Overview of Big Data Based Space Situational Awareness

Weigui Zhou[1,a], Hong Ao[1,b], Quan Zhou[1,c], Yuan Gao[1,3,d] and Yi Li[2,e,*]

[1]Xichang Satellite Launch Center, Sichuan, China

[2] The High School Affiliated to Renmin University of China, Beijing, China

[3] China Defense Science and Technology Information Center, Beijing, China

[a]wgzhou@nudt.edu.cn, [b]aohong76@aliyun.com, [c]zhouquanxslc@foxmail.com
[d]yuangao08@gmail.com, [e]liyi@rdfz.cn, *Corresponding Author

Keywords: Space, Situational Awareness, Big Data, Architecture

Abstract. Space situational awareness is the development of space resources, to protect its space assets and space to master control over the implementation of key military space operations. US government and military decision-makers aware of the importance of capacity building for space situational awareness. From the Cold War began to build space surveillance system. With the development of science and technology and their space surveillance systems become more advanced and complete, the development of a complete system of space situational awareness. In this paper, combined with China's space situational awareness Situation research project, planned by big data tools, built on space situational intelligence dynamic sensing system big data platform to provide space situational analysis, fusion prediction function, and based on the perception of network sensors phase combined to provide accurate information space, to provide assistance and services to national strategic planning.

Introduction

The rapid development of information society, data has become an important production factor in all walks of life, people and the use of massive data mining, heralded the emergence and development of new technologies. In the economic sphere, large data help finance economic restructuring, promote the upgrading of information consumption; in the field of culture, big data to promote the integration and sharing of cultural information resources; in the military field, data-led military decisions became key military management. In March 2012, Obama will be "big data strategy," [1] a national strategy that big data is "future new oil." In the "era of big data," a book that thinking companies use big data to discover the existence of the data potential commercial value and new opportunities. How to effectively use mass data big data environments, forecast analysis for a particular field of technology development, new opportunities and new challenges of the current information science research methods in the field of application technology forecasting facing.

This paper attempts to answer the research questions the following aspects: ① what is the next big data environment prediction technology is characterized ② big data environment how to use research methods of information science technology to predict ③ Combination of low-carbon technologies, from how big data analysis development environment frontier and hot topics in the field. Through predictive analytics technology to provide a new perspective for the next big data environment intelligence analysis method in the technical field of research questions above answers and practical applications, and analyze large data space field for our low-carbon technologies catch play a positive role [2-4].

Characteristic of Big Data Based Space Situational Awareness

Big data, suggests that large-scale data sets. Big Data is the core technology to establish prediction method based on correlation analysis. 2013 Professor Desouza think big data having Volume (metric), Velocity (speed), Variety (breadth), Viscosity (viscosity), Variability (variability) and Veracity

(reliability) characteristics. Technology forecast big data environments will combine a variety of data sources, the use of new information technologies and information resources, thereby providing predictive analytics technology more intelligent. Large data environment technology forecasting the following features.

Integrated Handle a Variety of Data Sources. Technology forecast big data environment requires utilization of scientific papers, patents, special reports, television, books, and other statistical data, and different data sources, to reveal a field of technology development in the context of the evolution from different angles, find the technical field of Science and Technology frontier research focus and direction. At the same time, due to the characteristics of different types of information it also shows the need for mutual supplement between them. Therefore, technology forecasting need to use effective tools Comprehensive information about a variety of data sources, insight into the problems in the technical field, in order to obtain valuable data while making the entire application technology to improve forecasting techniques.

The Use of New Information Technologies. Large data environments, in the face of massive data, automation technology is essential. Technology forecasting data source retrieval more support is the use of new technologies retrieval system, rather than manual information collection, fusion more dependent data warehouse and data mining technology information needs the cloud, big data parallel processor (MapReduce) and other technologies to increase the data processing speed. Meanwhile, in the analysis phase, through the "public cloud" and "hybrid cloud" technology, technical development trend analysis to predict.

The Use of Intelligent Processing Method. Predict the next big technical background data, the need for more intelligent processing requirements, the need for more intelligent application of advanced technology, complex information processing and analysis to help intelligence researchers freed from the complex manual labor. In the complex big data environment, technology forecasting data may come from sources such as Twitter or Facebook social media, it may come from it, and the Dow Jones stock market or market data, and therefore the use of intelligent processing methods to improve data collection effectiveness.

Method of Big Data Based Prediction

Technology forecast big data environments can be integrated using different methods of analysis and prediction and Information Sciences. Among them, there are more commonly used visual analysis, cluster analysis, citation analysis and market forecasting method [5].

Aerospace Big Data technology applications, NASA has been at the forefront of the world, the establishment of open-source Integrated Cloud Services Platform Nebula cloud computing platform iRODS, application MapReduce model processing lunar photographs Hadoop cluster, elastic storage and parallel processing Mars orbit task aerospace and other large data flood. Big Data and cloud machine manufacturing aerospace hospital in our country in the space data archives are large areas of aerospace data and cloud computing professional useful exploration.

Space for the characteristics and processing requirements of large data, combined with the advantages of a distributed file system, distributed databases, in-memory computing, hybrid computing model and MapReduce parallel framework and other large data processing technique, the applicability of Big Data technology in the aerospace field [6].

Distributed File System. A distributed file system to store aerospace unstructured data, is data stored on multiple storage nodes physically dispersed, these resources unified management and allocation, and provide users with access to the file system interface. It can solve the local file system in file size, number of files and other documents limited number of open questions.

