ISSN 2959-6130

The role of big data technology in project schedule management: the case of Amazon

Rui Liu

Department of Accounting, Beijing Normal University - Hong Kong Baptist University United International College, Zhuhai, China
*Corresponding author: r130006111@mail.uic.edu.cn

Abstract:

With the rapid development of the Internet, big data technology has received widespread attention in the field of management. Traditional project schedule management methods cannot adapt to the complex and changing environment, and some researchers believe that big data technology has become a key tool to improve and optimize project management. However, there is a lack of mature theoretical systems and practical experience to guide the application of big data technology. Based on the case of Amazon, this paper analyses how the company uses big data technology to optimize project schedule management. It is found that the application of big data technology in project management brings significant benefits. It is able to provide empirical and innovative support for project schedule management. This paper discusses the application of big data technology in project schedule management, highlighting its capability to enable timely adjustments. Real-time monitoring and forecasting facilitate proactive management of project directions and potential risks. Moreover, it leverages real-time data to optimize resource allocation and enhance resource utilization. Furthermore, the paper addresses the challenges and proposes countermeasures associated with big data technology in this context. To tackle the issue of uneven data quality, a robust data collection mechanism should be established.

Keywords: Project schedule management; big data technology; Amazon.

1. Introduction

With the development of the Internet, the data and information in the society are becoming more and more complicated. All kinds of data are constantly being generated and stored, penetrating into various industries and jobs [1]. Massive amounts of data form large and complex datasets that are often impossible to obtain and process efficiently by traditional data tools, known as Big Data [1]. It has the following five characteristics. Volume: The huge volume of data is the most intuitive characteristic of Big Data, and the volume of data is often measured in terabytes or even petabytes [2]. Velocity: The data is generated and transmitted at a very high speed (e.g., real-time data streaming) [2]. Variety: The data types are diverse, including structured data (e.g., database tables), semistructured data (e.g., JSON, XML), unstructured data (e.g., data in the form of tables), and unstructured data (e.g., images, videos) [2]. Value: Big Data contains a lot of useless information and therefore has a low-value density, which needs to be refined and extracted from the useful information [2]. Veracity: the authenticity of the data is also an important characteristic of Big Data, which relates to the quality and trustworthiness of the data [2]. The emergence of big data as a phenomenon raises the need for new types of data processing technology, and big data technology is a concrete solution to meet this need. As a result, Big Data technologies have gained increasing attention in the broad field of management. Managers and business leaders try to utilize big data technologies in the field of business and management. They specialize all the data in the dataset for the purpose of acquiring, storing and analyzing the data within a reasonable time. It is also being used to organize the cluttered data into effective information that helps in business decision-making and adds value to the data. Many studies have concluded that big data analysis has become a key factor for modern enterprises to improve their core competitiveness [3].

Project schedule management is a specialized subset of project management, focusing on tracking and controlling the progress of each phase throughout a project's lifecycle. This includes the detailed planning, execution, and monitoring of the project timeline. Effective project schedule

management is essential to ensure timely project completion and optimal resource allocation [4]. Traditional project management approaches rely heavily on experience and static data. As the complexity of projects increases, the limitations of this approach in coping with changes gradually become apparent. Combined with these characteristics, such problems are likely to be effectively solved if big data is reasonably used for project management. Big data technology is a method based on the characteristics of big data for data processing. It is able to feed real-time data through cloud platforms and so on, and at the same time carry out analyses and predictions. It is also capable of dynamically adjusting recommendations as real-time data changes, thus enabling intelligent decision-making and dynamic and refined management of data. Therefore, the significant effect of big data technology to cope with changes is important for optimizing project schedule management.

Unfortunately, although a part of scholars believe that the advantages of big data technology can bring help to project schedule management, the theoretical research on big data technology in project schedule management is relatively small. There is a lack of a mature theoretical system to guide practice. Moreover, there are few application cases of big data technology in project schedule management. Insufficient real-world experience exists to support the growth and advancement of the idea. This study aims to explore how big data technology can optimize project schedule management and provide theoretical and practical references for the application of big data technology in project schedule management. Through the combination of case study and literature research, this study will take Amazon as an example to deeply explore the application and effect of big data in project schedule management in Amazon.

