BIG DATA APPLIED TO CYBER-PHYSICAL LOGISTIC SYSTEMS: CONCEPTUAL MODEL AND PERSPECTIVES

Enzo Morosini Frazzon¹, Moisés Lima Dutra¹, William Barbosa Vianna¹

¹ Federal University of Santa Catarina (UFSC) – Florianópolis, SC, Brazil

Abstract

The management of supply chains and their subsystems embodies strategic and contemporaneous challenges. The competitiveness of supply chains depends increasingly on their flexibility. Logistics as a supply chain function has to compensate increasing requirements regarding flexibility by means of a higher information density and the employment of innovative techniques for data acquisition and analytics. Indeed, the flow of information that commands the physical flow embodies a critical issue for the management of supply chains. Hence, by means of a theoretical and conceptual approach, this paper aims to propose to research management perspectives and propose a model for the application of innovative techniques for data acquisition and analytics in Cyber-Physical Logistic Systems.

Keywords: cyber-physical systems, logistics, big data, business analytics, supply chains

INTRODUCTION

The application of new information, communication and sensor technologies, the capture and analysis of data as well as the proper use of resulting information can contribute for the emergence of more adaptable supply chains. For consolidating this potential is necessary to concatenate involved technologies, planning and control methods and the human-machine interaction. In fact, the availability of suitable information at the place, time and format for the decision making is fundamental for improving the efficiency of physically distributed production processes. The management of supply chains and their subsystems embodies strategic and contemporaneous challenges, involving a combination of people, hardware, software, communication networks, computation capabilities, control systems, data, policies and procedures that store, retrieve, transform and disseminate data and information (O’Brien and Marakas, 2007). Logistics can be considered the subsystem of supply chains responsible for the movement and stock of goods and raw materials from order creation till its delivery, crossing several stages such as raw materials acquisition, production, internal and external displacement.

The competitiveness of supply chains depends increasingly on their flexibility. The best supply chains are not only fast and efficient, but also agile and self-adaptable (Lee, 2004). For example, shorter development times, production of variants and individually configurable goods are suitable strategies to satisfy the customers and to react as soon as possible to changing market demands (Ritcher, 2007). Hence, logistics has to compensate increasing requirements regarding flexibility by means of a higher information density and the employment of innovative techniques for data acquisition and analytics. Indeed, the flow of information that commands the physical flow represents a critical issue for the management of supply chains (Dutra et al., 2013) and has impelled huge investments in supportive information technologies (Laseter et al., Oliver, 2005; Bandeira et Maçada, 2008).

Furthermore, the concept of a system fusing physical and information-related perspectives has been intensively argued (Broy, 2010). The capabilities of so called cyber-physical systems (CPS) enable the division of systems into entities, which are able to communicate, to recognize the environment and to make decisions. A CPS accrues the benefits of embedded systems and the possibility to communicate via a broad range of communication technologies. They allow the connection of the information...
and material flows in logistic processes (Hribernik et al., 2010). CPS combines the cybernetic aspects of computing and communication with the dynamics and physics of physical systems operating in the real world (Rajkumar, 2012). Along with the high relevance of information management and analysis for the planning and control of logistic systems, the concepts and tools of Big Data and business analytics have emerged and been investigated.

The combination of physical objects with cybernetic intelligence allowed by the application of Cyber-Physical Systems (CPS) concepts and techniques, which embrace new potential for improved efficiency, accountability, sustainability and scalability in logistic systems. In fact, they are seen by researchers as well as practitioners as one of the major drivers and challenges in industry in the future. Regarding the aforementioned integration, it is paramount that information management supports the development of suitable decision making supportive systems (Dutra et al., 2013). The capability of being aware of the existing context is one of the foundations of cyber-physical systems (Weider et al., 2013; Kannengiesser et Müller, 2013). It is possible to aggregate information of distributed systems of logistics-related data through the combination of existing systems and standards (Hans et al., 2008). By embedding CPS into communication networks, the interaction between physical and cyber world is fostered (Poovendran, 2010).

In order to contribute to a better understanding of the potential employment of big data in Cyber-Physical Systems, the present research paper addresses this issue from an application-oriented perspective. The paper aims to propose a conceptual model and to investigate promising perspectives for the application of big data acquisition and analysis techniques in Cyber-Physical Logistic Systems. The paper is organized as follows: Section 2 proposes a conceptual model for this application and Section 3 describes management perspectives and challenges.

