Intelligent Big Data Processing for Wind Farm Monitoring and Analysis Based onCloud-Methodologies and Digital Twins A quantitative approach

[1]CHANDAN KUMAR PANDA, *Gandhi Institute of Excellent Technocrats, Bhubaneswar,India*
[2]SEETAL KUMAR PRADHAN, *Jagannath Institute of Engineering and Technology, Cuttack, Odisha, India*

Abstract— Checking and investigation cycles of enormous information are regularly bound either to a specialized or business setting and, surprisingly, further isolated into constant or discrete goal. In our methodology a sound semantical incorporation of these different data types is introduced and applied to a substantial use-case, the worldwide observing and examination of wind ranches. Moreover, another UI style in light of expanded the truth is introduced permitting - even to a fledgling - a natural and fast comprehension of perplexing examination results.

Keywords-big data processing; augmented reality; digital twin; cloud computing; internet of things; real-time processing; micro services; wind farms; smart control station

I. INTRODUCTION

Monitoring of wind farms is a normal and well-known task these days. But at a closer look, the focus is usually put on the analysis of technical data resulting from sensors in different discrete time intervals. The complex dependency structures between the data streams from different sensor types and business information are not integrated and cannot be understood intuitively. Most control stations ignore the related business information, because it is not needed from atechnical point of view. But several use-cases suffer from this lack of integration e.g. this may lead in most of the casesto unfair decisions with respect to certain owners of the equipment.

To avoid this, a cohesive and sound big data processing approach is presented which transparently integrates all relevant information from various data sources. At the beginning of the process they are categorized and separated in different equivalence classes. Thereafter they are homogenized with respect to minimal time units. In a following staging process semantic objects are built. They are the basis for the analyzing algorithms which generate strategic material. Finally, all the information is presented in a user-friendly model: a so called digital twin [1]. The underlying system landscape allows the user a real-time viewand analysis of a complete wind farm.

II. DATA SOURCES

The term big data describes mass data that can no longerbe processed with traditional databases and ETL-tools due toits size, diversity and fast pace of data generation [2].

Sources that contribute to big data are diverse. Examples are several IT-systems such as Enterprise Resource Planning- (ERP), Customer Relationship Management-(CRM) or Supply Chain Management- (SCM) systems. Mass data can also be extracted from different social media like facebook, Twitter or even YouTube. In the context of "Internet of Things" (IoT) technologies [3], [4] almost any device can be linked to the internet and therefore send data about itself. In our use-case, wind energy converters (WEC) are equipped with various sensors that collect information about the facility, its components and environment.

As shown in Figure 1, data from the described sources can be categorized into business and technical information and also into discrete and continuous data streams. Obviously, what kind of data is gathered depends on the industry or rather the specific use-case. For a WEC manufacturer or service provider business master data about customers, contracts and materials are available. Data streams of business information could be electricity rates of different stock exchanges or different key performance indicators. A WEC manufacturer also owns massive amounts of technical master data such as model descriptions, dimensions and weight of their product line and also first commissioning dates of the installed WECs. A service provider may improve this data with continuous maintenance information [5]. The sensors on the WEC itself create continuous data streams by continuously measuring many attributes like wind speed, humidity, vibration or spindle temperature.

The logical, qualitative problem of the integration ofcontinuous and discrete data sources is solved via the Message Queuing Telemetry Transport protocol (MQTT). It serializes all incoming data packages from continuous sources. The IoT-service relates a local timestamp to the object. If two data sources with different granularity have to be combined e.g. a numerical algorithm based on cubic spline interpolation is applied, generating new statisticalvirtual data items.

