Preprints of the 18th IFAC World Congress
Milano (Italy) August 28 - September 2, 2011

# Improved Correlation Analysis and Visualization of Industrial Alarm Data

**Fan Yang*,**, Sirish L. Shah*, Deyun Xiao**, Tongwen Chen*****

**Dept. of Chemical & Materials Engineering, University of Alberta, Edmonton AB T6G 2G6 Canada (Tel: 780-492-5162; e-mail: {fyang.cme, Sirish.Shah}@ualberta.ca).*
***Dept. of Automation, Tsinghua University, Beijing 100084 China (e-mail: {yangfan, xiaody}@tsinghua.edu.cn).*
****Dept. of Electrical & Computer Engineering, University of Alberta, Edmonton AB T6G 2V4 Canada (e-mail: tchen@ece.ualberta.ca).*

**Abstract:** The problem of multivariate alarm analysis and rationalization is complex and important in the area of smart alarm management due to the interrelationships between variables. Capturing and visualizing the correlation information, especially from historical alarm data directly, is beneficial for further analysis. In this paper, the Gaussian kernel method is applied to generate pseudo continuous time series from the original binary alarm data. This can reduce the influence of missed, false and chattering alarms. By taking into account the time lag between alarm variables, a correlation color map of the transformed or pseudo data is used to show the cluster of correlated variables with the alarm tags reordered to better group the correlated alarms. Thereafter statistical methods such as singular value decomposition techniques can be applied within each cluster to find the redundant alarm tags. This improved method is shown to be better than the alarm similarity color map when applied in the analysis of industrial alarm data.

## 1. INTRODUCTION

During routine operation of industrial systems, abnormal situations show up in the control room as alarms. Most processes, however, have too many alarm tags configured on process variables and other state variables mainly due to safety consideration. The number of alarmed tags is large enough to overwhelm even experienced operators. Because of this situation, alarm management has been recognized as an important problem in the area of system monitoring and fault detection (Izadi *et al.*, 2009). New and revised guidelines and standards have been proposed from different viewpoints to tackle this problem (Holifield & Habibi, 2007; EEMUA, 2007; ISA, 2009).

The focus of many current projects in industry is on 'alarm rationalization', that is, to reduce the number of alarms and yet at the same time be able to detect all potential abnormal situations. Most alarms are analyzed in a univariate framework. However it is well known that most processes are multivariate and thus the alarms are not independent. Therefore a good strategy for the analysis and visualization of alarm data should be based on a multivariate framework. This type of multivariate information can be easily captured from process data and multivariate statistics can be applied to generate efficient alarm strategy (Kondaveeti *et al.*, 2009) or optimize alarm limits (Yang *et al.*, 2010a). Such methods can be categorized as indirect methods because we resort to the process data only, i.e., measured values of process variables in engineering units. However, one may have many alarm tags associated with a single process variable and often alarm tags are without any association with any specific process variables (e.g. digital alarms to indicate a specific event or situation). All of these necessitate newer methods to mine the alarm data directly.

There are two ways to capture correlation from alarm data: one is to employ Pearson's correlation coefficients as done for continuous data (Yang *et al.*, 2010a); the other is to introduce similarity measures based on binary data (Kondaveeti, 2010). By computing the correlation coefficient or similarity measure for each pair of variables, a matrix is constructed. However, this matrix which is composed of specific numbers is not generally convenient for inspection or cursory analysis. Therefore the correlation data is converted into a colour map and this correlation color map offers better visualization capability; it uses different colors to show different degrees of correlation (Tangirala *et al.*, 2005; Zhang, 2000). The colors are usually discretized into several levels according to the color code. Through the reordering and clustering of variables, the color-coded correlation map can show the clusters of alarm tags intuitively. In this way, the problem of a large number of variables is separated into smaller sub-problems with much fewer variables.

