CONET Newsletter

INFSO-ICT-224053 www.cooperating-objects.eu **Issue nr.15, October 2011**

Contents

Editorial

Welcome to the 15th issue of the CONET newsletter. CONET is the EU FP7 network of excellence on Cooperating Objects, merging the fields of embedded systems for robotics and control, pervasive computing and wireless sensor networks. CONET focuses on establishing the field of Cooperating Objects within the research and industrial community, thus strengthening the position of Europe in the research landscape.

This issue presents a guest column from University of Texas at Dallas (UTD) about data mining in Body Sensor Networks.

The issue's "member profile" presents Telecom Italia Group and some of its activities related to wireless sensor networks and healthcare.

Then an article from a CONET partner (University of Duisburg-Essen) presents the vision and the goals of the PLANET (PLAform for the deployment and operation of Heterogeneous NETworked cooperating objects) project.

Finally a short report of the 3rd CONET summer school is presented.

If you are interested in obtaining up-to-date information about the CONET project please visit our website at: http://www.cooperating-objects.eu

We hope you will enjoy this issue ■

A Mining Technique for Body Sensor Network Data Repository

By Vitali Loseu and Roozbeh Jafari, University of Texas at Dallas

Body Sensor Networks are becoming an increasingly popular field of research for a variety of applications ranging from fall and posture detection and tele-medicine to rehabilitation and sports training. These systems are composed of lightweight wearable sensors that capture different physiological data from the human body. This physiological data may include an inertial description of human movements, Electrocardiograph (ECG) readings of the human heart, Electromyography (EMG) readings of the muscle activity, skin conductance level, blood pressure, and many more. All of these signals can be observed by sensors mounted on wearable devices. However, the modern sensing platforms are not perfect. Along with the useful physiological information they also capture noise and other data collection artifacts. Data collection artifacts are abnormalities in the signal that can be introduced by the specific sensor deployment conditions. For example, the type of a strap used to attach a sensor to the body can significantly affect the recorded observation. Additionally, while on the high level similar movements may look the same, the specific movement execution can introduce variations to the senor observations of those movements. For example, a *sit-to-lie* movement can be performed smoothly, or the subject can throw themselves on the bed and briefly bounce on the mattress. While both movements achieve the same goal, from the inertial sensor perspective they do not look exactly the same.

Need for BSN Data Repository

BSN platforms are desirable because they provide a relatively inexpensive way to collect realistic and, more importantly, quantitative data about the subjects without constraints of the lab environment. A problem that has not received sufficient attention is storing and tracking the collected data. The data collected from these wearable systems are especially valuable in the cases of medical observations. The ability to search and compare BSN observations can potentially shed light on diseases such as Parkinson's Disease, which do not have a cure or even a quantitative, objective diagnostic process. Parkinson's Disease is a neurological disorder, however many of its symptoms, such a slow automatic movements (such as blinking), inability to finish some movements, impaired balance while walking, muscle rigidity, and varies tremors, severely affect human movements and can be observed with the help of inertial sensors. The task is aided by the fact that many devices in our daily lives, such as cell phones, already have inertial sensors built-in. Furthermore, the seamless nature of BSN nodes allows their deployment prior to serious health problems to monitor the onset of the condition. The following example demonstrates the usefulness of the idea. A person can be monitored for an extended period of time (e.g. multiple years) with the help of a few sensor nodes. All of the data is simply collected and stored in the data repository. After some time, this person is diagnosed with a disease that involved gait abnormalities. It would be beneficial to analyse old data and extract gait parameters for disease evaluation, examination of disease progress, and treatment. A data mining approach would be able to identify movements of interest, in this case walking, so that the raw data of movements can be used to extract the required gait parameters.