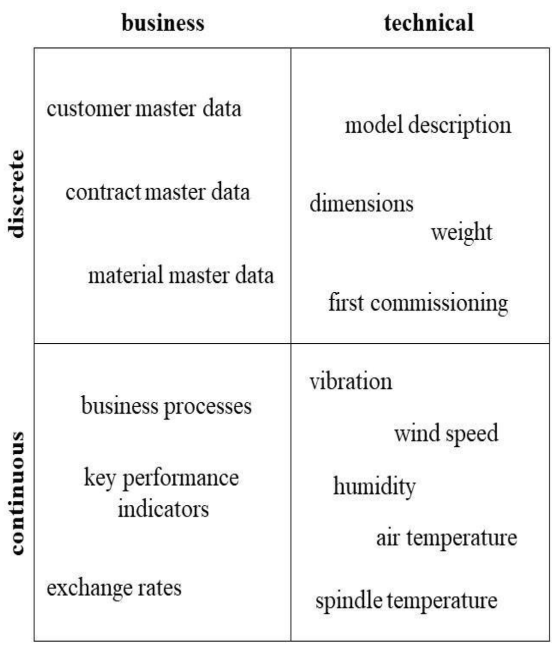

Figure 1. Type identification of data sources for wind farm monitoring

III. Data Processing and Dynamic Semantic ObjectGeneration

Data processing and the dynamic building processes for business objects are based on an event driven approach. Eachinformation item is interpreted in its own semantical contextand triggers certain actions depending on the origin of the event and its setting. This can happen successively several times until a semantical fixed point is reached. It is a bit likecomputing the transitive closure of a complex dependency graph. This virtual graph structure is not only constructed in a hierarchical, tree-like structure but also in a general directed graph structure allowing dependencies like multipleinheritance approaches in object-oriented programminglanguages.

For instance, the number of revolutions of the rotor of a certain WEC is measured by a sensor in a cycle time of 50 milliseconds. When the connected Raspberry Pi sends this information to the cloud based IoT-System the first semantic envelope is put around the measured data by the local hardware device, e.g. date-, time- and GPS-information. Thisis the smallest, initial semantic object (SO^0) in our theory. The SO^0 creates an event in the central IoT-engine andinitiates a cascade of consecutively computations (e.g. n- iterations) in which the SO^0 is enlarged with more and more context information ending in SO^n. In the case of sensor data,related business content is added. This algorithm is not runtime critical and can be parallelized, because each computation instance generates information based on disjunct data environments. The underlying architecture allows a full micro service approach ensuring an almost linear scalability of this process.

From an abstract point of view, this procedure transforms each single e.g. sensor information (a numerical information)into an object in a m-dimensional data space with technical and business information. Each dimension has its own syntactical domain which is linked to a semantic domain.

In our approach each data source can have its own data space and dimensions. The combination of arbitrary semantic objects SO_i und SO_j ($i \neq j$) can only be established on dimensions with the same sematic domains – independently of the data space and the number ofdimensions.

Most of our analysis algorithms are based on the computation of a metric between certain semantic objects. The “sematic neighborhood relation” $SNR(SO_i,SO_j)$ is derived from the underlying data space of the data items. It can also be derived from a subset of dimensions as $SNR(\pi_i(SO_i), \pi_j(SO_j))$ (also allowed π = identity). When a relation is established between different data spaces than a logical cross product is built on the joined semantic domain property and the SNR suitably enhanced.

IV. System Landscape

The underlying system landscape is based on cloud technologies, because of the enormous and simple to handle scalability options in the dimensions of compute power, main memory and storage. Furthermore, administration efforts for this kind of complex system landscape can be reduced to a minimum – as predicted via the PaaSparadigm.The system landscape and data flow can be reenact in Figure2.

As stated before, each WEC is equipped with several different sensors. Raspberry Pi-computers are implemented on each WEC to gather data streams from those sensors and to send a single continuous data stream via MQTT-protocol to the cloud IoT-interface of a SAP Cloud Platform (SCP). The Raspberry Pi computers bundle the data sent from various sensors and send only a single data stream. However,it is also possible to send several unique data streams to the SCP as shown by the example of SCADA data in the figure.

