, Arturo Locatelli e Augusto Ferrante e svoltosi a Milano nel periodo Aprile, Maggio 2001 (20 ore). *Valutazione:* ottimo.
- Corso “Introduzione all’analisi e al controllo dei sistemi non lineari” organizzato da Alessandro Astolfi, Sergio Savaresi e svoltosi a Milano nel periodo Novembre 2001 - Gennaio 2002 (20 ore). *Valutazione:* eccellente.
- Corso “Calcolo stocastico” organizzato da Alberto Barchielli, Marco Fuhrman e svoltosi a Milano nel periodo Luglio 2001 - Marzo 2002 (40 ore). *Valutazione:* eccellente.

- Corso “Metodi innovativi di analisi dei dati: filtraggio e stima” organizzato da Sergio Bittanti, Marco Campi, Maria Prandini e svoltosi a Milano nel periodo Aprile, Maggio 2002 (20 ore). *Valutazione*: eccellente.
- Corso “Algoritmi randomizzati per il controllo robusto” organizzato da Roberto Tempo, Giuseppe Calafiore, Fabrizio Dabbene e svoltosi a Milano nel periodo Ottobre 2002-Dicembre 2002 (20 ore). *Valutazione*: eccellente.
- Corso “Tecniche di controllo Data-Based” organizzato da Sergio Savaresi e svoltosi a Milano nel periodo Aprile-Giugno 2003 (20 ore). *Valutazione*: eccellente.
- Corso “Tecniche di Lyapunov per il controllo dei sistemi dinamici” organizzato da Franco Blanchini e svoltosi a Milano nel periodo Ottobre-Dicembre 2003 (20 ore). *Valutazione*: ottimo.
- Corso “Principles of Nonparametric Learning” organizzato da CISM (International Center of Mechanical Science) - L. Gyorfi (coordinatore) e svoltosi a Udine nel periodo 9-13 Luglio 2001 (35 ore).
- Corso “V Scuola Nazionale CIRA: Controllo Automotive e Controllo Predittivo” organizzato da L. Glielmo, R. Scattolini (coordinatori) e svoltosi a Bertinoro (FO) nel periodo 16-21 Luglio 2001 (45 ore).
- Corso “VI Scuola Nazionale CIRA: Sistemi lineari - approccio lineare e approccio geometrico” organizzato da Patrizio Colaneri, Giovanni Marro (coordinatori) e svoltosi a Bertinoro (FO) nel periodo 15-17 Luglio 2002 (20 ore).
- Corso “Lèvy processes” organizzato da: CNR – IMATI e Jean Bertoin e svoltosi a Milano nel periodo 7-11 Ottobre 2002 (10 ore).
- Corso “Controllo di sistemi dinamici stocastici” organizzato da: Politecnico di Milano e Gianmario Tessoro svoltosi a Milano nel periodo Marzo-Maggio 2003 (20 ore).
- Corso “Mini-workshop on Convex Optimization Techniques for Control Systems Analysis and Design” organizzato dall’Università di Siena e svoltosi a Siena il 30 Giugno 2003 (5 ore).
- Corso “VII Scuola Nazionale CIRA 2003: Sistemi lineari - Analisi e controllo di sistemi ibridi” organizzato da A. Bicchi e M. Di Benedetto (coordinatori) e svoltosi a Bertinoro (FO) nel periodo 17-19 Luglio 2002 (20 ore).
- Corso “Convex Optimization” organizzato dall’Università di Siena e Stephen Boyd e svoltosi a Siena nel periodo 8-11 Settembre 2003 (16 ore).

ATTIVITÀ DI RICERCA

- Da *Gennaio 2005*: svolge attività di ricerca presso il Dipartimento di Elettronica ed Informazione del Politecnico di Milano come ricercatore nel settore disciplinare ING-INF/04.
- *Agosto 2006 - Settembre 2006*: è vincitore di una borsa CNR per la mobilità di breve durata dei ricercatori e svolge come “visiting scholar” attività di ricerca presso il “Department of Mechanical and Aerospace Engineering – University of California San Diego”, San Diego, CA, USA, in collaborazione con il prof. Robert Bitmead.
- *Marzo 2004 - Dicembre 2004*: svolge attività di ricerca come assegnista di ricerca presso il Dipartimento di Elettronica e Informazione del Politecnico di Milano.
- *Agosto 2003*: svolge attività di ricerca come ospite presso il “Lund Institute of Technology”, Lund, Svezia, sotto la supervisione del Prof. Rolf Johansson.
- *Marzo 2001 - Febbraio 2004*: svolge attività di ricerca al Dipartimento di Elettronica e Informazione del Politecnico di Milano come allievo del corso di Dottorato di Ricerca in Ingegneria dell’Informazione (XVI ciclo).

Principali attività di ricerca

Ottimizzazione mediante algoritmi randomizzati:

Numerosi problemi della teoria dei sistemi e del controllo possono essere riformulati come problemi di ottimizzazione vincolata e conseguentemente possono essere risolti usufruendo di tutti i risultati teorici e numerici della teoria dell’ottimizzazione. La presenza di elementi incerti nei dati di un problema (ad esempio, un parametro incerto nella funzione di trasferimento di un sistema) può tuttavia portare alla presenza di un’*infinità* di vincoli, uno per ogni valore che gli elementi incerti possono assumere. Si pensi, a titolo di esempio, al problema di verificare se un dato controllore stabilizzi contemporaneamente tutti i sistemi in un dato insieme di confidenza.

Tale problema può essere riformulato nel seguente modo: minimizzare k , con k maggiore o uguale della parte reale degli autovalori di tutti i sistemi di controllo ottenuti a partire dall'insieme di confidenza; verificare poi se il k ottenuto è minore di zero. Problemi di questo tipo vengono detti problemi di ottimizzazione robusta.

Nonostante la relativa facilità con cui i problemi possono essere riformulati, trovare la soluzione esatta di un problema di ottimizzazione robusta può essere in molti casi impraticabile, sia analiticamente che numericamente, a causa della numerosità eccessiva di vincoli da rispettare. È stato provato in letteratura che la risoluzione di un problema di ottimo robusto può essere addirittura NP-hard.

L'idea fondamentale che è stata sviluppata in questo contesto consiste nel formulare un nuovo problema di ottimizzazione che abbia un numero finito di vincoli, questi ultimi estratti randomicamente dall'insieme infinito iniziale di vincoli, [2,7]. In questo modo, la soluzione del nuovo problema di ottimizzazione può essere calcolata utilizzando metodi standard e ad un costo computazionale relativamente basso. Inoltre tale soluzione possiede buone proprietà di "generalizzazione" in quanto è stato possibile dimostrare come la soluzione del nuovo problema di ottimizzazione soddisfi non solo i vincoli estratti, ma anche una buona porzione di quelli che non sono stati presi in considerazione. In particolar modo, in [3] si è stabilito quante estrazioni casuali sono strettamente necessarie al fine di garantire che solo una data frazione dei vincoli iniziali non sia soddisfatta, dando così all'utente la possibilità di fissare a priori il livello di robustezza della soluzione ottenuta mediante randomizzazione.

Sebbene questi primi risultati, lo studio dei metodi randomizzati per l'ottimizzazione robusta è solamente agli inizi e questo lavoro di ricerca si propone di sviluppare l'approccio randomizzato in modo da renderlo una tecnologia flessibile e matura per la risoluzione di problemi altrimenti ritenuti intrattabili. L'attività di ricerca si sta sviluppando lungo due fronti. Da un lato, si sta cercando di affinare l'analisi teorica del metodo proposto e di alcune sue varianti. Lungo questa direzione, oltre al già citato articolo [3] dove la robustezza della soluzione randomizzata viene valutata in modo esatto, sono stati prodotti gli articoli [5,8] dove si considera oltre alla randomizzazione dei vincoli la rimozione di alcuni di essi, affrontando così il problema della modulazione della robustezza in favore delle prestazioni (di filtraggio, di controllo, ecc.) della soluzione ottenuta.