Design Challenges

BSN sensor nodes are highly constrained in terms of memory, processing resources, and battery lifetime. This means that all of the collected data cannot be stored on the wearable device, communicated wirelessly for an indefinite amount of time, or processed with complicated and possibly slow computational approaches on the device itself. At the same time, they have a potential to produce very large data sets over time. This suggests that the data representation approach needs to significantly reduce the complexity of the data, while maintaining the characteristic structure of the signal. This task is further complicated by the possibility of errors in the signal and inter-subject variability in movement performance. This problem can be solved by applying limited processing to the sensor data, as it is being collected, to reduce its size and complexity. This step, however, needs to preserve the structural parameters of the signal. This can be achieved by applying limited processing that exclusively focuses on identifying transitions in the signal that uniquely characterize each movement. For this step to be successful, it is essential for the system to extract the properties of the signal capable of capturing such characteristic transitions. While in other systems, redundancy may be acceptable and even desirable, the resource and time constraints of the BSNs demand that the considered set of signal properties to be minimal. That means that some of the machine learning and signal processing techniques may not be suitable for implementation on the sensor nodes.

System Operation

We propose a system that consists of a set of wearable nodes placed on the body to collect inertial observations of the human movements, and a computer that maintains the BSN repository and facilitates data organization and mining (as demonstrated in Figure 1). The wearable nodes are connected to the computer via wireless radios. It is desirable to shift the computer functionality to the wearable nodes, and avoid using the battery for wireless communication due to its overwhelming power cost.

The system begins operation by sampling the local sensors of the wearable devices. The local sensor data is communicated to the computer for processing. Information, relevant to a given application, is then extracted from the data on the computer side. The extracted information is used to construct a query to the BSN repository that can return observations already stored in the repository that most resemble the observed data.

Specifically, we first come up with efficient data representation that can reduce the complexity of the data. We then select a distance metric for the new data representation. Finally, we define a classifier in the context of the distance metric selected and the new data representation.

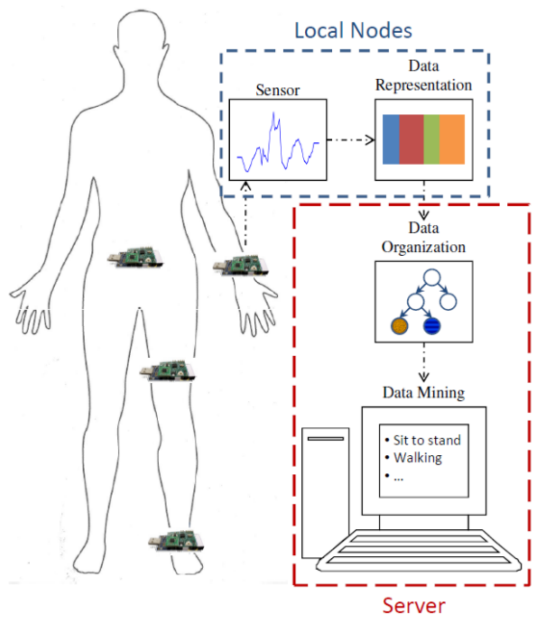

Figure 1: System Operation

Data Representation

A physical movement can normally be represented as a sequence of shorter motions. Capturing the structure of the movement involves capturing these shorter motions and timing relationships between them. This can be done by identifying motion primitives. This task is not trivial as it has to be done without any prior knowledge of the movement itself but with respect to the observed signal. A common way for unsupervised data grouping is clustering. We follow the idea introduced in [1] and use a clustering technique for primitive generation. We extract features from the signal and cluster the resultant feature set, which means that the clustering outcome is dependent on the perspective that the features can provide. This adds flexibility to the system because different feature sets can characterize the signal from different perspectives.

Data Clustering

Clustering is a very effective method of grouping similar data points, and distinguishing between different data points. When trying to cluster BSN data, a clustering approach is normally applied to feature vectors extracted from the original signal. There is a variety of features that can be extracted from inertial data. During our study we tried to identify a small and simple feature set that would produce good results. The resultant clusters should be able to identify enough transitions in the signal so that each movement of interest would be characterized with a unique subset of such transitions.

There is a wide array of clustering techniques that includes hierarchical, partitional, conceptual, and density based approaches. For our analyses we consider expectation maximization in Gaussian Mixture Models (GMM)[5]. This approach tries to identify the centers of natural clusters of the data instead of artificially selecting points in the training set as cluster centers. GMM based clustering computes the probability that any given point is assigned to every individual cluster and makes an assignment that maximizes its likelihood of such assignment.

