Project Newsletter

Issue 3 - December 2013

Our open-source policy

Sharing the software implementation of some core results complementing what we published in top international journals, is one of i-Sense project dissemination policies. As such we offer as open and free software cores of the algorithms developed within i-Sense: around those you can build your own embedded solution.
Our open source policy is aligned with the digital open access guidelines to EU funded research oriented towards the dissemination of results and outcomes.
We also provide a public repository of some i-Sense benchmarks in the spirit of cooperative scientific progress. You are free to download from our website the datasets for non-commercial research and educational purposes.
Although, the contents of this newsletter focus on theses public results, we invite you to visit our project website *www.i-sense.org* to find out our innovative methodologies and algorithms for cognitive fault diagnosis and fault tolerant control.

In the sequel you will find descriptions of:

- The **open-source library** with the following items:
  - ⇒ The Matlab-CONTAM Toolbox which is an expandable research simulation tool that facilitates the implementation of various algorithms related to contaminant event monitoring
  - ⇒ The Hierarchical ICI-based Change Detection Test. That by extracting a set of features from available data assesses the stationarity of process generating the data
- The **open-source i-Sense Robotic Platform**
- The **open-source benchmarks and datasets.** In particular,
  - ⇒ Environmental Monitoring Benchmarks
  - ⇒ Smart Buildings and Contaminant Events Benchmarks

In this issue:

i-Sense Public Results

Matlab-CONTAM Toolbox for Contaminant Event Monitoring in Intelligent Buildings

An intelligent building should take all the necessary steps to provide protection against the dispersion of contaminants from sources (events) inside the building which can compromise the indoor air quality and influence the occupants' comfort, health, productivity and safety. Such events could be the result of an accident, faulty equipment or a planned attack. Multi-zone models and software, such as CONTAM (open-source), have been widely used in building environmental studies for predicting airflows and the resulting contaminant transport.

The developed Matlab-CONTAM Toolbox allows the creation of datasets from running multiple scenarios using CONTAM by varying the different problem parameters. The Matlab-CONTAM Toolbox is an expandable research tool which facilitates the implementation of various algorithms related to contamination event monitoring. In particular, we describe the implementation of state-of-the-art algorithms for detecting and isolating a contaminant source. The use of the Toolbox is demonstrated through a building case-study.
The goal of the Toolbox is to serve as a common programming framework for research, which facilitates simulating multiple contamination events under varying conditions as well as to store computed data in data structures so that they can be reused by different algorithms.
The Matlab-CONTAM Toolbox features a user-friendly Graphical User Interface (GUI) and a modular architecture. It allows the creation of multiple scenarios by varying the different problem parameters (wind direction, wind speed, leakage path openings, source magnitude, evolution rate and onset time) as well as the storage of the computed results in data structures. The data from these scenarios are further analyzed by the developed algorithms for determining solutions for contaminant event monitoring and sensor placement.

Our open Matlab Library available at

http://www.i-sense.org/open_library.html

i-Sense Public Results

The Hierarchical ICI-based Change Detection Test

Intelligent systems meant to operate in nonstationary environments have to detect possible changes to proactively interact with the environment and adapt to evolving nonstationary conditions.

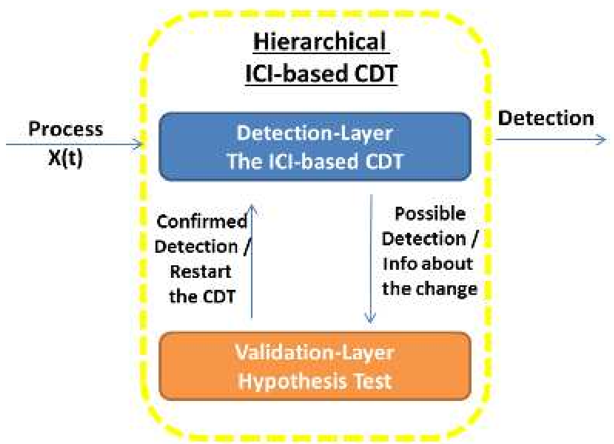
One of the most appealing solutions consists in monitoring the statistical behavior of data by means of nonparametric sequential Change-Detection Tests (CDTs). These tests assess the stationarity of the data-generating process without requiring any a-priori information about the process or change.