Dall'altra parte, l'approccio randomizzato permette di introdurre paradigmi innovativi nell'ambito della teoria dei sistemi e del controllo, oltre che in altri ambiti quali la statistica e l'analisi dei segnali per citarne alcuni. Nella fattispecie, in [4] è stato proposto un metodo innovativo per la rilevazione di disturbi in presenza di saturazioni dell'azione di controllo, mentre in [1,6,9] sono stati introdotti dei modelli alternativi per la predizione, i cosiddetti modelli intervallari, ed è stata sviluppata una teoria per l'identificazione di tali modelli a partire dai dati basata sulla risoluzione di problemi di ottimizzazione randomizzati.

1. M. Campi, G. Calafiore, S. Garatti, "Interval predictor models: identification and reliability", **Automatica (regular paper)**, **45**(2): 382-392, 2009.
2. M. Campi, S. Garatti, M. Prandini, "The scenario approach for systems and control design", **Annual Reviews in Control**, **33**(2):149-157, 2009.
3. M.C. Campi, S. Garatti, "The exact feasibility of randomized solutions of uncertain convex programs", **SIAM Journal on Optimization**, **19**(3): 1211-1230, 2008.
4. M. Prandini, M. Campi, S. Garatti, "Controller design through random sampling: an example", in: "**Advance in control theory and applications**" (eds. C. Boniventi, A. Isidori, L. Marconi, C. Rossi) – Springer Verlag, LCNIS series, 2007.
5. M. Campi, S. Garatti, "Chance-constrained optimization via randomization: feasibility and optimality", **Optimization Online** (http://www.optimization-online.org/DB_HTML/2008/09/2092.html), 2008.
6. S. Garatti, M.C. Campi, "L-inf layers and the probability of false prediction", **Proceedings of the 15th IFAC Symposium on System Identification (SYSID)**, Saint-Malo, France, 2009.
7. M. Campi, S. Garatti, M. Prandini, "The scenario approach for systems and control design", **Proceedings of the 17th IFAC world congress** (stampato inoltre nel volume "Plenary Papers, Milestone Reports & Selected Survey Papers", pp. 180-188), Seoul, Korea, 2008.
8. M. Campi, S. Garatti, "Modulating robustness in robust control: making it easy through randomization", **Proceedings of the 46th IEEE Conference on Decision and Control**, New Orleans, Louisiana USA, 2007.
9. M. Campi, G. Calafiore, S. Garatti, "New results on the identification of interval predictor models", **Proceedings of the 16th IFAC world congress**, Prague, Czech Republic, 2005.

Stima di parametri in modelli a scatola bianca

Uno dei problemi basilari nella teoria dell'identificazione è quello di stimare il valore di un parametro incognito che appare nelle equazioni di un modello matematico dato, a partire dalle misure dei segnali di ingresso ed uscita. A questo proposito, i metodi esistenti basati sulla teoria del filtraggio alla Kalman e dell'identificazione a minimizzazione dell'errore di predizione dell'uscita possono soffrire di alcuni inconvenienti, quali la non convergenza della stima da un lato e la presenza di minimi locali dall'altro. Questo fa sì che tali metodi debbano essere manualmente ritirati ogni volta che il problema di stima si pone.

In certe situazioni la ritaratura manuale dello stimatore non è possibile. Si pensi a titolo di esempio il problema di stimare i parametri caratteristici di uno pneumatico di un'automobile a partire dalle misure dell'accelerazione laterale dell'autoveicolo, dove lo stimatore deve essere implementato come un componente elettronico, [1]. E' chiaro che durante la vita dell'autoveicolo la tipologia di pneumatico montato può cambiare sensibilmente e tuttavia lo stimatore deve continuare a fornire delle stime adeguate senza richiedere ritarature della centralina elettronica.

In questo lavoro di ricerca, [1,2,3,4], si vuole proporre un nuovo approccio per la stima dei parametri, basato su un uso intensivo del simulatore. L'idea di base è quella di generare per un numero di possibili valori del parametro incognito i corrispondenti andamenti delle variabili di ingresso ed uscita. Partendo da tali dati in simulazione, poi, la relazione sussistente tra gli andamenti di ingresso e uscita e i valori del parametro incognito viene ricostruita fuori linea e una volta per tutte. Ovviamente quest'ultimo problema può essere alquanto complesso se affrontato direttamente e, per questa ragione, in questo lavoro viene proposto un metodo in due passi: in una prima fase intermedia l'informazione contenuta nei segnali di ingresso e uscita viene compressa in dati artificiali, riducendo in questo modo la complessità del problema; successivamente, viene stabilito il legame tra dati artificiali e i possibili valori del parametro, ricorrendo a tecniche di identificazione di funzioni non lineari, ad esempio mediante reti neurali. La cascata della prima fase e della seconda risolve il problema iniziale di ricostruzione del legame tra andamenti delle variabili e valori del parametro.

Alcune simulazioni al calcolatore hanno mostrato che l'approccio a due passi può offrire significativi miglioramenti rispetto agli approcci tradizionali e l'obiettivo attuale è quello di verificare la validità del metodo in problemi applicativi reali, come il problema della stima dei parametri di uno pneumatico precedentemente citato e la stima di parametri di un impianto industriale.

1. S. Garatti, S. Bittanti, "Parameter estimation in the Pacejka's tyre model through the TS method", **Proceedings of the 15th IFAC Symposium on System Identification (SYSID)**, Saint-Malo, France, 2009.
2. S. Bittanti, S. Garatti, "Revisiting the basic issue of parameter estimation in system identification – a new approach for multi-value estimation", **Proceedings of the 47th IEEE Conference on Decision and Control**, Cancun, Mexico, 2008.
3. S. Garatti, S. Bittanti, "Estimation of white-box model parameters via artificial data generation: a two-stage approach", **Proceedings of the 17th IFAC world congress**, Seoul, Korea, 2008.
4. S. Garatti, S. Bittanti, "Parameter estimation via artificial data generation with the 'two-stage' approach", **Proceedings of the 7th World Congress on Intelligent Control and Automation**, Chongqing, China, 2008.

Teoria dell'identificazione e stima della qualità dei modelli identificati:

Stimare la qualità di un modello è un problema importante nella teoria dell'identificazione dei modelli. Come infatti è noto un modello identificato è di poca utilità nelle applicazioni se non viene fornita anche una valutazione del suo campo di validità.

Uno degli strumenti più usati in questo contesto è la teoria asintotica dell'identificazione, mediante la quale è possibile calcolare delle regioni di confidenza ellissoidali per i parametri del sistema. Va tuttavia notato che la teoria asintotica è corretta solo quando il numero di dati a disposizione tende all'infinito, mentre in pratica si dispone di dati in numero finito. Di conseguenza, la stima dell'attendibilità dei modelli identificati mediante la teoria asintotica ha validità esclusivamente euristica e, nonostante i risultati ottenuti siano corretti in molte situazioni, può capitare in certi casi (specialmente in presenza di ampia incertezza – dati poco informativi) che la regione di confidenza stimata sia del tutto irrealistica e ingannevole. Questa limitazione della teoria asintotica è piuttosto grave poiché la presenza di un alto grado di incertezza è tipica di molte applicazioni reali (ad esempio, quando l'identificazione deve essere compiuta in anello chiuso, con banda passante limitata). Più in generale, si può anche affermare che valutare la qualità del modello è tanto più importante quando la discrepanza fra il sistema e il modello è elevata.

L'obiettivo di questo studio è la discussione critica delle condizioni per l'applicabilità della teoria asintotica dell'identificazione. L'attenzione è stata rivolta in particolare ai metodi basati sulla minimizzazione dell'errore di predizione (PEM) e sul metodo della variabile strumentale (IV) - si veda [3,6] e [2,5] rispettivamente. In entrambi i casi, sono state analizzate in modo approfondito le condizioni che garantiscono un corretto utilizzo della teoria asintotica anche in presenza di dati poco informativi e, a tal fine, si è reso necessario dimostrare alcune proprietà asintotiche dell'identificazione sotto ipotesi non standard.