APPLICATION CONCEPTUAL MODEL

The capabilities of CPS enable the sharing of processes in standalone and modular systems, able to communicate, recognize the context in which they live and make decisions. The generic structure of a CPS is constituted by: an embedded system; human-machine interface; connection to other systems. The embedded system includes: the sensors and actuators; electronic hardware; software. CPS combines cyber aspects of processing and communication with the dynamic aspects and physical systems (Rajkumar, 2012). The supply chain competitiveness is directly related to the physical and information flows (Frazzon, 2009).

Multiple definitions of Big data have been presented, referring to the storage and analysis of data, considering high velocity or/and complexity and special processing technics (Khouri, 2014). Waller et Fawcett (2013) present a list of applications of Big data for the principal agents in a typical supply chain: transporter, producer, re-seller. Also, they presented a list of applications according to some logistic perspectives: Inventory Management, Shipping Management, Customer and Supplier Relationship Management. In the collection of data and the use of business Analytics remains a potential to aid planning, operation and control of supply chains (Sethurman et Kunadharaju, 2013).

Regarding to the necessary hardware to enable the aforementioned approaches, we can mention the application of new communication technologies, information and positioning in distributed systems. In parallel, the adoption of these technologies and the implementation of new approaches to programming and control, leverage positive impacts on the competitiveness of industrial and service organizations.

Davenport (2009) highlights the use of RFID tags (Radio Frequency Identification) for displacement and location of the product. Many cases were reported of successful use of this technology, which costs have become more affordable, with increasingly wide applications, covering, among others, inventory management, lot tracking and pass information along the chain.

Availability of data and analyzes throughout the supply chain to manufacturers and suppliers is essential. For instance, the Collaborative Planning, Forecasting and Replenishment model (CPFR) provides joint visibility of key metrics of supply chain and iterative planning and forecasting the point of view of all participants (Davenport, 2009). As collaboration consequence, some initiatives have been boosted, for example: Vendor-managed inventory (VMI), efficient consumer response (ECR) and the collection of retail point of sale data (Fiedler, 2003). Also, wireless communication, internet of things (IoT) and sensors in networks made possible the development of advanced manufacturing systems, as well as tracking and tracing solutions in production (Schuh et al., 2007). Different formalisms are used between the components of systems. In order to allow the network synchronization and to provide a certain distributed real-time semantic, Eidson et al., (2011) proposed the principles for programming temporally integrated distributed embedded systems.

In the midst of an increasing relevance of information management for the planning and control of logistic systems, the concept of Big Data and related techniques has emerged. According to Krishnan (2013), Big Data consists of a vast amount of data available in different complexity levels, created by humans or machines at diverse paces and presenting several levels of ambiguity, so that they cannot be
computed by using traditional technologies, communication devices, processing methods, algorithms or any off-the-shelf solution. According to Katal et al., (2013), Big Data refers to a large amount of data that requires new technologies and architectures to make possible to extract value from it by capturing and analysis process.

In addition, business analytics are defined as an application of various advanced analytic techniques to data in order to answer questions or solve problems (Trkman, 2010) in a particular area of knowledge or market sector, e.g. for supporting supply chain and logistic systems management. This definition refers to the use of quantitative and predictive models and the fact-based support to managerial decision making (Davenport, 2009). The meaning of Business Analytics covers data mining, predictive analytics, applied analytics and statistics. Thus, Business Analytics meaning and definition has similarities with Big Data concept. Both search information from data and turn them into commercial advantage, as McAfee and Brynjolfsson (2012) explain. As Ward and Barker (2013) reported after analyzing the different suggestions of defining Big Data, all allude to at least the following aspects: (i) size: the volume of data is a critical factor; (ii) complexity: the structure, it behavior and permutations made between the data sets are also a critical factor; and (iii) technology: the tools and techniques used to process the data sets are essential for the concept.

However, in line with the definition given above, three aspects differentiate Big Data from Business Analytics: volume, velocity and variety. In other words, the data from Big Data and the technology to capture and process them differs from Business Analytics because of their larger values in these three aspects. Furthermore, in some cases, the speed of data creation is even more important than the volume. Real-time or nearly real-time information allow a company to have valuable competitive advantages (McAfee et al., 2012). In regard to a Big Data typology, it exists in numerous forms and possesses different characteristics (Figure 1).

