www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 3 Issue 2 February, 2014 Page No. 3996-4001

DATA REPORTING OF TRACING EVENTS BY EVENT DRIVEN INCIDENCE MATRIX

***S.Keerthana*[1], *Mr.P.Sathishkumar M.E*[2]**
[1]department of computer science and engineering,
K.S.Rangasamycollege of technology,
skskeerthana@gmail.com
[2]department of computer science and engineering,
K.S.Rangasamycollege of technology,

Abstract

A platform for online Sensitivity Analysis (SA) that is applicable in large scale real-time data acquisition (DAQ) systems. Supervisory Control and Data Acquisition (SCADA) sensors and actuators connected to monitor the processes of manufacturing and its transmittable operations as a case study for resistant of concept. It deploys the Rank Order Clustering (ROC) method to automatically group all existing data sensors and actuators of the system to the Key Performance Indicators of the system. The sensors and actuators data collected shapes the input data for measuring the performance. The Event Cluster algorithm is located in inside the control centre of the SCADA system to assess the influence of each input to the overall key performance indicators of the process. This method progresses the quality of data analysis and reduces computation overhead on the control system.The flexibility to adapt can only be assured if data is succinctly interpreted and translated into corrective actions in a timely manner. Every single or combination of events could subsequently results in a change to the system state. The Proposed Event-Driven Incidence Matrix is designed based on sorting the rows for inputs and columns for key performance indicators (outputs). Incidence matrix elements can take a value of 0 and 1.

Keywords:Sensitivity Analysis,EventTracker,DATA acquisition,Rank Order Clustering.

1 INTRODUCTION

DATA acquisition (DAQ) systems that deal with large quantities of input variables and have higher sampling frequencies result in high bandwidth communication and place a heavy computational load on the higher tier data processing and information systems within their hierarchy. Industrial data acquisition systems are usually designed to manage large quantities of input variables with high sampling frequencies. The complex Information Systems (IS) infrastructure built within and around these systems normally struggle with the large flow of raw input data (clogging communication channels) and the complex interpretation and algorithms. The computational effort thus increases exponentially – leading to higher energy and time-to-action costs. One of the areas of focus for researchers and engineers in this area has been eliminating input variables that have the least impact on the system to minimize the costs [1],[2],[3]. The data filtering and sensitivity analysis techniques strive to focus on the most valuable information that has significant impacts on state/behavior of the system. Knowing the importance of your input data can help to improve the accuracy of performance measures. Industries faced with challenges of the ever more competitive global markets have realized the importance of accurate and timely account system state. The knowledge of the state of the system may include energy consumption, environmental impact, product and process quality control, inventory, and so on. Therefore efforts should be made to improve the accuracy and timeliness of the important input variables that feed into performance measuring algorithms. The term “important” in variable selection has been interpreted into two separate notions of “usefulness” and “relevance” that are explained [4],[5]. This paper introduces an event clustering method that could be categorized as an Input Variable Selection (IVS) technique. The purpose of IVS technique is to maximize the quality of data acquisition and interpretation. In this context, input variables determine performance parameters. Therefore, the cause-effect relationship between the input variables and performance parameters generate the knowledge about the system. Traditional IVS techniques normally rely on historical knowledge (e.g. statistics) or heuristics to interpret input data and system state information. A comprehensive review has been done on IVS literature review in [5]. The reason for the proposed Event-Clustering platform is that modern industrial systems are assumed to be able to capture data in real-time and adjust to changing system requirements [4]. An important factor that facilitates data interpretation and information modeling is an appreciation of the effect system inputs have on each output within a time frame. Many existing IVS methods are time consuming and sluggish due to their reliance on historical data. Sensitivity analysis techniques help system analysts to focus on the most valuable information, information that most significantly impacts on system behaviour.

2. RELATED WORK

The majority of sensitivity analysis methods attempt to determine the impact of changes in one variable in relation to others by means of analytical models that describes the relationship that exists between them. Methods such as Differential Analysis, Coupled/Decoupled

The Statistical Distribution of Input Variables

The sensitivity indices of a system are normally influenced by the distribution of the input data series. For example, nonlinear relationships between input and output series in a model cannot be recognized by correlation-based sensitivityanalysis methods alone. Variance based and Entropy-based indices are expected to be more sensitive to heteroscedastic data [8], while the homoscedasticity of data series can be higher in discrete signals and much higherbetween binary signals.

Sensitivity analysis is a computationally hungry process. In domain-wide sensitivity analysis methods, large batches of input variables are captured in a specified periodic time interval and subsequently values of sensitivity are determined using historical data analysis. For example, sampling based methods need to generate new and equivalent sized batches of sample values for both output and input data regardless of the original sampling rates. The magnitude of resources required by such algorithms and their associateddata processing requirements are comparable to the expected savings resulting from their application. In the following sections, after a brief introduction to existing SA methods, a detailed description of event-driven data types and their impact on sensitivity analysis is provided. The proposed EventTracker method and itsapplication in a case study are discussed. The advantages and application of EventTracker in an industrial case studyis presented in the penultimate section.