Il risultato di questa analisi ha messo in luce che la teoria asintotica può portare a risultati attendibili o ingannevoli a seconda della classe di modelli utilizzata nell'identificazione. Per questa ragione, si sono poi studiate le classi di modelli standard (ARX, ARMAX, Box Jenkins, ecc.) per capire in quali casi la stima dell'incertezza è valida anche in condizioni di scarsa informazione contenuta nei dati. Il risultato ottenuto è che in diversi casi la teoria asintotica può essere usata in sicurezza in quanto in grado di stimare correttamente l'incertezza presente. Allo stesso tempo, però, si è mostrato come nel caso di certe classi di modelli (come ad esempio la classe Box Jenkins) l'utilizzo della teoria asintotica richieda una certa cautela in quanto i risultati forniti possono essere completamente inaffidabili.

Al fine di superare gli inconvenienti citati, i lavori [1,4], infine, propongono l'analisi e lo studio di tecniche di tipo *resampling*, come il subsampling, Jackknife e Bootstrap, per la ricostruzione garantita a partire da numero finito di dati dell'intera funzione di distribuzione di probabilità dei parametri identificati con tecniche di tipo PEM.

1. S. Garatti, R.R. Bitmead, "On resampling and uncertainty estimation in linear system identification", **Automatica (regular paper)**, 46(5): 785-795, 2010.
2. S. Garatti, M. Campi, S. Bittanti, "The asymptotic model quality assessment for instrumental variable identification revisited", **Systems & Control Letters**, 55(6): 494-500, 2006.
3. S. Garatti, M. Campi, S. Bittanti, "Assessing the quality of identified models through the asymptotic theory - When is the result reliable?", **Automatica (regular paper)**, 40(8): 1319-1332, 2004.
4. S. Garatti, R.R. Bitmead, "On re-sampling and uncertainty estimation in linear system identification", **Proceedings of the 15th IFAC Symposium on System Identification (SYSID)**, Saint-Malo, France, 2009.
5. S. Garatti, S. Bittanti, M. Campi, "Model quality assessment for Instrumental Variable methods: use of the asymptotic theory in practice", **Proceedings of the 42nd IEEE Conference on Decision and Control**, Maui, Hawaii USA 2003. pp. 6015-6020.
6. S. Bittanti, M. Campi, S. Garatti, "New results on the asymptotic theory of system identification for the assessment of the quality of estimated models", **Proceedings of the 41st IEEE Conference on Decision and Control**, Las Vegas, Nevada USA 2002. pp. 1814-1819.

Metodi iterativi per il progetto del controllore:

Il controllo iterativo è una metodologia di recente sviluppo per il progetto di controllori ad alte prestazioni per sistemi a dinamica a priori non nota. L'idea principale può essere così descritta: progettare il controllore mediante una sequenza di iterazioni, ciascuna delle quali costituita da due fasi: la prima consiste nell'identificazione del sistema, mentre quest'ultimo lavora in anello chiuso con il controllore determinato all'iterazione precedente; nella seconda fase invece il precedente controllore viene aggiornato sulla base del nuovo modello identificato. In questo modo è possibile aumentare progressivamente, man mano che la conoscenza del sistema viene aggiornata, le prestazioni del sistema di controllo.

Pur portando in generale a buoni risultati in termini di prestazioni finali, le tecniche iterative standard hanno un principale difetto: basandosi unicamente sul modello identificato, ad ogni iterazione il controllore deve essere aggiornato in modo cauto per evitare la destabilizzazione del sistema. Questo fatto a sua volta comporta che per ottenere il controllore desiderato sia in genere necessario un numero eccessivo di esperimenti sull'impianto, rendendo così i metodi iterativi troppo onerosi per molte applicazioni.

In questo lavoro di ricerca è stato proposto uno schema di controllo iterativo che tenga esplicitamente conto dell'incertezza del modello identificato (controllo iterativo robusto - [1,2,3,4]). Questo nuovo approccio prevede che ad ogni iterazione sia stimato non solo un modello nominale, ma anche la sua incertezza. Corrispondentemente l'aggiornamento del controllore viene effettuato anche tenendo conto del livello di incertezza rilevato. Così facendo, è possibile spingere al massimo il miglioramento delle prestazioni ad ogni iterazione, preservando tuttavia la stabilità robusta del sistema di controllo. Ne consegue un algoritmo che, a differenza di quelli standard, permette di ottenere un controllore ad alte prestazioni sulla base di un numero limitato di esperimenti sul sistema.

Lo studio della problematica sopra descritta è stato sviluppato fino ad arrivare alla messa a punto di un nuovo algoritmo iterativo robusto. A tal fine, si è reso necessario da un lato studiare e risolvere i problemi connessi alla stima dell'incertezza del modello durante le iterazioni e dall'altro sviluppare un'adeguata metodologia di controllo robusto da utilizzare per ogni progetto intermedio del controllore. Nella fattispecie, si è fatto uso della teoria del controllo robusto in media prima, [3,4] e in probabilità poi [1,2]. Entrambe le metodologie possono essere implementate ad un basso costo computazionale ricorrendo ai cosiddetti algoritmi randomizzati.

1. S. Garatti, M. Campi, S. Bittanti, "Iterative robust control: speeding up improvement through iterations", **Systems & Control Letters**, **59**(2):139-146, 2010.
2. S. Bittanti, S. Garatti, M. Campi, "Introducing robustness in iterative control", **Proceedings of the 44th IEEE Conference on Decision and Control**, Seville, Spain, 2005.
3. S. Bittanti, M. Campi, S. Garatti, "Some critical implementation issues in iterative robust control design", **Proceedings of the 8th IFAC workshop on Adaptation and Learning in Control and Signal Processing**, Yokohama, Japan 2004.
4. S. Bittanti, M. Campi, S. Garatti, "An iterative controller design scheme based on average robust control", **Proceedings of the 15th IFAC World Congress**, Barcelona, Spain 2002.

Analisi dei dati mediante tecniche di data-mining:

Negli ultimi anni è emersa una nuova problematica nel campo dell'identificazione, quando si debbano ricercare relazioni nascoste a partire da grandi basi di dati, come ad esempio quando si deve classificare il comportamento dei clienti di un supermercato a partire dalle rilevazioni dei loro acquisti.

Le basi di dati analizzate in questo lavoro di ricerca sono state due: i dati relativi ad una comunità virtuale di Internet e dati bioinformatici generati mediante DNA-microarrays. Principalmente si è fatto uso di metodi di partizione non supervisionata i quali, a partire solamente da una nozione di distanza tra campioni, permettono di suddividere i dati a disposizione in un numero ridotto di categorie rappresentative. In tal modo è stato possibile estrarre informazioni significative dai dati a disposizione, nonostante la loro elevata dimensionalità. Al fine di ottimizzare i risultati ottenuti, è stato anche necessario rivolgere particolare attenzione alla preparazione dei dati, nonché all'analisi dei risultati e all'assimilazione delle informazioni ottenute.

Per quanto concerne le applicazioni considerate, nel caso delle comunità virtuali in Internet (ovvero gruppi di persone che comunicano mediante Internet) l'analisi è stata effettuata su due insiemi di dati differenti ([1,2,3,4]). Da un lato, si è considerato il data-base dei profili degli utenti (informazioni personali compilate durante la registrazione alla comunità virtuale) e da questo insieme di dati si è estrapolato un numero ridotto di prototipi di utenti. Dall'altro, si è considerato il data-base degli accessi (pagine visitate) al sito Internet che fornisce i servizi di comunicazione per la comunità virtuale. Dopo aver ricostruito i percorsi di navigazione degli utenti, l'insieme delle sessioni di navigazione è stato suddiviso in un numero ridotto di categorie di sessioni di navigazione. Grazie ai risultati ottenuti, è stato infine possibile determinare la correlazione esistente tra tipologie di utenti e di sessioni, così da ottenere una descrizione significativa del comportamento degli utenti della comunità virtuale considerata. Per quel che riguarda invece l'applicazione di natura bioinformatica, i dati a disposizione erano relativi ad un certo numero di pazienti affetti da una malattia tumorale. Più precisamente per ogni paziente è stata misurata l'attività dei geni presenti nelle cellule affette dalla malattia. Grazie ai metodi di partizione non supervisionata è stato possibile individuare e classificare differenti patologie della medesima malattia sulla base della sola conoscenza dell'attività genica dei pazienti, [5,6,7,8].

