United Airlines In-Flight Wi-Fi Health Management: Revolutionizing Aircraft Connectivity through Real-time Prognostics and Big Data Analysis

Ehsan Rahimi[1], Shuang Ling[2], Luis Mesen[3]

[1,2,3]*Technical Operations Data Analytics (TODA), United Airlines, Chicago, Illinois, 60606, United States of America*

ehsan.rahimi@united.com
shuang.ling@united.com
luis.mesen@united.com

ABSTRACT

This paper highlights an innovative initiative, focusing on Prognostics and Health Management (PHM) to enhance in-flight Wi-Fi performance by proactively identifying aircraft component failures. We propose a novel metric, the Normalized Wi-Fi Health Score (NWiHS), alongside a corresponding alerting mechanism, which together represents a significant advancement in the evaluation and improvement of in-flight Wi-Fi connectivity. To achieve this goal, we utilized big data consisting of millions of historical Wi-Fi heartbeats (HBs) received from each aircraft over the past three years. These HBs refer to periodic data packet transmissions sent from United's aircraft to ground stations, providing crucial real-time insights into the Wi-Fi system's status. Leveraging that data, we utilized advanced statistical methods to estimate a NWiHS - a robust indicator of aircraft-level connectivity performance, which quantifies the percent of missing Wi-Fi HBs normalized to exclude the effect of Wi-Fi provider performance and global coverage.

1. INTRODUCTION

Monitoring internet connectivity in ground-based systems, like home internet, is quite distinct from doing so in in-flight systems due to various environmental and technological factors. As highlighted by Rula et al. (2018), ground-based systems typically benefit from stable, wired, or wireless networks with limited mobility, leading to lower latency and greater reliability. On the other hand, in-flight connectivity must function at high altitudes and speeds, often relying on satellite connections or air-to-ground communication. These conditions cause increased latency and lower bandwidth compared to ground networks.

In the rapidly evolving field of commercial aviation, maintaining reliable in-flight Wi-Fi connectivity poses unique challenges. These include the dynamic nature of aircraft environments, the complexity of onboard systems, and the need for uninterrupted service amidst variable coverage areas. Our research addresses these challenges by applying PHM strategies specifically to In-Flight Connectivity (IFC) systems, utilizing advanced analytics to predict and prevent Wi-Fi system failures. In modern aviation, the integration of onboard Wi-Fi in aircraft stands as a groundbreaking leap, transforming the inflight passenger experience while fostering seamless data communication. As it is defined by Zio (2022), PHM is a data-driven approach that integrates physical insights, information, and operational data of structures, systems, and components to facilitate the identification of abnormalities, diagnose faults, and assess the degradation of equipment and processes. A research study conducted by Kordestani et al. (2023) highlights the importance of PHM in managing the complexities and interconnected subsystems of aircraft. This study emphasizes the need for advanced prognostic strategies to maintain aircraft safety and reliability.

In the context of IFC systems, the application of PHM strategies is pivotal. These systems are critical for enabling onboard internet access, real-time communication, and entertainment services. Therefore, it would be required to have robust monitoring and maintenance to ensure uninterrupted service. While specific research directly linking PHM and IFC systems in commercial aviation is scarce, the key principles of PHM applied in aviation can be extrapolated to manage the health and performance of IFC systems. This includes leveraging data-driven prognostic techniques, and real-time analysis to expect and mitigate potential aircraft-level Wi-Fi system failures. It ensures optimal performance and minimizes disruptions to passengers' connectivity.

This study underscores the potential of integrating big data and advanced analytics to improve onboard Wi-Fi performance through rapid identification of aircraft-level IFC system failures. More specifically, it introduces the NWiHS as an innovative metric designed for Wi-Fi health management across United Airlines' fleet. Building on this metric, this study also unveils an algorithm—an alerting mechanism— aimed at early detection of IFC system failures, thus enabling technicians and engineers to proactively address those issues. Furthermore, the application and effectiveness of this alerting mechanism is proved through its deployment on the Airbus 320 fleet.

2. DATA

This section presents the data forming the backbone of the analysis. As discussed earlier, the aim is to continuously monitor the IFC system and detect failures as they occur. To achieve this, we used big data consisting of historical Wi-Fi HBs collected from each aircraft over the past three years. These HBs are automatically generated every 5 minutes by the aircraft's onboard broadband controller or file server, where the United portal is installed. Each HB is then transmitted in real-time to United's ground stations, conveying essential Wi-Fi connectivity status information.

The transmitted data includes the HB timestamp, departure and arrival stations, aircraft details, the volume of data bytes transmitted and received, and the Wi-Fi service provider information. A flight is considered "healthy" if all expected HBs are received during its duration. Conversely, if HBs are missing at the expected intervals, this shows a potential connectivity issue or failure. Importantly, the structure of the HB data remains consistent across all IFC-equipped aircraft, ensuring uniformity in monitoring and analysis.

