Advances in Computer Vision and Pattern Recognition

For further volumes:
http://www.springer.com/series/4205

Chris Aldrich • Lidia Auret

Unsupervised Process Monitoring and Fault Diagnosis with Machine Learning Methods

Springer

Chris Aldrich
Western Australian School of Mines
Curtin University
Perth, WA, Australia

Department of Process Engineering
University of Stellenbosch
Stellenbosch, South Africa

Lidia Auret
Department of Process Engineering
University of Stellenbosch
Stellenbosch, South Africa

Series Editors
Sameer Singh
Research School of Informatics
Loughborough University
Loughborough
UK

Sing Bing Kang
Microsoft Research
Microsoft Corporation
Redmond, WA
USA

ISSN 2191-6586 ISSN 2191-6594 (electronic)
ISBN 978-1-4471-5184-5 ISBN 978-1-4471-5185-2 (eBook)
DOI 10.1007/978-1-4471-5185-2
Springer London Heidelberg New York Dordrecht

Library of Congress Control Number: 2013942259

Printed on acid-free paper

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

Preface

Although this book is focused on the process industries, the methodologies discussed in the following chapters are generic and can in many instances be applied with little modification in other monitoring systems, including some of those concerned with structural health monitoring, biomedicine, environmental monitoring, the monitoring systems found in vehicles and aircraft and monitoring of computer security systems. Of course, the emphasis would differ in these other areas of interest, e.g. dynamic process monitoring and nonlinear signal processing would be more relevant to structural health analysis and brain–machine interfaces than techniques designed for steady-state systems, but the basic ideas remain intact. As a consequence, the book should also be of interest to readers outside the process engineering community, and indeed, advances in one area are often driven by application or modification of related ideas in a similar field.

In a sense, the area of process monitoring and the detection and analysis of change in technical systems are an integral part of the information revolution, as the use of data-driven methods to construct the requisite process or systems models becomes dominant over first-principle or higher knowledge approaches. This revolution has changed the world as we know it and will continue to do so in as yet unforeseen ways.

Rightly or wrongly, there is a perception that the mining engineering environment is conservative as far as research spending is concerned, reluctant to embrace future technologies that do not have an immediate proven impact on the bottom line, also as far as process automation is concerned. However, this is rapidly changing, with large mining companies investing considerable sums of money in the development of advanced process automation systems with no immediate benefit. These new automation systems will have to sense changes in their environment and be able to react to these changes, consistently, safely and economically. Apart from the development of advanced sensors, process monitoring technologies would play a central role in the success of these automated mining systems. For example, in underground mining, these systems would have to be able to differentiate between mineral and the surrounding gangue material in real time or be able to differentiate between solid rock and rock that might be on the verge of collapse in a mining tunnel. Humans have mixed success in these tasks, and current automation systems are too rudimentary to improve on this.

These new diagnostic systems would have to cope with the so-called *Big Data* phenomenon, which will inevitably also have an impact on the development and implementation of the analytical techniques underpinning them. In many ways, *Big Data* can simply be seen as more of the same, but it would be unwise to see it simply as a matter that can be resolved by using better hardware. With large complex data sets, the issues of automatically dealing with unstructured data, which may contain comparatively little useful information, become paramount. In addition, these data streams are likely to bring with them new information not presently available, in ways that are as yet unforeseen. Just like video data can simply be seen as a series of images, if taken at a sufficiently high frequency, these data can reveal information on the dynamic behaviour of the system that a discontinuous series of snapshots cannot. It is easy to see that in some cases this could make a profound difference on our understanding of the behaviour of the system.

In the same way that *Big Data* can be seen as data, just more of it, machine learning can arguably be seen as statistics, simply in a different guise, as in many ways it is without a doubt. However, looking into the future, as systems rapidly grow in complexity, the ability of machines to truly learn could also be influenced in unforeseen ways. By analogy, one could consider a novice chess player, who has learnt the rules of chess and knows how to detect direct threats to his individual pieces on the board. However, it is only by experience that he learns to recognize the unfolding of more complex patterns or emergent behaviour that would require timely action to avoid or exploit.

Perth, WA, Australia Chris Aldrich

Acknowledgements

In many ways, this book can be regarded as a product of the Anglo American Platinum Centre for Process Monitoring and the research work of a large number of postgraduate students that have passed through the Process Systems Engineering group at Stellenbosch University over the last decade or more. The collaboration between academia and industry has been especially productive in this respect.

