CSIMQ
Complex
Systems
Informatics
and
Modeling
Quarterly

Information Logistics Solutions in Healthcare: From Data to Demand Fulfilling Information

Sven Meister, Wolfgang Deiters, Sven Schafer, Valentin Stahlmann

Fraunhofer Institute for Software and Systems Engineering, Emil-Figge-Str. 91, 44227 Dortmund, Germany

sven.meister@isst.fraunhofer.de, sven.schafer@isst.fraunhofer.de,
valentin.stahlmann@isst.fraunhofer.de

Abstract. The extensive usage of information and communication technologies (ICT) within healthcare, like electronic health records, Ambient Assisted Living (AAL) and telemedicine, results in a continuously increasing amount of heterogeneous data. Data, on the one hand, is not tantamount to information on the other. Data has to be transformed into demand fulfilling information to cope with one of the biggest problems in information science: data and information overload. The objective of this paper is to introduce technologies like Complex Event Processing (CEP) to develop approaches to cope with the problem of information overload in healthcare according to the principles of Information Logistics (ILOG). We will present a solution that is called TiEE (Telemedical ILOG Event Engine) to process vital signs in real-time and aggregate them to relevant pattern, resulting in demand fulfilling information.
Keywords: CEP, Telemedicine, Decision Support, Trend Recognition.

1 Introduction

The usage of information and communication technologies (ICT) like electronic health records plays an important role to improve healthcare supply by optimizing the communication between the stakeholders of the public health sector. Lack of adequate information could lead to inefficient medical treatment or to erroneous decisions [18]. Most of the discussion related to healthcare supply and its anticipators as well as existing solutions are focussed on the therapy process. Within the following years the demographic change will cause two major problems:

- The demographic change causes a shift within the age structure. The increase of the amount of the elderly amplifies the need for personalized concepts within rehabilitation and prevention [9].
- The increased demand can only be satisfied by expanding the staff [20].

The quality of medical treatments and medical care extensively relies on the knowledge and experience of the leading stakeholder. Putting the two major problems mentioned above together, it is obvious that the need for those will increase, while in relation expert know how will decrease. To meet the need, medical institutions have to compensate the missing resources by deploying unskilled staff. It would make costs explode to use senior experts only. The missing knowledge and experience of unskilled staff could negatively affect the quality of healthcare supply and endanger the health status and thus also the success of a therapy.

Therefore new technologies are needed, like the proposed by the research areas of Ambient Assisted Living (AAL) and telemedicine. They enable a continuous health monitoring, decision support and knowledge transfer but they also force the amount of data in a way that is not manageable anymore. Especially implementations based upon so called Cyber Physical Systems

(CPS) or other kind of sensor networks make it necessary to process a large amount of data on-time. Thus, we have to cope with an increasing amount of heterogeneous data, leading to the term Big Data. There is no common definition of this term but Gartner proposed its 3V's model to describe todays situation in information society [21]: Increasing data **volumes** demand for new storage and analysis technologies. The **variety** of data makes it hard to transform those into decision supporting information. Last but not least information systems today have to cope with the **velocity** of data input and output. Bonnie [10] stated that healthcare data will explode from 500 petabytes in 2012 up to 25.000 petabytes in 2016, also forced by the extensive usage of IP-based devices. Remote monitoring and the documentation of health related data will become easier. BITKOM confirms an annual growth of data of around 40-50 % [3]. Frost& Sullivan currently estimated an amount of 1 billion terabytes of health related data within hospitals and forecasted an amount of 1.8 zetabytes in the following years [11]. The expansion rate for Big Data is assumed to be up to 36 % and there is a recognizable chance to positively affect the cost efficiency within the health care sector by integrating and processing different kinds of data [3]. It is expected that the market for Big Data solution will increase from currently € 198 million up to € 1.6 billion in 2016.

Thus, the interlink of information technologies and healthcare offers a great opportunity for a better healthcare supply on the one hand but on the other hand challenges in data and information processing. From a physician's point of view all data has to be aggregated and integrated to be able to use it as a decision supporting information. Physicians therefore complain about the problem of information overload, which means that the information processing requirements exceed the information processing abilities [16].

In this paper we will investigate typical challenges in telemedicine originating from high data volumes and present selected concepts for tackling them with approaches and techniques from information logistics. In particular, we will show that solutions exploiting data processing technologies, like Complex Event Processing (CEP), have the potential to contribute to demand fulfilling supply with information in health related scenarios, which is demonstrated with TiEE. We well describe a method to model information demand in a way that the formalized demand can be mapped on CEP processing concepts. The paper will discuss core ideas of TiEE and present first experiences as well as result from a short evaluation.
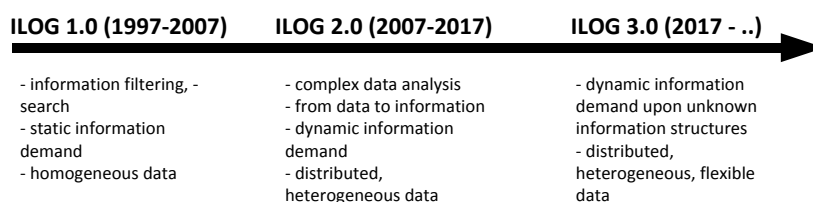
2 Related Work

Today's world is full of data. Especially within the healthcare sector sensors, electronic patient records or intelligent sensor networks play an important role. A forthcoming technology is called Cyber Physical Systems (CPS) which is an integration of sensor systems, data capturing techniques, data processing method and real-life processes [2], [19]. A large amount of continuously produced data has to be delivered to recipients, like physicians. Unfortunately it is impossible for those recipients to process and react to all of these data. The important thing is called: Demand fulfilling information. Only such information can improve decision quality within healthcare processes.

Investigating solutions means to bear the following three aspects in mind: The **quantity** of information duplicates every 19 years [38]. Also the **time** for searching and processing medical information is very limited. Searching for information should be limited to 30 seconds to be acceptable for a physician [26]. The **characteristics** e.g. the complexity of medical information changes with the progress of medical technologies, so physicians need profound data processing abilities [14]. Therefore, approaches aiming to solve the problem of information overload have to reduce the amount of data at first. Afterwards, the time for acquiring information has to be reduced by distributing relevant data along the levels of escalation. Third, the complexity and heterogeneity of data have to be reduced.