One of the main drawbacks of CDTs are false positives, i.e., detections not corresponding to an actual change in the data-generating process, as these induce intelligent systems to raise false alarms.

The hierarchical ICI-based CDT belongs to the family of ICI-based CDTs, which extract a set of features from data and assess the stationarity of these features by means of the ICI rule.

The hierarchical ICI-based CDT is a nonparametric CDT characterized by two levels:

- The Detection Layer, which operates on-line and performs a sequential analysis of the observation stream. The purpose of this layer is to provide a prompt detection of possible nonstationarities; for this layer the ICI-based CDT is considered.
- The Validation Layer, which is activated after a detection raised by the detection layer. The purpose of this layer is to validate the prospective nonstationarity by means of an hypothesis test on features extracted by the ICI-based CDT.



Such a hierarchical CDT significantly reduces the number of false positives and the detection delays, since the validation layer allows for tuning the ICI-based CDT to provide prompter detections. This solution has been tested in a distributed solution meant for wireless and hybrid sensor networks.

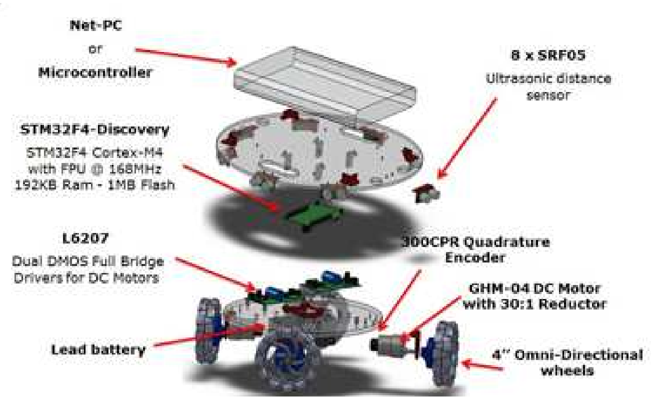
The **open-source Matlab package** which contains both functions implementing the hierarchical ICI-based CDT and dataset generation is available on the project website.

The i-Sense Robotic Platform

A further public result of i-Sense is the open-source i-Sense Robotic Platform: a complete project of an **omni-directional mobile Robot**.

The material made available contains:

- Both off-the-shelf and custom mechanical parts (CAD drawings)
- Electronic circuits schematics
- Low level firmware running on STM32F4 microcontroller
- High level control software running in Matlab for the real-time control of the robot.



Smart Buildings and Contaminant Events Benchmarks

The dispersion of contaminants from sources (events) inside a building can compromise the indoor air quality and influence the occupants' comfort, health, productivity and safety. Such events could be the result of an accident, faulty equipment or a planned attack. Under these safety-critical conditions, immediate event detection should be guaranteed and the proper actions should be taken to ensure the safety of the people. In this work, we consider an event as a fault in the process that disturbs the normal system operation. This places the problem of contaminant event monitoring in the fault diagnosis framework of detection and isolation.

The description of the datasets generated by the Matlab-CONTAM Toolbox for Contaminant Event Monitoring in Intelligent Buildings are classified in:

⇒ Smart Building Datasets with Contaminant Event
⇒ Smart Building Datasets with Sensor Fault
⇒ Smart Building Datasets with Contaminant Event and Sensor Fault
⇒ Class Imbalance in Fault Detection Benchmarks
⇒

In our website you can **download the following datasets** generated by the Matlab-CONTAM Toolbox:

Smart Building Datasets with Contaminant Event

Considering the Holmes House building scenario (see figure 1), we assume a single source of the contaminant of interest (i.e. CO2) that is released at time 1 hour and the simulation lasts for 24 hours. Furthermore, we assume sensors in all 14 of the building zones.