The classic correlation color map only uses the correlation coefficients. It may be ineffective when there is a time lag between two variables. Hence the correlation coefficients should be lag-adjusted to take into account the delays between each pair of variables. The alarm similarity color map (ASCM), which is specially designed for alarm data analysis (Kondaveeti *et al.*, 2010), has proved to be an effective method for visualization. It captures time-delayed similarity information from the binary data. However it has some disadvantages. Firstly, in order to weaken the sensitivity caused by the time shift, each unique alarm is padded with extra 1's to enrich the data. This step is short of physical evidence and as result of this individual false alarms and chattering alarms may be magnified unreasonably. In this paper we suggest another method in which each unique alarm is replaced by a Gaussian distribution along with its neigh-

bors (Control Arts Inc., 2010). It has some similar problems as the padding method; however, it provides a better way of transforming binary alarm data into continuous data that can be analyzed by statistical approaches. Another disadvantage of the ASCM method is that it takes into account all the alarm tags separately regardless their physical relationship. If typical analog alarms (HI/LO/HH/LL) associated with the same process variable are grouped together, the result should be expected to be more reasonable. Thirdly, the similarity measure only indicates the distance but loses the direction (positive or negative correlation), while the correlation coefficient indicates both the similarity and the direction.

This paper is organized as follows. The basic method for correlation analysis using the kernel method and the visualization by color maps are proposed in Sections 2 and 3 respectively. Section 4 discusses implementation issues. A case study is given in Section 5 followed by concluding remarks.

## 2. BASIC METHOD FOR CORRELATION ANALYSIS

### *2.1 Gaussian Kernel Method for Data Preprocessing*

Alarm data are binary data with '1' and '0' to denote abnormal and normal states respectively. Since alarm data are obtained and discretized from continuous process data according to alarm limits, some information is lost in this process. However, alarm data do have advantages as well: 1) binary data are easy for storage and convenient for statistical analysis; 2) they have much higher sampling frequencies than process data and thus include detailed information; and 3) they include more types of data such as digital alarms showing the status of some elements. Thus we can use alarm data directly to analyze the correlation.

For a particular variable, one alarm point can be regarded as a sample of the time series. In order to estimate the time series, the kernel method can be used in the temporal domain (Silverman, 1986), which is a nonparametric method, to fit the function with any shape. Here the Gaussian kernel function is used because of the smoothness. At each alarm point, a Gaussian kernel function is superimposed around this time instant. The function is defined as:

$$K(t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{t^2}{2\sigma}} \tag{1}$$

where $\sigma$ is the standard deviation. If all the alarm points are considered, a continuous time series can be obtained as the superposition of all the time-shifted Gaussian kernel functions. Thus the resulting time series is given as

$$P(t) = \sum_{i=1}^{N} K(t - t_i) \tag{2}$$

where $N$ is the number of alarm points at the time instants $t_i$, $i$=1,2,…,$N$. This time series $P(t)$ can be regarded as the estimation of the corresponding process data although there may not exist such a physical process variable; hence we call it pseudo data. Fig. 1 illustrates the principle of the generation.

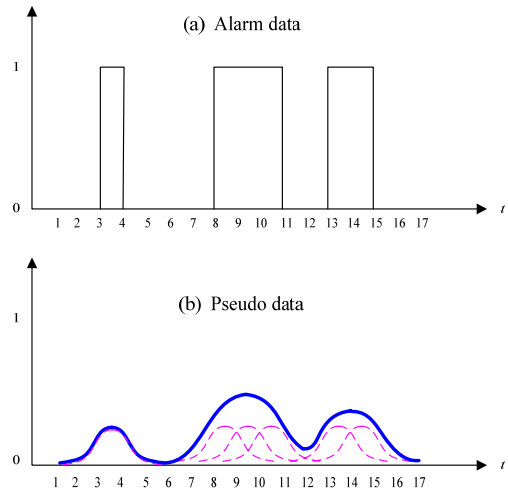



Fig. 1. Pseudo data generation from alarm data.

The pseudo data have the following properties:

- It is continuous and smooth because it is approximated by superimposed Gaussian functions.
- For consecutive alarm points lasting a long time, the pseudo data are also 1's because the integral of the Gaussian kernel function over the whole time domain is 1. This is different from the estimation of probability density function because the samples here are at different time instants.
- With appropriate variance of the kernel function, the magnitude of the kernel function is small; thus non-consecutive and sparse alarm points cannot result in spikes in the pseudo data. This property is particularly important for the filtering of chattering alarms so that the proposed method can be used directly on the data set with chattering alarms.

Consider the example of alarm data as shown in Fig. 2 (a) in which there is a consecutive alarm period of 400 seconds including a short (5 seconds) break at around the 510th interval. At both the beginning and the end of this period, there is some chattering. From this data set, we can generate the pseudo data with standard deviation of 30 seconds as shown in Fig. 2(b). Here the chattering and the gap are sufficiently smoothed and at the same time the alarm period prevails. The effects of missed alarms and false alarms are lessened because they usually only exist over a short time duration.