Nowadays, approaches are often related to the real big thing called Big Data [11]. Methods and technologies are focussed on processing and filtering data but there is a lack between data and information – especially demand fulfilling information. Also a discrepancy between the amount of information and the information processing capacities of a human being directly affects the quality of decision making. Information overload could lead to a loss of quality within medical decision making. Therefore, Information Logistics is strongly related to the topics of information overload and information need or information demand. According to Wilson and others information overload expresses „that the flow of information…is greater than can be managed effectively" [16, 36]. To cope with the problem of information overload a variety of technologies could be used to process data. In general one has to distinguish between real-time processing on the one hand and retrospective processing of large datasets on the other.

The situation is not only caused by an improvement of technology but also related to organizational, personal and external factors. An intelligent information supply requires a clear definition of information need. Line distinguishes five subcategories: Need, want, demand, user and requirement [22]. It is important to give a clear distinction between need and demand. The former describes all pieces of information which are needed to answer the given question and the latter the amount of inquired information which are assumed to be useful to answer the given question [33].

At present Information Logistics is viewed as a detached research area to deliver the right information, in the right format at the right time to the right place [7], [35], [37] and is partially used for information filtering or with context-models to optimize communication, e.g. in the healthcare domain [30]. Haftor et al. give a broad overview of the state-of-the art research in Information Logistics by analyzing 102 scientific publications [12]. One can sketch up the past, present and future of Information Logistics like shown below. Big Data plays an important role within ILOG 2.0 as an enabler for processing large amount of data but ILOG goes a step beyond by formalizing and mapping information demand onto data processing capabilities to deduce demand fulfilling information. Data processing technologies, like Data Mining or Machine Learning as well as Complex Event Processing, are enabler for a better information supply by clustering data to relevant patterns. In the following we will focus on the usage of CEP to process high data volumes in real-time (see Figure 1).

ILOG 1.0 (1997-2007)	ILOG 2.0 (2007-2017)	ILOG 3.0 (2017 - ..)
- information filtering, - search - static information demand - homogeneous data	- complex data analysis - from data to information - dynamic information demand - distributed, heterogeneous data	- dynamic information demand upon unknown information structures - distributed, heterogeneous, flexible data

Figure 1. History and future of Information Logistics

Basic definitions and concepts of CEP were developed and defined by Luckham, Chandy and Bates [1], [4], [25]. Citing them, an event is „an object that is a record of an activity in a system". A detailed overview about open questions and the current state of research in CEP is discussed in the Dagstuhl Seminar on „Event Processing" in 2010 [5]. In [24], [34] Lowe and Weber present ongoing work on applying CEP to health data using the event processing engine Esper in the context of the Stride project „Stanford Translational Research Integrated Database Environment". One basic open question is how to cope with the problem of heterogeneity of medical data to achieve an overall processing.

3 Motivating Use Case

To give the reader a short insight into telemedicine we set up a consistent example that will also be used in the following to describe the main concepts and to realize a short evaluation (see chapter 4.4).

Cardiac diseases, like hypertension or cardiac insufficiency, are forthcoming especially in the industrial countries. In this context, exercise is a basic instrument for a therapeutic modification of the lifestyle. In many cases, physicians will prescribe a therapy combining medicine and sports. The choice and control of intensity of sports applied as intervention mechanism against cardiac diseases requires a continuous monitoring of different vital signs. Continuous means high-frequent data (e.g. every second) on the one hand as well as long-term persistent measurement (e.g. once per day/week).

Hypertension (high blood pressure) is one of the most common worldwide chronic health conditions. Hypertension requires intensive monitoring of the following vital parameters:

- Blood pressure: The measurement and evaluation of the blood pressure value, considering its development, is the most significant vital parameter for optimizing the therapy.
- Pulse: Observing the pulse is also vital, especially when certain heart rate modulating drugs are prescribed. A simple, though little precise formula for determining the optimal pulse is: "180-Age".

In common parlance, cardiac insufficiency refers to a reduced pumping capacity as well as disturbed filling of the heart. The symptoms may include irregular heartbeat, tachycardia, hypertension, and water retention. Deficient therapy may lead to continuously reduction of the capacity of the cardiovascular system. Early and permanent medication is therefore essential for long life free of complaints. The following two parameters play a basic role for the monitoring:

- Weight: Changes in weight is an important indicator for water retention in the lungs and limbs. Especially the weight is a subject to fluctuations, which is caused amongst others by certain medicine (e.g. diuretics). Left ventricular insufficiency often causes pulmonary edema, while right ventricular insufficiency encourages water retention in the limbs.
- Blood pressure: A significantly increased blood pressure, especially in combination with weight gain can be a sign for an aggravating cardiac insufficiency.

It is imperative to detect relevant or critical patterns as soon as possible. The human-based continuous monitoring of patients suffering from cardiac diseases like described above seems to be impossible. The amount of data would exceed the processing capacities of, e.g. a physician. In the next chapter we will show how processing of real-time vital sign data, bearing information demand in mind, could be realized using TiEE.

4 TiEE – The Telemedical ILOG Event Engine

The reduction of information overload in telemedical scenarios requires to process telemedical information in the sense of aggregation, filtering as well as analysis of causal and temporal relationships. There are three basic requirements for the processing of vital signs in telemedical scenarios one should take into account [29]:

- Sensors for measuring vital signs act in a highly distributed manner. Every type, e.g. blood pressure or blood sugar concentration, is a result of a single sensor or telemedical application. The prospective solution should be able to process telemedical values from different sources to achieve an overall monitoring and decision making. This requires an overall description of a telemedical value, in sense of a vital sign, as well as methods to modular process different types of telemedical values depending on the actual medical situation of a patient.
- Monitoring of vital signs in telemedical scenarios produces a high amount of data which has to be processed in real-time. Not every single vital sign represents an important

medical situation. The relevance depends upon the temporal ordering as well as the coincidence of different types of vital signs. So, a solution has to aggregate a set of those to higher order decisions.