For analytical precision and due to operational differences among Wi-Fi providers, our study focuses exclusively on airborne HBs, discounting any data transmitted while the aircraft is on the ground. This includes HBs from the moment of takeoff to landing, ensuring that data reflects the connectivity status under operational flight conditions. Figure 1 illustrates the selection process for HB data corresponding to a single flight, providing a visual representation of the data filtering method. To provide additional context on the data structure and as an example, an aircraft in a healthy state is expected to transmit 12 HBs to the ground station during a flight with a 60-minute airtime. A deviation from this expected HB transmission frequency may signal an unhealthy state linked to connectivity issue. In the next section, we discuss how these HBs are utilized as a base for monitoring Wi-Fi connectivity performance and to detect anomalies in the IFC system.

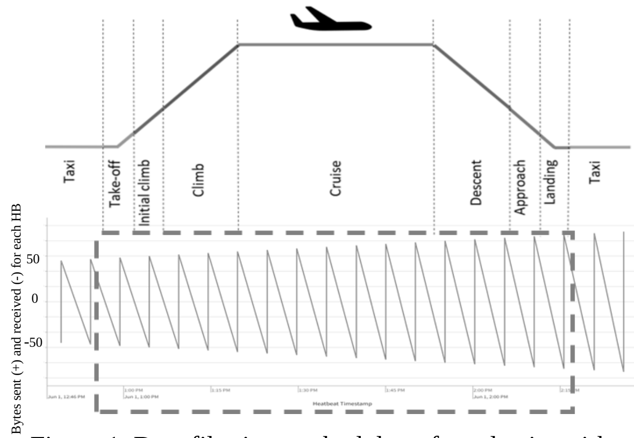


Figure 1. Data filtering methodology for selecting airborne heartbeat data per flight.

3. METHODOLOGY AND RESULTS

This section aims to present core algorithms that transform raw HB data from an aircraft into a real-time NWiHS, a crucial metric for monitoring IFC health. Subsection 3.1 delves into the NWiHS metric, elaborating on the estimation method employed. Subsequently, subsection 3.2 explores the alerting mechanism, showcasing its implementation within the Airbus 320 fleet.

As outlined in Section 2, aircraft equipped with an IFC system are anticipated to send a HB to the ground station every five minutes of flight time, provided the connectivity is in a healthy state (i.e. aircraft maintains Wi-Fi connectivity while airborne). To quantify the disparity between the actual and expected number of HBs, the Missing Heartbeat Percentage (MHP) metric was introduced, as defined in Equation (1). MHP quantifies the percentage of a flight where no heartbeats are received, with 100% indicating no connectivity throughout the entire flight and 0% indicating complete connectivity.

$$MHP_{flight} = \left(1 - \frac{Actual\ number\ of\ HBs_{flight}}{Expected\ number\ of\ HBs_{flight}}\right) \times 100 \tag{1}$$

The primary objective here is to pinpoint failures within the aircraft's Wi-Fi system and its components, not other external factors. To this end and as explained by Rosenbaum, P. R. (2002), an initial step involves identifying and controlling for potential confounding variables that could affect the main variable of interest, the aircraft-level Wi-Fi system, as depicted in Figure 2. This identification process was conducted using exploratory data analysis and correlation studies aimed at uncovering variables linked to the MHP and the likelihood of failures in the aircraft-level IFC system. The results of our correlation analysis indicated that over 70% of the system noise might be attributed to route-based connectivity coverage, which encompasses the connection availability provided along the flight route's geographical coordinates and the satellite's bandwidth. Figure 3.a depicts the MHP for an aircraft over the past three years, illustrating considerable noise in the metric and the challenge of discerning a clear pattern. As seen in the figure, there are several occurrences of high MHP values. The key question, however, is whether these high values are indicative of issues within the aircraft's system itself, or if they stem from factors such as provider performance and route coverage. To address this, we aim to decompose these contributing factors to better identify potential issues within the aircraft system. The next section is devoted to answering this question and discussing our goal of decoupling aircraft-related issues from other factors affecting connectivity performance.

3.1. Normalized Wi-Fi Health Score (NWiHS)

In the realm of big industry, selecting an optimal method for denoising time series data, as elucidated by Shumway et al. (2017) and Baumann et al. (2023), is a critical decision that significantly impacts operational outcomes. The chosen method must not only excel in accuracy, ensuring precise isolation and analysis of the factor of interest but also align with the industry's demand for simplicity and efficiency, particularly in environments handling real-time operational data (George et al., 2014).

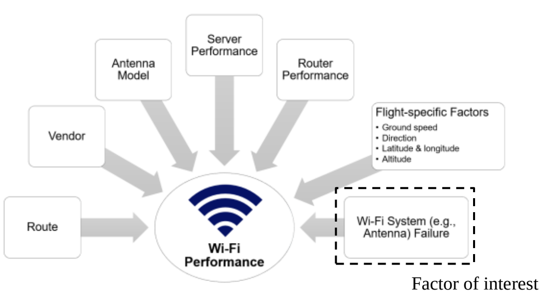


Figure 2. Factors influencing Wi-Fi performance.