Certain IoT Micro Services – as explained in chapter III -embedded in the SCP, receive the data stream sent from eachWEC of wind farms around the world and load them into a SAP HANA in-memory database [6]. Java Micro Services, which are also embedded in the SCP, extract data from the HANA database and provide them to the Gateway Server by using OData services. At the end, the Gateway Server transmits the information to the SAP Enterprise Central Component (ECC) backend. Both Gateway Server and SAP ECC backend are located in the same cloud.

Various IT-systems and external cloud services (for example weather forecast, flight of birds, exchange rates) are connected to the ECC Backend and serve as virtual data sources for both technical and business information. Even cloud-based Business Intelligence (BI) solutions can be integrated. That way insights and knowledge gained from BIanalyses can be added to the database of the DT.

The data is visualized for monitoring and analysis of a wind farm by using augmented reality (AR) [7], [8]. A smart glass interface like the Microsoft HoloLens [9] is needed forthe user to view the visualization. The HoloLens is connected to the IT-landscape in the cloud via wireless LAN.

For security reasons it is highly suggested to implementat least one firewall between the SCP and the own cloud and another one between the own cloud and external systems thatact as external, virtual data sources. The figure does not include these firewalls to ensure readability.

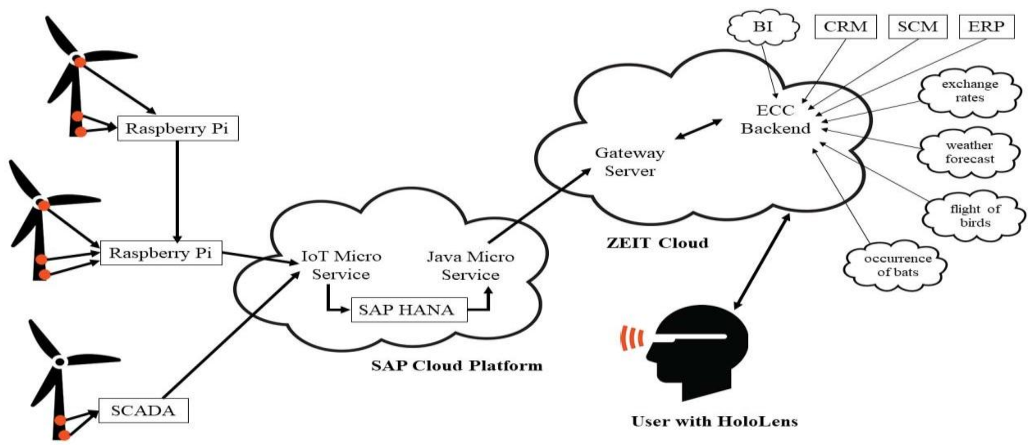

Figure 2. System landscape

(GIS) and has information about the coordinates of each

V. MONITORING OF WIND FARMS VIA DIGITAL TWIN

As of now, no reliable and formal definition of a “digitaltwin” (DT) can be found in literature. In this approach the term is used to describe a virtual representation based on AR-technology [8] of a physical object which can or alreadyexists in the real world.

Smart glasses enable the user to access the AR [10], [11]and to interact with the DT to monitor and analyze single WECs and also entire wind farms. The use of AR and smart glasses also allows the cooperation of several users. Multiplesmart glasses can be connected and all wearers can then access the same DT.

The business processes of a WEC manufacturer or a WEC service provider can be divided logical into research and development, sales, production, construction and service.The DT is adapted to the area of application or setting in which it is used. During research and development phases for example, the DT might be used to visualize CAX-drawings and to simulate the WEC operation during critical weather conditions, e.g. wind load. Figures 3, 4 and 5 show the DT that is used during the construction of a windfarm. The DT is connected to a geographical information system

WEC which allows for accurate placement of the WECs on the hologram-map shown in Figure 3. The current construction progress of a WEC is also visualized and can be reenact in Figure 4. Furthermore, the DT is able to representthe construction progress, the usage of needed tools such as cranes as well as current weather conditions.

Rain, snow andfog can be displayed as realistic animations.