2. Case analysis

Amazon is the world's largest e-commerce and cloud computing company with a huge logistics network. The complexity of the business brings many challenges to Amazon, such as the dynamic changes in the logistics network, the complexity of inventory management, and the rapid changes in customer demand [5]. These factors put higher demands on Amazon's project schedule management, requiring Amazon to be able to monitor and adjust the project schedule in real-time.

In recent years, Amazon has used big data technologies to optimize its project schedule management, including logistics optimization, market forecasting and product recommendation. They use big data technology to collect and analyze massive amounts of data on orders, inventory, shipping and customer behavior, managing and analyzing the data. In addition, they forecast order volumes, optimize inventory management and dynamically adjust delivery routes based on the results of data analysis. The ability of big data technology to process information quickly provides technical support for obtaining low-latency data and dealing with emergencies. It realizes real-time monitoring and dynamic adjustment of project progress, successfully coping with various challenges brought about by business complexity, which significantly improves the operational efficiency of the enterprise. The case of Amazon has a high reference value for the progress management of other projects.

3. Specific applications of big data technologies in project schedule management

3.1 Data mining, collection and storage

Amazon uses a variety of means to collect data. One of them is the big data technology provided by Amazon Web Services (AWS), the cloud technology branch of Amazon. AWS provides Amazon S3 and Amazon EBS technologies that can be used to mine, collect and store massive amounts of data. They enable Amazon to collect data not only within the platform, but also from external sources and partners, providing the underlying data to support Amazon's project progress management. Amazon S3 is a scalable, secure and cost-effective big data technology. It meets the company's storage needs for data of all sizes. Amazon EBS is easy to use and supports the storage of high-performance data at any scale, making it ideal for companies that require frequent read-and-write operations. For data analysis and processing, AWS introduced Amazon EMR and Amazon Redshif Cloud Big Data Platform. Amazon EMR is the industry-leading big data platform for when Amazon processes data using open-source tools such as Hadoop, Spark, and HBase. It simplifies big data processing and facilitates users to process and analyze large-scale datasets. Amazon Redshift, on the other hand, is a fully managed petabyte-scale data warehouse that allows companies to analyze data with ease as it enables automated provisioning of resources.

These big data technologies support Amazon's project schedule management with experience and innovation. These types of big data technologies store data information of the whole project lifecycle. At the end of the project, it provides comprehensive and accurate review information for the project progress summary. The project team can therefore deeply analyze the reasons for the success or failure of the project in terms of schedule

management, from which the experience can be extracted to provide reference for future management activities. In addition, Amazon generates many best practices, lessons learned, and technical solutions during each project. Big data technology systematically organizes and stores these valuable knowledge assets to form a project knowledge base within Amazon. The knowledge base facilitates knowledge sharing and learning within the project team and provides strong empirical support for project schedule management. For areas that have never been attempted before, Amazon uses external data acquired and stored by big data technology to provide guidance to project teams based on the cases of others. This can significantly reduce the number of failures to manage a brand-new project schedule and encourage innovation within the organization.

3.2 Real-time monitoring and forecasting

The big data technology used by Amazon enables 24/7 tracking and analysis of the progress of the project. The company mainly uses data from dashboards and notifications from early warning systems to identify problems in a timely manner and resolve them at a manageable stage to ensure that the project progresses as per the plan. Amazon Kinesis Data Analytics is one of the key tools that has powerful real-time processing capabilities [6]. It handles the latency of streaming data and ensures that incoming data is analyzed and responded to in real time, making it an easy way to respond to business and customer needs in a timely manner. The tool's in-built visualization capabilities enable project progress to be displayed in real-time in dashboards. For example, resource utilization, cost runs and schedule adherence. Team members and project managers can view project updates at any time. The tool also compares real-time data with historical data to assess whether project schedule performance is improving or declining. The early warning system monitors the project based on set thresholds. In the event of an anomaly or deviation from the plan for a particular piece of data, or a significant drop in project performance, an alert is automatically sent to notify the project manager of the specific situation. This real-time data feedback and early warning system notification help team members adjust their work plans and event priorities in a timely manner. They enable project managers to identify problems in advance and respond quickly to avoid project delays.