2.1EVENT TRACKING SENSITIVITY ANALYSIS

The event tracking SA method uses an input and output occurrence. This matrix is populated at predefined time intervals. The current platform is designed to allow a user to set the initial system update time interval. For example, in safety sensitive systems such as power plant reactor monitoring, the rate of populating the data tables will be a short interval. Whereas in scenarios that employ less time critical systems, such as finance, then the interval will be longer. This matrix is designed to map the relationships between causes that trigger event and the data that describes the actual events.In this way the “Event Tracker” method is able to construct a discrete event framework where events are loosely coupled with respect to their triggers for the purpose of sensitivity analysis.

Trigger Data and Event Data

Any input variable whose value results in the registration of an event is defined as Trigger Data in our DES. The series of data that represent the state of the system at a given time is described as

Eventdata.

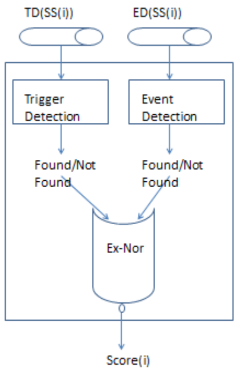

Fig. 1.Trigger-event detection functionality on each search Slot.

It is possible that the numbers of EDs and TDs in a system are different. For example, a number of TD series may be responsible for changing a single ED series.It should be noted that various TD series could have differing impact on specified ED series. This is because individual or combination of input variables may have different effects on different system outputs.

Discrete Event Systems

As opposed to continuous systems, a Discrete Event System is defined by the disparate occurrence of events in a specified time span. In other words, the state of the system changes when the input variables and consequently the outputs of the system change. Each state transition of the system is called an event. Therefore, in DES, only the attributes that represent the occurrence of an event are considered. These attributes are discussed in the following section.

Methods and Parameters for Event Tracking

The Event Tracker platform is based on four functional parameters that are initialized by a user with domain knowledge. The Search Slot (SS) and the Analysis Span (AS) parameters are about tracing the values of the acquired data series. Whereas the remaining two parameters Event Threshold (ET) and Trigger Threshold (TT) are about the magnitude of transition detection and the overall system state analysis. Subsequently, these parameters are automatically optimized by the Event Tracker platform.

- **Search slot:** The SS is a fixed time slot within which batches of TD and ED are captured. It can also be described as the scan rate. The scan rate is determined by a system expert.
- **Analysis span:** The AS is the time span within which a period of sensitivity analysis occurs.
- **Event threshold:** The fluctuations in the ED series that are interpreted as triggers are determined in comparison with the Event Threshold. This value is expressed as a proportion of the overall range of ED series values occurring in an AS. It is expressed as a percentage.
- **Trigger threshold:** The fluctuations in the TD series that are interpreted as triggers are determined in comparison with the Trigger Threshold.

Entropy-Based Epistemic Sensitivity

Analysis In order to determine sensitivity indices then one only needs to establish the values of independent input variables (denoted by X) and dependent output variables (denoted by Y) [4]. The sensitivity indices using the Entropy method. The method replaces the time consuming sample generation of X and evaluation of Y by Simple-Random Sampling (SRS) using piecewise uniform density function estimations. Krzykacz-Hausmann demonstrates the feasibility of the estimation approach in a test case with 15 independent and two dependent variables. Reasonable results were achieved with far lower computational cost.

2.2 Event Tracker Algorithm

The algorithm is designed to respond quickly and in essence has a life cycle that is equivalent to an AS.

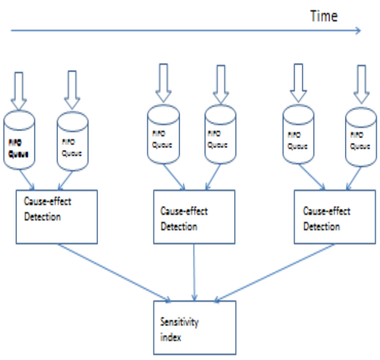

Fig. 2. Functionality diagram of EventTracker algorithm.

This life cycle is divided into several SS. Within each slot, TDs and EDs are captured from two time series and used to provide a value which is translated into a sensitivity index. This index is then added to the indices of subsequent search slots. At the end of each AS, the sensitivity indices of all data series are linearly normalized.The main steps of the algorithm are as follows:

Stepwise Scan

A First-In-First-Out queue is allocated for every batch of data in a search slot. The size of the queues is unbounded. The content of the queues are flushed at the end of each search slot. The data is then passed to the Event Tracker detection and scoring algorithm. The next search slot continues to fill the queue immediately. Using this technique no data is lost.