- The delivery of those decisions mentioned above should be done in an intelligent way. So, a derived decision should be transported at the right time to the right place, e.g. prior to a physician a notification should be emitted to a family member.

With TiEE, the Telemedical ILOG Event Engine, Fraunhofer ISST investigated methods for an event based processing of vital signs in telemedical scenarios considering the requirements mentioned above. We will start by giving a broad overview about the basic requirements to realize an Information Logistics processing within TiEE. Afterwards we will describe the three basic concepts telemedical event, TIL and TIL-Profile in more detail.

4.1 Requirements for a flexible vital sign processing

The conceptualization of TiEE started with a meta-analysis of ten papers, coping with the identification of requirements for data processing in the area of telemedicine and Telemonitoring applications. As a result we documented 11 basic requirements as follows:

- **R1 Integrated and continuous communication:** Telemedical systems should have harmonized and standardized interfaces as well as convergent communication channels.
- **R2 Intuitive Information representation:** Information should be presented in a way that they make a decision supporting without any necessity to be processed or transformed by a human.
- **R3 Recognition of critical patterns:** Analysis of incoming data to recognize pattern of interest.
- **R4 Demand-driven processing:** Processing of vital signs has to be realized according to users demand.
- **R5 Modular processing:** Existing systems for vital sign processing are very specialized to process data only for one use case. There is a need to modularize processing fragments and to stick them together depending on any possible use case.
- **R6 Flexible and situation-based algorithms:** A system for vital sign processing needs more than monitoring thresholds. Modularization and complex pattern recognition are key-factors for a telemedical data processing system.
- **R7 Extensibility:** In terms of Plug&Play new types of vital signs have to be assigned to an analytical infrastructure without high efforts.
- **R8 Aggregation of vital signs:** Processing of vital signs has to be structured patient specific. Different types of vital signs have to be aggregated.
- **R9 Real-time processing:** Depending on the situation telemedical data could be time critical data that has to be processed on-time to detect critical or relevant patterns.
- **R10 Connectivity:** Vital signs could be measured in a highly distributed manner. Thus, a flexible and modular system has to offer clearly defined interfaces to gather vital signs.
- **R11 Temporal processing:** The semantics of vital signs and patterns is given by a time-based correlation of signal data items. Short-term as well as long-term measurements and high-frequency data as well as low-frequency data should be processable, according to real-life telemedicine.

In summary, we have to define a method to be able to model information demand in a way that the model can be mapped onto processing unites. In the following we also have to conceptualize those processing units upon existing concepts of CEP. We therefore have to take into account the temporal relationship within a stream of vital signs. To be able to realize an overall processing of different types of vital signs the input has to be standardized in terms of a telemedical event specification.

4.2 Information Logistics Processing

We started our discussion about information overload in telemedical scenarios by stating that it is not possible for a physician to deduce relevant information out of a stream of thousands of vital signs. Thus, realizing a support for clinical decision making one has to reduce the amount of data and deduce higher level, relevant information. Delivered relevant information should fulfill a given information demand, thus a physician needs a possibility to express its demand.

The modeling of information demand for a given telemedical questions requires extensive expert knowledge. Regarding the motivating use case in chapter 3 one has to define the dependencies between pulse, blood pressure and weight. Thus, at the end we will need a set of rules and patterns processable and detectable by algorithms within TiEE. Expert knowledge can be gathered from humans or also medical guidelines. Such a guideline is diagnosis depending and documents best-practices to treat a set of given medical conditions.

To cope with requirement R4, with TiEE we investigated a graphical demand modeling approach (Demand Modeling Language) upon the Clinical Algorithm Standard (CAS) [32] which is well known in medicine. Within CAS we have five elements (see Figure 2): a start node (oval), a condition (rhombus), an activity (rectangle), a terminal node (circle) and a connector (arc). A condition is the evaluation of a given expression, like pulse higher than 150. The logical AND concatenation is modeled by writing conditions from left to right and the OR concatenation is realized by writing conditions down. The result of the evaluation of a set of conditions could be the terminal node, thus nothing will happen, or a red or green colored activity. An activity symbols the necessity to inform about a relevant situation. Relevance is divided into critical (red) and uncritical (green) situations.
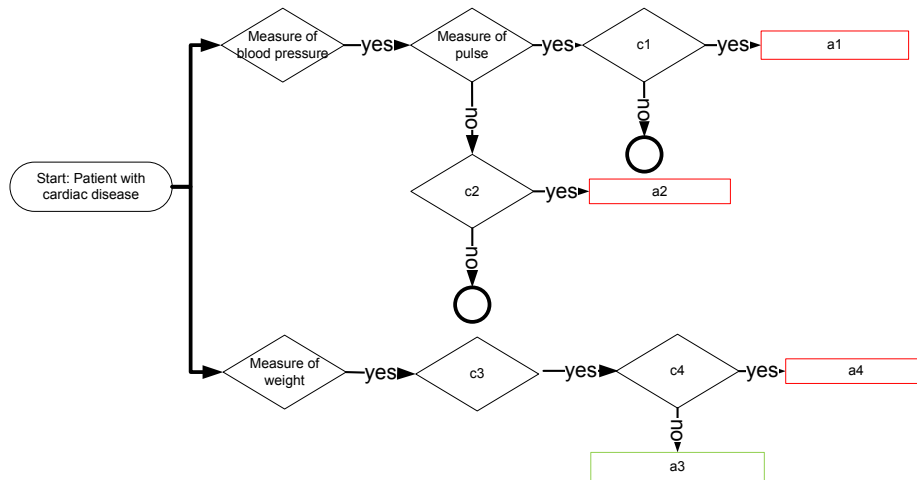
Figure 2. Demand modeling using the Clinical Algorithm Standard

Regarding the figure shown above we have a patient coping with a cardiac disease, like described in the motivating use case. He has to measure weight, pulse and blood pressure. Condition $c2$ is only valid in case that the pulse is not measured otherwise condition $c1$ is valid which would lead to action $a1$ or a terminal node.

