Information-Driven Monitoring of Production Process: A Semantic Data Model

Oksana Petrina, Sergei Marchenkov, Dmitry Korzun
Petrozavodsk State University (PetrSU)
Petrozavodsk, Russia
{petrina, marchenkov, dkorzun}@cs.petrsu.ru

Abstract—**Recent technologies of industrial monitoring provide highly fragmented information. Large amounts of multi-parameter sensed data on production process and equipment operation are stored in disparate structures (databases). Effective use of the collected information requires data fusion within an information-driven monitoring system. In this short paper, we design a semantic data model to create a unified information space that fuses events derived from real-time sensed data streams. Our semantic data model considers a hierarchy of production equipment nodes and supports rules for identifying and composing events in the production process under monitoring.**

I. Introduction

A software and hardware system for information-driven monitoring of production processes and equipment operation is developed in Petrozavodsk State University (PetrSU) [1]. The monitoring system implements continuous (real-time) observation of the object under monitoring. The sensors are attachable (mounted) equipment. A monitored object is such production equipment as metalworking centers, presses, gas turbines, etc.

Industrial monitoring provides highly fragmented information, when large amounts of multi-parameter sensed data are stored in disparate structures (databases). Effective use of the collected information requires data fusion within an information-driven monitoring system. In this short paper, we design a semantic data model to create a unified information space [2] that fuses events derived from real-time sensed data streams. The semantic data model provides fusing the data from production process monitoring.

The semantic data model is developed for a graph type non-relational database management system. We apply the well-known solutions for the data presentation in Industrial Internet. Our semantic data model considers a hierarchy of production equipment nodes (units) and supports rules for identifying and composing space-time related events in the production process under monitoring.

The rest of this paper is organized as follows. Section II presents the studied class of data presentation models for industrial monitoring. Section III introduces our semantic data model for representing space-time events and operating conditions of production equipment based on multiparameter sensed data. Section IV discusses our implementation of the semantic data model and early experiments. Finally, Section V concludes the paper.

II. Related Work

Effective use of the collected information requires data fusion within an information-driven monitoring system. An efficient way is needed to store, manage, and retrieve information about the monitoring object and high events to build recommendations for work on repair and maintenance. One of the key approaches to development of monitoring systems is an ontologically driven methodology for describing and presenting objects in the subject area, integrating mixed information, and formalizing data and knowledge [3].

Work [4] presents a model for monitoring deformations of structures of potentially dangerous objects, which implements the integration of heterogeneous data and calculation modules into an integral distributed information processing system. Paper [5] suggest using databases adapted for time series analysis for the system of archiving and data analysis of the monitoring system. Data is stored in a circular record — stale data is filtered to keep the aggregated data smaller.

Industry 4.0 project [6] aims at automation and communication in creating smart factories that maximize production capabilities. Modeling and implementation of components are implemented using a model Resource Description Framework (RDF) using semantic technologies. The execution of production processes depends not only on their internal state and interaction with the user, but also on the context of their execution to provide additional information to improve monitoring.

Work [7] presents the few existing ontologies for Inductry 4.0. Ontologies can provide the solution by formalizing the smart manufacturing knowledge in an interoperable way to promote adoption of a coherent approach for the semantic communication in between multiple intelligent systems, which include human and artificial (software or hardware) agents.

Work [8] proposes applying ontology-oriented models to represent knowledge on production processes and equipment operation in a machine interpreted form. The proposed ontology in [9] formalizes domain knowledge related to condition monitoring tasks of manufacturing processes.

Work [10] presents a semantic approach to modeling automation systems using ontology and its applicability in industrial automation networks using OPC UA. The Automation I4.0 ontology provides a semantic tool for representing the concepts of high, medium and low-level automation systems and offers a way to directly integrate into software applications and communication protocols based on semantics.

Work [11] presents analysis of typical standardized Web of Things (WoT) ontologies covering various levels of WoT. Some basic ontologies describing measurements of physical parameters and sensors of the system are identified: SOSA ontology [12], SSN ontology [13], DUL ontology [14]. The use of those ontologies in the development of a binding model is considered in the next section.