### *2.2 Effects of Time Lags*

There may exist a time lag between two correlated alarms due to the propagation or response time. This lag reduces the correlation coefficient and thus may mask the correlation between these variables. To eliminate this influence, different time lags should be assumed and the maximal correlation coefficient can be computed; this would be regarded as the real correlation (Bauer & Thornhill, 2008).

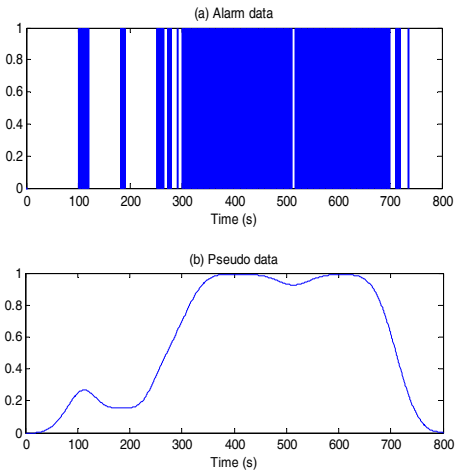



Fig. 2. Original alarm data and generated pseudo data.

Assume that $x$ and $y$ are time series of $n$ observations with means $\mu_x$, $\mu_y$ and variances $\sigma_x$, $\sigma_y$ respectively, then the cross-correlation function (CCF) with an assumed lag $k$ is:

$$\phi_{xy}(k) = \frac{\mathrm{E}[(x_i - \mu_x)(y_{i+k} - \mu_y)]}{\sigma_x \sigma_y}, \; k = -n+1, \cdots, n-1 \qquad (3)$$

The expectation can be estimated as the sample CCF by:

$$\hat{\phi}_{xy}(k) = \begin{cases} \dfrac{1}{n-k}\displaystyle\sum_{i=1}^{n-k}(x_i - \mu_x)(y_{i+k} - \mu_y) \Big/ \sigma_x \sigma_y, & k \geq 0 \\ \dfrac{1}{n+k}\displaystyle\sum_{i=1-k}^{n}(x_i - \mu_x)(y_{i+k} - \mu_y) \Big/ \sigma_x \sigma_y, & k < 0 \end{cases} \qquad (4)$$

A value of the CCF is obtained by assuming a certain time lag for one of the time series. Thus the absolute maximum value can be regarded as the real cross-correlation and the corresponding lag as the estimated time lag between these two variables. For mathematical description, one can compute the maximum and minimum values $\phi^{\max} = \max_k \{\phi_{xy}(k), 0\} \geq 0$ and $\phi^{\min} = \min_k \{\phi_{xy}(k), 0\} \leq 0$, and the corresponding arguments $k^{\max}$ and $k^{\min}$. Then the time delay from $x$ to $y$ is:

$$\lambda = \begin{cases} k^{\max}, \text{ if } \phi^{\max} \geq -\phi^{\min} \\ k^{\min}, \text{ if } \phi^{\max} < -\phi^{\min} \end{cases} \qquad (5)$$

(corresponding to the maximum absolute value) and the actual time delayed cross-correlation is $\rho = \phi_{xy}(\lambda)$ (between -1 and 1). If $\lambda$ is less than zero, then it means that the actual delay is from $y$ to $x$. Thus the sign of $\lambda$ provides the directionality information between $x$ and $y$. The sign of $\rho$ indicates whether the correlation is positive or negative. Note that this directionality does not mean causality because there may exist another common cause of the relationship between $x$ and $y$ (Yang *et al.*, 2010b).

### *2.3 Finding Redundant Alarms Based on Correlation*

If several alarms are highly correlated, we should check if there is redundancy in them, i.e., if some alarms can be obtained by linearly combining other alarm signals. Singular value decomposition (SVD) is a well-developed method to do this by finding the singular values of the set of alarm data and thus enabling one to identify colinear columns. If there are several large singular values that possess the dominant proportion of all the values, then the number of large or dominant singular values can be regarded as the number of independent alarm tags and the residual number is the number of redundant alarms. However, SVD cannot tell us which alarm is redundant because we are dealing with the transformed or pseudo data to generate a set of new data. It only provides an opportunity for one to check if there is an alarm that can be removed or replaced by the linear combination of other alarm signals. In addition, the independent series in the new data may not have the approximate binary property, making these values inappropriate to be taken as generated alarm signals. One should also consider the physical meaning of all alarms. This ambiguity in using SVD is much severe if the number of alarms is large. In this case, it is very hard to identify the independent and redundant alarms merely based on the result of SVD. Therefore, we need to separate the variables into several groups and perform SVD on each group of variables. The clustering process based on visualizing the result of the correlation analysis would help one in only analyzing SVDs of smaller group of variables.