The graphical model has to be transformed into processable data structures, thus we defined a set of structures upon the Extended Backus-Naur Form (EBNF) in such a way that the Demand Modeling Languages $DML \subset EPL$ is a subset of an Event Processing Language and further can be mapped to the extending concepts like TILs and TIL-Profiles. Below we sketched up the EBNF specification of the formalize demand and afterwards in Table 1 examples according to Figure 2.

```
demand   = "DEMAND" (condition | condition "→" condition)   CRLF ;
             "ACTION" activity "ELSE" activity ";" CRLF ;
```

The expression in Table 1 shows that condition c_1 is only valid if blood pressure as well as pulse is increased. In such a case a physician has to be informed on triggering activity a_1 because the criticality of the situation is defined with 90 (range: 0-100) and the relevance is defined with 80 (range: 0-100). The combination of condition and activity results in demand d_1.

Table 1. Expressions for condition c_1, action a_1 and the demand d_1 according to the example shown in Figure 2.

ID	Expression
c1	bloodpressure.BP1 = INCREASING AFTER pulse.BP1 = INCREASING
a1	INFORMATION Critical increase of blood pressure and pulse. Please stay in contact with the patient.; RELEVANCE 80; CTITICALITY 90; RECIPIENT ID1 | Physician | JMS | service_in | DIRECT;
d1	**DEMAND** c1 **ACTION** a1 **ELSE** NIL;

Now, the above expressed and formalized demand has to be mapped on the concepts of TILs and TIL-Profiles bearing the Event Processing Language used within TiEE in mind. At first, an incoming telemedical event will be processed by a patient-specific TIL-Profile, like described in chapter 0. Afterwards a set of TILs is trying to detect characteristic patterns within the stream of telemedical events, so called Complex Trend Pattern Events (CTPE). A CTPE is an abstraction of a set of telemedical events (see also Figure 3) and characterizes the progression of the measured values. Based upon the research of Charbon et al. [6], [13] we distinguish five basic types of trend pattern: slope, slope reverse, saltus up and saltus down steady.
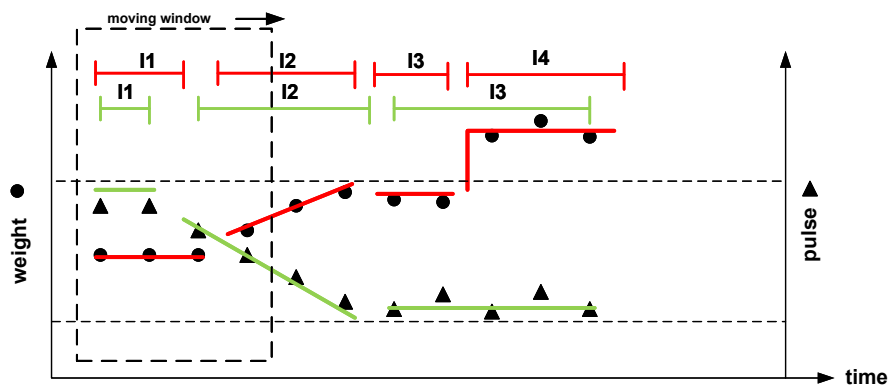
Figure 3. Abstraction of incoming telemedical events by building intervals in Terms of CTPEs

Thus, we derive an abstraction, the pattern, from a set of underlying measurements to reduce the amount of data and cope with the problem of information overload according to the requirements R11 and R3. The derived pattern is described using different types of parameters, e.g. the statistical spread or the amount of increase/decrease. As a basic feature of the TIL concept one can use any kind of algorithm for trend calculation as long as it fulfills the formal

definitions of a TIL. CUSUM (cumulative sum) or ARIMA (autoregressive integrated moving average) based approaches are the examples for processing time related data.

In summary, every TIL derives CTPEs for a specific telemedical event, i.e. a type of vital signs, and forwards them up to the referring TIL-Profile. Now, this TIL-Profile has to detect higher order, demand fulfilling patterns within the set of forwarded CTPEs. By using the formalized demand the processing within a TIL-Profile is organized as follows:

- Trends of same underlying types of vital signs: The repeated increase, decrease etc. of a set of vital signs could be abstracted to a single trend pattern.
- Trends of different underlying types of vital signs: It is obvious that there is a relation between weight and blood pressure in cases of cardiac decompensation. A TIL-profile has to detect the increase of both during a given time window and derive a new abstraction, emitting a new trend pattern.

Upon rules registered in the TIL-Profile, Information Logistics decisions are made to generate and send relevant information to a person, like required in R2.

4.2.1 Telemedical event and HL7 Telemedical Event Format

Within telemedical scenarios you will find a lot of different sensors from various manufacturers to measure vital signs. The overall processing of vital signs using TiEE requires some concept to standardize the input, like proposed within requirements R1, R7 and R10. While TiEE is based upon the idea of complex event processing, every vital sign should be interpreted as an event. Therefore we defined the term telemedical event as a *measurement of a telemedical value and an instance of a telemedical event type, formatted in the HL7 Telemedical Event Format* [28]. We distinguish four segments within a telemedical event:

1. Type segment: First of all we cover a possibility to specify the type of a telemedical event. On the one hand this information is related to the level of inheritance and on the other hand to the measured vital sign.
2. Processing segment: Most of the attributes will be related to ILOG and CEP processing. This segment allows for grouping those.
3. Medical content segment: This segment contains information about the measured vital sign which caused the change of state and is the nucleus of the following event processing.
4. Payload segment: Sometimes there is a need to store attributes that are not directly related to the genesis of the telemedical event but important to assure a consistent event processing. One example is the storage of runtime attributes that are specific to the event processing engine.

With respect to previous work of Luckham and Etzion [8], [25] there are some common parameters necessary to process events, e.g. time, eventID or producer. Especially time is important for ILOG processing, but one has to distinguish on a semantic level between the time of generation, time of arrival, time of processing etc. From the ILOG point of view there are also parameters like the receiving application or person. From a conceptual point of view Table 2 gives an overview of the most important attributes that should be covered by a telemedical event type.

The HL7 Telemedical Event Format is a message format we investigated to achieve an interoperable transportation of Telemedical Events. The refinement of this format is done by combining elements of the HL7 standard with such from the IEEE 11073 standards. HL7 is a widespread international standard for data exchange in the healthcare sector [15]. All HL7 V3 data types are based upon the HL7 Reference Information Model (RIM). In turn IEEE 11073 is a family of standards to harmonize the output of sensors using the IEEE 11073 Domain Information Model (DIM) [17]. Using both standards we modeled a format that takes all attributes for complex event processing and ILOG processing of vital signs into account.