Our special thanks therefore to Dr. J.P. Barnard and Ms. Corné Yzelle for making available the Centre's *Process Diagnostics Toolset* software without which the methods outlined in Chap. 6 in the book could not have been implemented.

In addition, we would also like to express our sincere gratitude to Dr. Gorden Jemwa, not only for his contributions to the Process Systems Engineering group over many years but also specifically for his major contribution as main author of Chap. 8 in the book.

Finally, it may be a cliché, but it does not make it less true that a book like this does not write itself, and the authors would like to make use of this opportunity to thank their families and friends for their understanding and active support in this respect.

Chris Aldrich and Lidia Auret

Contents

Acronyms

Acronym	Description
ACF	Autocorrelation function
ADALINE	Adaptive linear element
AHPCA	Adaptive hierarchical principal component analysis
AID	Automatic interaction detection
AKM	Average kernel matrix
AMI	Average mutual information
AR	Autoregressive
ARL	Alarm run length
ARMA	Autoregressive moving average
ARMAX	Autoregressive moving average with exogenous variables
AUC	Area under curve
BDKPCA	Batch dynamic kernel principal component analysis
BDPCA	Batch dynamic principal component analysis
BZ	Belousov–Zhabotinsky
CART	Classification and regression trees
CHAID	Chi-square automatic interaction detection
COW	Correlation optimized time warping
CSTR	Continuous stirred tank reactor
CUSUM	Cumulative sum
CVA	Canonical variate analysis
DD	Detection delay
DICA	Dynamic independent component analysis
DISSIM	Dissimilarity
DKPCA	Dynamic kernel principal component analysis
DPCA	Dynamic principal component analysis
DTW	Dynamic time warping

(continued)

(continued)

Acronym	Description
EEMD	Ensemble empirical mode decomposition
ELM	Extreme learning machine
EMD	Empirical mode decomposition
EWMA	Exponentially weighted moving average
FAR	False alarm rate
FS	Feature samples
ICA	Independent component analysis
INLPCA	Inverse nonlinear principal component analysis
IOHMM	Input–output hidden Markov model
JITL	Just-in-time learning
k-DISSIM	Kernel dissimilarity
KICA	Kernel independent component analysis
KKT	Karush–Kuhn–Tucker
KPCA	Kernel principal component analysis
KPLS	Kernel partial least squares
LCL	Lower control limit
MA	Moving average
MADALINE	Multiple adaptive linear element
MAID	Multiple or modified automatic interaction detection
MAR	Missing alarm rate
MCEWMA	Moving centre exponentially weighted moving average
MEB	Minimum enclosing ball
MHMT	Multi-hidden Markov tree
MICA	Multiway independent component analysis
MKICA	Multiscale kernel independent component analysis
MPCA	Multiway principal component analysis
MPLS	Multiway partial least squares
MSDPCA	Multiscale dynamic principal component analysis
MSE	Mean square error
MSKPCA	Multiscale kernel principal component analysis
MSPC	Multivariate statistical process control
MSSA	Multichannel singular spectrum analysis
MSSPCA	Multiscale statistical process control
MSSR	Mean sum of squared residuals
MVU	Maximum variance unfolding
MVUP	Maximum variance unfolding projection
NIPS	Neural information processing systems

(continued)

(continued)

Acronym	Description
NLPCA	Nonlinear principal component analysis
NN	Neural network
NOC	Normal operating conditions
OOB	Out of bag
PAC	Probably approximately correct
PCA	Principal component analysis
PDPCA	Partial dynamic principal component analysis
PLS	Partial least squares
RBM	Restricted Boltzmann machine
RF	Random forest
ROC	Receiver operating curve
RQA	Recurrence quantification analysis
SBKM	Single batch kernel matrix
SI	Subspace identification
SOM	Self-organizing map
SPC	Statistical process control
SPE	Squared prediction error
SPM	Statistical process monitoring
SSA	Singular spectrum analysis
SSICA	State space independent component analysis
SVD	Singular value decomposition
SVDD	Support vector domain description
SVM	Support vector machine (1-SVM one class SVM)
SVR	Support vector regression
TAR	True alarm rate
THAID	Theta automatic interaction detection
TLPP	Tensor locality preserving projection
UCL	Upper control limit
VARMA	Vector autoregressive moving average
VC	Vapnik–Chervonenkis