To interact with the DT no additional data glove is necessary. By air-tapping – a gesture where the user taps together thumb and index finger – on a single WEC, both technical and business information are presented to the user. Using hand gestures, the elements of the DT can be scaled insize and arranged as needed [12].

The user can navigate hierarchically through the semantic objects that were previously described. Data about maintenance needs, ownership and conditions of external service providers are displayed graphically [13].

These information can be used by managers for several different tasks such as planning maintenance works [14], coordinating contracts and customers as well as dealing withliability claims.

Figure 3. Digital twin of a wind farm

Figure 4. Detailed view of the digital twin

Figure 5. Integration of business and technical information

VI. CONCLUSION

The case study has proven that building a control station based on cloud technologies in combination with big data analysis algorithms [15] on optimized data structures is real-time capable. We have presented an approach for the integration of technical and business data in one single digital twin implemented via AR.

We have applied this technology to several wind farms in different countries around the world. We were able to monitor and operate all wind farms in parallel.

The smallest granularity in our cloud-based system landscape is actual 40 milliseconds in environment of 2,000

ISSN: 2278-4632
Vol-10 Issue-09 No.03 September 2020

sensors and up to ten ERP systems. A maximum reaction time of 190 milliseconds in the complex and worldwide distributed system landscape could be guaranteed (time slice starting from the incoming sensor data moment, including the analysis & interpretation phase as well as time needed tosend steering signals to the WEC).

REFERENCES

[1] H. Pargmann, D. Euhausen, and R. Faber, "Applied Augmented Reality as an integration approach for ERP- and CAX-systems in manufacturing and operating processes for wind farms", in: LNBIP Post-Proceedings of ERP Future conference 2017, Innsbruck,

[2] Springer Publishing, 2017 – in press

[3] D. Fasel and A. Meier, "Big Data: Grundlagen, Systeme und Nutzungspotenziale", Springer Vieweg Publishing, 2016

[4] H. Chaouchi, "The Internet of Things: Connecting Objects to the Web", ISTE Ltd. and John Wiley & Sons Inc., 2010

[5] F. Sprenger and C. Engemann: "Internet der Dinge: Über smarteObjekte, intelligente Umgebungen und die technische Durchdringung der Welt", transcript Publishing, 2015

[6] T. Bauernhansl, M. ten Hompel, and B. Vogel-Heuser: "Industrie 4.0 in Produktion, Automatisierung und Logistik: Anwendung, Technologien, Migration", Springer Vieweg Publishing, 2014

[7] P. Preuss: "In-Memory-Datenbank SAP HANA", Springer GablerPublishing, 2017

[8] A. B. Craig, "Understanding Augmented Reality: Concepts and Applications", William Andrew Publishing, 2013

[9] R. Dörner, W. Broll, P. Grimm, and B. Jung, "Virtual und Augmented Reality (VR / AR): Grundlagen und Methoden der Virtuellen und Augmentierten Realität", Springer Publishing, 2014

[10] Microsoft, "Windows Mixed Reality Dev Center",

[11] https://developer.microsoft.com/en-us/windows/mixed-reality/, 2017

[12] A. Mehler-Bicher and L. Steiner, "Augmented Reality: Theorie und Praxis", de Gruyter Publishing, 2014

[13] P. Milgram, H. Takemura, A. Utsumi, F. Kishino, "Augmented reality: A class of displays on the reality-virtuality continuum", in: Proceedings of SPIE, 1994

[14] M. Tönnis, "Augmented Reality: Einblicke in die Erweiterte Realität",

[15] Springer Publishing, 2010

[16] M. Billingshurst, T. Starner, "Wearable devices: new ways to manageinformation", in: Computer, vol. 32., no.1,1999

[17] T. Kaufmann, "Geschäftsmodelle in Industrie 4.0 und dem Internet der Dinge: Der Weg vom Anspruch in die Wirklichkeit", Springer Vieweg Publishing, 2015

[18] R. Reichert: "Big Data: Analysen zum digitalen Wandel von Wissen, Macht und Ökonomie", transcript Publishing, 2014