## 3. VISUALIZATION OF CORRELATION

In order to visualize the correlation matrix, the correlation color map is developed, in which the order of variables is rearranged in a ranked order where variables that are highly correlated with each other appear together with prominent colors or shades, called clusters.

### *3.1 Ordering and Clustering of Alarm Tags*

Based on the correlation coefficient, the distance measure between two alarms is called the similarity measure $S$ (Lesot *et al.*, 2009) which has the properties of positivity ($S(x,y) \geq 0$), symmetry ($S(x,y)=S(y,x)$), and maximality ($S(x,x) \geq S(x,y)$). There are many similarity measures available for binary data (Choi *et al.*, 2010); hence if we compute the similarity measures based on original alarm data, we can choose one of them, for example Kondaveeti *et al.* (2010) have used the Jaccard similarity measure. However we are dealing with pseudo data; so generally we can use the covariance, Euclidean distance (L2), Kendall's $\tau$, Pearson's correlation coefficient, Spearman's rank, City-block (L1), etc (Wu *et al.*, 2010). In this paper, we use Pearson's correlation coefficient and define the following measure or similarity $S$:

$$S(x,y) = 1 - \left|\rho_{xy}\right| \qquad (6)$$

which lies in the interval [0,1].

During clustering, one should employ both the similarity measure between two alarms and also the measure between two clusters of alarms should be employed. The latter can be defined by single-, complete-, and average-linkage methods (Wu *et al.*, 2010). The definition of single-linkage is

$$S_S(X,Y) = \min_{x \in X, y \in Y} \{S(x,y)\} \quad (7)$$

where $x$ and $y$ are alarm tags, and $X$ and $Y$ are clusters.

Using the similarity measure between each pair of alarms or their clusters, all alarms can be clustered based on various clustering algorithms such as agglomerative hierarchical clustering, the methodology for which is shown as a dendrogram. This process has been illustrated in (Kondaveeti et al., 2010). Other ordering and clustering algorithms can also be used such as the ellipse ordering. For this purpose, the data analysis software GAP is a useful tool (Wu *et al.*, 2010).

### *3.2 Correlation Color Map*

The correlation matrix with rearranged rows and columns is color coded to transform it into a color map (Tangirala *et al.*, 2005). The matrix is symmetric and both the vertical and horizontal axes are alarm tag names. Each grid ($x$,$y$) shows the corresponding correlation between the two alarm tags $x$ and $y$. The color is coded to show the correlation (-1 to +1) or its absolute (0 to 1). To explain the meaning of each color, a color bar legend is placed beside the plot. The order of the alarm tags is determined by the above algorithm. The color map is symmetric about the diagonal which means the self-correlation coefficients are 1's. The clusters are shown as blocks located along the diagonal with similar color codes. From this map, one can acquire the approximate correlation between each pair of alarms, and easily find clusters and the corresponding alarm tags. By matching this map with the physical meaning, one can locate the possible redundant alarms.

## 4. IMPLEMENTATION ISSUES

In the above methods there are several issues to be noted.

### *4.1 Computational Effort*

Although the number of the samples is generally large due to the high frequency of sampling rate, the alarm data set includes large amount of normal data resulting in the data set being quite sparse. In addition, when an alarm occurs consecutively, the corresponding pseudo data is a set of consecutive 1's according to the property mentioned earlier in Section 2; such periods can be ignored in the computation by letting them unchanged. Therefore the computation only concentrates on the inconsecutive alarm points and the boundary of the consecutive alarm points.

In SVD, the vectors of the pseudo data can be very long, making the computation infeasible. We can either lower the sampling rate, or use the technique of sparse matrix analysis (Pissanetzky, 1984).

### *4.2 Sampling Rate of the Alarm Data and Pseudo Data*

The alarm data are recorded by DCS, PLC or other smart devices; the sampling rate can be very high. This is necessary sometimes because the alarms occur with a very high frequency due to oscillation or chattering. Generally speaking, we can record them in the database as frequently as possible, say one per second. A proposed rule of thumb is a 15 seconds sample interval as a lower bound (Control Arts Inc., 2010).