III. Semantic Data Model

The data model describes the monitoring object (equipment units, mounted sensors and video cameras, events) and the surrounding context (operating conditions, employee profiles). The data model describes the monitoring object, as well as the hierarchy of equipment nodes with reference to the sensor. At the same time, a shared information space for the functioning of the object is created with the digitalization of its main aspects. The change in technical state is determined by an event-driven model using basic and composite events.

The semantic data model design of the model was developed based on several well-known ontologies. Ontology of sensor, observation, sampling and actuator (SOSA) provides a formal but lightweight universal specification for modeling interactions between objects involved in observation, triggering, and sampling [12]. Semantic Sensor Network (SSN) ontology [13] describes sensors and observations and related concepts, but does not describe domain concepts, time, location, etc. Together, SSN and SOSA ontologies describe systems of sensors and actuators, observations, procedures used, objects and their properties that are observed or acted upon, samples and the sampling process, and so on. DOLCE+DnS Ultralite (DUL) ontology [14] is the basic ontology for representing events to simulate physical or social. DUL ontology connects time and space, objects and people. The main entities of the linking model are shown in Fig. 1.

The class Machine describes the main object of monitoring, for example, a machine tool or other production equipment. The equipment consists of nodes described by the subclass Node, which are also monitored. The model takes into account the hierarchy of monitoring objects. Class System is abstraction for pieces of infrastructure as sensor, linked with the class Node. Both classes have parameters for installing a device on a node, described by the class Deployment. Deployment describes the Deployment of one or more Systems for a particular purpose (a temperature Sensor deployed on wall).

Class System includes two subclasses Sensor and DAS (DataAcquisitionSystem). Both classes have System Capability that describes normal measurement, actuation, sampling properties such as accuracy, range, precision. The System continues to operate as defined using SystemCapability.

A sensor makes observations. Class Observation is represented by two parameters: timestamp and result. The result of an Observation contains a value representing the value associated with the observed Property. Class ObservableProperty describes a physical quantity—observable quality (property, characteristic).

A sensor is connected to data acquisition system. Class DAS includes data acquisitor that implements Procedure. Procedure is workflow, protocol, plan, algorithm, or computational method specifying how to make an Observation. Class Procedure is recursive, has input and output, which are represented by the corresponding classes. Class Output saves data in the class TSD (Time Series Database).

Class DataAcquisitor is a subclass of class SoftwareAgent. A software agent implements service and detects basic events.

The Class Event describes any physical, social, or mental process, event, or state. An event in multiparameter monitoring systems is determined by a finite time interval at which an atomic physical phenomenon occurs. A physical phenomenon usually occurs when an observation object changes its state at a certain point in time. Moreover, an event occurring in a time interval between two consecutive points is considered to have occurred at the time of the end point of the interval.

A fragment of the model describing events in the monitoring system and related entities is shown in Fig. 2. Class Event has two subclasses: BasicEvent and CompositeEvent. A basic event is an event, the composition and detection algorithm of which are predefined and implemented in the monitoring system. The definition of a basic event in the system is set by the following parameters:

1) eventID in an identifier of event;
2) timestamp;
3) level of event importance: 0 is no problem, 1 is weak problem, 2 is medium problem, 3 is strong problem;
4) class DataEvent linked class Sensor;
5) class TypeEvent.

A composite event consists of several basic events. A composite event is an event that is generated and detected by applying a set of different operators (logical, causal, etc.) to a set of basic and composite events. The occurrence time of a composite event is the occurrence time of its latest basic event. The features of events in multiparameter monitoring systems are as follows: events occur at a high speed, while multiple occurrences of the same type of event only clarify the previous data values; events are causally dependent events; composite events are detected along a moving time window (trend analysis and forecasting). These features are taken into account in the descriptions of composite events using operators.

A composite event is described with operators in class EventOperator. Disjunction, when one of two events happens. It is taken into account that events in the system cannot occur simultaneously. Conjunction, when two events happen regardless of the order. Sequence of two events is a composite event that occurs when the second event occurs after first event. Periodic and aperiodic events are also taken into account.

A composite event includes operators based on an event specification language (termed Snoop) [15] for active databases. Snoop supports temporal, explicit, and composite events in addition to the traditional database events. The applicability of this concept of composite events to the industrial monitoring context requires additional research.

