GREASE Framework
Generic Reconfigurable Evaluation and Aggregation of Sensor Data

Matthias Vodel, Rene Bergelt, and Wolfram Hardt
Dept. of Computer Science
Chemnitz University of Technology
Chemnitz, GERMANY
Email: { vodel | berre | hardt }@cs.tu-chemnitz.de

***Abstract*—The proposed research work represents a generic, energy-efficient concept for the synchronised logging, processing and visualisation of any kind of sensor data. The concept enables a chronological coordination and correlation of information from different, distributed sensor networks as well as from any other self-sufficient measurement systems. Based on the achieved relation between the several sensor sources, the information quality can be increased significantly. Therefore, the system-wide description of the monitoring scenario and each data set is realised in XML. Accordingly, the aggregated, heterogeneous sensor information are convertible into multiple output formats. Dependent on the application specific requirements for the visualisation, we are able to consider additional meta-information from the test environment to optimise the data representation. The definition of advanced data fusion techniques and pre-processing mechanisms allows a selective data filtering to shrink the network load. To evaluate of basic usability requirements and the efficiency of the proposed concept, an automotive sensor network represents a capable test system for the proposed framework. Within the demonstrator, the available on-board measurement systems were extended by high-precision sensor nodes, which establish wireless sensor network topology. Afterwards, the correlated measurement information were converted and visualised for several professional data analysis tools, e.g., jBEAM, Google Earth, and FlexPro.**

***Keywords*-Data Aggregation; Data Fusion; Data Synchronisation; Heterogeneous Wireless Sensor Network; Sensor Actuator Systems.**

I. INTRODUCTION

Actual research projects in the field of wireless sensor networks operate on different, proprietary hardware platforms and contain multifaceted types of sensors. Currently, each measurement scenario consists of several application-specific and independent operating processes for the data-collection, -storage and -analysis.

It does not exist any uniform synchronisation techniques between the autonomous sensor systems. Accordingly, a detailed and target-oriented post-processing of the data sets within a shared knowledge base is not feasible. In consequence, we are not able to create unique relations between the different measurement information. Due to these missing relations, it is very hard to create a common primary index for the given, heterogeneous sensor platforms.

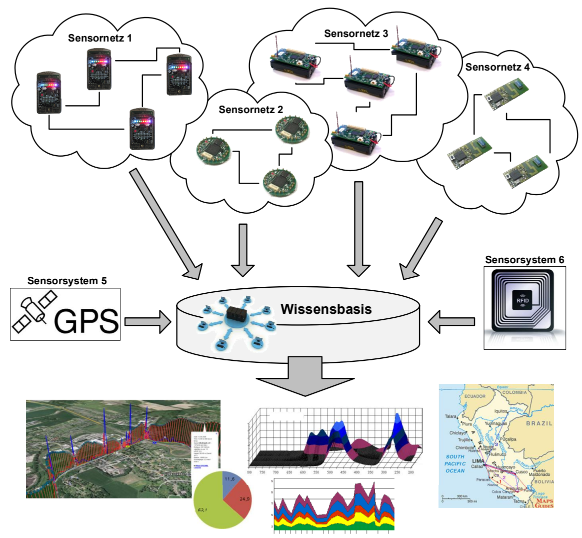

Figure 1. GREASE - A synchronised data-processing framework allows an efficient integration of different, autonomous sensor networks.

To solve this problem, we present *GREASE* - a Generic Recon-figurable Framework for the Evaluation and Aggregation of heterogeneous Sensor Data (see *Figure* 1). In order to introduce this integrated data processing concept, this paper is structured as follows: After this introduction, section *II* provides an overview about heterogeneous, distributed sensor environments, the data processing flow and respective challenges. The proposed GREASE framework is introduced in Section *III*, including conceptual fundamentals, basic requirements, system parameters, and the top level structure (Section *IV*). Accordingly, Section *V* provides implementation details of the GREASE software architecture as well as the overall application flow within the framework. Section *VI* specifies the application scenarios with all integrated components and the environmental conditions. The respective data analysis is described and discussed in Section *VII*. Finally, the paper concludes with a summary and an outlook for future work in this research project.

