Expert Systems with Applications xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A robust unsupervised consensus control chart pattern recognition framework

Siavash Haghtalab [a], Petros Xanthopoulos [a,*], Kaveh Madani [b]

[a] Department of Industrial Engineering and Management Systems, University of Central Florida, 4000 Central Florida blvd., Orlando, FL 32816, USA
[b] Centre for Environmental Policy, Imperial College London, 14 Prince's Gardens, London SW7 2AZ, UK

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Consensus clustering
Unsupervised learning
Control chart pattern recognition
k-Means
Spectral clustering
Graph partitioning

ABSTRACT

Early identification and detection of abnormal patterns is vital for a number of applications. In manufacturing for example, slide shifts and alterations of patterns might be indicative of some production process anomaly, such as machinery malfunction. Usually due to the continuous flow of data, monitoring of manufacturing processes and other types of applications requires automated control chart pattern recognition (CCPR) algorithms. Most of the CCPR literature consists of supervised classification algorithms. Fewer studies consider unsupervised versions of the problem. Despite the profound advantage of unsupervised methodology for less manual data labeling their use is limited due to the fact that their performance is not robust enough and might vary significantly from one algorithm to another. In this paper, we propose the use of a consensus clustering framework that takes care of this shortcoming and produces results that are robust with respect to the chosen pool of algorithms. Computational results show that the proposed method achieves not less than 79.10% *G*-mean with most of test instances achieving higher than 90%. This happens even when in the algorithmic pool are included algorithms with performance less than 15%. To our knowledge, this is the first paper proposing an unsupervised consensus learning approach in CCPR. The proposed approach is promising and provides a new research direction in unsupervised CCPR literature.

1. Introduction

Time series analysis is an area of research with numerous application in many fields of science and engineering (Box, Jenkins, & Reinsel, 2013). In manufacturing, for instance, time series pattern recognition is important since slide alterations might be indicative of a malfunction that requires a course of appropriate corrective actions (e.g. maintenance). Manual monitoring is tedious and requires specialized personnel's undistracted attention. For this, machine learning based automated algorithms, also known as control chart pattern recognition (CCPR) algorithms, have been proposed to detect abnormal behaviors. The term was originally coined by Shewhart (1931). An early taxonomy of the patterns was presented in an early publication of the western electric company (Company, 1958). Fig. 1 depicts six of the most common abnormal patterns studied in the literature.

These different abnormal patterns are usually related to a specific malfunction and their early detection can provide useful insights for corrective actions and thus improve systems reliability. In the crank case manufacturing operations, up trend and down trend patterns reveal tool wear and malfunction (El-Midany, El-Baz, & Abd-Elwahed, 2010a). Shift patterns might be associated with variation related to operator, material or machine instrument (Davy, Desobry, Gretton, & Doncarli, 2006; El-Midany et al., 2010a). Cyclic patterns are associated with voltage variability (Kawamura, Chuarayapratip, & Haneyoshi, 1988) but they can also appear in manufacturing processes like frozen orange juice packing (Hwarng, 1995). In the car manufacturing industry certain anomalies in the automotive body assembly process appear as up/down trends, cyclic, and systematic patterns (Jang, Yang, & Kang, 2003). Up/down trend patterns can be used in order to detect abnormal stamping tonnage signals (Jin & Shi, 2001). Finally up/down trend signals appear in paper making industry (Chinnam, 2002; Cook & Chiu, 1998) whereas uptrend patterns by itself can be used for detecting fault states in end-milling process (Zorriassatine, Al-Habaibeh, Parkin, Jackson, & Coy, 2005).

During several years, different pattern recognition algorithms have been studied in the literature with the proposed approaches ranging over a broad spectrum of machine learning algorithms. The majority of the proposed schemes follow the supervised

* Corresponding author.
E-mail addresses: s.haghtalab@gmail.com (S. Haghtalab), petrosx@ucf.edu (P. Xanthopoulos), k.madani@imperial.ac.uk (K. Madani).

http://dx.doi.org/10.1016/j.eswa.2015.04.069
0957-4174/© 2015 Published by Elsevier Ltd.

