User Centred Networked Health Care
A. Moen et al. (Eds.)
IOS Press, 2011

doi:10.3233/978-1-60750-806-9-460

A Nomenclature for the Analysis of Continuous Sensor and Other Data in the Context of Health-Enabling Technologies

Matthias GIETZELT[a,1], Klaus-Hendrik WOLF[a], Reinhold HAUX[a]
[a]*Peter L. Reichertz Institute for Medical Informatics, University of Braunschweig – Institute of Technology and Hannover Medical School*
Germany

Abstract. Due to the progress in technology, it is possible to capture continuous sensor data pervasively and ubiquitously. In the area of health-enabling and ambient assisted technologies we are faced with the problem of analyzing these data in order to improve or at least maintain the health status of patients. But due to the interdisciplinarity of this field every discipline makes use of their own analyzing methods. In fact, the choice of a certain analyzing method often solely depends on the set of methods known to the data analyst. It would be an advantage if the data analyst would know about all available analyzing methods and their advantages and disadvantages when applied to the manifold of data. In this paper we propose a nomenclature that structures existing analyzing methods and assists in the choice of a certain method that fits to a given measurement context and a given problem.

Keywords. Continuous sensor data, health-enabling technologies, ambient assisted living, analysis methods, nomenclature

1. Introduction

In research we face a large variety of possible questions. Every problem needs the right tool to produce a useful solution matching the desired question. Especially in dealing with continuously recorded sensor data there is an enormous amount of methods to analyze the data and new analysis methods are often developed with the availability of new data.

In the area of health-enabling and ambient assisted technologies there are various use cases for continuously recorded sensor data. In the context of a patient-centered care these sensor data should be seen in combination with data e.g. from a medical health record in order to increase the information gain [1]. Health-enabling technologies are used to detect emergency situations, to give a feedback about the health status, or to assist in daily living [2]-[4]. Due to the interdisciplinarity of this field (e.g. informatics, electrical engineering, medical science and psychology) every

[1] Corresponding Author: Matthias Gietzelt, E-mail: Matthias.Gietzelt@plri.de; Peter L. Reichertz Institute for Medical Informatics, University of Braunschweig – Institute of Technology and Hannover Medical School, Mühlenpfordtstr. 23, D-38106 Braunschweig, Germany.

discipline makes use of their own analyzing methods. In fact, the choice of a certain analyzing method often solely depends on the set of methods known to the data analyst. It would be an advantage if the data analyst would know about all available analyzing methods and their advantages and disadvantages when applied to the manifold of data. It is desirable to know which analyzing method is most appropriate to handle a certain problem.

2. Objectives

Our fundamental question was how to choose the most suitable analysis method(s) based on a certain measuring context and a certain problem. To the author's knowledge, there is no tool that supports the choice regarding the most suitable method(s) and no systematization of methods for analyzing continuous sensor data.

3. Nomenclature for the Analysis of Continuous Data

Since a nomenclature is an established approach for systematization (e.g. in economics a nomenclature is used to make an environmental analysis [5]), we developed an open three-axial mono-hierarchical nomenclature in order to structure analysis methods for continuous sensor data. This nomenclature is considered to assist in the selection of one or more appropriate analysis methods in a certain context and with a certain problem. The subsequent semantic dimensions appeared to be reasonable:

- Context: description of the situation that was measured;
- Problem: underlying problem to be solved;
- Analysis method: a scheme of steps for analyzing the measured data

3.1. Context Axis

The first axis is the context axis. The context axis gives us information about the situation in which the sensor data was collected.

At first, we have to consider the object to be measured. In the area of health-enabling technologies the object is primarily a person. But there are also cases, in which a certain room or an electrical device is primarily measured. It is also necessary to identify the reference system in which the data was collected. The reference system can be a person, a certain room, a flat or a car. E.g., in case of on-body sensors, the reference system is the person measured. For the analysis of the data of some on-body sensors it is crucial to provide information about the wearing position and orientation.

In addition, the context axis describes the data source used. It is important to know the specifications of the sensor used to get a deeper insight and a better comprehension for the data. Wolf et al. developed a classification system for sensor-based data sources which describes relevant properties of a sensor [6]. This scheme was refined in [7]. This classification scheme was adopted in our nomenclature. But there may be other (non-sensor-based) data sources based on questionnaires, results of physical examinations, or data from an (electronic) medical record. These data can also be seen as a possible continuous data source and therefore, we added it to the proposed nomenclature.