It is generally unnecessary to record process data at such a high rate because the process always has a time constant with a magnitude of minutes or even longer. In real applications, we often record such samples at one sample per minute. Different from alarm data and process data, the pseudo data are generated from alarm data but behave as continuous process data. Thus we can record them at the same frequency as original alarm data, but we can also properly decrease the sampling rate to reduce the computational burden.

### *4.3 Variance of Gaussian Kernel Function*

The variance of the kernel function affects the robustness of the pseudo data at each alarm point. If the variance is very small, each alarm point generates a spike in the pseudo data making the method very sensitive to an individual sample or during a short period of alarms. On the other hand if the variance is large, the magnitude of each kernel function is very small; hence it needs quite a few alarm points during a period of no alarms or quite a few 'holes' during a period of consecutive alarms to change the trend of the pseudo curve. Although this can reduce the influence of the missed alarms, false alarms and chattering alarms due to its robustness to individual changes, the latency increases. This is a trade-off between large and small variances. To our best knowledge, most variables in the process industry have a time constant that ranges from several seconds to several minutes and even hours; thus the standard deviation of the Gaussian kernel function can be chosen over this range. If there is a step (from no alarm to alarm for example) in the alarm data, it needs several minutes for the pseudo data to completely change the value (from 0 to 1).

### *4.4 Data Requirement*

The computation of correlation requires sufficient data that contain alarm points because the normal points do not provide information in this sense. A short time period of simultaneous alarms is a much better data set than a long time period of rare alarms. In particular, the alarm data during an alarm flood is a very good source for study.

Aiming at one process variable, there are usually four alarms associated with it: HI, LO, HH, and LL alarms. If these alarm tags are treated separately, the relationship between them is lost. Instead, these four alarms can be combined together by assigning them different values based on their directions and degrees, for example, HI as +1, LO as -1, HH as +2, and LL as -2. This treatment makes the alarm data have several states instead of binary values. When generating the pseudo data,

these values should be taken as the weights of the Gaussian kernels.

### *4.5 Choosing Color Code for the Color Map*

We can either use the same color with different shades or different colors. Nevertheless, the number of color scales is very important, which determines the interpretability of the map. There is a trade-off between the sensitivity and the interpretability. A general recommendation is to use fewer number of codes first and increase gradually to suit your eyes.

In our work we use warm colors to describe positive correlation and cold colors to describe negative correlation. The four bands of values are chosen as (0-0.25), (0.25-0.5), (0.5-0.75), and (0.75-1) respectively. Please see the color bar in Fig. 4(b) for the color assignment.

## 5. CASE STUDY

To illustrate the practicality and utility of the method suggested in this work, consider data from a real industrial process. There are 10 analog alarms with HI and LO settings that have alarms over a period of one week. The sampling rate is one sample per second. By assigning HI and LO alarms +1 and -1, the alarm data are shown as a high density plot in Fig. 3 where alarm tags 3 and 7 have both HI and LO alarms. If we use the method proposed in (Kondaveeti *et al.*, 2010), the ASCM obtained is shown in Fig. 4(a) where the padding length used is 5 seconds. We find that alarms 1 and 2 are correlated, and alarms 4, 5, and 6 are grouped in another cluster.

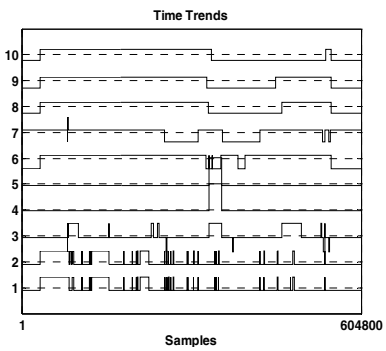
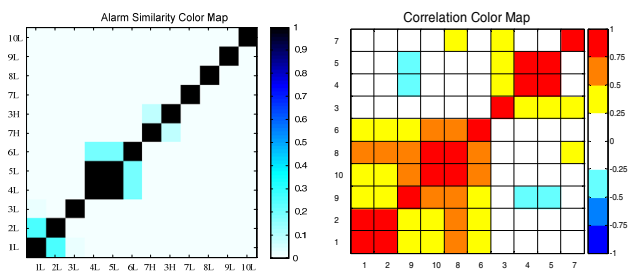



Fig. 3. Original alarm data for the case study.