IV. Discussion

In the current implementation, industrial monitoring service data is stored in two databases implemented using MongoDB

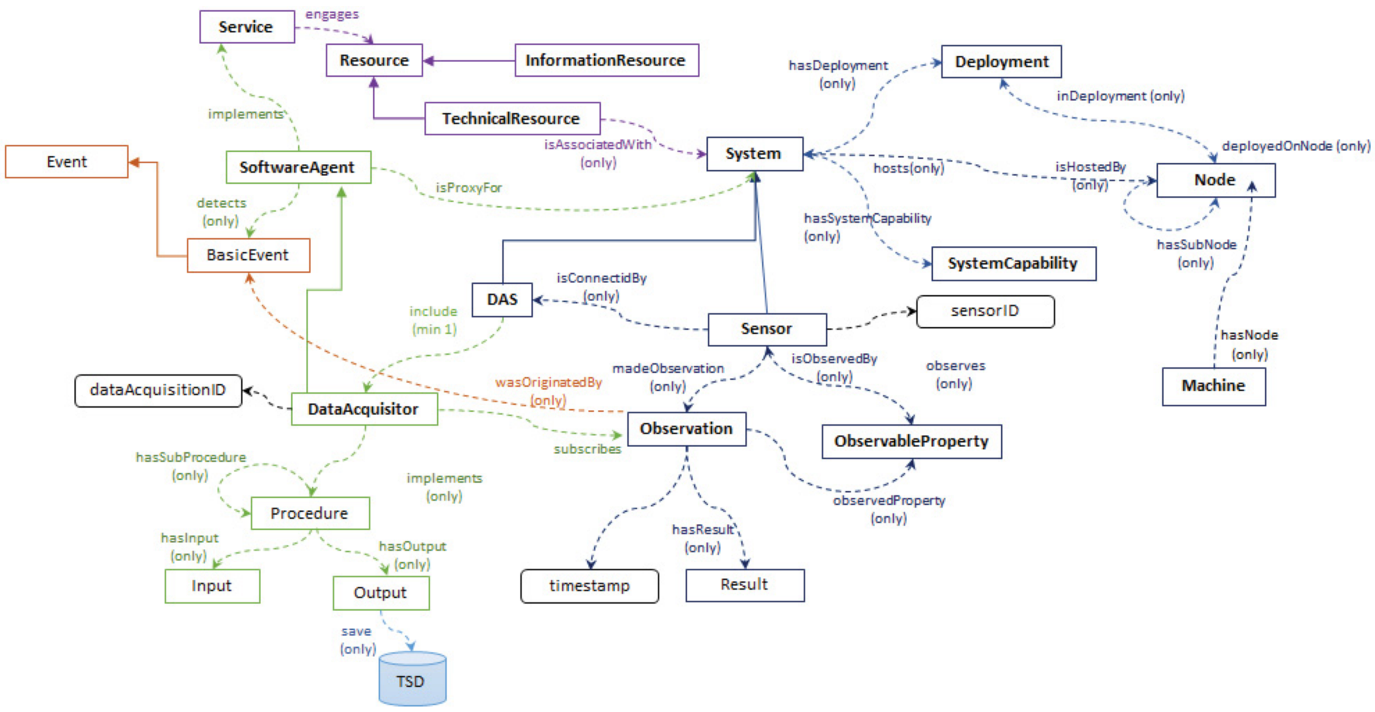

Fig. 1. Semantic Data Model: Main entities and relations

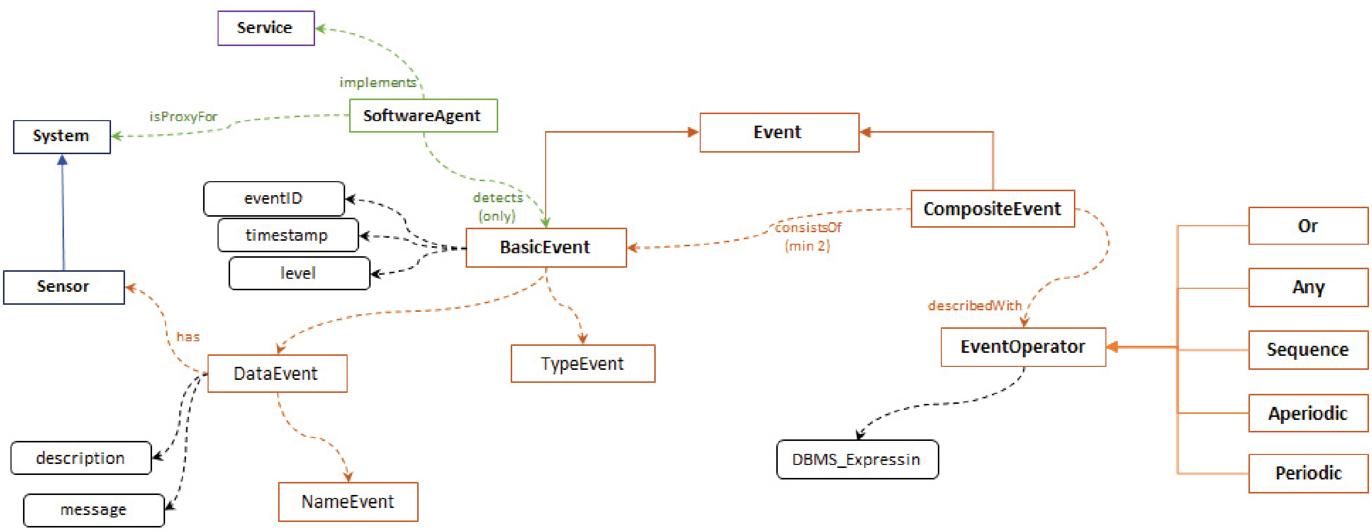

Fig. 2. Semantic Data Model: Composite Events

and Clickhouse. Each service is represented by one or more software modules. MongoDB database stores module collection data. Each module stores its data in its own collection.

Streaming data from sensors is saved to the database Clickhouse. The paper [16] presents the concept for storing and indexing numeric time series that allows creating compact data representations optimized for cloud storages and perform typical operations (uploading, extracting, sampling, statistical aggregations, and transformations) at high speed.

The current database implementation does not support semantic relations. First, the hierarchy of production equipment nodes cannot be stored. Second, the sensors and data streams are not tied to specific nodes of the production equipment. Third, no support is available for implementing rules to identify composite events.

There is necessary to restructure and transfer data to another DBMS to implement the proposed semantic data model. We plan to use the graph database that provides better performance, flexibility, and agility compared to non-relational databases. Paper [17] presented an analysis of five the most commonly used graph databases: AllegroGraph, ArangoDB, InfiniteGraph, Neo4J and OrientDB. Authors conclude that Neo4J and ArangoDB offer the best functionalities to implement a graph database with Neo4J for standing out for its simplicity and due to its powerful query language named Cypher. Neo4j is available in a GPL3-licensed open-source ”community edition”.