Trigger-Event Detection

The batch of TD values is searched for fluctuations greater than the specified TT threshold, and ED values similarly checked for changes larger than the ET threshold. A cut off threshold (CT) is defined for each series of indices within an ED series. Their values lie between the minimum and maximum index values for that range, as in

$CT = Min(SI_{ED}) + CR^{*}(Max(SI_{ED})-Min(SI_{ED}))$,

where, CR is the Cut off Ratio in the range 0 <=CR<= 1. For example, if CR is 0.5, then the value of the cut off thresholds are all in the middle of their associated sensitivity indices range.

Two-Way Matching Score

In each SS the simultaneous existence or nonexistence of a change in each pair of data batches is scored as þ1, otherwise the score is -1. This operation is similar to a weighted logical Exclusive-NOR and is shown in Table 1. This approach is adopted to better emphasize the impact of inputs on a given output rather than simply scoring +1 for existence and 0 for non-existence.

3. PROPOSED EVENTS

Event clustering method for use in educating the quality of real time data in SCADA systems. Event Cluster does not require preceding information of the analytical or algebraic relationship that may be present among input and output variables. The reduction in cost complexity of any related SCADA systems. A key feature of the procedure is its ability to quickly generate an event driven occurrence matrix and measure degree of encouragement of input sensors on performance displays. Moreover Event Cluster does not require preceding information of the analytical or algebraic relationship that may be present among input and output variables. The technology will not only have important impact on the strengthen industry but with minor differences can be functional to a whole range of developed continuous and or separate production processes. One key advantage of the method is the reduction in cost complexity of any related SCADA systems.As a upcoming work the projected Event Cluster could be industrialized to measure stricture sensitivity analysis and be associated with other sensitivity analysis techniques Event Tracker in computational efficiency and accuracy.

4 METHODS

4.1 Clustering Methods

It uses simple matrix manipulation methods to rearrange the row and columns of a matrix in an iterative manner that will ultimately and in a finite number of steps, result in a matrix form in which both the rows and columns are arranged in order of decreasing value. It is an effective algorithm to determine clusters of occurrence as block diagonal format. These approaches are limited in that they are based on the assumption that groups of data are highly similar and will be placed into mutually exclusive blocks.

4.2 Key Performance Factors

Key performance factors (KPF) are factors of a system are normally determined by the experts in that system and are related to their influence on profit margins of the system, provided that the system is an industrial one. Other systems have their own KPF.

Step1: Populate the Sensor-KPI event coincidence
matrix with binary weighting (use Exclusive NOR function).

Step2: Sort rows of the binary matrix in decreasing
order of the corresponding decimal weights.

Step3: Repeat the preceding two steps for each column.

Step4: Repeat the preceding steps until the position of each element in each row and column does not change.

With sufficient iterations, the final grouping of Sensors KPI is achieved. For example, Fig. II illustrates Sensor Group {3, 2, 6} is in the KPI-Group {C, E, I}, as in the KPI-Group {C, E, I}.

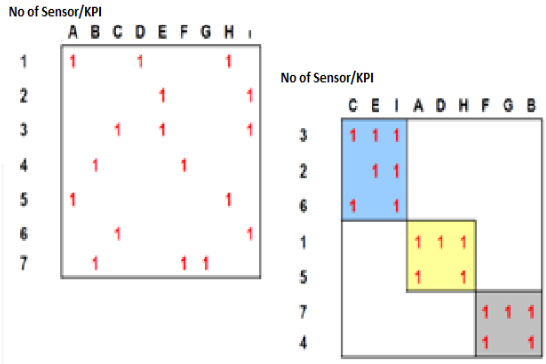

Fig 3. Sensors-KPI incidence matrix before (left) and after (right) implementation of ROC algorithm

4.3 Event Cluster and Data Grouping Techniques

Event Clustering method is that the state of a system during its life span can be broken down into series of consecutive discrete events. The change in the state of a system can be triggered by events. These events are instigated by change in the state of the input variables (sensors & actuators). A group of events that coincide with a change in system state demonstrate an important role.

This in real-time can help us to group those important events with the performance indicators of the system. It is important to realize that the discrete event here implies that the system is not aware of the previous event. For definition of unaware systems.

A. Rank Order Clustering (ROC)

ROC methods was first proposed by King [6] .It uses simple matrix manipulation methods to rearrange the row and columns of a matrix in an iterative manner that will ultimately, and in a finite number of steps, result in a matrix form in which both the rows and columns are arranged in order of decreasing value. It is an effective algorithm to determine clusters of occurrence as block diagonal format. These approaches are limited in that theyare based on the assumption that groups of data are highly similar and will be placed into mutually exclusive blocks. In the cluster analysis method a group of data values are “similar” according to a “similarity criteria”. They can be either replaced by a new value representing the group (clumping) or assigned a unique type of label (partitioning) [7].