1. S. Garatti, S.M. Savaresi, S. Bittanti, L. La Brocca, "On the relationships between user profiles and navigation sessions in virtual communities: a data-Mining approach", **Intelligent Data Analysis**, **8**(6): 579-600, 2004.
2. S. Savaresi, S. Garatti, S. Bittanti, "Modeling the relationships between the user DB and the web-log file of a large virtual community", **Proceedings of the 13th IFAC Symposium on System Identification**, Rotterdam, The Netherlands 2003. pp. 219-224.
3. S. Savaresi, S. Garatti, S. Bittanti, L. La Brocca, "Data-Mining of a large virtual community: relationships between the users DB and the web-log file", **Proceedings of the 3rd SIAM International Conference on Data Mining**, San Francisco, California USA 2003. pp. 299-303.
4. S. Garatti, "Finding relationships between the users DB and the web-log file of a large Internet community: a data mining approach", **Rapporto Interno N° 2002.48** – Politecnico di Milano, DEI, 2002.

5. S. Garatti, S. Bittanti, D. Liberati, A. Maffezzoli, "An unsupervised clustering approach for leukaemia classification based on dna micro-arrays data", **Intelligent Data Analysis**, 11(2): 175-188, 2007.
6. D. Liberati, S. Bittanti, S. Garatti, Z. Zhao, M. Pappalettera, "Classificazione di leucemie mediante analisi di dati da microarray", **Automazione e Strumentazione** LIII(10):79-86, 2005.
7. S. Bittanti, S. Garatti, D. Libearati, "From DNA microarrays to disease classification: an unsupervised clustering approach", **Proceedings of the 16th IFAC world congress**, Prague, Czech Republic, 2005.
8. D. Liberati, S. Garatti, S. Bittanti "Unsupervised mining of genes classifying Leukaemia", in: "**Encyclopedia of Data Warehousing and Mining**", (J. Wang ed.) – Idea Group Publishing, 2005.

ATTIVITÀ DIDATTICA

Lezioni

- I/II semestre a.a. 04/05: docente del corso "Progetto di identificazione ed analisi dei dati", Politecnico di Milano. (2.5 cfu)
- I/II semestre a.a. 05/06: docente del corso "Progetto di identificazione ed analisi dei dati", Politecnico di Milano. (2.5 cfu)
- I semestre a.a. 05/06: docente del corso "Progetto di I.M.A.D.", Politecnico di Milano. (2.5 cfu)
- I semestre a.a. 05/06: docente del corso "Sistemi Adattativi ed Apprendimento", Politecnico di Milano. (5 cfu)
- II semestre a.a. 05/06: docente del corso "Identificazione dei Modelli ed Analisi dei Dati", Politecnico di Milano – Polo di Como. (5 cfu)
- I/II semestre a.a. 06/07: docente del corso "Progetto di identificazione ed analisi dei dati", Politecnico di Milano. (2.5 cfu)
- I semestre a.a. 06/07: docente del corso "Progetto di I.M.A.D.", Politecnico di Milano. (2.5 cfu)
- I semestre a.a. 06/07: docente del corso "Sistemi Adattativi ed Apprendimento", Politecnico di Milano. (5 cfu)
- II semestre a.a. 06/07: docente del corso "Identificazione dei Modelli ed Analisi dei Dati", Politecnico di Milano – Polo di Como. (5 cfu)
- I/II semestre a.a. 07/08: docente del corso "Progetto di identificazione ed analisi dei dati", Politecnico di Milano. (2.5 cfu)
- I semestre a.a. 07/08: docente del corso "Progetto di I.M.A.D.", Politecnico di Milano. (2.5 cfu)
- I semestre a.a. 07/08: docente del corso "Sistemi Adattativi ed Apprendimento", Politecnico di Milano. (5 cfu)
- II semestre a.a. 07/08: docente del corso "Identificazione dei Modelli ed Analisi dei Dati", Politecnico di Milano – Polo di Como. (5 cfu)
- I semestre a.a. 08/09: docente del corso "Progetto di Automatica – parte 2", Politecnico di Milano. (2.5 cfu)
- I semestre a.a. 08/09: docente del corso "Sistemi Adattativi ed Apprendimento", Politecnico di Milano. (5 cfu)
- II semestre a.a. 08/09: docente del corso "Model Identification and Data Analysis", Politecnico di Milano – Polo di Como. (5 cfu)
- I semestre a.a. 09/10: docente del corso "Progetto di Automatica – parte 2", Politecnico di Milano. (2.5 cfu)
- I semestre a.a. 09/10: docente del corso "Sistemi Adattativi ed Apprendimento", Politecnico di Milano. (5 cfu)
- II semestre a.a. 09/10: docente del corso "Model Identification and Data Analysis", Politecnico di Milano – Polo di Como. (5 cfu)

Esercitazioni

- II semestre a.a. 00/01: Seminari didattici per il corso "Automatica I", prof. Sergio Bittanti, Politecnico di Milano. (20 ore)

- Il semestre a.a. 01/02: Seminari didattici per il corso “Automatica I”, prof. Sergio Bittanti, Milano. (28 ore)
- I semestre a.a. 02/03: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati” (V.O.), prof. Sergio Bittanti, Politecnico di Milano. (32 ore)
- Il semestre a.a. 02/03: Seminari didattici per il corso “Automatica I”, prof. Sergio Bittanti, Politecnico di Milano. (10 ore)
- I semestre a.a. 03/04: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano. (10 ore)
- I semestre a.a. 03/04: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati 2”, prof. Sergio Bittanti, Politecnico di Milano. (20 ore)
- Il semestre a.a. 03/04: Seminari didattici per il corso “Segnali e Sistemi”, prof. Patrizio Colaneri, Politecnico di Milano. (16 ore)
- Il semestre a.a. 03/04: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano – sede Bovisa. (20 ore)
- Il semestre a.a. 03/04: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Savaresi, Politecnico di Milano – sede Como. (8 ore)
- I semestre a.a. 04/05: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano. (20 ore)
- I semestre a.a. 04/05: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati 2”, prof. Sergio Savaresi, Politecnico di Milano. (10 ore)
- Il semestre a.a. 04/05: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Savaresi, Politecnico di Milano – sede Como. (16 ore)
- Il semestre a.a. 04/05: Seminari didattici per il corso “Segnali e Sistemi”, prof. Patrizio Colaneri, Politecnico di Milano. (20 ore)
- I semestre a.a. 05/06: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano. (20 ore)
- I semestre a.a. 05/06: Seminari didattici per il corso “Sistemi Adattativi ed Apprendimento”, prof. Simone Garatti, Politecnico di Milano. (8 ore)
- Il semestre a.a. 05/06: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Simone Garatti, Politecnico di Milano – sede Como. (20 ore)
- I semestre a.a. 06/07: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano. (20 ore)
- I semestre a.a. 06/07: Seminari didattici per il corso “Sistemi Adattativi ed Apprendimento”, prof. Simone Garatti, Politecnico di Milano. (8 ore)
- Il semestre a.a. 06/07: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Simone Garatti, Politecnico di Milano – sede Como. (20 ore)
- I semestre a.a. 07/08: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano. (20 ore)
- I semestre a.a. 07/08: Seminari didattici per il corso “Sistemi Adattativi ed Apprendimento”, prof. Simone Garatti, Politecnico di Milano. (8 ore)
- Il semestre a.a. 07/08: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Simone Garatti, Politecnico di Milano – sede Como. (20 ore)
- I semestre a.a. 08/09: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano. (20 ore)
- I semestre a.a. 08/09: Seminari didattici per il corso “Sistemi Adattativi ed Apprendimento”, prof. Simone Garatti, Politecnico di Milano. (8 ore)
- I semestre a.a. 09/10: Seminari didattici per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano. (20 ore)

- I semestre a.a. 09/10: Seminari didattici per il corso “Sistemi Adattativi ed Apprendimento”, prof. Simone Garatti, Politecnico di Milano. (8 ore)

Attività di laboratorio e tutoraggio

- Il semestre a.a. 02/03: Tutor per lo svolgimento dei progetti integrativi del corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano.
- Il semestre a.a. 02/03: Assistente di laboratorio per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano. (4 ore)
- I semestre a.a. 03/04: Assistente di laboratorio per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano. (6 ore)
- Il semestre a.a. 03/04: Assistente di laboratorio per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano – sede Bovisa. (6 ore)
- I semestre a.a. 04/05: Assistente di laboratorio per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano. (6 ore)
- I semestre a.a. 05/06: Assistente di laboratorio per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano. (6 ore)
- I semestre a.a. 06/07: Assistente di laboratorio per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano. (6 ore)
- I semestre a.a. 07/08: Assistente di laboratorio per il corso “Identificazione dei modelli ed analisi dei dati”, prof. Sergio Bittanti, Politecnico di Milano. (6 ore)

Attività di supporto a laureandi

Tesi di Laurea Specialistica seguite come relatore:

- “Stima dei parametri incerti di un motore a induzione : tecniche di identificazione a confronto”. Autore: Mattia Sarati. Relatore: Simone Garatti. Correlatori: Sergio Bittanti. Dicembre 2009.