Data can come in the form of updates in social networks (Facebook, Instagram), remote sensing maps (Google Maps, GPS), photos (Flickr), videos (YouTube), Internet search tools (Google) and so on. These forms of data are collectively known as unstructured data. Each person today is potentially a walking data generator. Most traditional technologies, structured databases, are not well suited for storing and processing Big Data. New technologies (storage, memory, processing and bandwidth) have emerged to reduce the costs of intensive approaches to data analysis (Arellano, 2013). Thus, the development of virtualization processes and distribution of products / services through collaborative networking requires, process integration and semantic interoperability of data and information, as well as the use of ontologies for facilitating such integration. There is a number of tasks that are potentially benefiting from the use of ontologies in the context of logistics integration in a supply chain management. Formally representing a process, product or service means allowing a fundamental step in order to remove ambiguity and to improve communication between different work involved groups. Moreover, a good ontological representation allows the inference of information that is not explicit at first, which makes it easy to identify any conflicting or inconsistent situations in the context of systemic relations. With a long tradition in the field of philosophy, ontologies provide support for knowledge engineering and for artificial intelligence regarding the modeling of knowledge domains in terms of concepts, attributes and relationships, generally classified into hierarchical relations of the type specialization/generalization (Noy et Mc Guinness, 2001; Simperl, 2009).

Modern organizations that are involved in the design and development of new products and services should adopt flexible methods of working to meet the numerous and varied demands of the global marketplace, if they intend to remain competitive. This flexibility should not only comprise the ability of offering good answers to their customers, but also the capability of detecting potential changes and future trends within the whole system. Considering the scenario of supply chains, this strategy means to combine and coordinate all involved system actors – humans and not-humans –, from raw materials production to final product delivery.

This combination relies on the complex process of integrating three essential problem dimensions: computation, control, and communication (Conti, 2012). Integrating these dimensions means considering the different participants and their viewpoints, different knowledge areas, different network topologies and equipment, different requirements, and different stages of global and local activity. Once this process is successfully set up it might produce a more agile and self–adaptable system.
The proposed model (Figure 2) works both reactively and proactively, since it processes not only feedback data but also forecast and trend information. In order to keep the system sustainable and scalable, it is considered that information sharing is a fundamental attribute, which provides human and non-human actors (e.g. software agents) with pre-processed data inputs, enabling them to make better decisions. While the feedback data will permit adjustments, corrections, and updates to the whole supply chain, forecast trends and future insights – reasoned out and produced through mechanical and non-mechanical techniques –, will be the basis for organizational actions towards innovation, anticipation, and value-creating strategies, which envisage to gain competitive advantage by outperforming its competitors. To accomplish such a vision, BigLogDatix will benefit from Big Data techniques and Data Analytics.

We highlight the efforts of BigLogDatix model to combine the computation, control, and communication aspects of Big Data, as the integration of these three domains is one of the main challenges in CPS research (Conti, 2012). BigLogDatix is structured upon distributed-networked modules. Each BigLogDatix module is instantiated for a specific supply chain scenario. In other words, the modules are customized for different logistic players, such as the shop floor control, the customer relationship manager and business analytics team.

Albeit the circumstances may be different, the methods and techniques used are quite similar. A Data Warehouse is proposed for integrating collected data from different sources. This central repository will be populated through an ETL (extract, transform, load) process, which transforms the extracted data from its previous form into the BigLogDatix form. Besides, we propose an ontological representation of information, in order to enable a reasoning process of knowledge. The outcome of this process feeds a Knowledge Base, which is used for discovering and recognizing tacit knowledge. Bohn and Short (2009) says that “There are many potential criteria for measuring the value of a stream of information, including subjective judgment, selling price, willingness to pay by consumers, development cost, and audience size. But there is no clear way of comparing value, especially when comparing information of different kinds”. Considering this assertion, we propose an ontological inference-engine to reason out and compare information value.

In this way, the conceptual model BigLogDatix intends to support the decision-making process in Big Data context. The application model is based on epistemological and formal aspects that need to be explained. In the case of BigLogDatix we consider that the problems do not have only a physical existence, but they depend on the intervention of one or several human and non-human actors that interact in communication, control and computing processes in the decision-making along Cyber-Physical logistic systems. In
order to put together the several actors and subsystems involved in the context of logistic systems and supply chains, we consider the BigLogDatix environment (Figure 3).