Objectives

- Detection of contaminant event
- Isolation and identification of contaminant

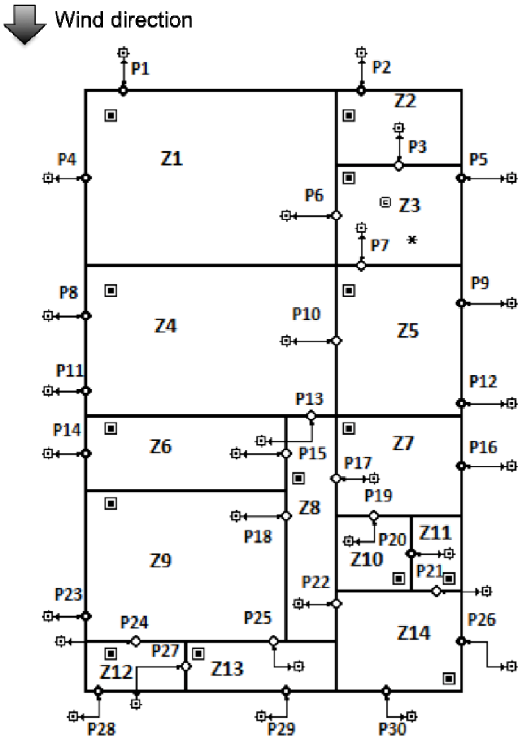


Figure 1: The Holmes House case study

Smart Building Datasets with Sensor Fault

We consider the Holmes House building scenario. The contaminant of interest (i.e. CO2) is present in the atmosphere with a mean concentration of 50 mg/m^3/sec modeled as a pseudorandom sequence. The time interval between the transitions (jumps) follows a Markov process with 0.1 transition probability while the magnitude of the sequence after a transition is a random number between [45, 55]. For the simulations, we assume wind direction 90deg (from the east), wind speed 10m/s and fully open leakage path openings. Furthermore, we assume sensors in all 14 of the building zones. The sensors are monitoring the contaminant of interest and a single, permanent, additive fault is introduced at time t=2000 sample in the sensor placed in the kitchen (Z5).

Objectives

- Detection of sensor fault
- Isolation and identification of sensor fault
- Data reconstruction for faulty sensor data

Smart Building Datasets with Contaminant Event and Sensor Fault

We use the Holmes House building scenario with the setup described in the Smart Building Datasets with Sensor Fault.

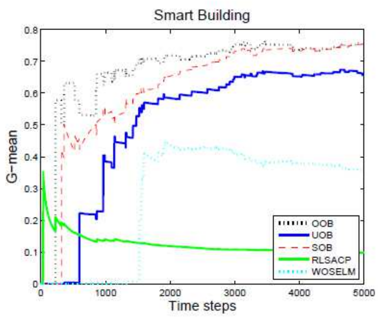
The sensors are monitoring the contaminant of interest and a single, permanent, additive fault is introduced at time t=2000 sample in the sensor placed in the kitchen (Z5). In addition to the sensor fault, a contaminant source of emission rate 100 mg/s is also placed in the kitchen (Z5) at time t=2000 sample.

Objectives

- Detection of sensor fault
- Isolation and identification of sensor fault
- Data reconstruction for faulty sensor data
- Detection of contaminant event
- Isolation and identification of contaminant event

Class Imbalance in Fault Detection Benchmarks

Using the Smart Building real-world project we explored the ability of existing online class imbalance methods on fault detection applications. It was considered with other five datasets that were highly imbalanced. We design and look into a series of practical scenarios, including not only data streams that are constantly imbalanced, but also data streams suffering short-term fluctuations of class imbalance status.



The open-source datasets available on our website regard the following four scenarios (one has a fixed imbalance rate, and the other three have class imbalance fluctuations):

1. The first scenario simulates stable systems with hidden faults that are hard to be captured.
2. In the second scenario all the faults occur at the beginning of the training data.
3. In the third scenario, all the faults occur at the end of the training data. It can happen when a key part of the system is broken suddenly.
4. In the last scenario, the faults become more and more frequent in the training data. It simulates the case when the damaged condition gets worse gradually.

i-Sense Public Results

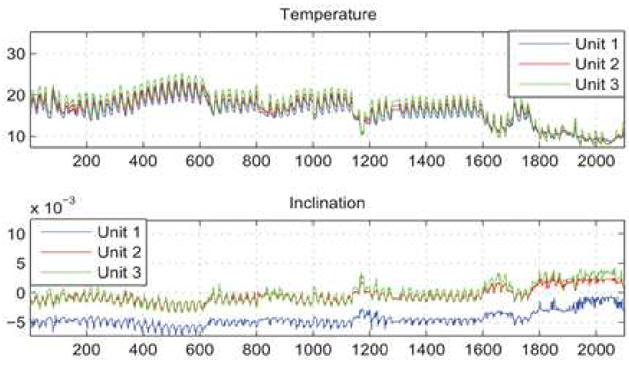
Environmental Monitoring Benchmarks

Sensor networks monitoring a real environment are prone to faults or aging phenomena, whose impact affects the overall system performance. In fact, permanent or transient faults can influence the sensors, the analog electronics, and the digital part of the embedded system inducing, in the best case, functional errors in the processing chain. In turn, erroneous information generates a strong side effect on the subsequent control chain, leading to wrong decisions and inappropriate control actions.