Table 2. Basic attributes which are important for a telemedical event type

Parameter	Segment	Description
eventTypeID	Type	An ID to specify the type of the event.
eventAnnotation	Type	Some kind of textual description of the event.
eventID	Type	An ID that uniquely identifies the event.
simpleEvent	Processing	Expresses whether the event is a simple or a complex event.
detectionTime	Processing	The time point of the detection of the event.
createTime	Processing	The time point of the creation of the event.
processTime	Processing	The time point of processing the event.
eventSource	Processing	A reference and description of the emitter of the event.
parentEvents	Processing	A reference to the parental events.
receivingResponsibility	Processing	The receiving party of the event.
timeEmitter	Processing	Regarding to heterogeneous event processing infrastructures it is important to specify a global clock.
content	Medical content	The event value that is recorded as a result of the state change of the system.
location	Medical content	The location of the measurement.
patientID	Medical content	An unique identifier for the patient.

Formally a telemedical event is an n-tuple $e_i^T := (e_i, HL7_{Trans})$ where:

- $E^T := \bigcup_{i=1}^{\infty} e_i^T$ is the set of all telemedical events.
- $e_i \in E$ is an event based upon a telemedical event type $et^T \in ET^T$.
- $HL7_{Trans}: E \rightarrow E^T$ is a transformation into the HL7 Telemedical Event Format.

Two telemedical events are identical $e_i^T \equiv e_j^T$ if and only if $e_i = e_j$ and $HL7_{Trans}(e_i) = HL7_{Trans}(e_j)$. $HL7_{Trans}$ is a function to transform a given event into the HL7 Telemedical Event Format.

Regarding the motivating use case we need three types of telemedical events to represent pulse, blood pressure and weight.

4.2.2 Telemedical ILOG Listener –Profile (TIL-Profile)

Every medical situation represents an individual moment in lifetime. An ILOG-infrastructure for telemedical scenarios therefore has to be modular and should be able to aggregate different kinds of vital signs, like stated in requirement R5 and R8. Thus, TiEE has to offer functionalities for a patient specific processing of incoming telemedical events, which very fast can be adapted to a new situation. Therefore we investigated the term TIL-Profile. A TIL-Profile realizes a patient specific filtering of telemedical events thus it reduces the amount of data. Afterwards the event is processed, depending on the type of vital sign, by one of the TIL's (see section 4.2.3) connected to this profile. For every type of vital sign one has to register one TIL. In the following the output of the TIL's is processed within the TIL-Profile by additional filtering, pattern detection and transformation into higher order decisions. That means that a TIL-Profile correlates the progression of different types of vital signs, e.g. blood pressure and weight, detects a medical situation of relevance and derives a higher order medical decision.

Formally a Telemedical ILOG Listener Profile is defined as an n-tuple as follows $til_{Profil} := (et_{in}, ET_{out}, f_{in}, TIL, VL)$:

- $TIL_{Profil} := \bigcup_{n=1}^{\infty} til_{Profil_n}$ set of all TIL-Profiles.
- et_{in}: The event type on which all instances $inst(et_{in}) = e_{in}$ are based upon. Initially this is the telemedical event type $et_{in} := et^T$.
- ET_{out}: Analogous to the definition of et_{in}, ET_{out} is a set of event types $|ET_{out}| \geq 1$ which are permitted to be emitted as output.
- f_{in}: The filter function is based upon a Boolean function. Given to functions f_{in_1} and f_{in_2} it is imperative that $f_{in_1} \equiv f_{in_2}$ are identical if and only if $patient\ (f_{in_1}) = patient\ (f_{in_2})$ that is, both functions relate to the same patient. Two mutually different functions $f_{in_1} \sqcap f_{in_2} = \emptyset$ are disjoint.
- TIL: The set of registered TIL's $TIL := \{til_1, \cdots, til_N\}$ in the TIL-Profile where $|TIL| \geq 1$, so at least one TIL has to be registered.
- VL: Every TIL-Profile consist of processing logic VL which is a set of rules in terms of:

$$VL := \left\{ R \mid R = \begin{cases} f: E^T \to (true, false), & Filter \\ p: E \to E, & Pattern \\ t: E \to E, Transformation \end{cases} \right\}$$

Two TIL-Profiles are identical $til_{Profil_1} \equiv til_{Profil_2}$ if and only if $et_{in_1} \equiv et_{in_2}$, $ET_{out_1} \equiv ET_{out_2}$, $f_{in_1} \equiv f_{in_2}$, $TIL_1 \equiv TIL_2$ and $VL_1 \equiv VL_2$.

With respect to the motivating use case we instantiated one TIL-Profile per patient containing one TIL for pulse, blood pressure and weight processing.

4.2.3 Telemedical ILOG Listener (TIL)

A TIL-Profile realizes a patient specific filtering of the incoming telemedical events. The second step of filtering the high amount of events is done within the concept of a TIL. Related to CEP a TIL is some kind of Event Processing Agent, specialized for processing one type of telemedical values e.g. blood pressure events [31]. It should address the requirements R5 and R6 by modularizing processing routines for a type of vital sign.

Besides the operation of filtering, a TIL encapsulates methods to detect patterns of interests in the stream of incoming events. Therefore different types of rules could be instantiated. Thus, a TIL is a modular piece of concept to encapsulate algorithms which are highly specialized to process one type of vital sign but is not specialized to a patient. That enables an easy reuse.