(a) (b)

Fig. 4. Alarm similarity color map (a) and correlation color map (b) for the case study for the case study.

Now we use the method proposed in this paper by setting the standard deviation of Gaussian kernel as 30 seconds. Shown below is the correlation matrix (with the elements below the diagonal removed due to symmetry).

$$\begin{bmatrix} 1 & 0.97 & -0.06 & -0.08 & -0.08 & 0.33 & 0.23 & 0.54 & 0.34 & 0.42 \\ & 1 & -0.03 & -0.07 & -0.07 & 0.34 & 0.23 & 0.55 & 0.35 & 0.43 \\ & & 1 & 0.47 & 0.47 & -0.21 & 0.26 & -0.08 & 0.04 & -0.12 \\ & & & 1 & 1.00 & -0.20 & 0.10 & -0.16 & -0.27 & -0.13 \\ & & & & 1 & -0.20 & 0.10 & -0.16 & -0.27 & -0.13 \\ & & & & & 1 & -0.25 & 0.64 & 0.34 & 0.67 \\ & & & & & & 1 & 0.25 & 0.17 & 0.06 \\ & & & & & & & 1 & 0.60 & 0.79 \\ & & & & & & & & 1 & 0.70 \\ & & & & & & & & & 1 \end{bmatrix} \quad (8)$$

Based on this matrix, Table 1 illustrates the clustering process. Initially each alarm tag is a cluster. In the first step, the similarity measure (1−absolute correlation) of each pair of alarm tags are compared, in which the one between 4 and 5 has the smallest value, meaning that they are most similar. So alarm tags 1 and 2 are grouped into a new cluster, 11, and the old clusters 4 and 5 are removed. Then we compute the similarity measure between each pair of new clusters. We take alarm tags 1 and 2 in the second step, and then 10 and 8 in the third step. In the fourth step, the measure between clusters 9 and 13 is the smallest, which is determined by the similarity between alarm tags 9 and 8 according to the single-linkage method. So clusters 9 and 13 become the new cluster 14 which contains three alarm tags. The algorithm continues until only one cluster is left. The corresponding dendrogram is shown in Fig. 5. Thus the new order of alarm tags is obtained and thereby we have the correlation color map shown in Fig. 4(b).

**Table 1. Clustering Process**

| Step | Clusters | | | | | | | | | | Smallest measure | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | Correlation | Pair |
| 0 | 1 | 2 | 9 | 10 | 8 | 6 | 3 | 4 | 5 | 7 | 1.00 | 4, 5 |
| 1 | 1 | 2 | 9 | 10 | 8 | 6 | 3 | **11** | | 7 | 0.97 | 1, 2 |
| 2 | **12** | | 9 | 10 | 8 | 6 | 3 | 11 | | 7 | 0.79 | 10, 8 |
| 3 | 12 | | 9 | **13** | | 6 | 3 | 11 | | 7 | 0.69 | 9, 8 |
| 4 | 12 | | **14** | | | 6 | 3 | 11 | | 7 | 0.67 | 6, 10 |
| 5 | 12 | | **15** | | | | 3 | 11 | | 7 | 0.55 | 2, 8 |
| 6 | **16** | | | | | | 3 | 11 | | 7 | 0.47 | 3, 4/5 |
| 7 | **17** | | | | | | | | | 7 | -0.27 | 4/5, 9 |
| 8 | **18** | | | | | | | | | | 0.26 | 7, 3 |

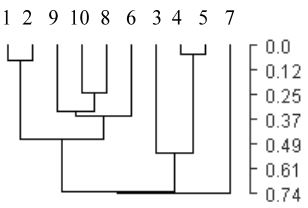



Fig. 5. Dendrogram for the 10 variables based on the single-linkage measure. The scale on the right indicates the similarity measure.

Apparently, alarms 1, 2, 9, 10, 8, and 6 form a single cluster with positive correlations, and alarms 3, 4, and 5 also have some correlations between them. This result clearly provides more information than the former one in Fig. 4 because it reveals some correlations that were not uncovered in the former one, such as the correlation between 8 and 9. Then we take the data of alarm tags 1, 2, 9, 10, 8, and 6 to perform SVD. The six singular values are 1720, 498, 347, 302, 199, and 165. Thus it is apparent that there exists redundancy in these tags. According to the user's preference and the corresponding physical meanings, some alarms may be removed.