B. Definition of Key Performance Factors and Indicators

Key performance factors (KPF) are factors of a system are normally determined by the experts in that system and are related to their influence on profit margins of the system, provided that the system is an industrial one. Other systems have their own KPF. For instance, KPF in a manufacturing system could be categorized as: levels of customization, productivity, resource utilization, efficiency, inventory management. These factors can be defined loosely and in general terms (i.e. subjective terms) which then need to be broken down into key performance indicators (KPI). The indicators are objective and normally metricized.

C. Basics of EventClustering Method

The basic assumption of EventClustering method is that the state of a system during its life span can be broken down into series of consecutive discrete events. The change in the state of a system can be triggered by events. These events are instigated by change in the state of the input variables (sensors & actuators). A group of events that coincide with a change in system state demonstrate an important role. This in real-time can help us to group those important events with the performance indicators of the system. It is important to realize that the discrete event here implies that the system is not aware of the previous event. For definition of unaware systems see [4]. In time the frequency of such coincidence will increase the role of these inputs on system state leading to a near real-time novel sensitivity. The sensitivity analysis is
outside the scope of this paper and is proposed as a potential for future work. To explain the implementation process, a short example of how the ROC is used for EventClustering. A coincidence matrix of Sensors (rows) and Key Performance Indicators (column) is drawn.

5CONCLUSION

A sensitivity analysis methodology for use in large scale “real-time” data.The method in comparison to Entropy-based SA technique was shown to be faster, more accurate, and less computationally more accurate,andlesscomputationallyburdensomeacqui sition The reason that ESA was used as the basis for comparison is that like Event Tracker, the ESA method is a SA method that relies least on historical data.

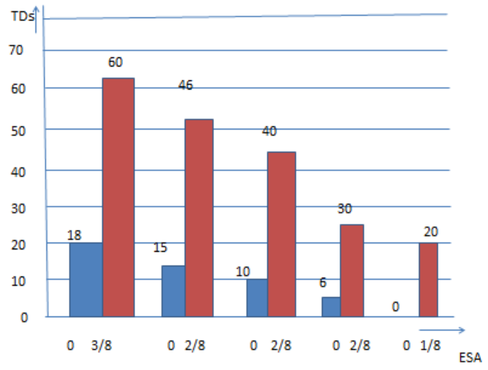

Fig. 5.Comparison of proportion of less important TDs with low false negative ratios on EventTracker and ESA methods.

Event Tracker is an event-driven sensitivity analysis method and not a probability-based approach. The process is deterministic in the sense that it is only instigated when an event with a predetermined threshold is detected. There is no reliance on statistical or model-based equations, only on the interpretation of transition between system states and in that sense the technique is completely“unaware”systems. One of the strengths of proposed method of Network sharing security is the freedom of choice it offers the user to specify a scan rate based on the very nature of the application itself. The platform in its current formprovides the flexibility for a system analyst to choose an appropriate value based on their experience and local knowledge.

6REFERENCES

1. Bernard Krzykacz-Hausmann(2006), “An approximate sensitivity analysis of results from complex computer models in the presence of epistemic and aleatory uncertainties”, Elsevier Reliability Engineering and System Safety, Vol:92,No.4

2. ChonggangXu, George ZdzislawGertner(2008), “A general first-order global sensitivity analysis method”, Elsevier Reliability Engineering and System Safety, vol:93,No.3

3. Darren P. Durkee, Edward(2009), “Sensitivity Analysis Of Availability Estimates To Input Data Characterization”, Qual. Reliab. Engng. Int, Vol:14 ,No.3

4. H. L. Cloke, F. Pappenberger(2008), “Multi-Method Global Sensitivity Analysis (MMGSA) for modelling floodplain hydrological processes”, IEEE Trans On Hydrol. Process, Vol:22, No.2

5. Huibin Liu, Wei Chen(2005), “Relative Entropy Based Method For Probabilistic Sensitivity Analysis”, IEEE Trans On Data Engineering, Vol:72, No.2

6. Martin Brown, Yizeng Liang(2007), “ Improving Data Fitting of a Signal Transduction Model by Global Sensitivity Analysis”, IEEE Transaction, Vol:105

7. Roger A. Cropp(2005), “The New Morris Method: an efficient second-order screening method”, Elsevier Reliability Engineering and System Safety, Vol:78,No.3

8. SiamakTavakoli, AlirezaMousavi(2013), “Event Tracking for Real-Time Unaware Sensitivity Analysis (EventTracker)”, Ieee Transactions On Knowledge And Data Engineering, Vol:25, No. 2