Tesi e tesine di Laurea Specialistica e Vecchio Ordinamento seguite come correlatore:

- “Modello di un generatore a vapore e identificazione delle caratteristiche di un flusso bifase a partire dal transitorio dell’impianto”. Autori: Lino Fiorita, Eleonora Porro. Relatore: Sergio Bittanti. Correlatori: Antonio de Marco, Simone Garatti, Carlo Sandroni. Settembre 2009.
- “Tecniche di stima parametrica con applicazione al problema dell’identificazione dei parametri di Pacejka di un pneumatico”. Autore: Carlo Sandroni. Relatore: Sergio Bittanti. Correlatore: Simone Garatti. Dicembre 2008.
- “Application of VRFT (Virtual Reference Feedback Tuning) and MPC to a two-fluid single-pass heat exchanger”. Autore: Paolo Varutti. Relatori: Sergio Bittanti, Frank Allgower. Correlatore: Simone Garatti. Aprile 2007.
- “Identificazione di parametri fisici in sistemi non lineari: un confronto fra tecniche tradizionali e un approccio black-box a 2 passi”. Autore: Filippo Aspesi. Relatore: Sergio Bittanti. Correlatore: Simone Garatti. Dicembre 2006.
- “La tecnologia microarray per lo studio di malattie tumorali: analisi dei dati mediante tecniche di data mining”. Autore: Andrea Maffezzoli. Relatore: Sergio Bittanti. Correlatori: Diego Liberati, Simone Garatti. Luglio 2003.
- “Metodi basati sull’approccio montecarlo per la valutazione del prezzo di un’opzione finanziaria”. Autore: Davide Ferraguti. Relatore: Sergio Bittanti. Correlatori: Danilo Tilloca, Simone Garatti. Febbraio 2003.

Elaborati di Laurea Triennale seguite come relatore:

- “Identificazione di modelli per lo studio di un fenomeno franoso in località Cortenova (LC)”. Autori: Vincenzo Guerrisi, Daniele Gianola. Relatore: Simone Garatti. Correlatore: Sergio Bittanti. Settembre 2009.

Elaborati di Laurea Triennale seguite come correlatore:

- “Stima di un parametro nel modello di una turbina: due metodi a confronto”. Autore: Matteo Garza. Relatore: Sergio Bittanti. Correlatore: Simone Garatti. Febbraio 2010.
- “Identificazione di un modello del rendimento settimanale del tasso di cambio dollaro USA-sterlina mediante i dati del COT report”. Autore: Sergio Iommi. Relatore: Sergio Bittanti. Correlatore: Simone Garatti. Aprile 2008.
- “Modelli PARMA per lo studio della concentrazione di PM10 nella città di Padova”. Autore: Gabriele Maggioni. Relatore: Sergio Bittanti. Correlatore: Simone Garatti. Settembre 2007.
- “Identificazione sull’indice del fatturato industriale italiano”. Autore: Matteo Bertasa. Relatore: Sergio Bittanti. Correlatore: Simone Garatti. Settembre 2007
- “Analisi di serie temporali di temperature terrestri globali dal 1880 ad oggi tramite sistemi lineari e reti neurali”. Autori: Simone Fontolan, Alessandro Garghetti. Relatore: Sergio Bittanti. Correlatore: Simone Garatti. Settembre 2007
- “Metodi di data mining per la classificazione del traffico telefonico di utenti di una compagnia telefonica”. Autore: Luca Sisler. Relatore: Sergio Bittanti. Correlatore: Simone Garatti. Luglio 2007.
- “Metodi di analisi di dati per la sintesi di nuovi farmaci: un caso di studio”. Autore: Leonardo Scamazzo. Relatore: Sergio Bittanti. Correlatori: Simone Garatti, Diego Liberati. Settembre 2005.
- “Data mining per dati genici da microarray”. Autore: Maria Pia Bazzano, Sabino Abruzzese. Relatore: Sergio Bittanti. Correlatore: Simone Garatti. Luglio 2003.

Attività di sviluppo di strumenti didattici

E’ co-autore di alcuni programmi software utilizzati a supporto didattico dei corsi di *Identificazione ed Analisi dei Dati* del Politecnico di Milano.

PREMI E BORSE DI STUDIO

Nel 2006 ha vinto una borsa CNR per la mobilità di breve durata dei ricercatori e ne ha usufruito come *visiting scholar* presso il *Department of Mechanical and Aerospace Engineering – University of California San Diego*, San Diego, CA, USA.

Nel 2005 è stato nominato “Outstanding reviewer” della rivista “Automatica”.

PARTECIPAZIONE A CONFERENZE E VISITE

Visite e soggiorni presso centri di ricerca stranieri

- Dicembre 2007: ciclo di seminari presso l’ *Operational Research Center – MIT*, Boston, MA, USA, e contestualmente presso il *Department of Electrical and Computer Engineering – NorthEastern University*, Boston, MA, USA, dal titolo “The exact feasibility of randomized solutions to robust optimization: theory and applications” (in collaborazione col prof. Marco Campi dell’Università di Brescia).
- Agosto 2006 – Settembre 2006: *visiting scholar* presso il *Department of Mechanical and Aerospace Engineering – University of California San Diego*, San Diego, CA, USA. Ospite del prof. Robert Bitmead.
- Agosto 2003: ospite presso il “Lund Institute of Technology”, Lund, Svezia, sotto la supervisione del Prof. Rolf Johansson.

Conferenze internazionali a cui ha preso parte presentando un lavoro di cui era autore o co-autore

- 15th IFAC World Congress, Barcellona, Spagna, 21-26 Luglio 2002.
- 41st IEEE Conference on Decision and Control, Las Vegas, Nevada, USA, 9-13 Dicembre 2002.
- 3rd SIAM International Conference on Data Mining, San Francisco, California, USA, 1-3 Maggio 2003.
- 13th IFAC Symposium on System Identification, Rotterdam, Olanda, 26-27 Agosto 2003.

- 16th IFAC World Congress, Praga, Repubblica Ceca, 4-8 Luglio 2005.
- 46th IEEE Conference on Decision and Control, New Orleans, Louisiana, USA, 12-14 Dicembre 2007.
- 6th International Conference on Computational Management Science, Ginevra, Svizzera, 1-3 Maggio 2009.
- 15th IFAC Symposium on System Identification (SYSID), Saint-Malo, France, 6-8 Luglio 2009.

Altre conferenze internazionali a cui ha preso parte

- 7th IFAC Workshop on Periodic Control Systems, Cernobbio-Como, 27-28 Agosto 2001.
- 7th IFAC Workshop on Adaptation and Learning in Control and Signal Processing, Cernobbio-Como, 29-31 Agosto 2001.

Conferenze nazionali a cui ha preso parte presentando un lavoro di cui era autore o co-autore

- Convegno nazionale CIRA, Modena, 10-12 Settembre 2003.
- Convegno nazionale CIRA, Tropea, 8-10 Settembre 2005.
- Convegno nazionale SIDRA, Genova, 10-12 Settembre 2007.

