APPLICATION OF BIG DATA SYSTEMS TO AVIATION AND AEROSPACE FIELDS; PERTINENT HUMAN FACTORS CONSIDERATIONS

Chang-Geun Oh, Ph.D.
Kent State University,
Kent, Ohio 44242

The aviation and aerospace are typical areas that can apply big data systems due to their scales. This paper identifies aviation/aerospace areas that can utilize big data infrastructures to enhance their operational performances, and lightens human factors considerations related to the use of big data. The NextGen's network-centric infrastructure defines sharing a huge amount of aeronautics, flight, and weather data under the system wide information management program. Sensors installed on aircraft components extract huge numbers of aircraft health and operational status data. All professionals who work in the different aviation sectors require this shared situational awareness information for their own distinctive purposes, and big data systems will enable the effective use of the information. The improved prediction model by the big data analytics will improve aviation safety, reduce flight delays, and save the time and cost for maintenance. The pilot behavior research can adopt the naturalistic study method to supplement limitations of simulation test. The naturalistic flying study needs to consider collecting and analyzing data through big data systems. Human factors research questions naturally arise as aviation/aerospace fields apply big data systems pervasively.

Introduction

The aviation field has encountered a drastic growth of air traffic demand and required the effective management of aviation systems. It is going to be more difficult to ensure the safety for passenger and cargo. Recently the data acquisition collected from aviation infrastructures (satellite systems, ground stations, and airport radar systems) and sensors installed on aircraft became to be shared to most people who work in various aviation fields or customers who use airports or other aviation services. The volume of data extracted from these systems and sensors is too big to handle using the traditional computing capabilities with databases. For example, the average flight data collected during a current flight operation is up to 1000 gigabytes (Wholey, Deabler, & Whitfield, 2014). This "big data" is considered as one solution to enhance the aviation safety for the increased traffic volume and produce higher revenues for airlines.

The big data is referred to a huge volume of data that cannot be managed by the traditional data management paradigm (Schroeck, Shockley, Smart, Romero-Morales, & Tufano, 2012). The structured and unstructured data in non-unified format collected from various machines and sensors can be stored and utilized to discover new correlations or hidden information (Schroeck, Shockley, Smart, Romero-Morales, & Tufano, 2012). Many business sectors are interested in constituting big data infrastructures in their business environments as decision-making aids (Schroeck, Shockley, Smart, Romero-Morales, & Tufano, 2012). Different from the traditional data management, the big data systems employ separate application software components for data collection, data store, data curation, data use, data analytics, data update, and data transfer to next levels in an independent operating system. Tremendous efforts are required to design and test these systematic big data architectures in the ad-hoc manner. The big data analytics even enable users to utilize the hidden unstructured data that never actively used for any purpose. Based on the big data's four properties (volume, variety, velocity, and veracity; Schroeck, Shockley, Smart, Romero-Morales, & Tufano, 2012), users and operators can discover patterns, relationships, and insights that had not been easily identified with a limited volume of data. Developing a big data environment and applying the big data analytics for the aviation field can provide valuable novel information and insights to pilots, air traffic controllers, dispatchers, maintenance, and business leaders to improve the safety and operational performance. The federal aviation administration (FAA), industry, and research organizations became interested in the big data infrastructures.

This study identified three different aviation areas that need or already developed the big data systems for their subject matters and examined human factors considerations while dealing with the big data in each area. Following sections specified histories and plans of the big data application for (1) aviation infrastructure, (2) aircraft, and (3) operator. To highlight the human factors professionals' role in the big data environment, human factors questions for aircraft pilots, air traffic controllers, aviation dispatchers, and aircraft maintenance staffs were created in each field.

Aviation Fields Considering Big Data Application

Aviation Infrastructure

Next Generation Air Transportation System defines a concept of network centric infrastructure. Under the net-centric infrastructure, every aircraft that install automatic dependent surveillance-broadcast (ADS-B) system have the access to all the aeronautical/flight/weather data for their precision flight operations (JPDO, 2011).