 ISBN: 978-1-61208-189-2

II. Related Work

During the last two decades, a couple of commercial tools for the measurement data recording and monitoring were developed. Unfortunately, most of them have functional or conceptual restrictions. Some of the vendors offer exclusive, hardware-specific analysis tools, which require special devices, predefined product series or vendors. Other sensor systems do not have any special software tools for extracting the measured data sets. There is no support for further post-processing steps.

In consequence, *LabView* from National Instruments [1], *jBEAM* from AMS [2] or FlexPro [3] offer multiple features to enhance the restricted vendor tools, which only provides a small set of general data recording an handling functions. These applications allow the interpretation of offline data from data bases or files as well as the live analysis of a given data source. Both jBeam and LabView operates platform-independent and all of these related tools support a lot of established data formats and communication interfaces. Especially jBeam, which integrates the *ASAM* standard *(Association for Standardisation of Automation and Measuring Systems)* [4], enables an easy and modular extension with user-defined components. Furthermore, FlexPro includes a lot of additional visual plugins and represents a complete visualisation framework for the given measurement data.

The very high system requirements of all the related software applications represent a critical disadvantage. Accordingly, these tools are not capable for resource-limited data recording environments. Thus, such frameworks have to be used in a second data processing step on dedicated workstations with sufficient hardware components. Hence, small and energy saving hardware system, which are used exclusively for collecting and storing multiple data from different sensor sources, are not able to use the data aggregation [5] and visualisation features of these software frameworks. Due to these circumstances, most of the ongoing sensor system projects use proprietary software solutions to organise and synchronise the collected measurement data [6]. In fact, there are many critical compatibility problems between such software tools. In consequence, modifications of the measurement scenario or the system configuration take a lot time and bind important resources. In conclusion, the user wants a universal software tool for collecting and analysing the entire pool of sensor information in an application-specific and resource-efficient way. Automated or semi-automated data visualisation techniques represent further essential requirements.

III. Concept

We are now looking for generic utilities and standards to route information from different sensor systems into a common data processing unit in a synchronised way, considering scenario-specific configuration schemes and sensor parameters. Hereby, synchronised time stamps for the heterogeneous sensor data sets are very important to allow correct correlations in the common knowledge base. Furthermore, such utilities allow us to define user-specific data analysis procedures during the measurement runtime and advanced data fusion techniques [7][8][9] to shrink the data volume directly within the sensor nodes.

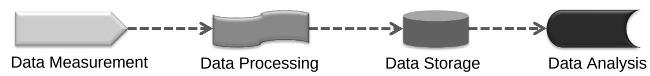

Figure 2. GREASE - Integrated data processing flow for heterogeneous measurement topologies.

To provide such features, GREASE represents a software framework based on a capable and lightweight data management concept, which is able to bypass the already mentioned disadvantages. It combines advanced sensor network configuration features with resource-efficient operating parameters. GREASE integrates the entire data processing flow in an efficient way, including all stages like the data measurement, the data processing, data storage and finally data analysis tasks. *Figure* 2 illustrates this flow.

The concept focuses on resource-limited systems and has to be feasible for a wide variety of application scenarios. Thus, the primary objective is a dynamic and flexible processing environment, which is adaptable to modifications in the configuration or in the analysis requirements. Furthermore, the data processing core has to be separable into two spatial, chronological and platform-specific operating modes. All components for the data measurement are working within the first mode. All relevant modules for the data analysis as well as possible visualisation plugins operate independent within the second mode. Based on this requirement, we are able to map different data processing functions to predefined configuration scenarios. In contrast, other related software tools do not separate the data handling process into different phases in an efficient way.

GREASE deals with a standardised data transport and definable synchronisation parameters for collecting information from several distributed sensor components. Accordingly, changes in the data analysis process have no effects on the components of the data recording. This feature provides significant benefits, especially for complex sensor systems or inaccessible measurement environments.

In addition, all GUI *(Graphical User Interface)* actions, which are accessible by the user, also have to be executable and controllable in an automated or semi-automated way. This feature represents another important difference to other related software tools, which not provide any script-based operating mode without GUI. But especially for continuous maintenance-free sensor measurement scenarios, the scripting of user-defined activities is essential.