A major problem to consider during unsupervised clustering is the number of clusters that produces the best results. To find the best solution we varied the possible number of clusters from 2 to the length of the shortest observation in the training set, while evaluating parameters of the GMM model. We used Expectation Maximization (EM) to find the best mixing parameters for GMM. The mixing parameters, such as the mean and covariance matrices, depend on the number of clusters. Once the GMM parameters are selected, the quality of clustering can be evaluated via Bayesian information criterion (BIC).

Comparison Metric

Once the BSN data is converted to motion transcripts the system requires an efficient way to classify and search them. Edit distance is a common approach to compare strings. However, edit distance does not perform well when the input data has noise and varies in length. Additionally, the edit distance calculation is very slow. While it may be an acceptable solution for a small application, its speed performance is not at all acceptable for a large data repository potentially containing terabytes of data. To resolve the issue of edit distance, the system can use the idea of n-grams[4], or substrings of length n, that can track transitions in motion primitives in linear time with respect to the trial length. The goal of the n-grams is to track important transitions between movement primitives in string transcripts. However, the task of identifying n-grams that represent important transitions is not simple. The difficulty of the task is increased since overlapping n-grams are extracted to improve the quality of the recognition. This means that potentially there is a very large number of n-grams that can be selected from any given transcript. Furthermore, we expect the data to be noisy; and it is important to train the system only on the n-grams that correspond to the characteristic transitions in the signal and not on the n-grams that represent noise. This task is achieved via n-gram selection.

N-gram Selection

The objective of this operation is to identify a small number of n-grams that can uniquely characterize the movement of interest, and provide means of distinguishing that movement from others in the repository. There is a variety of ways to select proper n-grams, once n-grams are extracted from all of the training data. Information gain (IG) has proven to be effective in the field of natural language processing. Specifically,

$$Gain(m_i, f) = H(m_i) - H(m_i|f)$$

where H(m) defines entropy of the movement set, and H(m|f) defines conditional entropy of the movement set with respect to feature f.

Information gain becomes complicated to compute and less effective when each evaluated feature can take a large number of values. However, in our experiment each n-gram has two possible values. A specific n-gram can be present in a motion trial and the value of '1' is assigned to it, or the n-gram can be absent with a value of `0' assigned.

IG can assess the effectiveness of a feature by tracking changes in the entropy after consideration of that feature. Once all the n-grams have an IG assigned to them for each movement, we can sort the list of IGs and select a subset of n-grams that have the best IG.

Classifier

Once the set of good n-grams is selected, an approach needs to be defined for fast movement classification and search. This approach also should not rely on the knowledge of the complete structure of the data, and be able to finish classification and search based on partial information. These properties are exhibited by suffix trees; more specifically, we used the Patricia tree in our implementation [3]. Patricia trees are used to represent sets of string by splitting them into substrings and assigning substrings to the edges. This idea fits naturally with n-grams that are substrings. Once all of the n-grams are selected for each movement, we combine them and assign the combined set to the edges of a Patricia tree. The paths from the root to all leafs corresponds to all the possible permutations of the combined n-gram set (as demonstrated in Figure 2).

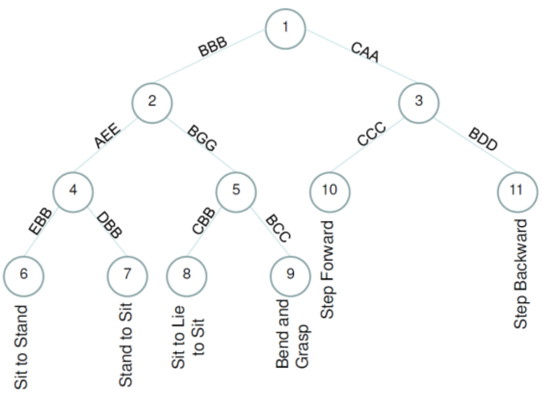

Figure 2: Patricia Tree

Data Mining Model

Based on the construct defined earlier we propose a data mining approach. The approach has two distinct parts: **Training** and **Query Processing**. During the training phase of the system execution, the system defines clusters, selects important n-grams, and computes the classifier based on the training data. During the query processing phase of the execution, the system uses the parameters learned during the training process and makes classification decisions.