The Rock Collapse Forecasting System scenario is an experiment which refers to a real-world dataset provided by a real-time monitoring system for rock fall forecasting designed by Politecnico di Milano group and deployed in the Alps. The benchmarks considered in the next sections are related to measurements coming from a new generation of intelligent clinometer sensors which have an internal thermal sensor to correct and compensate online the measurements.

Stuck-at fault Benchmark

The Stuck-at fault Benchmark has been acquired by the fully-wireless monitoring system for landslide forecasting that has been deployed in Northen Italy (Towers of Rialba).

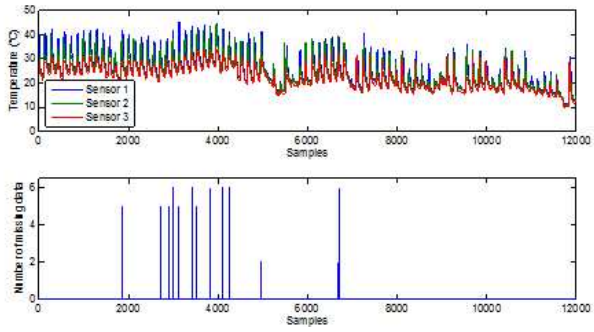


The system is composed by 3 units and 14 sensors: Unit 1 and 2 are endowed with 5 sensors (Int./ Ext. Temperature, Biaxial Clinometer, Crackmeter), while Unit 3 is endowed with 4 sensors (Int./ Ext. Temperature, Biaxial Clinometer). The length of the benchmark is 2100 samples (from August 1, 2011 to October 31th, 2011), while the sampling frequency is 1 sample/hour. The real fault present in the benchmark is a stuck-at fault affecting all the units of the network (between samples 1280-1305).

Dataset Description
Number of units: 3
Unit 1: 5 sensors (Int./Ext. Temperature, Biaxial Clinometer, Crackmeter)
Unit 2: 5 sensors (Int./Ext. Temperature, Biaxial Clinometer, Crackmeter)
Unit 3: 4 sensors (Int./Ext. Temperature, Biaxial Clinometer)
Length of the dataset: August 1, 2011 until October 31th, 2011 (2100 samples)
Sampling frequency: 1 sample/hour
Type of fault: Stuck-at fault affecting all the sensors

Missing data Benchmark

The Missing data Benchmark has been acquired by the a hybrid wireless-wired monitoring system for rock-collapse forecasting that has been deployed in Northern Italy (Towers of Rialba). The number of units of the system is 3. In this benchmark only temperature measurements have been considered. The length of the dataset is 12000 samples (from Jul. 25, 2012 to Oct. 17, 2012), while the sampling frequency is 6 samples/hour. The type of real existing fault is missing data induced by communication problems.



Dataset Description
Number of units: 3
Sensors per unit:
Temperature sensors
MEMS accelerometer/Geophone/Inclinometer/ Crackmeter (not included in the benchmark)
Length of the dataset: Jul. 25, 2012 until Oct. 17, 2012 (12000 samples)
Sampling frequency: 6 sample/hour

Some i-*Sense* partners joint publications*

Adaptation and Learning

► Soares, R. G. F., Chen, H., Yao, X. (2013) "*A Fully Semi-supervised Ensemble Approach to Multi-class Classification*" submitted to IEEE Transactions on Neural Networks and Learning Systems, 2013 .