This information-sharing program has been evolved from Aircraft Situation Display to Industry (ASDI) to system wide information management (SWIM) program. Feeding the ASDI data stream including aircraft in-flight location, flight plan, altitude, airspeed, destination, estimated time of arrival, designated identifier to all airliners and aviation organizations was initiated by the department of transportation (DOT) in 1990s (Ayhan, Pesce, Comitz, Sweet, Bliesner, & Gerberick, 2013). Many airline industries subscribed to this program to access the datasets for their businesses. The performance of this program stayed limited since the aircraft that want any data only connected to the data source remotely on demand that required complex procedures (Verma, 2016). The SWIM program is the modernized one that solved the point-to-point access problem. The SWIM employs a centralized common data platform of national airspace system (NAS) data that connects all data sources and users easily and rapidly (Verma, 2016). The FAA publishes all data stream in the SWIM so that all users with the FAA permission can have the access to whatever data they want (Verma, 2016). The data list expanded to a larger dataset adding airport operational status, weather information, status of special use airspace, and NAS restrictions. Stored in the cloud, the SWIM data is expected to increase the common situation awareness among all aviation communities during their operations since the ASDI was decommissioned at 2016 (Verma, 2016).

The dataset of SWIM is maintained best using the big data system since its volume is very big and it contains a mixture of structured and unstructured data. Users need to make a decision on which two heterogeneous data lists will be relevant to extract any valuable information. As conducting updated analytics with accumulated data, the insight becomes more accurate. Applying machine learning, automated flight management systems can recommend better alternative paths when encountering bad weather ahead of ownship based on the better prediction as machines themselves accumulate the information of rerouting recommendation for pilots (Akerkar, 2014). Ayhan, Pesce, Comitz, Sweet, Bliesner, and Gerberick (2013) demonstrated the actual vs. planned route based on the flight computer's big data information. Kasturi, Prasanna, Vinu, and Manivannan (2016) proposed an airline route profitability-optimization model based on big data analytics. This best recommended rerouting path are shared with air traffic controllers for shared situation awareness. The big data from the SWIM infrastructure is shared to everyone who want to know the flight information that affect the airplane delay or cancellation. Airline passengers and airport limousine services utilize the flight information using applications on their mobile devices (e.g. FlightAware).

Human Factors Challenges: human factors professionals may have these questions related to the utilization of big data systems for aviation infrastructure.

- How do we indicate predictive information or insights for any specific flight operation on the limited cockpit display screen?
- How can a pilot evaluate the information accuracy for their situation awareness? (recent but old information vs. near real-time information vs. real-time information)
- Which level of SWIM big data analytics information should be allowed to access and interact with for pilots and for air traffic controllers?
- Will the big data analytics reduce the workload for pilots and air traffic controllers?
- How does cockpit displays visualize multiple variables of information for pilots?
- How does air traffic control displays visualize multiple variables of information?

Aircraft

Recent aircraft install very high number of sensors on engines, avionics, or electrical components. Airbus A380-1000 model is expected to have 10,000 sensors in each wing; the number of sensors and the captured data using the sensors will further increase in the future (Marr, 2015). The purpose of installing these sensors on the aircraft parts is to monitor the aircraft health and extract status information during specific operational stages (Bellamy, 2014). This data enables the predictive maintenance – identifying what components are in bad conditions and repairing the components before they fail. Like the state-of-the-art automobile technology, aircraft also can monitor the fuel consumption in real-time. Accumulating the fuel consumption data in different operational stages, the smarter fueling decision can be made (Wholey, Deabler, & Whitfield, 2014). As well as the smarter operational performance prediction, this aircraft monitoring strategy may increase reliability and help accident investigations. Maintenance staffs can integrate the spare part-supply status data into the monitoring part status data to make a quick maintenance decision (Wholey, Deabler, & Whitfield, 2014). Identified component vulnerability results analyzed by the sensor data may also provide insights about the aircraft component design and development (Wholey, Deabler, & Whitfield, 2014). Since the quantity of updated data is a huge amount, the aircraft sensor data should be managed in big data systems. Many aerospace manufacturers developed big data architectures for diagnosis of their products (Chen et al., 2016).

Human Factors Challenges: human factors professionals can consider following questions related to the use of big data of aircraft sensors.