All central requirements for an synchronised data logging, processing and visualisation framework, specially in the field of heterogeneous sensor network systems can be summarised as follows:

- Synchronisation of different, autonomous sensor systems
- Modular extension with plugins and an easy modification / adaptation
- Using *XML ((Extensible Markup Language)* as common data exchange format instead of proprietary data types
- Offline and live data analysis from files, network file systems or databases
- Graphical User Interface for configuration and maintenance
- Automated or semi-automated data analysis and data representation mechanisms

Based on the proposed concept, the developed framework act as coordinator between application-specific components. The framework itself operates as a generic coordinating unit and includes no application-specific logic.

IV. Structure

As already mentioned, the structure of the proposed concept is divided into two operating modes. The first one encapsulates the data recording, synchronisation and correlation. A second mode processes the data analysis and generates a user-defined representation of the information sets. Due to the modular operating concept, the sensor net framework is completely independent from the given sensor configuration. Therefore, the environment uses an end-to-end communication design, called the *hourglass architecture*, which enables maximum interoperability between the several components. It means, that multiple sensor units and corresponding data processing units are connected by a dedicated *SensorController*. The sensor controller ensures a specific, universal mapping of the sensor information into a predefined format (see *Figure* 3). For the data output, the *SensorLogReader* component provides different modules for the information representation. The result is a very high diversity on both data input and data output components. In contrast, the binding middle part shows a strict uniformity.

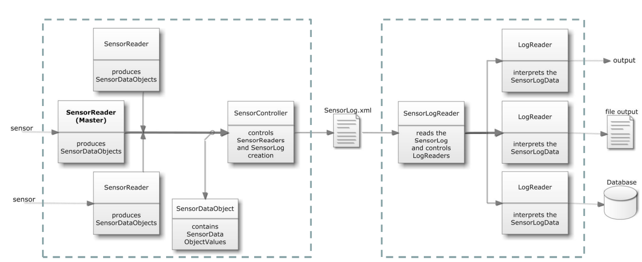

Figure 3. Sensornet framework structure. The separation into two operation modes for the data collecting (left) and data analysis (right) is described.

This structure fulfils the central requirement of a separated data processing core for gathering and analysing the sensor measurements. Hereby, the common XML representation of the entire scenario is essential and allows us to transfer the information for any kind of application. All internal and external parameters of the environment as well as special meta-information regarding the measurement scheme are correlated together with the data sets. Accordingly, researchers are able to reconstruct the whole test scenario with synchronised data, time stamps and a detailed system configuration. The reuse factor, for instance in the field of automotive testing scenarios, increases substantially.

Furthermore, we have the ability to modify existing sensor configuration in a time efficient way. The structure allows us to add or remove single components without changing the data flows within the overall monitoring system. In principle, it is also possible to create a direct interconnection between data collecting and visualisation. This increases the runtime significantly and reduces the used resources. But without the common XML representation, is not possible to capture the entire measurement scenario in a reusable way.

V. Implementation & Application Flow

To enable a platform-independent operation, the proposed concept was implemented in *Java*. Thereby, the resulting framework is also able to include external components, which are written in another programming language. A precondition is the common interface specification within the SensorController. To provide such a feature, the proposed framework implementation contains a *central core library*, which encapsulates the entire data processing logic. This library can also be used for the development of further modules. The communication between the SensorController and the SensorLogReaders with its set of modules uses predefined protocols. All of these protocols are designed as generic as possible to allow a universal usage. This flexibility also simplifies the integration of third party modules and ensures the compatibility during further developments [10]. Due to the fact, that the sensor configurations and all kinds of scenario parameters are also transmitted within the XML representation, possible enhancements for customer-specific applications can be done in an easy way.

Regarding to the application flow, the initialisation of the SensorController starts with loading a application-specific configuration file. This file contains all information for the actual project as well as the structure of all corresponding *SensorReader* components. Accordingly, the SensorController loads and activates all necessary sensor components and starts the data recording. Each SensorReader module operates simultaneous as a dedicated thread. When a SensorReader receives a data set, one *SensorDataObject* will be generated, which is predefined by the framework configuration scheme. Additional meta-information, for example physical measurement units or special indicators, are also included. Afterwards, the SensorController receives the respective signal for the completed object. The object will be transmitted. During the following data processing, the controller analyses all correlations to information from other SensorReaders. Finally, the correlated and classified data set will be stored as an XML data structure.