Experimental Verification

During the experimental verification, we demonstrated that our approach can achieve an average 96% precision, and 99% recall for a pilot application. For the detailed results, discussion, and trade-off evaluation see the full paper [2].

[1] H. Ghasemzadeh, V. Loseu, and R. Jafari, "Collaborative signal processing for action recognition in body sensor networks: a distributed classification algorithm using motion transcripts," in Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks. ACM, 2010, pp. 244–255.

[2] Vitali Loseu, Hassan Ghasemzadeh, Roozbeh Jafari, A Mining Technique Using N-grams and Motion Transcripts for Body Sensor Network Data Repository, accepted for publication in *Proceedings of the IEEE (*PrIEEE*)*.

[3] S. Inenaga, H. Bannai, A. Shinohara, M. Takeda, and S. Arikawa, "Discovering best variable length don't care patterns," in Discovery Science. Springer, pp. 169–216.

[4] C. Shannon, "Communication theory of secrecy systems," MD Computing, vol. 15, no. 1, pp. 57–64, 1998.

[5] H. Friedman and J. Rubin, "On some invariant criteria for grouping data," Journal of the American Statistical Association, pp. 1159–1178, 1967.

ⓘ http://www.essp.utdallas.edu/

Member Profile: Telecom Italia

By Roberta Giannantonio, Telecom Italia

The Telecom Italia Group is a major Italian enterprise and a key European strategic ICT player. Driven by technological innovation and a commitment to service excellence, Group companies operate in fixed-line and mobile telecommunications, Internet & Media, Information Technologies.

Most of the R&D activities of the Group are performed inside the Innovation and R&D Departments of Telecom Italia and involve around 4500 researchers and technicians with an average investment of 650 Meuro per year.

TILab is the department whose remit is the supervision of technological innovation for the Group, scouting for new technologies and engineering operations for services and network platforms.

The work carried out by the R&D Departments is the outcome of a strategic partnership with the main manufacturers of telecommunications equipment and systems, and with centres of excellence in research at the most highly qualified national and international academic institutions. On the international level Telecom Italia has pledged a substantial commitment to the task of standardisation and has been involved from the beginning in the European Union Framework Programmes starting with the first pilot projects of the ESPRIT programme in 1983 and continuing as one the primary European collaborators in terms of both finance and the number of projects. In the Seventh Framework Programme Telecom Italia is at present involved in 24 different Projects.

Main research areas are: the evolution of mobile communication, from third generation mobile systems to a variety of overlapping wireless networks increasing access flexibility; the diffusion of broadband bandwidth, studying affective techno-economic solutions to deploy optical fibers; the dissemination of identification and localization systems embedding tagging technologies within telecommunication functionality.

Examples of specific research projects are the following: Wireless Sensor Network Applications, Context Awareness/Ambient Intelligence Platforms & Services, Innovative Services and Applications, Connected Car, 3D Multimedia Technologies, Software Defined Radio, e-Tourism.

Among the numerous research activities of Telecom Italia, it is worth mentioning an activity related to Body Wireless Sensor Networks.

Wireless Body Sensor Networks (BSNs) possess enormous potential for changing people's daily lives. They can enhance many human-centered application domains such as m-Health, sport and wellness, and also people-centered applications such as physical/virtual social interactions and emotion recognition.

In this project a framework has been developed for simplifying the design of Wireless Sensor Networks applications, with a particular attention for the area of health care and energy management. This project is particularly active in the Scientific Community, by means of: many publications about healthcare technologies; the promotion of international conferences; the participation to European collaborative projects and International Standardization Bodies. In particular, among the EU-funded projects under ICT, Framework 7, Telecom Italia participates to CONET, the Cooperating Objects Network of Excellence, where it manages the REWSN (Recognizing Emotions using Wireless Sensor Networks) research cluster. The aim of this activity is to define, implement and test algorithms aimed at recognizing emotional conditions, based on parameter values detected by sensors either available in the environment or worn by a subject. Wearable sensors (or Body Sensor Networks BSN) are the main focus of an Open Source project called SPINE.

SPINE (Signal Processing In Node Environment) is an open source domain-specific framework designed to support flexible and distributed signal processing for wireless body sensor network (WBSN) systems.