Our recent plan is transforming the data from the database MongoDB to the graph database Neo4j. Previously, we made a series of queries to retrieve events from the database MongoDB. After transferring the data to Neo4j, we plan to repeat the series of queries and compare the execution efficiency. The database structure is modified to take into account the proposed semantic date model and re-measure.

V. CONCLUSION

This paper introduced our semantic data model to collect information about the production process under monitoring. The semantic data model describes monitored objects and subjects (equipment nodes, mounted sensors and video cameras, events) including the context (operating conditions, employee profiles). The data model takes into account the hierarchy of production equipment nodes and sensors. At the same time, a unified information space is created to digitize the production process. The change in technical state is determined by an event-driven model using basic and composite events. The proposed semantic data model is oriented to implementation using graph type non-relational database management system.

ACKNOWLEDGMENT

This research is implemented in Petrozavodsk State University (PetrSU) with financial support of the Ministry of Science and Higher Education of Russia within Agreement no. 075-11-2019-088 of 20.12.2019 on the topic “Creating the high-tech production of mobile microprocessor computing modules based on SiP and PoP technology for smart data collection, mining, and interaction with surrounding sources”. This research is supported by RFBR (research project # 19-07-01027) in part of fundamental methods of smart spaces for ontology modeling. The work is in collaboration with the Artificial Intelligence Center of PetrSU.

REFERENCES

[1] D. Kirienko, P. Lunkov, V. Putrolaynen, S. Aryashev, and M. Belyaev, “Modular hardware platform for the development of iot devices implemented using multi-chip packaging technology,” in *IOP Conference Series: Materials Science and Engineering*, vol. 862, no. 3. IOP Publishing, 2020, p. 032012.