COLLABORAZIONE CON AZIENDE E ENTI DI RICERCA

- Collaborazione nell'ambito del contratto di ricerca "Comunità virtuali in Internet: ricerca delle relazioni tra i profili degli utenti e i loro percorsi di navigazione mediante tecniche di data-mining" stipulato il 5 febbraio 2002 con *Tiscali SpA*.
- Collaborazione con l'Istituto dei Tumori di Milano nell'ambito del tema di ricerca "Data-Mining dei dati generati mediante DNA-microarrays".
- Collaborazione nell'ambito del contratto di ricerca "Tuning di modelli comportamentali per apparecchi elettrodomestici" stipulato il 26 novembre 2004 con *Whirlpool SpA*

ATTIVITÀ IN PROGETTI DI RICERCA NAZIONALI ED INTERNAZIONALI

- Partecipazione al progetto europeo N.HPRN-CT-1999-00046 "NACO: Nonlinear and Adaptive Control-Tools and Algorithms for the User" negli anni 2001, 2002, 2003, 2004.
- Partecipazione al progetto co-finanziato MIUR "Nuove tecniche per l'identificazione e il controllo adattativo nei sistemi industriali", 2001- 2002.
- Partecipazione al progetto co-finanziato MIUR "Tecniche innovative per l'identificazione e il controllo adattativo nei sistemi industriali", 2003-2004.
- Partecipazione al progetto co-finanziato MIUR "Tecniche innovative per l'identificazione e il controllo adattativo nei sistemi industriali", 2005-2006.
- Partecipazione al progetto co-finanziato MIUR "Tecniche ed applicazioni innovative di identificazione e controllo adattativo", 2007-2008.
- Partecipazione al progetto co-finanziato MIUR "Nuovi algoritmi ed applicazioni di identificazione e controllo adattativo", 2009-2010.

ATTIVITÀ DI REVISIONE

Ha svolto attività di revisione per numerose riviste internazionali (Automatica, IEEE Transactions on Automatic Control, IEEE Transactions on Circuits and Systems I, IET Control Theory and Applications, International Journal of Control and Signal Processing, International Journal of Robust and Nonlinear Control, Journal of complexity, Linear

Algebra and its Applications, Mathematical Programming , SIAM Journal on Control and Optimization) e per conferenze internazionali (EUCA European Control Conference, IEEE American Control Conference, IEEE Conference on Decision and Control, IEEE Multi-conference on Systems and Control, IFAC Symposium on System Identification, IFAC World Congress). Nel 2005 è stato nominato "Outstanding reviewer" della rivista "Automatica".

ELENCO E PRESENTAZIONE DEI 10 ARTICOLI PIÙ SIGNIFICATIVI

1. M. Campi, G. Calafiore, S. Garatti, "*Interval predictor models: identification and reliability*", **Automatica (regular paper)**, **45**(2): 382-392, 2009

In questo lavoro viene considerata una nuova classe di predittori, i cosiddetti predittori intervallari, e viene affrontato il problema della costruzione di un predittore intervallare a partire da dati sperimentali. Differentemente dai predittori normalmente utilizzati, i quali restituiscono un sol valore come predizione della variabile d'uscita, i predittori intervallari sono funzioni a multi-valori che restituiscono direttamente un intervallo di predizione e, in questo modo, sono automaticamente predisposti per fornire anche una valutazione della qualità della predizione restituita. La scelta di un modello intervallare a partire dai dati è guidata dal principio che il modello debba correttamente descrivere tutti i dati raccolti (ovvero gli intervalli di predizione restituiti devono contenere le misure dell'uscita). Inoltre, tra tutti i predittori intervallari consistenti con i dati visti, viene scelto quello che restituisce intervalli di ampiezza più piccola così da fare predizioni dei futuri valori della variabile d'uscita le più stringenti possibili. In questo articolo viene mostrato come il modello così ottenuto possiede delle doti di affidabilità (misurata come la probabilità che futuri valori dell'uscita cadano negli intervalli di predizione restituiti dal modello) che possono essere quantificate a priori mediante una formula non asintotica che, differentemente da approcci più tradizionali, non dipende in alcun modo dalla struttura del meccanismo incognito mediante il quale i dati sono stati generati. In questo modo l'affidabilità del modello viene relazionata in modo fondamentale all'ammontare dell'informazione disponibile e cioè al numero di dati a disposizione.

2. M. Campi, S. Garatti, "*The exact feasibility of randomized solution of uncertain convex programs*", **SIAM Journal on Optimization**, **19**(3): 1211-1230, 2008.

Numerosi problemi della teoria dei sistemi e del controllo possono essere riformulati come problemi di ottimizzazione vincolata e conseguentemente possono essere risolti usufruendo di tutti i risultati teorici e numerici della teoria dell'ottimizzazione. La presenza di elementi incerti nei dati di un problema (ad esempio, un parametro incerto nella funzione di trasferimento di un sistema) può tuttavia portare alla presenza di un'infinità di vincoli, uno per ogni valore che gli elementi incerti possono assumere. Problemi con un'infinità di vincoli vengono chiamati problemi semi-infiniti ed è ben noto che trovare la soluzione esatta di un problema di ottimizzazione semi-infinito può essere in molti casi impraticabile, sia analiticamente che numericamente. E' stato provato in letteratura che la risoluzione di un problema di ottimo robusto può essere addirittura NP-hard. L'idea fondamentale che è stata sviluppata in questo contesto consiste nel formulare un nuovo problema di ottimizzazione con un numero finito di vincoli, mediante randomizzazione dell'insieme infinito iniziale di vincoli. In questo modo, la soluzione del nuovo problema di ottimizzazione può essere calcolata utilizzando metodi standard e ad un costo computazionale relativamente basso. Inoltre tale soluzione possiede buone proprietà di "generalizzazione" nel senso che la soluzione del nuovo problema di ottimizzazione soddisfa non solo i vincoli estratti, ma anche una buona porzione di quelli che non sono stati presi in considerazione. Quest'ultimo aspetto è stato analizzato in modo esaustivo in quest'articolo dove si è stabilito il numero esatto di estrazioni casuali strettamente necessarie al fine di garantire che solo una frazione data dei vincoli iniziali sia non soddisfatta. Grazie a questo risultato, viene data all'utente la possibilità di fissare a priori il livello di robustezza desiderato e giustifica ad un livello profondo l'uso di metodi randomizzati per ottenere soluzioni garantite.

3. M. Campi, S. Garatti, M. Prandini, "*The scenario approach for systems and control design*", **Annual Reviews in Control**, **33**(2):149-157, 2009.

L'approccio a scenario è una recente metodologia introdotta per risolvere problemi di ottimizzazione convessa con infiniti vincoli. Tali problemi non sono affatto infrequenti in quanto modellizzano situazioni in cui si vuole essere (ottimamente) robusti in presenza di parametri incerti e l'infinità di vincoli corrisponde all'infinita varietà di valori che il parametro incerto può assumere. La tecnologia a scenario si basa sulla randomizzazione dei vincoli e rappresenta uno strumento potente per trovare ad un basso costo computazionale delle soluzioni approssimate la cui robustezza, però, può essere garantita a-priori. Questo articolo di "survey" si propone di illustrare ad un livello elementare la metodologia a scenario, concentrandosi in particolar modo sugli aspetti algoritmici dell'approccio. Inoltre, mediante

una serie di esempi concreti, ne viene illustrata la versatilità per la risoluzione di problemi di filtraggio e controllo robusti. Una versione precedente di questo articolo è stata selezionata per essere pubblicata nel volume "Plenary Papers, Milestone Reports & Selected Survey Papers" del diciassettesimo IFAC world congress.

4. M. Campi, S. Garatti, "*Chance-constrained optimization via randomization: feasibility and optimality*", **Optimization Online** (http://www.optimization-online.org/DB_HTML/2008/09/2092.html), 2008 (sottomesso per la pubblicazione).