The main goal of SPINE is to provide WBSN developers with support for rapid prototyping of signal-processing applications. In SPINE, sensors and common processing blocks, such as math aggregators and threshold-based alarms, can be configured independently and connected together arbitrarily at run-time based on external controls. Such an approach allows heterogeneous applications to be built atop the same basic software components, enhancing code reusability and, more importantly, removes the need for redeploying the node side code based on a particular application.

Concerning standardization, Telecom Italia is active in many standardization bodies.

In particular, it has the leadership of ZigBee Telecom Services in the ZigBee Alliance. ZigBee is the only standards-based wireless technology designed to address the unique needs of low-cost, low-power wireless sensor and control networks in just about any market. ZigBee Telecom Services offers a global standard for interoperable products enabling a wide variety of value-added services, including information delivery, mobile gaming, location-based services, secure mobile payments, mobile advertising, zone billing, mobile office access control, payments, and peer-to-peer data-sharing services.

ⓘ http://www.telecomitalia.com

■

PLAform for the deployment and operation of Heterogeneous NETworked cooperating objects

By Chia-Yen Shih, University of Duisburg-Essen

The continuous miniaturization of embedded devices as well as the convergence of communication and control has been considered by many future information and communication technologies (ICT) [1] [2]. Such convergence will provide the ability to build large-scale, heterogeneous, networked systems that can be deeply embedded in the physical world. The underlying technology that serves as the basis for this field is Cooperating Objects (COs), which do not only have the capability to sense, but also can act upon the environment.

In general, the deployment of very large-scale and complex CO systems involves in dealing with a variety of heterogeneous COs (such as unmanned vehicles and wireless sensor/actuator networks) that collaborate with each other in a seamless fashion and integrate with pre-existing infrastructure in a transparent way. Thus, a solution that is applicable to a wide range of application scenarios requires the development of new distributed architectures and integration platforms that have the following characteristics:

- The platform should be able to cope with heterogeneity of devices in either hardware or capability such as sensing and mobility, etc.
- The platform should support the planning, deployment and maintenance of real-world applications.
- The platform should support security natively (not just as an add-on) regarding the algorithms used in the deployment as well as the planning and maintenance of the network.

The goal of the PLANET project is to provide an integrated planning and maintenance platform that enables the deployment, operation and maintenance of heterogeneous networked COs in an efficient way. Figure 3 depicts the PLANET's approach in offering iterative procedures in support deployment planning and network management. Furthermore, PLANET particularly emphasizes the capability of the platform in supporting deployment and operation strategies for large-scale systems composed of unmanned ground and aerial vehicles cooperating with wireless sensor networks.

Key Directions

In summary, the scientific and technical objectives of PLANET are as follows.

- Integrated Platform: Design and develop an integrated framework for monitoring of very large-scale systems. The architecture will be general, modular and extensible, and will integrate large number of multiple heterogeneous systems and smart objects including unmanned systems and a variety of field sensors for surveillance and monitoring of dynamic events.
- Adaptive Network Deployment: This objective deals with the development of new methods for optimal and adaptive network deployment. This will include the initial planning, deployment and testing, as well as strategies for ad-hoc autonomous deployment by means of the cooperation of mobile objects.
- Autonomous Systems for Deployment: This objective will include unmanned aerial vehicles, unmanned ground vehicles and combinations of both in order to perform the deployment, redeployment and retrieval of sensor nodes in inaccessible sites.
- Network-centric Computing: This objective deals with the development of techniques for distributed network-centric computing with dynamic resource discovery and management. This will include optimal data collection with autonomous vehicles (UAVs and UGVs)