If the *SensorLogReader* is reading and analysing an XML log file in a reverse process, each data set will be converted into a predefined, framework-conform object and accordingly provided to the data analysis components. The data sets can be reused for multiple representation or visualisation output formats. An essential advantage of this framework system is the fact that the data source is not restricted to local log files. Also data base systems or network file systems can act as data input for further analysis and post-processing steps.

In respect of the communication tasks in the sensor environment, we also have to discuss security features. Due to the fact, that GREASE focuses on research & development environments, we actually do not consider further security aspects for the distributed handling and storage of the sensor data. Within the different development stages of a given system, engineers design and implement complex test environments for getting valid and high-quality measurement results. Accordingly, the risks, which result from general communication threats are negligible. Anyway, we actually cooperate with related German car manufacturers in regards to this weak point. Several research projects focus on the development of advanced, energy-efficient security features for embedded, resource-limited sensor network topologies. In this context, the main challenge is the maintenance of a lightweight software architecture, which provides stable and flexible modules for diversified application scenarios. Here, advanced security mechanisms have a direct impact on the data throughput and the resource consumption. Accordingly, our goal is to find a good trade-off between runtime performance and security capabilities within GREASE.

VI. Application Scenarios

The proposed concept was developed to manage several sensor net scenarios at our computer engineering department. To clarify functional aspect of the implemented framework, we describe the data processing flow by a real-world automotive measurement system [11]. For this monitoring scenario, the existing sensor components of a given research vehicle were upgraded with high-definition sensor nodes. These nodes are placed at predefined positions to monitor the entire environment and provide independent measurement data about the current temperature, light intensity as well as the acceleration in two axes and the magnetic field strength. Thus, the established wireless sensor network provides meta-information about the measurement environment and external parameters. *Figure* 4 illustrates the measurement scenario.

The wireless sensor communication infrastructure is based on the *IEEE 802.15.4* and *ZigBee* [12][13]. Additionally, mobile sensor nodes are worn by the passengers. In order to realise localisation features for these nodes, they are equipped with *nanoPAN* ultra-low power network interfaces [14], which provide *RSSI*-based *(Received Signal Strength Indication)* distance information. Both communication technologies are using the 2.4 GHz frequency spectrum for the data transmission. A multi-interface, multi-standard data sink is able to handle both communication standards simultaneously. Robust communication stacks with adapted layer 2 and layer 3 protocols minimise interference-based influences on the communication behaviour.

In addition, we integrated a high-resolution *GPS (Global Positioning System)* sensor, which enables the correlation between absolute positioning information, speed, altitude and the available on-board vehicle data.

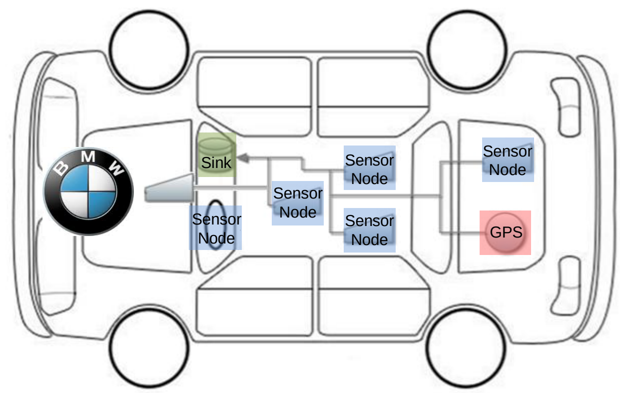

Figure 4. Measurement system - All data from the sensor nodes and the GPS module are transmitted to the data sink in the vehicle, represented by the proposed framework.

By providing a synchronised knowledge base of all sensor information, a detailed analysis of specific driving situations and the driver behaviour is possible. Thereby, the GPS data allows a verification of these situation based on available track information. Accordingly, we are able to calculated and predict driver profiles. The results are used to adjust and to optimise the characteristics of the entire vehicle, for instance, the engine management system or the suspension dynamics. Furthermore, an analysis of the wear measuring quantity provides interesting statements about the vehicle lifetime.