Furthermore, the model considers the limitation of objectivity adoption in decision-making processes and the need for interaction between the objective and subjective elements. Thus it becomes impossible to deny the importance of subjective factors and leave them aside in an attempt to use an entirely objective approach (Roy et Vanderpooten, 1996). The general model also considers the interpenetration of objective and subjective elements and their inseparability, since a decision-making process is a system of relations between elements of objective nature related to actions, and elements of a subjective nature related to the value system of actors.

Although the search for objectivity is a major concern, it is also crucial to consider that decision-making is first of all a human activity, sustained on the concept of value. The subjectivity is ubiquitous, so that the model needs to support such condition (Bana et Costa, 1993). For the case of BigLogDatix, the “decision context” can be understood as an abstraction of a given set of real-world events or as a conceptual entity that is perceived by the relevant group of actors involved, therefore the way the real world events are perceived by the involved actors lead to a particular perception of this decision context (Oral et Kettani, 1993).

DISCUSSION

Big Data definition is typically subjective. In order to be recognized as such, it incorporates a moving definition of how big a dataset needs to be. According to Manyika et al., (2011), “Big data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze.” We share the assertion of O’Neil (2013), who argues that people doing business analytics on a larger and larger scale are not actually performing big data. A true big data setup should mean that the human factor has been put outside the machine and the machine is doing the job.

However, we must take into account that Big Data by itself is not going to be the answer to all our problems. Some drawbacks must be pointed out, since this paradigm clearly has limitations. Lohr (2012) states that “Big Data has its perils. With huge data sets and fine-grained measurement, statisticians and computer scientists note there is increased risk of ‘false discoveries’.” Big Data “supplies more raw material for statistical shenanigans and biased fact-finding excursions. It offers a high-tech twist on an old trick”. In other words, still according to Lohr (2012), “A model might spot a correlation and draw a statistical inference that is unfair or discriminatory, based on online searches, affecting
the products, bank loans and health insurance a person is offered, privacy advocates warn”. In order to avoid such a scenario and to mitigate eventual issues, BigLogDatix combines both Big Data extraction techniques and Business Analytics people into one combined conceptual framework.

There are many technical challenges in the use of Big Data, mainly related with the use and management of vast volumes of raw data (McAfee and Brynjolfsson, 2012). Furthermore, management is still not prepared to deal with such a large variety and quantity of information. The awareness and identification of benefits that a good and meaningful analysis of these data could bring to business competitiveness is still starting. The proper appropriation of this competitive advantage is a challenge that does not involve only changes in technology, but adaptation of processes to the change of management and data analysis.

It is necessary to consider that, when the volume of available data is relatively small, difficult-to-obtain and/or not available in digital format, it is usual that people at high levels of organizational hierarchy take decisions based on their experience, what they usually do taking into account patterns of relationships that they have internalized throughout their careers (McAfee et Brynjolfsson, 2012).

In this case, the focus is directed to the specific interested in solving the problem, thus the interpretations given by other actors involved in the problem are considered secondary and partial ones (Landry et al., 1985).

On the other hand, challenges also arise when the data set size is beyond the software tools ability to capture, store, manage, and analyze typical databases (Manyika et al., 2011). According to Manyika et al., (2011), “in some cases, decisions will not necessarily be automated but augmented by analyzing huge, entire datasets using big data techniques and technologies rather than just smaller samples that individuals with spreadsheets can handle and understand”. In this sense, the decision-making process is strategic to deal with large volumes of data, analyze the information and determine how working with this volume and variety of information at proper speed. Organizations will not reap the benefits of a transition with the use of Big Data, unless they are able to manage the change effectively.

We consider being on the brink of a huge change in the way decisions are made inside the organizations. We may be witnessing a time where real-time micro-segmentation of citizens and costumers will reach its zenith, through the evolution and use of Big Data techniques. Sophisticated analytics strategies can substantially improve decision making, minimize risks, and unfold important information, which would otherwise remain unknown.

ACKNOWLEDGEMENTS

This work was partially supported by CNPq – The Brazilian Council for Research and Scientific Development – in the scope of call MCTI/CNPq/MEC/CAPES – 43/2013.

REFERENCES