[2] D. Korzun, E. Balandina, A. Kashevnik, S. Balandin, and F. Viola, *Ambient Intelligence Services in IoT Environments: Emerging Research and Opportunities*. IGI Global, 2019.

[3] P. Okhtilev, A. Bakhmut, A. Krylov, and M. Okhtilev, “Podkhod k otsenivaniyu strukturnykh sostoyanii slozhnykh organizatsionno-tekhnicheskikh ob”ektov na osnove obobshchennykh vychislitel’nykh modelei,” *Naukoemkie tekhnologii v kosmicheskikh issledovaniyakh zemli*, no. 5(9), pp. 73–82, 2017.

[4] T. Lapina *et al.*, “Postroenie sistem monitoringa parametrov sluchainykh protsessov,” *Modelirovanie, optimizatsiya i informatsionnye tekhnologii*, vol. 23, no. 4, pp. 484–493, 2018.

[5] O. Panarin and I. Zakharov, “Osobennosti monitoringa mobil’nykh sistem obrabotki informatsii,” *Nauchnyi servis v seti Internet: trudy XXI Vserossiiskoi nauchnoi konferentsii*, pp. 551–560, 2019.

[6] I. Grangel-Gonzlez *et al.*, “An rdf-based approach for implementing industry 4.0 components with administration shells,” in *In. Proc. 2016 IEEE Emerging Technologies and Factory Automation*, 2016.

[7] V. R. S. Kumar, “Ontologies for industry 4.0,” *The Knowledge Engineering Review*, vol. 34, 2019.

[8] F. Giustozzi, J. Saunier, and C. Zanni-Merk, “Context modeling for industry 4.0: an ontology-based proposal,” *Procedia Computer Science*, vol. 126, p. 675–684, 2018.

[9] Q. Caoa, F. Giustozzia, C. Zanni-Merka, F. de Bertrand de Beuvronb, and C. Reichc, “Smart condition monitoring for industry 4.0 manufacturing processes: An ontology-based approach,” *CYBERNETICS AND SYSTEMS: AN INTERNATIONAL JOURNAL*, vol. 50, no. 2, pp. 82–96, 2019.

[10] S. Gil and G. D. Zapata-Madrigal, “Semantic automation systems, a suitable approach for automation networks in the industry 4.0,” in *In. Proc. 2019 IEEE 4th Colombian Conference on Automatic Control (CCAC)*, 2019.

[11] W. Li, G. Tropea, A. Abid, A. Detti, and F. Le Gall, “Review of standard ontologies for the web of things,” in *In. Proc. 2019 Global IoT Summit (GIoTS)*, 2019.

[12] K. Janowicz, “Sosa: A lightweight ontology for sensors, observations, samples, and actuators,” *Journal of Web Semantics*, vol. 56, 2019.

[13] A. Haller and other, “The modular ssn ontology: A joint w3cand ogc standard specifying the semanticsof sensors, observations, sampling, and actuation,” *Semantic Web*, vol. 10, p. 9–32, 2019.

[14] A. Scherp, T. Franz, T. Franz, and C. Saathoff, “F - a model of events based on the foundational ontology dolce+dns ultralite,” in *In Proc. of the 5th International Conference on Knowledge Capture (K-CAP 2009)*, 2009, pp. 1–8.

[15] S. Chakravarthya and D. Mishra, “Snoop: An expressive event specification language for active databases,” *Data Knowledge Engineering*, vol. 14, no. 1, pp. 1–26, 1994.

[16] A. Visheratin, A. Struckov, S. Yufa, A. Muratov, D. Nasonov, N. Butakov, Y. Kuznetsov, and M. May, “Peregreen – modular database for efficient storage of historical time series in cloud environments,” in *In. Proc. 2020 USENIX Annual Technical Conference*, 2020, pp. 589–601.

[17] D. Fernandes and J. Bernardino, “Graph databases comparison: Allegrograph, arangodb, infinitegraph, neo4j, and orientdb,” in *In. Proc. DATA 2018 - 7th International Conference on Data Science, Technology and Applications*, 2018, pp. 373–380.