Abstract: Semi-Supervised Classification (SSC) is the task of learning from unlabelled data, along with a few labelled data, to predict labels of test instances. In this paper, we present a study on the usefulness of employing unlabelled data at both ensemble and base learner levels, comparing it to using such data at the ensemble level only. We propose Clusterbased Boosting (CBoost), a multi-class classification algorithm with cluster regularisation. This method extends the ClusterReg algorithm [1] and, unlike other semi-supervised ensembles in the literature, is composed of semi-supervised base classifiers. CBoost is able to learn from the clustering neighbourhood structure of pseudo-labels assigned by the ensemble, which leads to better generalisation when compared to learning the exact pseudolabels individually. CBoost can overcome incorrect pseudo-label assignments used in the training of a new base classifier. CBoost is robust to the position of labelled data within a cluster and is able to handle the potential presence of overlapping classes. .

► Wang, S., Minku, L.L., Ghezzi, D., Caltabiano, D., Tino, P. and Yao, X. (2013) "*Concept Drift Detection for Online Class Imbalance Learning* ", in Proceedings of IJCNN (International Joint Conference on Neural Networks), Dallas, Texas, USA, 4-9 August, 2013.

Abstract: Concept drift detection methods are crucial components of many online learning approaches. Accurate drift detections allow prompt reaction to drifts and help to maintain high performance of online models over time. This paper studies the concept drift problem for online class imbalance learning. We look into the impact of concept drift on single-class performance of online models based on three types of classifiers, under seven different scenarios with the presence of class imbalance. The analysis reveals that detecting drift in imbalanced data streams is a more difficult task than in balanced ones. Minority-class recall suffers from a significant drop after the drift involving the minority class. Overall accuracy is not suitable for drift detection. Based on the findings, we propose a new detection method DDM-OCI derived from the existing method DDM. DDM-OCI monitors minority-class recall online to capture the drift. The results show a quick response of the online model working with DDM-OCI to the new concept.

► J. Quevedo, H. Chen, M. A. Cuguero, P. Tino, V. Puig, D. Garcia, R. Sarrate, X. Yao (2013) "*Combining Learning Fault Diagnosis in Model Space with Data Validation/ Reconstruction in Critical Infrastructure Systems: Application to the Barcelona Water Network*" to appear in Engineering Application of Artificial Intelligence, 2013.

Abstract: The proposed methodology is implemented in a two-stage approach. In the first stage sensor communication faults are detected and corrected, in order to facilitate a reliable dataset to perform system fault diagnosis in the second stage. On the one hand, sensor validation and reconstruction are based on the combined used of spatial and time series models. Spatial models take advantage of the (mass-balance) relation between different variables in the system, while time series models take advantage of the temporal redundancy of the measured variables by means of Holt Winters models. On the other hand, fault diagnosis is based on the learning-in-model-space approach that is implemented by fitting a series of models using a series of signal segments selected with a sliding window. In this way, each signal segment can be represented by one model. To rigorously measure the 'distance' between models, the distance in the model space is defined. The deterministic reservoir computing approach is used to approximate a model with the input-output dynamics that exploits spatial-temporal correlations existing in the original data. Finally, the proposed approach is successfully applied to the Barcelona water network.

► Rotondo D., Reppa V., Puig V., Nejjari F. (2013) "*Adaptive observer for switching Linear Parameter-Varying (LPV) systems*" submitted to the 19th World Congress of the International Federation of Automatic Control, Cape Town, South Africa, August 24-29, 2014.

Abstract: In this paper, the problem of joint state and parameter estimation in switching Linear Parameter-Varying (LPV) systems is considered. The proposed solution relies on an adaptive observer, which is designed finding a solution to a system of Trilinear Matrix Inequalities (TMIs). It is shown that, under certain assumptions, the TMIs can be reduced to Linear Matrix Inequalities (LMIs) that can be solved using available software. An example of a four wheeled omnidirectional mobile robot subject to unknown offsets in the motor voltages is used to illustrate the efficiency of the proposed approach.