E' un fatto risaputo che le tecniche standard di controllo robusto possono portare a controllori conservativi in quanto tutta l'enfasi è data alla salvaguardia delle prestazioni del sistema di controllo a fronte di ogni possibile occorrenza, anche quelle più improbabili, dell'incertezza che affligge il sistema. Le basse prestazioni del controllore così ottenuto sono tipicamente il prezzo da pagare per avere una totale garanzia di robustezza. Sebbene in certi casi avere garanzie di robustezza nel 100% dei casi è strettamente necessario, in molte altre applicazioni, tuttavia, non lo è e accettando un piccolo compromesso (ad esempio accettando una robustezza al 99%) si può spesso ottenere un consistente aumento delle prestazioni. Il baratto tra robustezza e prestazioni in generale può essere difficile da ottenere in quanto non sempre è ovvio quali siano le istanze dell'incertezza da sacrificare e al momento non esistono algoritmi computazionalmente trattabili per risolvere tale problema. Questo lavoro si propone di aprire una nuova strada ad un alto livello di generalità per la risoluzione del problema del baratto tra robustezza e prestazioni attraverso una tecnica randomizzata che estende in modo significativo il cosiddetto approccio a scenario per la risoluzione di problemi di ottimizzazione robusta. Le garanzie per il nuovo controllore così progettato possono essere ottenute mediante una nuova teoria sviluppata per la prima volta in questo articolo.

5. S. Garatti, R.R. Bitmead, "*On resampling and uncertainty estimation in linear system identification*", **Automatica (regular paper)**, **46**(5): 785-795, 2010.

Al fine di valutare la qualità di un modello costruito a partire dai dati, in questo lavoro si considera il problema della stima della funzione di distribuzione della probabilità dei parametri identificati mediante delle tecniche a minimizzazione dell'errore di predizione. Al fine di superare i problemi legati all'uso della teoria asintotica, evidenziati in precedenti lavori (articoli 6 e 7 in questo elenco), vengono considerati schemi di ricampionamento come il Subsampling, Jackknife e Bootstrap. L'obiettivo è quello di costruire una distribuzione empirica di probabilità che approssimi quella vera, mediante la generazione di un elevato numero di sequenze di dati artificiali, ricostruiti a partire dalla sequenza originale di dati misurati. Il contributo principale di questo lavoro è stato quello di sviluppare una nuova teoria per il metodo Subsampling, al fine di dimostrarne la consistenza e valutare contestualmente l'errore di ricostruzione della funzione di distribuzione. Quest'ultimo, è un risultato non-asintotico, aspetto rilevante al fine dell'utilizzo dello stesso in casi reali. Un confronto dei vari metodi su un esempio in simulazione è inoltre proposto al fine di evidenziarne pregi e criticità.

6. S. Garatti, M. Campi, S. Bittanti, "*The asymptotic model quality assessment for instrumental variable identification revisited*", **System & Control Letters**, **55** (6): 494:500, 2006.

In questo lavoro si è considerato il problema della valutazione di regioni di confidenza per i modelli identificati mediante tecniche basate sulla variabile strumentale (identificazione IV). In certi contesti caratterizzati da poca informazione (debole eccitazione e numero di dati limitato), la teoria asintotica dell'identificazione può portare a risultati inattendibili se usata nella pratica (tale risultato è stato mostrato per la prima volta nell'articolo 7 in questo elenco). In questo lavoro è stato messo in evidenza come nel contesto dell'identificazione mediante variabile strumentale la teoria asintotica per la valutazione del modello identificato mostri una sorta di robustezza così da fornire risultati attendibili anche in presenza di poca informazione. Tale risultato è possibile alla luce di alcuni nuovi risultati asintotici specifici dell'identificazione IV, anch'essi presentati nell'articolo in considerazione.

7. S. Garatti, M. Campi, S. Bittanti, "*Assessing the quality of identified models through the asymptotic theory – When is the result reliable?*", **Automatica (regular paper)**, **40** (8): 1319-1332, 2004.

Un tipico approccio per valutare la qualità di un modello identificato consiste nel ricorrere alla teoria asintotica dell'identificazione mediante la quale si possono trovare delle regioni di confidenza per i parametri da stimare. Nonostante la teoria asintotica sia uno degli strumenti più utilizzati in questo contesto, in questo lavoro è stato messo in luce come in certe situazioni, caratterizzate dalla presenza di poca informazione, la teoria asintotica dell'identificazione a minimizzazione dell'errore di predizione (PEM) possa portare a risultati del tutto inattendibili. Alla luce di questo fatto, nell'articolo in considerazione sono state analizzate le condizioni teoriche per la validità della teoria asintotica e sono state individuate, tra quelle standard (ARX, ARMAX, Box Jenkins, ecc.), le classi di modelli per cui la teoria asintotica dell'identificazione PEM può essere utilizzata in sicurezza per valutare la qualità di un modello

identificato. Questo risultato è stato possibile grazie ad una nuova dimostrazione dei risultati asintotici dell'identificazione PEM sotto ipotesi non standard. I risultati ottenuti sono di ampio interesse per numerose applicazioni, come ad esempio il controllo iterativo.

8. S. Garatti, M. Campi, S. Bittanti, "*Iterative robust control: speeding up improvement through iterations*", **Systems & Control Letters**, 59(2):139-146, 2010.

Negli ultimi quindici anni le tecniche di controllo iterativo sono state ampiamente studiate in letteratura come una metodologia efficace per il progetto di controllori ad alte prestazioni per sistemi dalla dinamica a priori non nota. L'idea principale è progettare il controllore mediante una sequenza di iterazioni, in ciascuna delle quali viene eseguita l'identificazione in anello chiuso del sistema e il pre-esistente controllore viene aggiornato sulla base della nuova informazione pervenuta. Tuttavia, le tecniche iterative esistenti hanno il difetto di essere molto onerose in quanto richiedono un numero eccessivo di esperimenti sull'impianto. Per questa ragione, in questo lavoro è stato proposto un nuovo metodo iterativo che tenga conto dell'incertezza durante la fase di identificazione: il controllore progettato ad ogni passo è il migliore relativamente all'attuale livello di incertezza. In questo modo è possibile spingere al massimo il miglioramento delle prestazioni del sistema di controllo ad ogni iterazione, preservando tuttavia la stabilità robusta. Lo schema iterativo ottenuto richiede pertanto solo un numero limitato di esperimenti sull'impianto e risulta essere più efficiente di quelli esistenti.

9. S. Garatti, S. Bittanti, D. Liberati, A. Maffezzoli, "*An unsupervised clustering approach for leukaemia classification based on dna micro-arrays data*", **Intelligent Data Analysis**, 11 (2): 175-188, 2007.

In questo lavoro viene presentata l'analisi mediante tecniche di data-mining dei dati relativi a pazienti affetti da leucemia. Più precisamente è stata considerata una base di dati pubblica riguardante l'attività genica dei pazienti (ovvero la produzione di proteine da parte di ogni gene di ogni singolo utente), misurata mediante la tecnologia dei DNA microarrays. Mediante metodi di partizione non supervisionata quali il PDDP (Principal Direction Divide Partitioning), è stato possibile ricostruire la presenza di due differenti tipologie di malattia nei pazienti considerati senza ricorrere all'utilizzo di ulteriori informazioni ricavate da indagini di tipo medico. Questo risultato apre nuove possibilità sull'individuazione e classificazione di malattie, mettendo contestualmente in evidenza le cause geniche delle stesse. A quest'ultimo proposito, al fine di estrarre l'informazione più significativa all'interno del database disponibile, particolare attenzione è stata rivolta alla fase preliminare di pre-processing dei dati e alla fase finale di ricostruzione dei geni responsabili della malattia.

10. S. Garatti, S.M. Savaresi, S. Bittanti, L. La Brocca, "*On the relationships between user profiles and navigation sessions in virtual communities: a data-Mining approach*", **Intelligent Data Analysis**, 8(6):579-600, 2004.