Additionally, Amazon mines historical project data through big data technology to discover key factors and potential risks affecting project progress. Combined with the progress of current projects, big data technology can predict possible future risk situations and develop appropriate response strategies in advance. For example, by analyzing data from historical purchasing projects, Amazon can predict which products may experience inventory backlogs or shortages, and thus adjust the schedule of purchasing projects in advance. By analyzing forecasts of market trends, Amazon can predict which products are likely to experience sales declines and thus adjust the course of marketing projects in advance. This allows Amazon to reduce the likelihood of project delays and helps Amazon's project management team to efficiently manage. So that the project team can analyze and utilize data to make scientific decisions and improve business decisions and market competitiveness [7].

3.3 Real-time data and resource allocation

Big data technology provides real-time data to Amazon. Real-time data increases the flexibility of resource allocation and resource utilization, solving problems such as insufficient manpower, shortage of supplies or misallocation of funds [8]. Based on real-time data, Amazon can adjust the use of human, material and financial resources to cope with the rapid changes in the market and the complexity of internal operations, to ensure that the progress of the project continues to progress [8]. Business data and market information are changing all the time. The real-time data provided by big data technology is a safety net to ensure that Amazon adapts to the dynamic market and maintains its competitive advantage in project schedule management. The project manager and team will assess the progress and effectiveness of the project based on the real-time data collection of processed and up-to-date business data and market information.

Before the project even begins, Amazon's project managers can plan for staffing and time based on the optimal configuration analyzed from big data. As the project progresses and the details of the project change, Amazon is able to accurately identify the resource bottlenecks that currently exist in the project with the help of real-time data. For example, by analyzing the efficiency of employees through task completion, project managers can identify the need for staff additions or training enhancements for specific positions to ensure that the number and efficiency of employees in the project reach the highest level. For logistics resources, based on real-time data to analyze the usage rate of transport means and distribution efficiency, the project manager can adjust the fleet size and distribution strategy in time to avoid waste of resources and operational inefficiencies. For warehouse resources, according to the real-time inventory quantity and future sales of the warehouse, the inventory cycle is analyzed. Then arrange a reasonable time to increase the inventory to ensure that the warehouse always controls the optimal number of inventory, and makes the best use of the in-

ventory resources. They make timely adjustments when they find that the original plan is no longer applicable or that a more efficient execution path exists. For example, during peak sales periods, Amazon can dynamically adjust warehouse staff and transport truck configurations based on real-time data on order volumes. Real-time data makes resource allocation more flexible, ensuring stable service quality while also ensuring that the project is carried out in an orderly manner, which reduces the irrational allocation of resources. Project managers can therefore make scientific decisions and control the whole process of the project.

4. Challenges and Countermeasures

Amazon, as the world's largest e-commerce platform, also has self-developed and put into use big data technology, making it a leader in modern project schedule management. Accurate data collection and efficient data analysis are the key factors to achieving project success. However, because of the nature and technical difficulty of big data, Amazon still has some challenges in its application of big data. Continuing to improve and optimize project schedule management methods is an effective way for the company to maintain its competitive advantage.

4.1 Establish a perfect data collection mechanism

Big data itself has certain disadvantages, and the variable quality of data is one of them [9]. Amazon carries out project management and collects information from a variety of channels, and it is very likely that inaccurate or even false data, or data with low relevance, will be mixed in the huge amount of information. The information processed using such data is likely to be misleading and unreliable, making the project manager make wrong decisions and thus affecting the whole project process. Cleaning and pre-processing such data will cost a lot of time and money, and the project schedule will be forced to lengthen.

In order to ensure the effectiveness of using big data for project schedule management, Amazon needs to improve its data collection mechanism. This means that the data collected at all stages and from all angles of the project needs to be comprehensive and accurate. One possible solution is for the company to establish a central data warehouse in which data from different departments and systems are stored and managed. This not only improves data consistency and reliability, but also facilitates cross-departmental data integration and analysis. Another option is that Amazon could develop a standard set of data collection processes and guidelines to ensure that all relevant teams collect and submit data according to a uniform methodology. Standardize the definition, format, source and frequency of data collection. By standardizing operations, data errors and omissions can be reduced and the overall quality of data can be improved.