Distributed Database. Distributed database is a database technology and network technology product of the combination, which is characterized by physical and logical whole dispersion, with a flexible architecture that can increase the data redundancy, high reliability, availability, and can be extended outburst and Ease of integration with existing systems. It can make existing structured data space shine with new vitality. Distributed database including transactional databases and analytical

database. The former includes key management system: Big Table and HBASE; document management systems: Mongo DB and Couch DB; Fig database: Neo4j;-based database memory: SQLFire and VoltDB like. The latter include Hive and HAEQ like.

Memory Computing. Space is the form most big data stream for transmission, in some systems, the speed of data flow is very fast, real-time processing convection and low latency has strict requirements. With the processor to reduce the cost of memory, the memory computing will become the most effective means to an end across large data computing performance barriers, high real-time, low-latency computing. A typical memory system calculates are: SAP's Harm-memory database, Microsoft map data computing system Trinity, UC Berkeley AMP Lab of Spark and the like.

Hybrid Computing Model. Currently, Hadoop's MapReduce parallel computing model is suitable for a large batch of data, it is difficult to adapt to large data aerospace complex data relationships and complex computing model. But Hadoop can still exist with the way other computing model mixed. For example, Hadoop clusters, MPI cluster, CPU + GPU heterogeneous, CPU + MIC and heterogeneous high performance computing architecture combines in one system can meet the needs of a large diversity of aerospace data analysis.

MapReduce Parallel Framework. Traditional Space For large data handler to the Hadoop cluster Java-based MapReduce architecture upgrade, the developer put forward higher requirements. However, the Hadoop Streaming feature that allows non-Java executable code to run in parallel on a Hadoop cluster, just to control the relationship between processes, can aid Hadoop parallel architecture of the traditional program accelerated.

Map Data Parallel Processing. MH370 lost contact event, look for the wreckage of suspected result was published always lags behind the shooting time of 1 to 2 days. This indicates that an object recognition technology in the large number of images need to achieve breakthrough performance. Strong correlation map data, a distributed processing system in FIG storage key question is "graph partition", that is under effective graph partitioning strategy, large-scale distribution of map data stored in different nodes, and each on the local sub-graph node parallel processing. This is similar to the concept of task parallelism and data parallelism. Now there are a lot of distributed computing system diagram, which is more typical of the system Google's Pregel, Facebook's open source implementation of the Pregel Giraph, Microsoft Trinity, and Power Graph.

Data Security. Space data security is a complex and important issue, military space key data encryption and decryption usually transmitted. The big data systems are typically distributed network system, though it has the advantage of efficient sharing of data, but in terms of data security, but there are some risks. Features large data encryption and decryption mode is bound to bring new challenges, such as the possible options, the use of low-value data density of data features, insert a small amount of cipher text in the same frame image continuous transmission, resulting in improved data safety.

Space TT & C System Platform Cloud. World Space Systems Interconnection big trend is the development of space technology. Currently, based on virtualization technology and Cloud Stack OpenStack cloud platform that can integrate IT distribution of resources for the realization of on-demand and self-controlled cloud platform aerospace systems preferred embodiment provides technical support.

The development of space big data professional does not follow the "big data" concept go, always concerned about the collection, storage, analysis and visualization method for large-scale data, to continue to accumulate, implementation techniques and mindset breakthrough, in order to make better use of advanced information means of obtaining high-value data [7][8].

Summary

Space is the cause of human exploration of space, planets and even galaxies build system-level simulation model or simulation is an important means of space exploration and understanding. Big Data technologies could help space agencies calmly used to build such a large mass of data analog system, and improve the way knowledge representation, helping scientists to visually verify the

theory, deduction changes, expanding human knowledge of the universe. NASA Big Data Project: Visualization solar system simulator can be used as a reference.

The main carriers include cloud computing, including big data technology, as a new generation of information technology, has not yet fully mature, the market is not the monopoly of multinational companies, which is conducive to domestic cloud computing technology and solution design and promotion programs. At present, cloud computing standards have not yet introduced, there is no technical monopoly. These are the external environment for the aerospace Big Data technologies to develop self-control direction and opportunities.

References

[1] Tübke A, Ducatel K, Gavigan J, et al. Strategic policy intelligence: Current trends, the state of play and perspectives[C]. IPTS, Seville, 2001.

[2] Intepe G, Bozdag E, Koc T. The selection of technology forecasting method using a multi-criteria interval-valued intuitionistic fuzzy group decision making approach[J]. Computers & Industrial Engineering, 2013, 65(2): 277-285.

[3] Eto H. The suitability of technology forecasting/foresight methods for decision systems and strategy: A Japanese view[J]. Technological Forecasting and Social Change, 2003, 70(3): 231-249.

[4] Fye S R, Charbonneau S M, Hay J W, et al. An examination of factors affecting accuracy in technology forecasts [J].Technological Forecasting and Social Change, 2013, 80(6): 1222-1231.

[5] Mahr A, Katsahian S, Varet H, et al. Revisiting the classification of clinical phenotypes of anti-neutrophil cytoplasmic antibody-associated vasculitis: a cluster analysis[J]. Annals of the rheumatic diseases, 2013, 72(6): 1003-1010.

[6] Nwangburuka C C, Kehinde O B, Ojo D K, et al. Morphological classification of genetic diversity in cultivated okra, Abelmoschus esculentus (L) Moench using principal component analysis (PCA) and single linkage cluster analysis (SLCA) [J]. African Journal of Biotechnology, 2013, 10(54): 11165-11172.

[7] Bornmann L. How to analyze percentile citation impact data meaningfully in bibliometric: The statistical analysis of distributions, percentile rank classes, and top‐cited papers[J]. Journal of the American Society for Information Science and Technology, 2013, 64(3): 587-595.

[8] Halal W E. Forecasting the Technology Revolution: Results and Learnings from the TechCast Project[J].Technological Forecasting and Social Change, 2013,80(8): 1635-1643.