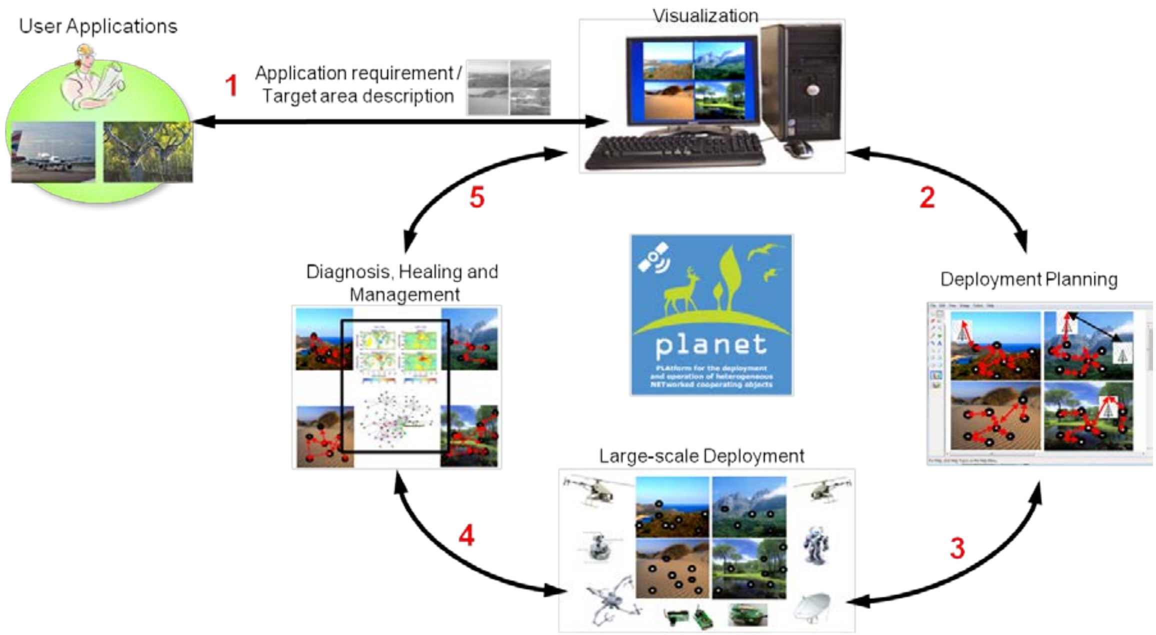

Figure 3. The PLANET Approach (Source of airplane: Phillip Capper via Wikimedia)

or through virtual sensing, distributed mission planning and task allocation, including distributed techniques for optimal management of mobile/static objects and optimal coverage.

- The objective is to investigate and define Security methods that, starting from real deployment scenarios, type of possible threats and architecture and resource constraints, make it possible to build security solutions for Cooperating Objects and WSNs. These methods need to accommodate the presence of a large number of heterogeneous, possibly mobile devices, support and promote modularity and allow dynamic adaptation to the changed operational conditions. The following issues will be considered: authentication and key management, security vs. performance/life-span, mechanisms for remote attestation and distributed algorithms for anomalous behaviour detection.

Experimental Validation

The PLANET approach will be validated in two realistic application scenarios. The first scenario is the environmental management system. The objective is to validate the PLANET platform in the Doñana Biological Reserve [3]. Here the need of cooperating objects and wireless sensor networks can be better justified in the case of especially valuable natural environments very sensitive to the impact of pollution, and the risk of natural disasters and other threats (catastrophic floods, droughts, spread of invasive plants, diseases, etc.). This requires the deployment of devices for the monitoring of pollution events, monitoring of natural disasters and inaccessible areas, localization and geo-referencing of objects (animals, etc.) and surveillance.

Figure 4. DBR Fauna

The second scenario is complementary to the first one. It is the highly automated airfield scenario in which security plays an important role and where the wireless communication and networking have different characteristics, involving significant radio-electric transmission problems and a large number of high velocity mobile objects. The objective is to validate the PLANET platform in the highly automated airfield scenario, which involves the cooperation among aerial systems, systems on board, ground vehicles, control systems and a variety of sensors for localization, local airfield measurements and in general, airfield resources condition monitoring. PLANET aims at providing support for airfield operations with cooperating objects, including the study of the required networking and communication infrastructure, management of ground operations, and emergency network deployment using unmanned vehicles.

Figure 5. Highly Automated Airfield (Source: Mario Roberto Duran Ortiz and Phillip Capper via Wikimedia)

Both scenarios, although complementary, have very different requirements regarding the QoS needed from the systems and, thus, enable the PLANET platform to show its adaptation and self-organization capabilities in diverse settings.