For this measurement environment, special SensorReader modules for the data sink communication interfaces were implemented. Incoming data from the sensor topologies are classified and converted into abstract data objects, which are transmitted to the controller. Another SensorReader module implements features for the GPS data input. Thereby, the *NMEA 0183 (National Marine Electronics Association 0183)* protocol for the positioning data is required. Accordingly, all kinds of GPS hardware, which supports the NMEA

 ISBN: 978-1-61208-189-2

protocol and the serial port as communication interface, are supported.

For the synchronisation of the several sensor data, one specific SensorReader has to be predefined during the initialisation of the measurement scenario.

In our exemplary case, the system provides a high-resolution GPS unit. Besides positioning information, this sensor also provides an accurate time signal. In consequence, the given time stamps from the GPS sensor represent the global *synchronisation master*. Thereby, the framework is not restricted for using a time stamp as master index. Especially for the integration of multiple, autonomous sensor systems and a missing central scheduling entity, a user-defined choice for the synchronisation master provides important benefits.

VII. Data Analysis

For post-processing the collected data, two data analysis components were implemented. The first one is an export module, which prepares the sensor data sets for the storage in a given database system and accordingly transmits the chosen information. A second module is responsible for converting the sensor data with dedicated visualisation plugins, e.g., for Google Earth. Hence, the data output of this module is a KML representation of all correlated sensor information. *Figure* 5 describes the data flow.

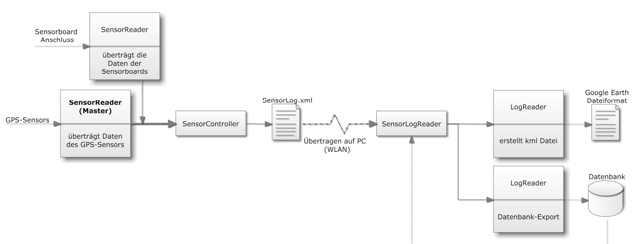

Figure 5. Application scenario of the proposed framework.

Within Google Earth, an additional 3D altitude track extends the visualised measurement curves (represented in *Figure* 6). The entire data processing flow integrates all proposed features for the data recording, handling and visual representation. All developed components are as generic as possible. This also includes a high compatibility level for both hardware and software environments [15][16].

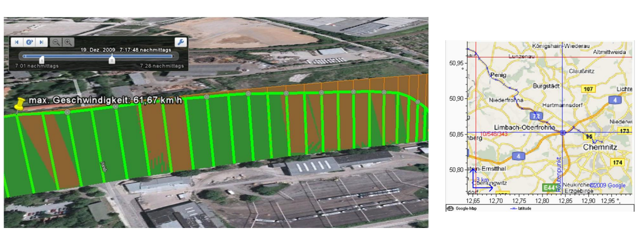

Figure 6. Data visualisation in Google Earth.

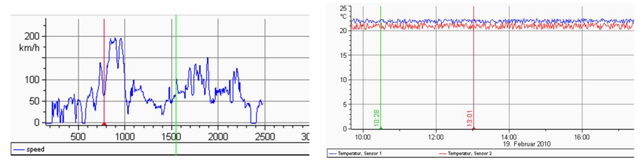

Figure 7. Data visualisation in FlexPro.

Other commercial software tools, e.g., FlexPro, or jBeam (see *Figure* 7, are able to import these information for the database for advanced, sectoral data post-processing tasks. Therefore, a dedicated *CSV (Comma Separated Values)* export module ensures a universal exchange interface.

VIII. Conclusion and Future Work

The proposed research work described the implementation of a comprehensive data processing environment for heterogeneous sensor systems. The basic concept provides generic structures for many further research projects in the field of novel data aggregation and data fusion techniques. For an easy data collecting and data analysis process, we are now able to synchronise and correlate the single data sets also on resource limited and embedded computer systems. The result is a common and extensive knowledge base, which integrates all information sources into complex data sets.