The last sub-axis describes the data source's type, an essential part to choose an appropriate analysis method. Some analysis methods require quantitative, some require qualitative data sources. In addition, one should know how many channels and how many dimensions each channel of the data source has. Since continuous sensor data are recorded in time, a channel can be interpreted as a single sensor measuring a certain physical, chemical or biological characteristic. Please note that each channel can have one or more dimensions measured at the same point in time and a defined sample rate, whereas different channels can have varying sample rates. An example for a continuous one-dimensional signal is a body scale and a triaxial accelerometer measures a three-dimensional signal. Multi-dimensional signals can be pictures or data of a computer tomography.

Figure 1 shows the context axis and selected sub-axes.

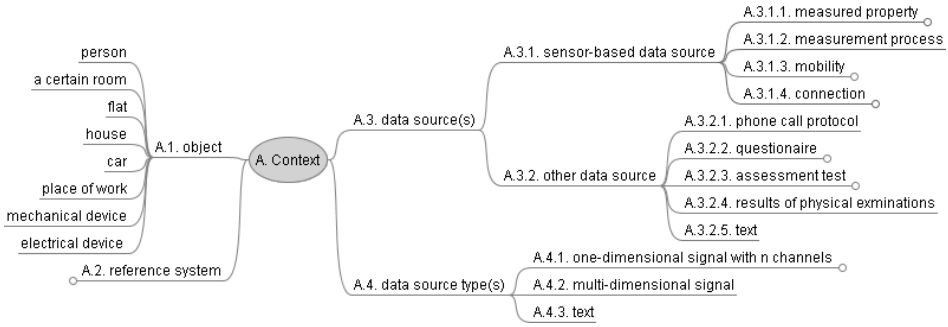

Figure 1. Mind map of the context axis.

3.2. Problem Axis

The second axis describes the underlying problem to be solved. Health-enabling technologies can be intended for a broad range of applications [1]. Thereby, a wide variety of aspects must be considered. An example for such a technology is a fall detector [8]. After a detected fall, the system has to initiate an alarm that has to be sent to persons who are able to help. Therefore, it is necessary to identify a strategy for an escalation and a de-escalation chain and to identify one or more suitable communication channels. It would be also helpful to give the helper auxiliary information, e.g. the localization of the affected person.

Besides emergency detection, health-enabling technologies can also be used for information and education purposes or even for wellness and sport [1]. Figure 2 shows some use cases for such technologies. This axis may help to go further into the problem, to define the functionality of the complete system, and therefore to define the outcome of the analysis.

3.3. Analysis Method Axis

The third axis is the analysis method axis. In this axis we structured methods for analyzing continuous sensor data. The sub-axes are structured in a way that they represent a typical procedure in analyzing continuous sensor data. At first, we have to extract candidate features from the data using filters or a frequency analysis. The second step is to select the most important features with the highest predictive capability. Thereby, we should avoid redundant or highly inter-correlated features. The

feature selection methods can be differentiated through their search behavior: if a method involves one feature at a time it is a single factor analysis (or uni-variate analysis) and otherwise a multi factor analysis (or multi-variate analysis). The third sub-axis contains the structure identification methods. These methods can be chosen using the information captured by the context and the problem axis. If a qualitative analysis is needed, one should prefer the classification and indexing methods; in case of a quantitative analysis one may choose a regression analysis.

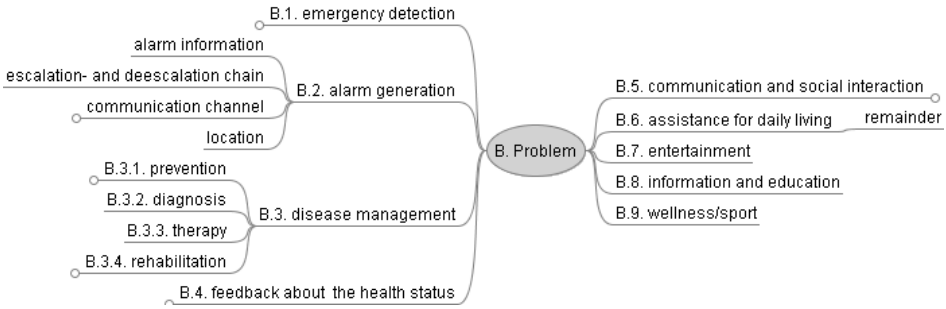

Figure 2. Mind map of the problem axis.