- How to design the interface of sensor data for technicians, engineers, and pilots?
- Is it required to integrate the sensor data to the SWIM infrastructure for the comprehensive management?
- How does a pilot maintain the SA of aircraft health even if the number of sensor increases?
- How to train engineers, safety managers and maintenance specialists to have the knowledge about big data for aircraft components?
- What are human factors considerations for precision maintenance based on the sensor data?

Operators (Pilots)

Like monitoring aircraft component statuses, operators' (pilots') behaviors can be monitored and the behavioral data can be collected to discover the potential human performance degrades or errors in specific operational stages. However, the environment of collecting human behavioral data is different from the aircraft condition data. Unlike the data from thermal, vibration, or pressure sensors, the sensor types to collect the human behavior are limited; video or audio sensors can be used, and the history of interaction with computer systems can be collected. To collect practical human behavioral data, it is important to make human operators comfortable while they are monitored to avoid the Hawthorne effect (i.e. the behavioral differences when participants are aware of being observed). Psychology fields defines this study methodology as the naturalistic study. The naturalistic study has been applied for the surface transportation. Virginia Tech Transportation Institute (VTTI) exploited "naturalistic driving study (NDS)" installing video sensors inside the car to monitor safety critical drivers' behaviors. The number of sensor-equipped car for the NDS was more than 100, and the period of time for data collection was several months to a year. The NDS experimenters have maintained separate storages for the data management and analysis.

The naturalistic study methodology can be applied to the aviation field for "naturalistic flying study (NFS)." Compared to car drivers, aircraft pilots have more list of safety critical task. Even the flight data including altitude, attitude, speed, and GPS signal should be recorded in line with the pilot behaviors. The NFS may have benefits to evaluate the pilot behaviors in the cockpit with multiple variables that was difficult to test in the simulated environment (Caponecchia, Wickens, Regan, Steckel, & Fitch, 2014). Researchers recently started the NFS. The collected data can apply the big data analytics to explore the hidden insights per specific stage of flight operation and pilot expertise level. To make the genuine big data system for the NFS, the experimenter should consider incorporate many external factors besides flight data, because the concept of big data for this matter is not merely an expansion of data volume of simulated study levels.

Human Factors Challenges: If the NFS passed their preliminary stages, several human factors research questions assuming more advanced testing environment may arise as follows.

- What is the privacy problem of videotaping pilot behaviors?
- Does the NFS validate the human-in-the-loop simulation test results for similar studies?
- Is it possible to integrate the NFS data into the SWIM infrastructure to create more comprehensive testing environment?

- Is it possible to integrate the NFS data into the aircraft sensor data to discover the relationship between aircraft sensor statuses and human behavior in specific operational stages?
- Is it possible to constitute real-time pilot behavior monitoring system in bigger aircraft?
- What is the security problem in implementing the NFS?
- Can the implications from the NFS with limited number of aircraft represent the larger pilot group in the same class?
- What kind of properties has been discovered while conducting the NFS compared with NDS?

Limitation of Big Data System for Aviation Applications

The FAA is interested in constituting the big data environment in air transportation system, but it has not been progressed as expected. There are some reasons for this. First, the big data system inherently demands connection with other dataset that is not directly related to the given dataset to create the hidden information. However, the investigation on which information should be discovered by connecting two information groups that are not directly related to each other, such as aircraft sensor data and meteorological data. Since connecting two datasets is a difficult task within a system, the obvious benefit by the connection should be found. Industries and aviation communities are still investigating the benefits and the current integration capability (Valeika, 2016).

Second, the data scientists often need to manipulate the dataset for analysis. However, the direct manipulation of scripting in the big data system is very difficult due to its scale (Fisher, DeLine, Czerwinski, & Drucker, 2012).

Third, the visualization techniques of analyzed information with high number of variables in the big data system needs to be studied. The visualization of huge statistically analyzed results may not fit in an average size screen and requires complex display techniques to understand (Fisher, DeLine, Czerwinski, & Drucker, 2012). Gorodov and Gubarev (2013) identified the problems of visualization in big data applications: visual noise, large image perception, information loss, higher performance requirements, and high rate of image change. This is also a human factors problem.