## 6. CONCLUSIONS

In this paper, the correlation matrix is plotted as a colour map visualizing the relationship between different alarm tags. Compared to the ASCM, the clustered correlation map of psuedo data map has the following three major advantages: 1) it is robust to missed, false, and chattering alarms; 2) the correlation provides directional (positive or negative) information in addition to the similarity; and 3) the pseudo data used in generating the color map can also be used in other statistical analysis that provides more potential. One disadvantage to be noted is that when computing correlation between two time series, the simultaneous 0's in both the series have little information because they are in the normal region, resulting in the reduction of correlation coefficients. Thus the period for analysis should be selected so that there are dense alarm signals, which in any case is the objective of this study.

The method in this paper should be developed to be user-friendly to tune the parameters. In the pseudo generation stage, the variance of Gaussian kernel and sampling rate of pseudo data are two parameters that require tuning. In the color map plotting stage, the clustering algorithm is an option. In the color map display period, one should be able to change the color code and the number of scales of colors. It is important to provide some degrees of freedom for the user because the plot is sensitive to these tuning parameters and an appropriate choice of parameters depends on the specific circumstances, requirement, and objectives of the analysis.

## ACKNOWLEDGEMENTS

This work was supported by NSERC (SPG and IRC) in Canada, NSFC (60736026 and 60904044) and the 973 Project (2009CB320600) in China.

## REFERENCES

Bauer, M. and N.F. Thornhill (2008). A practical method for identifying the propagation path of plant-wide disturbances. *J. Process Control*, **18**, 707–719.

Choi, S., S. Cha and C.C. Tappert (2010). A survey of binary similarity and distance measures. *J. Systemics, Cybernetics and Informatics*, **8**, 43–48.

Control Arts Inc. (2010). *Alarm System Engineering*. http://www.controlartsinc.com/Support/Publications.html.

EEMUA (2007). *Alarm Systems: A Guide to Design, Management and Procurement.* EEMUA Standard 191-2007.

Hollifield, B. and E. Habibi (2007). *The Alarm Management Handbook: A Comprehensive Guide*. PAS, Houston, TX.

ISA (2009). *Management of Alarm Systems for the Process Industries*. ANSI/ISA Standard 18.2-2009.

Izadi, I., S.L. Shah, D. Shook and T. Chen (2009). An introduction to alarm analysis and design. In: *Proc. 7th IFAC Symp. Fault Detection, Supervision and Safety of Technical Processes*. 645–650. Barcelona, Spain.

Kondaveeti, S.R., I. Izadi and S.L. Shah (2009). Application of multivariate statistics for efficient alarm generation. In: *Proc. 7th IFAC Symp. Fault Detection, Supervision and Safety of Technical Processes*. 657–662. Barcelona, Spain,.

Kondaveeti, S.R., I. Izadi, S.L. Shah and T. Black (2010). Graphical representation of industrial alarm data. In: *Proc. 11th IFAC/IFIP/IFORS/IEA Symp. Analysis, Design, and Evaluation of Human-Machine Systems*. Valenciennes, France.

Lesot, M.J., M. Rifqi and H. Benhadda (2009). Similarity measures for binary and numerical data: a survey. *Int. J. Knowledge Engineering and Soft Data Paradigms*, **1**, 63–84.

Pissanetzky, S. (1984). *Sparse Matrix Technology*. Academic Press, London, UK.

Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. 34–48. Chapman and Hall, London and New York.

Wu, H.M., Y.J. Tien and C.H. Chen (2010). Gap: a graphical environment for matrix visualization and cluster analysis. *Computational Statistics and Data Analysis*, **54**, 767–778.

Tangirala, A.K., S.L. Shah and N.F. Thornhill (2005). PSCMAP: a new tool for plant-wide oscillation detection. *J. Process Control*, **15**, 931–941.

Yang, F., S.L. Shah and D. Xiao (2010a). Correlation analysis of alarm data and alarm limit design for industrial processes. In: *Proc. 2010 American Control Conf.*. 5850–5855. Baltimore.

Yang, F., S.L. Shah and D. Xiao (2010b). SDG (Signed Directed Graph) based process description and fault propagation analysis for a tailings pumping process. In: *Proc. 13th IFAC Symp. Automation in Mining, Mineral and Metal Processing*. Cape Town, South Africa.

Zhang, H. (2000). *Statistical process monitoring and modeling using PCA and PLS* [M.S. thesis]. Dept. Chem. Mat. Eng., University of Alberta, Edmonton, Canada.