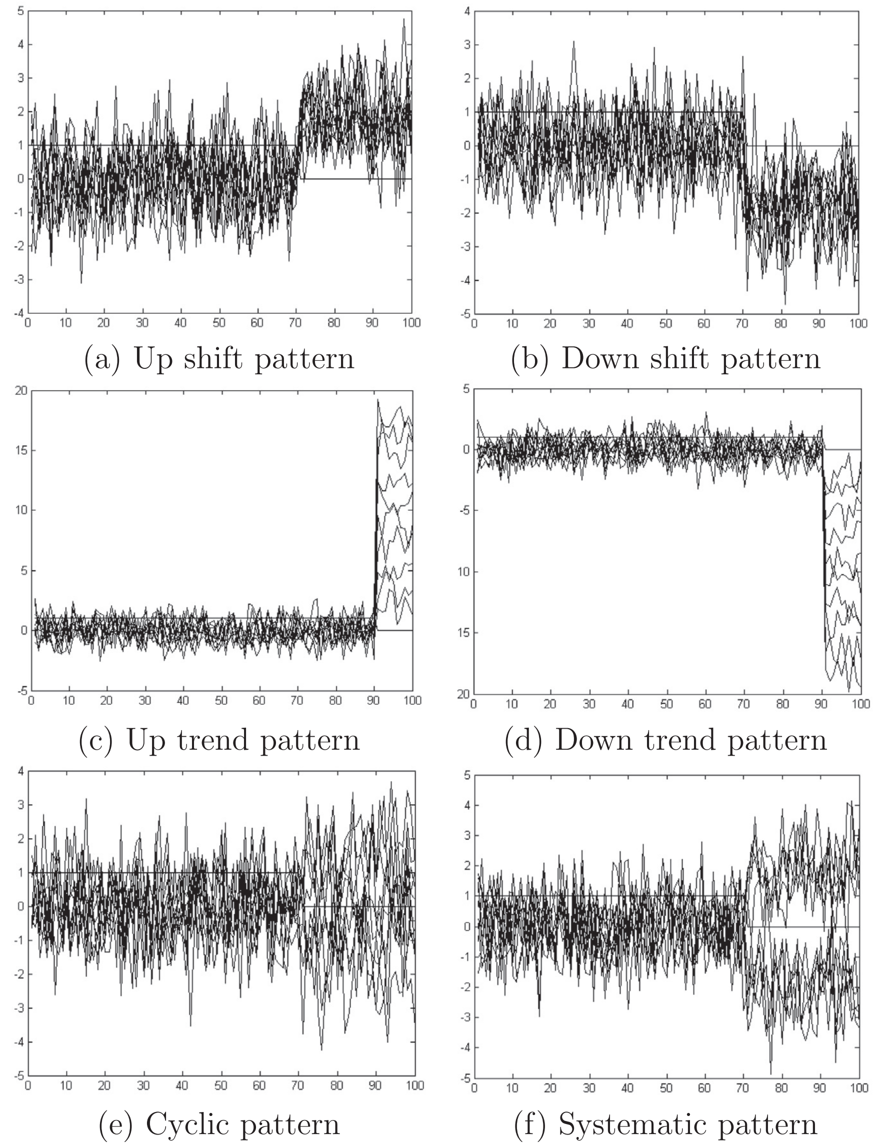

Fig. 1. Examples of six basic abnormal patterns.

learning framework, in which a model is trained with a historical dataset and then the trained model is used for prediction on an unknown testing data set. Some representative algorithms under this category include knowledge-based expert systems and artificial neural networks (El-Midany, El-Baz, & Abd-Elwahed, 2010b; Hwarng, 1995; Hwarng & Hubele, 1992, 1993a, 1993b; Guh & Hsieh, 1999; Kim, Jitpitaklert, Park, & Hwang, 2012; Perry, Spoerre, & Velasco, 2001; Wu & Yu, 2010; Yu & Xi, 2009), Bayes classification (Adam et al., 2011), and support vector Machines (SVM) (Camci, Chinnam, & Ellis, 2008). In more recent literature decomposition techniques are used as a preprocessing step before classification. Some examples include wavelets (Du, Huang, & Lv, 2013), independent component analysis (Cheng & Huang, 2013; Kao, Lee, & Lu, 2014) and extreme-point symmetric mode decomposition (Yang, Zhou, Liao, & Guo, 2015). In another recent study Wu, Liu, and Zhu (2014) proposed the combined approach of classification trees and SVM. For a comprehensive literature review we refer the reader to Hachicha and Ghorbel (2012) and Veiga, Mendes, and Lourenço (2015).

On the other hand, research on unsupervised CCPR algorithms is relatively limited. Unsupervised learning assumes no prior information and aims to categorize the data samples based only on their features (properties) (Warren Liao, 2005). The first unsupervised approach for CCPR was proposed by Al-Ghanim (1997) who developed an unsupervised self-organizing neural paradigm. Al-Ghanim and Kamat (1995) presented a CCPR technique using correlation analysis on trend, systematic and cyclic patterns and presented results with evaluation methods. Wang and Kuo (2007) used three different fuzzy clustering algorithms on CCPR for six patterns and compared their performance.

Unsupervised learning techniques have the profound advantage of not requiring prior labeling knowledge for prediction. On the other side, however, their behavior can be instable and sometimes inconsistent across algorithms or even across different runs of the same algorithm. In the clustering literature this shortcoming is normally addressed through *ensemble* or *consensus* learning schemes. Under this approach a number of clustering with different results is combined to a single clustering that is more robust according to some optimization criteria (Vega-Pons & Ruiz-Shulcloper, 2011; Xanthopoulos, 2014). However this idea has not been implemented yet for the CCPR problem. We anticipate that consensus framework will provide CCPR robust methodologies

Please cite this article in press as: Haghtalab, S., et al. A robust unsupervised consensus control chart pattern recognition framework. *Expert Systems with Applications* (2015), http://dx.doi.org/10.1016/j.eswa.2015.04.069