* For a complete list of journals and Conferences publications refer to the project website www.i-sense.org

Some i-*Sense* partners joint publications*

Fault Diagnosis

► Boracchi, G., Puig, V., Roveri, M. (2013) "*A Hierarchy of Change-Point Methods for Estimating the Time Instant of Subtle Leakages in Water Distribution Networks*" in Proceedings of LEAPS, Paphos, Cyprus, Sept. 30 - Oct. 2, 201, pp 615 – 624

Abstract: Leakages are a relevant issue in water distribution networks with severe effects on costs and water savings. While there are several solutions for detecting leakages by analyzing of the minimum night flow and the pressure inside manageable areas of the network (DMAs), the problem of estimating the time-instant when the leak occurred has been much less considered. However, an estimate of the leakage time-instant is useful for the diagnosis operations, as it may clarify the leak causes. We here address this problem by combining two change-point methods (CPMs) in a hierarchy: at first, a CPM analyses the minimum night flow providing an estimate of the day when the leakage started. Such an estimate is then refined by a second CPM, which analyzes the residuals between the pressure measurements and a network model in a neighborhood of the estimated leakage day. The proposed approach was tested on data from a DMA of a big European city, both on artificially injected and real leakages. Results show the feasibility of the proposed solution, also when leakages are very small.

► Alippi, C., Boracchi, G., Puig, V., Roveri, M. (2013) "*An Ensemble Approach to Estimate the Fault-Time Instant*", in IEEE Proceedings of ICICIP 2013, International Conference on Intelligent Control and Information Processing, June 9 - 11, Beijing, China.

Abstract: Since systems are prone to faults, fault detection and isolation are essential activities to be considered in safety- critical applications. In this direction, availability of a sound estimate about the time instant the fault occurred is a precious information that a diagnosis system can fruitfully exploit, e.g., to identify information consistent with the faulty state. Unfortunately, any fault-detection system introduces a structural delay that, typically, increases in correspondence of subtle faults (e.g., those characterized by a small magnitude) with a consequence that the fault-occurrence time is overestimated. In this paper we propose an ensemble approach to estimate the time instant a fault occurred. We focus on systems that can be described as ARMA models and faults inducing an abrupt change in the model coefficients.

► Boracchi, G., Michaelides, M., Roveri, M. (2014) "*A cognitive monitoring system for contaminant detection in intelligent buildings,*" The 2014 International Joint Conference on Neural Networks (IJCNN 2014), submitted.

Abstract: Intelligent buildings are equipped with sensing systems able to measure the contaminant concentration in the different building zones for safety purposes. The aim of these systems is to promptly detect the presence of a contaminant so that appropriate actions can be taken to ensure the safety of the people. At the same time, these sensing systems, which operate in real-world conditions, may suffer from noise and sensor degradation faults. This paper proposes a novel cognitive monitoring system that can be combined with real- time point-trigger sensors for performing contaminant detection in intelligent buildings. The proposed system reduces the occurrence of false alarms by means of a layered architecture and employs cognitive mechanisms to discriminate between the presence of a real contaminant source and a degradation fault affecting the sensors of the sensing system. In addition, the proposed system is able to isolate the building zone containing the contaminant source (or the faulty sensor) and estimate the onset time of the release (or the fault).

► Quevedo, J., Alippi, C., Cuguero Â, M.A., Ntalampiras, S., Puig, V., Roveri, M., Garcia, D., (2013) "*Temporal/Spatial Model-Based Fault Diagnosis vs. Hidden Markov Models Change Detection Method: Application to the Barcelona Water Network*", Control & Automation (MED), 2013, vol., no., pp.394,400, 25-28 June 2013

Abstract: This paper deals with a comparison of two different fault diagnosis frameworks. The first method is based on a temporal/spatial model-based analysis by exploiting a-priori information about the system under study, so fault detection is based on monitoring the residuals of combined spatial and time series models obtained from the network. The second method aims at characterizing and detecting changes in the probabilistic pattern sequence of data coming from the network. Relationships between data streams are modelled through sequences of linear dynamic time-invariant models whose trained coefficients are used to feed a Hidden Markov Model (HMM). When the pattern structure of incoming data cannot be explained by the trained HMM, a change is detected. Here, the performance obtained from this two distinct approaches is examined by using a dataset coming from the Barcelona water transport network.

* For a complete list of Journals and Conferences publications refer to the project website www.i-sense.org

i-Sense Dissemination: Special Sessions

IEEE Computational Intelligence Society

2014 IEEE WORLD CONGRESS ON COMPUTATIONAL INTELLIGENCE

July 6-11, Beijing International Convention Center, Beijing, China

WCCI 2014

2014 IEEE World Congress on Computational Intelligence (WCCI'2014)

Special Session on Computational Intelligence for Cognitive Fault Diagnosis

The proliferation of wireless sensor networks, machine-to-machine communication and Internet of Things (IoT) has made it possible to build large scale, distributed autonomous systems. These systems which are also referred to as Cyber-Physical Systems, typically involve a large number of individual components such as sensors and actuators, communication links as well as complex software and controllers.