PLANET addresses the design methodology and development of the platform as well as the algorithms required to support the deployment and maintenance of heterogeneous systems with mobile and static nodes. The most challenging algorithms that will form the core functionality of the PLANET platform are: optimal planning to achieve optimal coverage, network-centric sensing and actuation, cooperative transportation and retrieval of nodes, data muling techniques, secure, non-intrusive monitoring of the network for failures and possible threats and failure recovery for the application system integrity. The PLANET solution aims at combining the capabilities of these different disciplines into an integrated solution to cope with the complexity of real world applications.

ⓘ www.planet-ict.eu

References

[1] B. Andersson, N. Pereira, W. Elmenreich, E. Tovar, F. Pacheco, and N. Cruz. A scalable and efficient approach for obtaining measurements in CAN-Based control systems. IEEE Transactions on Industrial Informatics, 4(2):80–91, 2008

[2] John Bicket, Daniel Aguayo, Sanjit Biswas, and Robert Morris. Architecture and evaluation of an unplanned 802.11b mesh network. In MobiCom '05: Proceedings of the 11th annual international conference on Mobile com-

puting and networking, pages 31–42, New York, NY, USA, 2005. ACM.

[3] Doñana Biological Reserve. [Online] http://www.ebd.csic.es/website1/reserva/reserva.aspx.

■

3rd CONET Summer School "Networked Embedded Systems: Humans in the Loop"

By Kay Romer (University of Lübeck, Germany and ETH Zurich, Switzerland), Friedemann Mattern (ETH Zurich, Switzerland), and Luca Mottola (Swedish Institute of Computer Science)

Continuing the tradition set in previous years, CONET successfully organized a summer school featuring a series of lectures by distinguished scholars addressed to a selected audience of students and researchers, from within and outside the EU. The school took place the last week of July 2011 at the University Residential Center in Bertinoro (Forlì, Italy).

This year's theme of the school concerned the role of humans in the design, implementation, use, validation, and deployment of systems of Cooperating Objects. The goal was to explore the manifold relationships between Cooperating Objects and humans as their creators, users, and subjects. Lectures covered topics as diverse as deployment, debugging, and verification of networked embedded systems; sensing and affecting the human behaviour; participatory sensing; human-computer interaction; as well as security and privacy. The subjects were then cast in various application domains, from health care to smart energy, giving attendees a concrete feeling of the real-world challenges at stake.

Throughout the school, lecturers from top institutions in Europe, Australia, and the US engaged the attendees in discussions about current hot research topics in the broad field of Cooperating Objects. The 60 participants, selected from 28 countries among a total of 125 applications received, also had the opportunity to experience the relevant technology during a hands-on sensor network programming course and a programming contest that gave them the opportunity to collaborate on concrete problems. Some selected students also had a chance to present their on-going research work to the other participants and to the lecturers in a participants' workshop, gathering further feedback. The program was complemented with a guided visit to the nearby town of Urbino, a remarkable witness of Italian Renaissance.

The CONET summer school is now established as a prestigious venue for postgraduate students, PhD students, and young researchers from universities and industrial labs to obtain feedback from distinguished scholars and to network with peers. As it happened in the past, we hope this year's summer school will lay the foundation for longer-term collaborations among the participants.

■

Announcements

EWSN 2012 – European Conf. on Wireless Sensor Networks

February 15 – 17, 2012, Trento, Italy
ⓘ http://ewsn12.disi.unitn.it/
Demo-Poster Submission Deadline: December 5, 2011

ROBOT 2011: Experimental Robotics

November 28 and 29.

ⓘ http://grvc.us.es/robot2011

RED-UAS 2011. Research, Development and Education on Unmanned Aerial System

30 November-1 December.

ⓘ http://www.red-uas.com

Latest News

Contiki 2.5 is finally out!

New in Contiki 2.5 are ContikiRPL, the Contiki implementation of the new IETF RPL IPv6 routing protocol, ContikiMAC, a radio duty cycling mechanism that allows routing nodes to keep the radios off for more than 99% of the time, Contiki Collect, a complete rewrite of Contiki's native data collection protocol, and an implementation of the IETF CoRE CoAP Protocol.

ⓘ http://www.contiki-os.org/