In comparison to other related software tools, the proposed framework fulfils essential requirements for a flexible usage, a resource-efficient runtime behaviour as well as a automated or semi-automated operating mode. The presented framework is used for several wireless sensor and actuator network projects at the Chemnitz University of Technology. We developed a standardised process for monitoring and archiving data from a heterogeneous sensor network topology in a synchronised way. Besides storing basic information from the sensor data sets, the system also integrates meta-information from the environment to increase the reuse factor of the measurement scenario. The universal XML data representation and a modular plugin system ensure a generic usage for all kind of sensor scenario. Multiple data input and output interfaces provides a high level of compatibility to other software tools and data formats.

Regarding the presented automotive application scenario, the proposed framework enables correlations between the measured sensor data from the test track and specific driver profiles. Accordingly, these information allow dynamic adaptations of the driving parameters within the vehicle. This offers novel and interesting possibilities to optimise a vehicle for the specific characteristics of its driver.

References

[1] National Instruments. LabView. http://www.ni.com/labview/, 2010. [Online, retrieved: January, 2012].

[2] AMS GmbH. jBEAM. http://www.jbeam.de/german/produkte/jbeam.html, 2010. [Online, retrieved: January, 2012].

[3] Weisang. FlexPro. http://www.weisang.com/, 2010. [Online, retrieved: January, 2012].

[4] ASAM Consortium. Association for Standardisation of Automation and Measuring Systems. http://www.asam.net/, 2010. [Online, retrieved: January, 2012].

[5] L. Krishnamachari, D. Estrin, and S. Wicker. The impact of data aggregation in wireless sensor networks. In *Proceedings of the 22nd International Conference on Distributed Computing Systems Workshops*, pages 575–578. IEEE Computer Society, November 2002.

[6] M. Vodel, M. Lippmann, M. Caspar, and W. Hardt. Distributed high-level scheduling concept for synchronised, wireless sensor and actuator networks. *Journal of Communication and Computer*, 11(7):27–35, November 2010.

[7] V. Gupta and R. Pandey. Data Fusion and Topology Control in Wireless Sensor Networks. *WSEAS Trans. Sig. Proc.*, 4(4):150–172, 2008.

[8] H. Qi, S. S. Iyengar, and K. Chakrabarty. Distributed Sensor Fusion - A Review Of Recent Research. *Journal of the Franklin Institute*, 338(1):655–668, 2001.

[9] H. Qi, X. Wang, S. S. Iyengar, and K. Chakrabarty. Multisensor Data Fusion In Distributed Sensor Networks Using Mobile Agents. In *Proceedings of International Conference on Information Fusion*, pages 11–16, August 2001.

[10] A. Brown. *Component-Based Software Engineering*. Wiley-IEEE Computer Society Press, 1996.

[11] M. Vodel, M. Lippmann, M. Caspar, and W. Hardt. A Capable, High-Level Scheduling Concept for Application-Specific Wireless Sensor Networks. In *Proceedings of the World Engineering, Science and Technology Congress*, pages 914–919. IEEE Computer Society, June 2010.

[12] IEEE Computer Society. Part 15.4: Wireless medium access control (MAC) and physical layer (PHY) specifications for low-rate wireless personal area networks (WPANs). http://standards.ieee.org/getieee802/download/802.15.4-2006.pdf, 2007. [Online, retrieved: January, 2012].

[13] Zigbee Alliance. Zigbee specification. http://www.zigbee.org/en/spec_download/zigbee_downloads.asp, 2007. [Online, retrieved: January, 2012].

[14] Nanotron Technologies. Nanotron's transceiver enables iso compliant real time locating systems. In *The International Organization for Standardization and the International Electrotechnical Commission*, volume 24730-5:2010. New standard for Real Time Locating Systems (RTLS), April 2010.

[15] M. Vodel, W. Hardt, R. Bergelt, and M. Glockner. Modulares Framework fr die synchronisierte Erfassung, Verarbeitung und Aufbereitung heterogener Sensornetzdaten. In *Proceedings of the Dresdner Arbeitstagung Schaltungs- und Systementwurf*, pages 67–72. Fraunhofer Institute for Integrated Circuits, May 2010.

[16] M. Vodel, R. Bergelt, M. Glockner, and W. Hardt. Synchronised data logging, processing and visualisation in heterogeneous sensor networks. In *Proceedings of the International Conference on Data Engineering and Internet Technology*. Springer, March 2011.