Product-Service Systems across Life Cycle

The challenges of cybersecurity frameworks to protect data required for the development of advanced maintenance

Jaime Campos[a], Pankaj Sharma[b], Erkki Jantunen[c], David Baglee[d], Luca Fumagalli[e]

[a] Department of Informatics, Linnaeus University, SE-35195 Växjö, Sweden
[b] Mechanical Engineering Department, IIT Delhi, New Delhi, India
[c] VTT Technical Research Centre of Finland, P.O.Box 1000, FI-02044 VTT, Finland
[d] Department of Computing, Engineering and Technology, University of Sunderland, UK
[e] Department of Management, Economics and Industrial Engineering, Politecnico di Milano, Italy

*Corresponding author . Jaime Campos. Tel.: +46-(0) 470-708829 E-mail address:jaime.campos@lnu.se

Abstract

The main objective of the paper is to highlight the important aspects of the data management in condition monitoring and maintenance, especially when the emergent technologies, such as the cloud computing and big data, are to be considered in the maintenance department. In addition, one of the main data management elements highlighted in the current work are the cybersecurity issues which might be one of the biggest obstacles hindering the development of cloud based big data for condition-based maintenance (CBM) purposes. Further, the benefits and current risks of storing a company's data in the cloud are highlighted. The authors discuss as well different data needs in various processes in the area of asset management. In addition, the challenges and issues to be addressed for the optimal use of the company data at the cloud together with the big data approach are addressed. This is seen as an important part in an effort to achieve sustainable information and communication technologies for the industry.

Keywords: Big data; cloud computing; maintenance engineering; asset management

1. Introduction

Internet came in civilian use at the beginning of the 1990's and its use has increased manifolds after that. Its popularity has led to various new technology developments such as the e-maintenance, the Internet of Things (IoT), Industry 4.0, etc., which are all based on the broad implementation of ICTs, especially web technologies having the objective to support different processes of a company. The new Industry 4.0 initiative aims to connect these new ICTs to create smarter factories resulting in a revitalized manufacturing sector, promising to be the next industrial revolution for manufacturing [1]. Consequently, for its successful use it is important to understand the characteristics of the new emergent technologies to be able to use them properly in a company. In the cloud computing the entire network - based computing over Internet is seen as a service. It has evolved from the comprehensive dissemination of data and information virtualization from client server and service oriented architectures, automatic to utility computing [2]. The security aspects are rather important when the cloud computing is used since the security strategies that have been developed since the 1980's are not applicable to the cloud computing. The prime reason for this increased importance of

security is that the servers' part of the Cloud is not in the same domain, i.e. the data owner and Cloud computing servers are normally in two different domains. Consequently, numerous efforts have been made considering this matter and evolving system models and security strategies are being debated and tried, explicit for the system with the features of cloud computing. These efforts omit, for example, cryptographic methods and data decryption keys [3, 4]. Many private users as well as organizations hesitate over the adaptation of cloud computing and its services because of the risks related to the security and privacy of these services [5]. Kumar et al. [6] believes that security and privacy are the two major factors that hinder the growth of real time business related cloud computing for business purposes. It is, therefore, important for the academia and industry to develop new ways, services and technologies to provide a secure way for their storage and communication when using the cloud computing services. In addition, with the ICTs developments a new concept and approach named big data has emerged in academia and industry and most definitions highlight its increasing technological ability to capture, aggregate, and process an ever larger volume, velocity and variety of data, i.e. the 3 Vs. The main objective of data mining is to turn a large collection of data into knowledge and find hidden patterns [7]. In connection with the above mentioned the management of the stored data for later use in different processes and for different purposes is crucial to handle in a proper manner to avoid the security aspects mentioned above. It becomes, therefore, important to understand the obstacles that might impede the successful development and use of the cloud based big data for purposes of condition based maintenance (CBM) in asset management. In section 2, the current paper discusses different aspects of the big data, followed by the cloud computing and its architecture in section 3. Section 4 emphasizes the cloud computing security issues. Section 5 highlights important aspects to consider with respect to big data and the cloud computing approach in maintenance, especially when CBM is applied and finally the conclusions are presented.

2. Big data

Big data is essentially a similar analytical technique, but it differs from these earlier attempts like KDD (Knowledge Discovery in Databases), Data Mining and other numerous analytical techniques in terms of the Volume, Velocity, Variety, Veracity and other characteristics of the data. *The volume* of the data collection from the machines has increased. There are greater numbers of sensors placed on equipment that is being monitored. An increased ease of data collection and transfer by the operator with the help of hand held computers has also resulted in higher volume of data being created. Higher volume of data also means that there is a higher risk of data theft. Larger volume of data stored centrally has an amplified technical impact (entirety of data in jeopardy rather than a subset of data) and other privacy related issues [8]. *The velocity* refers to not only the speed of data being received, but also the speed at which data is being processed and analyzed. Data is being created, stored, processed and analyzed in real or near real time. What complicates the issue even further is that this data can arrive and require processing at different speeds. Data arrival at peak hours poses an opportunity to the malicious seeker to steal it. It is during such periods that organizations may lack internal capacity and tools to manage and protect information. A related point is that the attractiveness as a crime target is high during such periods [9]. While for some applications, the arrival and processing of data can be performed in batch, other analytics applications require continuous and real-time analyses, sometimes requiring immediate action upon processing of incoming data streams [10]. Less critical machines can resort to batch processing of data, whereas more critical ones like those deployed in nuclear and military applications may have to undergo real time processing of data for possible diagnosis/prognosis. This characteristic the data is referred to as variability in some literature [9]. *The variety* has to do with the large amount of historical data available regarding condition monitoring of machines. In addition, a variety of sensors capture different types of data from the machines. Some of this data is more structured; other that is picked up from the users is more unstructured. Most organizations lack capability to manage unstructured data, which arguably contains more sensitive information [9]. A large variety of information would make it more difficult to detect security breaches, react appropriately and respond to attacks (freepatentsonline.com, 2013). Data from various sources has different formats thereby making it difficult for the analysts to integrate this data. There is a need to have standard formats like Machinery Information Management Open System Alliance (MIMOSA) databases. *The veracity* is described as the trustworthiness of data. There may be instances of incorrect feeding of data by the operator or wrong inferring from the user's response. The problem assumes greater significance because of the sheer volume of data and hence difficulty in finding out the mistakes in the analysis. In addition, other characteristics can be defined in terms of exhaustive (capture entire population), fine grained resolution; uniquely indexical in identification; relational and flexible with extensionality and scalability [11].

3. The cloud computing and its architecture

The associated problem in dealing with this deluge of data was that it required higher computing resources and processing power to analyze this data to arrive at decisions. This invited higher spending in acquiring this computing power for the enterprises, which makes it economically unviable. This led to the development of cloud computing where the resources like networks, applications and servers can be hired for use. Organizations now have higher computing power available to them without having to establish and maintain such cost prohibitive infrastructure in their own premises. Industries are moving towards the Cloud due to the efficiency of services provided by the pay-per-use pattern based on the resources such as processing power used, transactions carried out, bandwidth consumed, data transferred, or storage space occupied [12]. The National Institute of Standards and Technology (NIST) channels its efforts into the standardization of the cloud computing, its definition, cloud