Le comunità virtuali sono gruppi di persone la cui interazione avviene mediante i servizi di comunicazione offerti da un sito Internet. L'obiettivo di questo lavoro è stato quello di determinare le relazioni esistenti tra il data-base dei profili personali degli utenti di una delle comunità virtuali più popolari in Italia e i loro percorsi di navigazione all'interno del sito, questi ultimi ricavabili dal data-base di tutti gli accessi al sito Internet della comunità. La ricerca delle relazioni esistenti tra queste due basi di dati eterogenee è in genere molto complessa a causa dell'elevata dimensionalità del problema. Per questa ragione, in questo lavoro, l'analisi dei dati è stata condotta mediante tecniche di data-mining che hanno permesso di suddividere il data-base dei profili in un insieme ridotto di tipologie di utenti e di raggruppare gli accessi al sito Internet in un insieme ridotto di categorie di sessioni di navigazione sul sito. Grazie ai risultati ottenuti, è stato quindi possibile analizzare la correlazione esistente tra le categorie di utenti e le categorie delle sessioni e, in questo modo, si è ottenuta una significativa modellizzazione (e classificazione) del comportamento degli utenti della comunità virtuale. Il lavoro è frutto della collaborazione tra il Politecnico di Milano e Tiscali SpA.

ELENCO COMPLETO DEI LAVORI SCIENTIFICI

A. Pubblicazioni su Riviste Internazionali

- [A.9] S. Garatti, R.R. Bitmead, "On resampling and uncertainty estimation in linear system identification", **Automatica (regular paper)**, **46**(5): 785-795, 2010.
- [A.8] S. Garatti, M. Campi, S. Bittanti, "Iterative robust control: speeding up improvement through iterations", **Systems & Control Letters**, **59**(2):139-146, 2010.
- [A.7] M. Campi, S. Garatti, M. Prandini, "The scenario approach for systems and control design", **Annual Reviews in Control**, **33**(2):149-157, 2009.
- [A.6] M. Campi, G. Calafiore, S. Garatti, "Interval predictor models: identification and reliability", **Automatica (regular paper)**, **45**(2): 382-392, 2009.
- [A.5] M. Campi, S. Garatti, "The exact feasibility of randomized solution of uncertain convex programs", **SIAM Journal on Optimization**, **19**(3): 1211-1230, 2008.
- [A.4] S. Garatti, S. Bittanti, D. Liberati, A. Maffezzoli, "An unsupervised clustering approach for leukaemia classification based on dna micro-arrays data", **Intelligent Data Analysis**, **11** (2): 175-188, 2007.
- [A.3] S. Garatti, M. Campi, S. Bittanti, "The asymptotic model quality assessment for instrumental variable identification revisited", **System & Control Letters**, **55** (6): 494:500, 2006.
- [A.2] S. Garatti, M. Campi, S. Bittanti, "Assessing the quality of identified models through the asymptotic theory – When is the result reliable?", **Automatica (regular paper)**, **40** (8): 1319-1332, 2004.
- [A.1] S. Garatti, S.M. Savaresi, S. Bittanti, L. La Brocca, "On the relationships between user profiles and navigation sessions in virtual communities: a data-Mining approach", **Intelligent Data Analysis**, **8**(6):579-600, 2004.

B. Capitoli di libri internazionali

- [B.2] M. Prandini, M. Campi, S. Garatti, "Controller design through random sampling: an example", in: "**Advance in control theory and applications**" (eds. C. Boniventi, A. Isidori, L. Marconi, C. Rossi) – Springer Verlag, LCNIS series, 2007.
- [B.1] D. Liberati, S. Garatti, S. Bittanti, "Unsupervised mining of genes classifying Leukaemia", in: "**Encyclopedia of Data Warehousing and Mining**" (ed. J. Wang) – Idea Group Publishing, 2005.

C. Pubblicazioni su Atti di Congressi e Seminari Internazionali

- [C.17] S. Garatti, R.R. Bitmead, "On re-sampling and uncertainty estimation in linear system identification", **Proceedings of the 15th IFAC Symposium on System Identification (SYSID)**, Saint-Malo, France, 2009.
- [C.16] S. Garatti, S. Bittanti, "Parameter estimation in the Pacejka's tyre model through the TS method", **Proceedings of the 15th IFAC Symposium on System Identification (SYSID)**, Saint-Malo, France, 2009.
- [C.15] S. Garatti, M.C. Campi, "L-inf layers and the probability of false prediction", **Proceedings of the 15th IFAC Symposium on System Identification (SYSID)**, Saint-Malo, France, 2009.
- [C.14] S. Bittanti, S. Garatti, "Revisiting the basic issue of parameter estimation in system identification – a new approach for multi-value estimation", **Proceedings of the 47th IEEE Conference on Decision and Control**, Cancun, Mexico, 2008.
- [C.13] S. Garatti, S. Bittanti, "Estimation of white-box model parameters via artificial data generation: a two-stage approach", **Proceedings of the 17th IFAC world congress**, Seoul, Korea, 2008.

- [C.12] M. Campi, S. Garatti, M. Prandini, "*The scenario approach for systems and control design*", **Proceedings of the 17th IFAC world congress** (stampato inoltre nel volume "Plenary Papers, Milestone Reports & Selected Survey Papers", pp. 180-188), Seoul, Korea, 2008.
- [C.12] S. Garatti, S. Bittanti, "*Parameter estimation via artificial data generation with the 'two-stage' approach*", **Proceedings of the 7th World Congress on Intelligent Control and Automation**, Chongqing, China, 2008.
- [C.10] M. Campi, S. Garatti, "*Modulating robustness in robust control: making it easy through randomization*", **Proceedings of the 46th IEEE Conference on Decision and Control**, New Orleans, Louisiana USA, 2007.
- [C.9] S. Bittanti, S. Garatti, M. Campi, "*Introducing robustness in iterative control*", **Proceedings of the 44th IEEE Conference on Decision and Control**, Seville, Spain, 2005.
- [C.8] M. Campi, G. Calafiore, S. Garatti, "*New results on the identification of interval predictor models*", **Proceedings of the 16th IFAC world congress**, Prague, Czech Republic, 2005.
- [C.7] S. Bittanti, S. Garatti, D. Liberati, "*From DNA microarrays to disease classification: an unsupervised clustering approach*", **Proceedings of the 16th IFAC world congress**, Prague, Czech Republic, 2005.
- [C.6] S. Bittanti, M. Campi, S. Garatti, "*Some critical implementation issues in iterative robust control design*", **Proceedings of the 8th IFAC workshop on Adaptation and Learning in Control and Signal Processing**, Yokohama, Japan 2004.
- [C.5] S. Garatti, S. Bittanti, M. Campi, "*Model quality assessment for Instrumental Variable methods: use of the asymptotic theory in practice*", **Proceedings of the 42nd IEEE Conference on Decision and Control**, Maui, Hawaii USA 2003. pp. 6015-6020.
- [C.4] S. Savaresi, S. Garatti, S. Bittanti, "*Modeling the relationships between the user DB and the web-log file of a large virtual community*", **Proceedings of the 13th IFAC Symposium on System Identification**, Rotterdam, The Netherlands 2003. pp. 219-224.
- [C.3] S. Savaresi, S. Garatti, S. Bittanti, L. La Brocca, "*Data-Mining of a large virtual community: relationships between the users DB and the web-log file*", **Proceedings of the 3rd SIAM International Conference on Data Mining**, San Francisco, California USA 2003. pp. 299-303.
- [C.2] S. Bittanti, M. Campi, S. Garatti, "*New results on the asymptotic theory of system identification for the assessment of the quality of estimated models*", **Proceedings of the 41st IEEE Conference on Decision and Control**, Las Vegas, Nevada USA 2002. pp. 1814-1819.
- [C.1] S. Bittanti, M. Campi, S. Garatti, "*An iterative controller design scheme based on average robust control*", **Proceedings of the 15th IFAC World Congress**, Barcelona, Spain 2002.

D. Pubblicazioni su Riviste Nazionali

- [D.1] D. Liberati, S. Bittanti, S. Garatti, Z. Zhao, M. Pappalettera, "*Classificazione di leucemie mediante analisi di dati da microarray*", **Automazione e Strumentazione LIII(10):79-86**, 2005.

E. Pubblicazioni su archivi in Internet, rapporti tecnici e relazioni

- [F.1] M. Campi, S. Garatti, "*Chance-constrained optimization via randomization: feasibility and optimality*", **Optimization Online** (http://www.optimization-online.org/DB_HTML/2008/09/2092.html), 2008.
- [F.2] S. Garatti, "*Finding relationships between the users DB and the web-log file of a large Internet community: a data mining approach*", **Rapporto Interno N° 2002.48** – Politecnico di Milano, DEI, 2002.