Formally a Telemedical ILOG Listener is defined as an n-tuple as follows $til := (et_{in}, ET_{out}, f_{in}, VL)$:

- $TIL := \bigcup_{n=1}^{\infty} til_n$ is the set of all TILs.
- et_{in}: The event type on which all instances $inst(et_{in}) = e_{in}$ are based upon. Initially this is the telemedical event type $et_{in} := et^T$.
- ET_{out}: Analogous to the definition of et_{in}, ET_{out} is a set of event types $|ET_{out}| \geq 1$ which are permitted to be emitted as output.
- f_{in}: The filter function is based upon a Boolean function $f_{in} \to E^T : (true, false)$.Given to functions f_{in_1} and f_{in_2} it is imperative that $f_{in_1} \equiv f_{in_2}$ are identical if and only if $vitaltype\ (f_{in_1}) = vitaltype\ (f_{in_2})$ that is, both functions relate to the same type of vital sign. Two mutually different functions $f_{in_1} \sqcap f_{in_2} = \emptyset$ are disjoint.
- VL: Every TIL consists of processing logic VL which is a set of rules in terms of:

$$VL := \left\{ R \mid R = \begin{cases} f: E^T \rightarrow (true, false), & Filter \\ p: E \rightarrow E, & Pattern \\ t: E \rightarrow E, Transformation \end{cases} \right\}$$

Two TILs are identical $TIL_1 \equiv TIL_2$ if and only if $et_{in_1} \equiv et_{in_2}$, $ET_{out_1} \equiv ET_{out_2}$, $f_{in_1} \equiv f_{in_2}$ and $VL_1 \equiv VL_2$.

According to the motivating use case we instantiated three types of TILs, one for pulse, weight and blood pressure. Each TIL has a set of parameters to enable a use case-specific configuration of the algorithms. Depending on the algorithm it could be imperative to parameterize it patient-specific or according to the time characteristics (e.g. frequency of incoming events). With respect to the evaluation we are able to fix the parameters as the setting is clearly defined.

4.3 Architectural insights into TiEE

The architecture of TiEE is based upon the event processing engine Esper which is commonly used in many commercial products. Figure 4 gives a broad overview about the main components and concepts like described in the past sections.
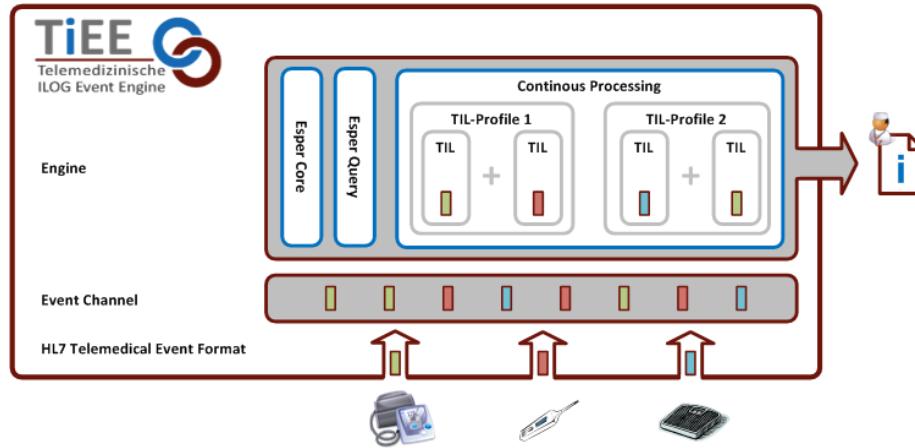
Figure 4. TiEE architectural overview based upon the event processing engine Esper

Starting at the bottom we have some kind of vital sign sensors from different manufactures. The sensors are connected to TiEE through Bluetooth HDP, supporting the IEEE 11073 standards family. To achieve an overall processing a single vital sign is interpreted as a telemedical event and will be encapsulated in the HL7 Telemedical Event Format. All events get transported to the event channel ordered by time. Above the channel we build the engine, introducing TIL-Profiles and TIL using Esper core and Esper queries of the Esper engine.

To realize a patient specific filtering like required by the TIL-Profile we extended Esper with the concept of PES, a patient-individual event stream. A PES formally represents an individual event stream that supports and bundles all types of events of one single patient. Thus there is a bijection between the patient and PES. Accordingly, any PES are pair wise different, in other words all PES are distinct, as each PES is characterized by a different patient, $PES_n \neq PES_m \, \forall \, n, m$ where $n \neq m$. The concept of a PES is based on the Variant Stream of Esper. All events of one single patient will be redirected in separate PES, with the help of a filter criterion, which checks for an individual patient identification. Likewise, a PES serves as the event source for all TILs in one TIL-Profile.

Furthermore we implemented methods and services to administrate TiEE. The developed services serve the purpose to modify the current system afterwards. So it is possible to add or remove main components like TIL and TIL-Profiles. In more detail the services allow to add, modify and remove patients, TILs and statements. Also, the services provide information about the components within the system. All services are designed as REST services and work with JSON objects.

4.4 Evaluation

The evaluation of TiEE is done upon two different data sets gained from two different projects. The first project reflects the motivating use case and is called FitPit, the Fitness Cockpit, a solution to optimize preventive and rehabilitative trainings developed at Fraunhofer ISST [27]. Pulse and oxygen saturation as well as weight and blood pressure will be measured at the beginning and at the end of the training. In total we recorded 2450 measurements gained from 10 patients in terms of a long term measurement. The second project is the Vital Signs Dataset of the University of Queensland [23]. They recorded over 10 vital sign parameters from 32 patients undergoing anesthesia. The data was recorded with a resolution of 10ms. Thus, in total we have around 240.000 measurements per patient, depending on the duration of the surgery.

Before we can apply the datasets above, we have to define the main questions that have to be evaluated using TiEE:

- Question 1: Is TiEE capable of reducing the amount of incoming data and process them according to the defined information demand?
- Question 2: Is TiEE capable to process long-term trends as well as high frequent data?
- Question 3: Does the usage of TiEE reduce the overall implementation efforts for an analytical infrastructure?