Fourth, large volume of data may not be always good. The provided big dataset should be evaluated if the dataset represents the larger group in many perspectives.

Fifth, the aviation fields generally require higher security level than other fields. Therefore, the higher security considerations should be applied when designing and developing a specific big data system for aviation. This could be a blocking factor to proceed human factors research activities.

Finally, any aircraft not equipping sensors will not reflect what happens in their components in the big data. Therefore, it is possible to have an inequality problem for representation of certain situation excluding the unequipped aircraft group (Wholey, Deabler, & Whitfield, 2014).

Conclusion

Employing big data systems to manage the data generated from the aviation infrastructure, aircraft sensors, and naturalistic flying study may provide benefits to discover hidden correlations and insights in all aviation sectors. Human factors professionals need to recognize challenges in these sectors including integrating two different datasets for the sake of users and comprehensible result visualizations when the big data systems are applied.

References

Akerkar, R. (2014). Analytics on big aviation data: turning data into insights. *International Journal of Computer Science and Applications*, *11*, 116-127.

Ayhan, S., Pesce, J., Comitz, P., Sweet, D., Bliesner, S., & Gerberick, G. (2013, April). Predictive analytics with aviation big data. In *Integrated Communications, Navigation and Surveillance Conference (ICNS), 2013* (pp. 1-13). IEEE.

Bellamy, W., III. (2014, December 01). Avionics Big Data: Impacting All Segments of the Aviation Industry. Retrieved December 01, 2016, from http://www.aviationtoday.com/2014/12/01/avionics-big-data-impacting-all-segments-of-the-aviation-industry/ Avionics Magazine Column

Caponecchia, C., Wickens, C., Regan, M., Steckel, R., & Fitch, G. (2014, September). Naturalistic Flying Studies: Innovation in Aviation Safety Research (Discussion Panel). In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 58, No. 1, pp. 69-71). Sage CA: Los Angeles, CA: SAGE Publications.

Chen, J., Lyu, Z., Liu, Y., Huang, J., Zhang, G., Wang, J., & Chen, X. (2016, April). A Big Data Analysis and Application Platform for Civil Aircraft Health Management. In *Multimedia Big Data (BigMM), 2016 IEEE Second International Conference on* (pp. 404-409). IEEE.

Fisher, D., DeLine, R., Czerwinski, M., & Drucker, S. (2012). Interactions with big data analytics. *interactions*, *19*(3), 50-59.

Gorodov, E. Y. E., & Gubarev, V. V. E. (2013). Analytical review of data visualization methods in application to big data. *Journal of Electrical and Computer Engineering*, *2013*, 22.

JPDO (2011). Concept of Operations for the Next Generation Air Transportation System (Version 3.2). Joint Planning and Development Office, Washington, DC.

Kasturi, E., Devi, S. P., Kiran, S. V., & Manivannan, S. (2016). Airline Route Profitability Analysis and Optimization Using BIG DATA Analyticson Aviation Data Sets under Heuristic Techniques. *Procedia Computer Science*, *87*, 86-92.

Marr, B. (2015, April 9). That's Data Science: Airbus Puts 10,000 Sensors in Every Single Wing! Retrieved November 30, 2016, from http://www.datasciencecentral.com/profiles/blogs/that-s-data-science-airbus-puts-10-000-sensors-in-every-single Data Science Central Blog Post

Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., & Tufano, P. (2012). Analytics: The real-world use of big data. *IBM Global Business Services*, 1-20.

Valeika, R. (2016, January 4). Opinion: 'Big Data' From Aircraft? Great. But For What? Retrieved December 2, 2016, from http://aviationweek.com/optimizing-engines-through-lifecycle/opinion-big-data-aircraft-great-what Aviation Week & Space Technology Editorial

Verma, A. (2016, April). SWIM consumer implementation. In *Integrated Communications Navigation and Surveillance (ICNS), 2016* (pp. 10B1-1). IEEE. JPDO, 2011

Wholey, T. J., Deabler, G., & Whitfield, M. M. (2014). *Commercial Aviation and Aerospace: Big Data Analytics for Advantage, Differentiation and Dollars* (Tech. No. GBW03316-USEN-00). Somers, NY: IBM Global Business Services.