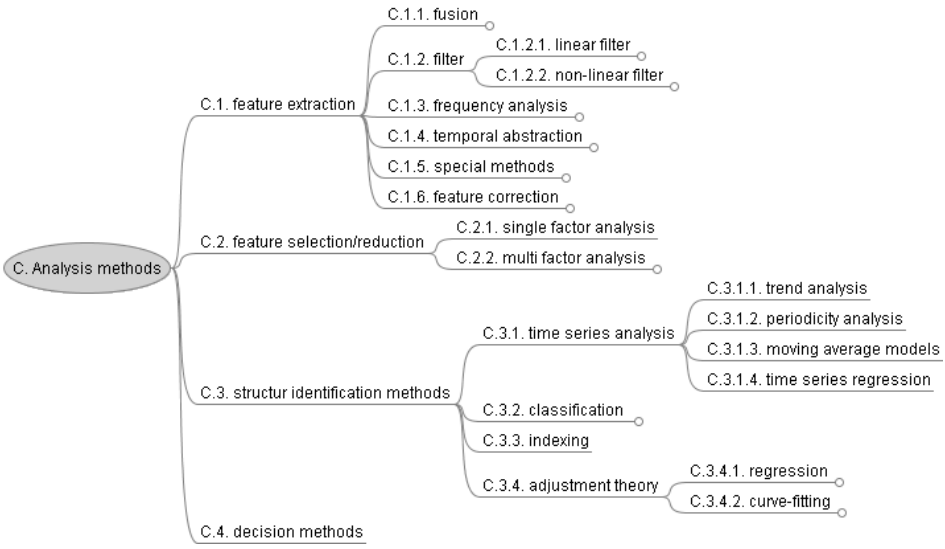

Figure 3. Mind map of the analysis method axis.

3.4. Statement Model

The statement model for the nomenclature summarizes all information that was derived from the situation measured: In a certain context x (if measuring the object x_1 with reference system x_2, using data sources x_3 which provide data of type x_4) and a given problem y, one should use the analysis method(s) z.

4. Discussion and Conclusion

In this paper we introduced an open three-axial mono-hierarchical nomenclature that may assists in the selection of one or more adequate analyzing methods for continuous sensor data. It covers the description of the context in which the data was collected, and the underlying problem to be solved. It was intentionally designed as an open nomenclature, so that new methods can be added. Within the proposed structure we also considered the typical procedure in data analysis.

This first step in the systematization we are conscious about that a nomenclature is not a sophisticated model in choosing analyzing methods. Our future research focuses in enhancing and refining the model in order to choose analyzing methods in a problem adequate manner.

4.1. Limitations

In the author's opinion there is, beside the development of new methods, an interdisciplinary and application oriented demand in research in structuring analyzing methods. But there is also a demand on establishing guidelines for choosing such methods. However, there are also a number of limitations to be stated. First, in spite of an intensive literature review, we did not find any results related to a systematization or a nomenclature for choosing analyzing methods for continuous sensor data. Second, the nomenclature presented is still work in progress and is supposed to be for discussion about its demand and content. Third, the proposed nomenclature has not been evaluated, yet. This will be done in our next step. Therefore, we will systematically analyze existing literature and make focus group discussions with experts in this field.

References

[1] Haux R, Howe J, Marschollek M, Plischke M, Wolf KH. Health-enabling technologies for pervasive health care: on services and ICT architecture paradigms. *Inform Health Soc Care* 33 (2008), 77-89.

[2] Saranummi N. IT applications for pervasive, personal, and personalized health, *IEEE Trans Inf Technol Biomed* 12 (2008), 1-4.

[3] Arnrich B, Mayora O, Bardram J. Pervasive or Ubiquitous Healthcare?, Methods Inf Med 49 (2010), 65-6.

[4] Demiris G. Smart homes and ambient assisted living in an aging society. New opportunities and challenges for biomedical informatics, *Methods Inf Med* 47 (2008): 56-7.

[5] Fleisher CS. Bensoussan BE. *Strategic and competitive analysis: Methods and techniques for analyzing business competition*, Prentice Hall, New Jersey (USA), 2002.

[6] Wolf KH, Marschollek M, Bott OJ, Howe J, Haux R. Sensors for health-related parameters and data fusion approaches. *Proceedings of the European Conference on eHealth ECEH 2007*:155–61.

[7] Koch S, Marschollek M, Wolf KH, Plischke M, Haux R. On health-enabling and ambient-assistive technologies. What has been achieved and where do we have to go? *Methods Inf Med* 48 (2009), 29-37.

[8] Bourke AK, van de Ven PW, Chaya AE, OLaighin GM, Nelson J. Testing of a long-term fall detection system incorporated into a custom vest for the elderly. *Conf Proc IEEE Eng Med Biol Soc. 2008*:2844-7.

[9] Klein LA. *Sensor and Data Fusion: A Tool for Information Assessment and Decision Making*. SPIE Publications, 2004.