What we will not try to answer at this point is, if TiEE is capable of being better in decision making compared to a physician. To answer the first question we started with modeling the information demand using CAS like mentioned before. Thus, we can prove that it is possible to map some kind of formalized demand to processing rules. To show the amount of data reduction we calculated the percentage of CTPEs per Minute. Within the FitPit scenario we had 1/11520 CTPEs per minute, respectively one trend within eight days. This very low value is caused by the long-term characteristic of the use case. Figure 5 demonstrates the ratio of incomings events and fired activities. Thus, there is a recognizable reduction of data or irrelevant information because activities are only fired according to a formalized information demand.
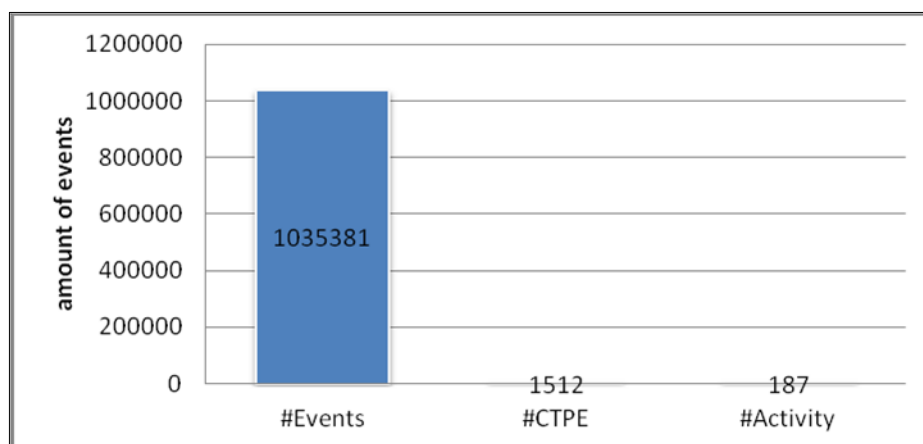
Figure 5. Ratio between incoming events generated CTPEs and fired activities of one patient of the University of Queensland data set

The high-frequent data of the second use case produces 8.3 CTPEs per minute. It is possible to optimize or modify the percentage of CTPEs by configuring the underlying algorithm. Thus, we can also show that TiEE is capable to process long-term as well as high-frequent data. We also evaluated the performance of TiEE including routing and algorithmic processing of the incoming events. The average duration is around 15ms per event. The third question is implicitly proven by the usage of TiEE within two different scenarios. We were able to reuse once defined TILs within both scenarios in terms of a repository. Thus, the only effort was to model the information demand and map it to a TIL-Profile. Furthermore with TiEE a basic infrastructure for communication I/O is already given. Summarizing we can point out that TiEE supports an optimization of information supply by reducing the amount of data.

5 Conclusion and Outlook

The extensive usage of ICT within the healthcare sector overstrains the information processing capabilities of the actors, like physicians. Large data sets, also known as Big Data, have to be processed and management to cope with the problem of information overload. Therefore data has to be transformed to demand fulfilling, relevant information to ensure high quality medical treatment. With TiEE we proposed a modular solution based upon Complex Event Processing to filter and aggregate data. We showed how data can be processed using well known technologies and how such data can be transformed to demand fulfilling information according to the principles of Information Logistics.

The evaluation and the instantiation of the motivating use case confirmed the suitability of TiEE for a variety of telemedical applications. There are different kinds of telemedical scenarios we have to cope with: on the one hand high-frequent short-term measurements and on the other hand long-term with a low frequency of incoming vital signs. We showed that using a TIL as a flexible processing unit, TIEE is enabled to be set-up for different processing requirements in a very fast way. CEP as a basic technology for data processing allows for a very fast processing and detection of patterns as well as the expression of temporal relationships. Complex events, like the CTPE, are a kind of compressed data capable to be transformed into decision supporting information. Processing is controlled by a formalization of the information demand. A graphical language allows the description of telemedical questions according to expert knowledge or medical guidelines.

Two discussing points for future work can be emphasized: at first the demand modelling, especially the formalization of conditions, is too complex to be used by any physician. Some kind of repository or supporting interface should be supplied to support the modeling process. There is also a lack between the automated transformation of the graphical demand notation and the representation in terms of EPL-statements. Another critical point is the parameterization of the underlying algorithms within a TIL. Depending on the patients' demographical parameters or the kind of telemedical scenarios (long term/short term, high/low frequency) there is sometimes a need to configure the trend pattern algorithms. This task is done manually at the moment but could be optimized by using machine learning technologies.

A transfer of the existing results could be done in a horizontal (width) or vertical (depth) manner. The vertical transfer or the transfer to other telemedical field would require the construction of a repository with a set of TILs, specialized for a wide range of vital signs and different algorithms. Fraunhofer ISST is just working on such a repository. The horizontal transfer would address the usage of the given concepts within other domains. Fraunhofer ISST has ongoing research within the monitoring of painting within museums. A painting, represented as a TIL-Profile, needs several circumstances thus it will not be damaged. Humidity, temperature and other factors will be monitored in terms of TILs.

Fraunhofer ISST is working on further concepts to enable the usage of CEP for an Information Logistics processing and will combine it with machine learning to allow an automated, situation-based adaption of processing.

References

[1]   J. Bates, et al., "Using Events for the Scalable Federation of Heterogeneous Components," in Proceedings of the 8th ACM SIGOPS European workshop on Support for composing distributed applications. ACM, New York, 1998, pp. 58–65. Available: http://dx.doi.org/10.1145/319195.319205

[2]   M. Broy, Cyber Physical Systems. Springer, 2010. Available: http://dx.doi.org/10.1007/978-3-642-14901-6

[3]   Bundesverband Telekommunikation und neue Medien e. V.: Big Data im Praxiseinsatz – Szenarien, Beispiele, Effekte. 2012.

[4]   K. M. Chandy, et al., "Towards a Theory of Events," in Proceedings of the 2007 inaugural international conference on Distributed event-based systems. ACM, New York, 2007, pp. 180–187. Available: http://dx.doi.org/10.1145/1266894.1266929

[5]   M. K. Chandy, et al., "10201 Executive Summary and Manifesto -- Event Processing," In: K. M. Chandy, et al. (Eds.) Event Processing. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, Germany, Dagstuhl, Germany, 2011.

[6]   S. Charbonnier, On-Line Extraction of Successive Temporal Sequences from ICU High-Frequency Data for Decision Support Information. AIME. 2003, pp. 1–10.

[7]   W. Deiters, et al., The Information Logistics Approach toward User Demand-Driven Information Supply. Cross-media service delivery : Based on papers presented at the Conference on Cross-Media Service Delivery - CMSD-2003. Kluwer Academic Publishers, 2003.

[8]   O. Etzion and P. Niblett, Event Processing in Action. Manning. 2010.