During normal operation, these components generate large volumes of data that need to be processed to extract useful information and to make decisions. At any point in time, it is likely that one or more components may fail and as a result the generated data may be inconsistent which may lead to wrong decisions.

The aim of this special session is to present recent advances in computational intelligence that can used in fault diagnosis.

Topics include but are not limited to the following areas:

- Computational intelligence for detecting faulty components
- Computational intelligence detecting and correcting inconsistent data
- Computational intelligence for fault detection/identification/isolation
- Learning and adaptation for fault diagnosis
- Learning from imbalanced data
- Fault diagnosis in time evolving environments
- Change detection test
- Applications

ORGANIZERS

Marios Polycarpou, Cyprus
Christos Panayiotou, Cyprus
Cesare Alippi, Italy

For more details please visit the website of the conference

http://www.ieee-wcci2014.org/

2014 International Joint Conference on Neural Networks (IJCNN 2014)

Special Session on Concept Drift, Domain Adaptation & Learning in Dynamic Environments

The proposed session on concept drift/nonstationary learning has three **main goals**:

1. Introduce the problem of concept drift, domain adaptation, nonstationary learning and more generally dynamic learning, and its associated issues, to the greater neural network & computational intelligence community who may not have been familiar with the topic, yet would like to familiarize themselves with the most recent approaches for solving this problem;
2. Provide a forum for researchers who have been actively working in this area to exchange new ideas with each other, as well as with the rest of the neural network & computational intelligence community
3. Present recent approaches to learning in dynamically changing environments from two perspectives: first, the more traditional and theoretical view of machine learning and computational intelligence; and second, from the more practical and application oriented view of using neural networks.

The **scope** of the proposed session includes but not limited to:

- Incremental learning / lifelong learning / cumulative learning
- Fault, change or anomaly detection algorithms
- Data mining from streams of data
- Domain adaptation, dataset shift, covariance shift
- Learning in non-stationary environments / concept –drift environments / dynamic environments
- Architectures / techniques / algorithms for learning in such environments
- Applications that call for incremental learning or learning in nonstationary environments
- Adaptive classifiers able to cope with concept drift and recurring concepts
- Development of test-sets / benchmarks for evaluating algorithms learning in such environments

ORGANIZERS

Robi Polikar, USA
Manuel Roveri, Italy
Giacomo Boracchi, Italy

News and Contacts

Visit our website

The i-Sense project website can be found at http://www.i-sense.org

In the website you can learn about the last research activities, read the abstracts of our publications, download open-source functions and benchmarks related to core algorithms.

Relevant events and project activities are there, waiting for you.

Click our QR code

Relevant events

1. IEEE International Symposium on Circuits and Systems, June 1 - 5, 2014, Melbourne, Australia, (ISCAS 2014)
2. The 22nd Mediterranean Conference on Control and Automation, June 16 - 19, 2014, Palermo, Italy (MED 2014)
3. International Conference on Machine Learning, June 21 - 26, 2014, Beijing, China (ICML 2014)
4. IEEE World Congress on Computational Intelligence, July 6 - 11, 2014, Beijing, China (WCCI 2014)
5. IEEE International Joint Conference on Neural Networks, July 6 - 11, 2014, Beijing, China (IJCNN 2014)
6. IEEE International Conference on Artificial Neural Networks, September 15-19, 2014, Hamburg, Germany (ICANN 2014)
7. IEEE Multi-Conference on Systems and Control, October 8 – 10, 2014, Antibes, Nice, France (MSC 2014)
8. IEEE Symposium Series on Computational Intelligence, December 9-12, 2014, Orlando, Florida, USA (SSCI 2014)
9. IEEE Conference on Decision and Control, December 15-17, 2013 Los Angeles, CA, USA (CDC 2014)

Contact us

Coordinator
Prof. Marios Polycarpou
mpolycar@ucy.ac.cy

Deputy Coordinator
Prof. Christos Panayiotou christosp@ucy.ac.cy

Visit our website
www.i-sense.org

i-Sense project partners