4.2 Enhancing data security and privacy protection

Big data technology has its own logical flaws, which can lead to threats to data security and personal privacy [10]. The principle of the application of big data technology is to collect and mine a large amount of data that records the browsing track, the number of browsing times, the time spent on the web page and responses. Seemingly single data are integrated and combed, analyzed, and by taking their overlapping parts they are able to get some sensitive data, identity information, age, and home address. This means that the security and privacy protection of big data is not only related to the security of information within the enterprise, but also involves the confidentiality of customer data. Given its vast repository of user data, Amazon faces significant risks beyond mere technical issues if a data breach occurs. Such an event could lead to legal consequences and adverse public opinion, highlighting the broader implications of data security for the company.

Therefore, in the process of using big data, Amazon needs to regard data security. The organization should privacy protection as the highest priority, adopting multi-level security measures. Using multi-layer encryption of data, access control and network security protection is also a valuable choice to ensure the compliance and security of data use. Regular data security training and drills are conducted for employees to improve their data security awareness and reduce the risk of data leakage due to operational errors.

5. Conclusion

This study analyzed the case of Amazon and found that big data technologies can be used to optimize project schedule management in three areas: 1. data mining, collection and storage 2. real-time monitoring and forecasting 3. real-time data and resource allocation. The data mined, collected and stored by big data technologies can provide past experience and form a knowledge base for project schedule management. It can also secure innovations in project schedule management from other organizations' data. The real-time tracking and prediction of big data technology provides a reference for project schedule adjustment and avoids project risks. Real-time data based on big data technology can rationalize resource allocation and improve resource utilization. However, the use of big data technology still has some challenges, including data

quality and personal privacy security. Challenges can be addressed by establishing a sound data collection mechanism and adopting multi-layered security measures to reduce the risk of project delays and data leakage.

As can be seen from the case of Amazon, big data technology is effective in the field of project schedule management and has great potential and broad application prospects. With the development of information technology and the improvement of data processing capability, more and more enterprises and organizations hope to use big data to improve project management. This paper summarises some experiences of applying big data technology to project schedule management based on Amazon's practice cases, which promotes the development and improvement of the theory. It fills some of the gaps in the field of project schedule management. There is some reference value for how enterprises use big data technology to optimize project schedule management.

This paper also has certain limitations, mainly because of the lack of research methods. Because there are fewer actual cases of project schedule management using big data technology, most of this paper is based on the analysis of a typical case, Amazon. As the case studies focus on specific organizations, the conclusions may not be applicable to project schedule management in all industries. In order to reduce the limitations of the study, future research can explore the specific application of big data technology in different types of projects through multiple cases. It can also combine data analysis to draw more general conclusions to ensure the objectivity of the study and increase the reliability of the results.

References

[1] Sandhu A K. Big data with cloud computing: Discussions and challenges[J]. Big Data Mining and Analytics, 2021, 5(1): 32-40.

[2] Mayer-Schönberger V, Cukier K. Big data: A revolution that will transform how we live, work, and think[M]. Houghton Mifflin Harcourt, 2013.

[3] Maroufkhani P, Wan Ismail W K, Ghobakhloo M. Big data analytics adoption model for small and medium enterprises[J]. Journal of Science and Technology Policy Management, 2020, 11(4): 483-513.

[4] Oburu A O. Effective project time management[J]. International Academic Journal of Information Sciences and Project Management, 2020, 3(6): 47-55.

[5] Archetti C, Bertazzi L. Recent challenges in Routing and Inventory Routing: E-commerce and last-mile delivery[J]. Networks, 2021, 77(2): 255-268.

[6] Tantalaki N, Souravlas S, Roumeliotis M. A review on big data real-time stream processing and its scheduling techniques[J]. International Journal of Parallel, Emergent and Distributed Systems, 2020, 35(5): 571-601.

[7] Wijayasekera S C, Hussain S A, Paudel A, et al. Data analytics and artificial intelligence in the complex environment of megaprojects: Implications for practitioners and project organizing theory[J]. Project Management Journal, 2022, 53(5): 485-500.

[8] Wang J, Liu Y, Ren S, et al. Evolutionary game based real-time scheduling for energy-efficient distributed and flexible job shop[J]. Journal of Cleaner Production, 2021, 293: 126093.

[9] Wu J, Wang J, Nicholas S, et al. Application of big data technology for COVID-19 prevention and control in China: lessons and recommendations[J]. Journal of medical Internet research, 2020, 22(10): e21980.

[10] Yang P, Xiong N, Ren J. Data security and privacy protection for cloud storage: A survey[J]. Ieee Access, 2020, 8: 131723-131740.