[9]   European Commission: Demography report 2010 — Older, more numerous and diverse Europeans. Publications Office of the European Union, Luxembourg. 2011.

[10]   B. Feldmann, et al., Big Data in Healthcare - Hype and Hope. 2012.

[11]   Frost & Sullivan: Drowning in Big Data? Reducing Information Technology Complexities and Costs For Healthcare Organizations. 2012.

[12]   D. M. Haftor and M. Kajtazi, What is Information Logistics? An Explorative Study of the Research Frontiers of Information Logistics, 2009.

[13]   I. Haimowitz, Knowledge-based trend detection and diagnosis. Massachusetts Institute of Technology, 1994.

[14]   A. Hall and G. Walton, "Information overload within the health care system: a literature review," Heal. Info Libr J. 21, 2, 2004, pp. 102–108. Available: http://dx.doi.org/10.1111/j.1471-1842.2004.00506.x

[15]   A. Hinchley, Understanding Version 3: A primer on the HL7 Version 3 Healthcare Interoperability Standard. Münch, A. 2007.

[16]   R. E. Hunt and R. G. Newman, Medical Knowledge Overload: A Disturbing Trend for Physicians. Heal. Care Manag. Rev. 22, 1997, pp. 70–75. Available: http://dx.doi.org/10.1097/00004010-199702210-00009 http://dx.doi.org/10.1097/00004010-199724000-00009   http://dx.doi.org/10.1097/00004010-199701000-00009

[17]   IEEE: ISO/IEEE 11073-10201:2004: Health informatics Point-of-care medical device communication Part 10201: Domain information model. 2004.

[18]   O. Koch and E. Rotaru, "Using Context to Improve Information Supply in the Medical Sector," in Abramowicz, W. et al. (Eds.) Business Information Systems Workshops - BIS 2010 International Workshops. Springer, 2010, pp. 192–203. Available: http://dx.doi.org/10.1007/978-3-642-15402-7_25

[19]   B. H. Krogh, et al.: Cyber-Physical Systems. 2008.

[20]   M. D. Landry, et al., Analysis of factors affecting demand for rehabilitation services in Ontario, Canada: A health-policy perspective. Disabil. Rehabil. 30, 24, 2008, pp. 1837–1847. Available: http://dx.doi.org/10.1080/09638280701688078

[21]   D. Laney, {3D} Data Management: Controlling Data Volume, Velocity, and Variety. 2001.

[22]   M. Line, Draft Definitions: Information and Library Needs, Wants, Demands and Uses. Aslib Proc. 26, 87, 1974. Available: http://dx.doi.org/10.1108/eb050451

[23]   D. Liu, et al., University of Queensland Vital Signs Dataset: Development of an Accessible Repository of Anesthesia Patient Monitoring Data for Research. Anesth. Analg. 2011.

[24]   H. J. Lowe, et al., STRIDE--An Integrated Standards-based Translational Research Informatics Platform. AMIA. 2009, pp. 391–395.

[25]   D. C. Luckham, The Power of Events: An Introduction to Complex Event Processing in Distributed Enterprise Systems. Addison-Wesley, Boston, 2002.

[26]   S. M. Maviglia, et al., KnowledgeLink: Impact of Context-Sensitive Information Retrieval on Clinicians' Information Needs. J Am Med Inf. Assoc. 13, 1, 2006, pp. 67–73. Available: http://dx.doi.org/10.1197/jamia.M1861

[27]   S. Meister, et al., Fitpit-Fitness Cockpit: Information System To Optimize Training Schedules In Rehabilitation. ECIS 2013 Proc. 106, 2013.

[28]   S. Meister, "Telemedical Events: Intelligent Delivery of Telemedical Values Using CEP and HL7," in L. Niedrite, et al. (Eds.) BIR Workshops. Springer, 2012, pp. 1-13. Available: http://dx.doi.org/10.1007/978-3-642-29231-6_1

[29]   S. Meister, "Usage of CEP and HL7 to Solve Information Logistics Problems in Telemedicine," in 4th International Workshop on Information Logistics and Knowledge Supply for Viable Enterprises (ILOG 2011), 2011.

[30]   S. Meister and O. Koch, "Using Complex Event Processing and Context for Intelligent Information Filtering and Supply," in K. Schilling and F. Leutert, (Eds.) 1st IFAC Conference on Embedded Systems, Computational Intelligence and Telematics in Control CESCIT, 2012.

[31]   S. Meister and V. Stahlmann, "Telemedical ILOG Listeners: Information Logistics Processing of Telemedical Values Using CEP and HL7," in R. Wichert and B. Eberhardt, (Eds.) Ambient Assisted Living - Advanced Technologies and Societal Change. Springer, 2012. Available: http://dx.doi.org/10.1007/978-3-642-27491-6_18

[32]   Society for Medical Decision Making: Proposal for Clinical Algorithm Standards: Society for Medical Decision Making Committee on Standardization of Clinical Algorithms. Med. Decis. Mak. 12, 2, 1992, pp. 149–154. Available: http://dx.doi.org/10.1177/0272989X9201200208

[33]   V. Vuori, Methods of Defining Business Information Needs. In: Maula, M. et al. (Eds.) Frontiers of e-Business Research ICEB + eBRF 2006 - Conference Proceedings. 2006, pp. 311–319.

[34]   S. Weber, et al., Implementing a Real-time Complex Event Stream Processing System to Help Identify Potential Participants in Clinical and Translational Research Studies. AMIA. 2010, pp. 472–476.

[35]   A. Willems, et al., Information Logistics Research Report - Frameworks in Healthcare Industry 2009.

[36]   T. D. Wilson, "Information Overload: Implications for Healthcare Services," Health Informatics J. 7, 2001, pp. 112–117. Available: http://dx.doi.org/10.1177/146045820100700210

[37]   R. Winter, et al., Integrierte Informationslogistik. Springer Berlin Heidelberg, 2008.

[38]   J. Wyatt, Information for Clinicians: Use and Sources of Medical Knowledge. Lancet. 338, 8779, 1991, pp. 1368–1373. Available: http://dx.doi.org/10.1016/0140-6736(91)92245-W