Our methodology transforms raw HB data into a real-time NWiHS, leveraging advanced statistical techniques to filter out noise and accurately assess Wi-Fi health. The noise in HB data is largely attributed to factors such as connectivity coverage, which varies depending on the route and the connectivity provider. NWiHS is designed to decompose these non-aircraft-related factors, including route and satellite bandwidth, so it can more accurately represent failures in the aircraft's IFC systems. Consequently, any anomalies detected within this refined metric may signal underlying issues with the aircraft's systems or components, necessitating maintenance intervention. Equation (2) explains the formulation of the proposed metric, illustrating its foundational structure in assessing aircraft IFC system health.

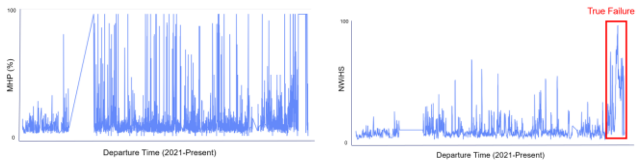
$$NWiHS_{flight} = \left(\frac{MHP_{flight}}{Route - based\ Coverage\ Adjuster_{flight-provider}}\right) + \varepsilon \qquad (2)$$

Where, $Route - based\ Coverage\ Adjuster_{flight-provider}$ is a calculated metric that represents the connectivity coverage, taking into account the specific route of the aircraft and the associated connectivity provider. ε denotes an error term, capturing the influence of other unobserved external factors. Subsequently, we will delve into the methodology employed to estimate the Route-based Coverage Adjuster, providing a comprehensive understanding of how this metric is derived and its relevance in assessing connectivity coverage.

The NWiHS metric is notably sensitive to variations in route-based coverage, underscoring the necessity for precise estimation of this coverage to ensure the metric's effectiveness. From another standpoint, data on coverage and provider performance is often elusive and proprietary, and such coverage is subject to fluctuations over time and with varying traffic volumes, influenced by constraints in satellite bandwidth. Considering these conditions and limitations, we have utilized HB data to construct a coverage map based on United Airlines' operation network. Below is an iterative algorithm to estimate the Route-based Coverage Adjuster for each combination of Origin-Destination-Provider (O-D-P):

1. **Initialization:**
   - **Step 0:** Initialize by setting $i = 0$ and $D_i = 0$.
2. **Iteration and Analysis Window Setup:**
   - **Step 1:** Increment the iteration: Set: $i = i + 1$
   - **Step 2:** Define the analysis window: Set $D_{i+1} = D_i + 30$
3. **Calculation and Condition Check:**
   - **Step 3:** Calculate the median for the current window: $M_i = median(MHP)_{D_i}$
   - **Step 4:** Compute the number of unique aircraft flying in the current window: AC_i
   - **Step 5:** Check if AC_i is sufficient (i.e., $AC_i \geq 10$). If not, return to Step 1. If yes, proceed to Step 6.
4. **Final Adjustment:**
   - **Step 6:** Set the Route-based Coverage Adjuster for the origin-destination-provider based on the current median: $Route - based\ Coverage\ Adjuster_{O-D-P} = M_i$

The algorithm ensures that the Route-based Coverage Adjuster is not only current and adaptable but also differentiates between the varying flight patterns of narrow-body aircraft on frequent routes and wide-body aircraft on seasonal or less frequent routes by expanding the analysis window. This differentiation bolsters the algorithm's robustness against anomalies, such as accounting for aircraft in suboptimal condition flying the same routes, enhancing the reliability of the NWiHS as an indicator of IFC system health. Figure 3.b illustrates that the effective implementation of NWiHS can substantially filter out system noise, thereby uncovering the actual failures within the aircraft's IFC system.



a) MPH time series b) NWiHS time series

Figure 3. MPH vs. NWiHS timeseries plots for an aircraft

3.2. NWiHS as a Tool for Real-Time Anomaly Detection: A Case Study with the Airbus A320 Fleet

Anomaly detection in time series data is pivotal for identifying irregular patterns that may indicate critical events, remarkably in industrial settings. The threshold-based anomaly detection method, particularly the standard deviation approach, is highly valued in industry for its straightforwardness and efficiency in spotting system failures as highlighted by Clark et al. (2018). In our study, we leveraged historical NWiHS time series data alongside IFC component failure data to ascertain the most effective approach for an alerting mechanism. Initially, we applied an aircraft-specific standard deviation method, determining that $\mu_{aircraft}(NWiHS) + 2 \times \sigma_{aircraft}$, where μ represents the mean and σ denotes the standard deviation, offers highest performance based on recall and precision. Our analysis further indicated that over 90% of the aircraft exhibited a threshold limit below 4.7 across the A320 fleet, as shown in Figure 4. Additionally, we observed that the NWiHS time series data are predominantly stationary, with the normal state exhibiting minimal fluctuations over time. This stability is attributed to the Route-based Coverage Adjuster, which effectively normalizes the NWiHS against various time-dependent factors, such as air-traffic volume (satellite's bandwidth) and provider network enhancements. Given our findings, which prioritize high recall and precision while avoiding undue complexity, we recommended the 90th percentile as the optimal threshold's base value. Based on historical data from 2021 to 2023, implementing this NWiHS time series threshold approach could achieve recall and precision rates of 88% and 78%, respectively, in detecting IFC system failures.

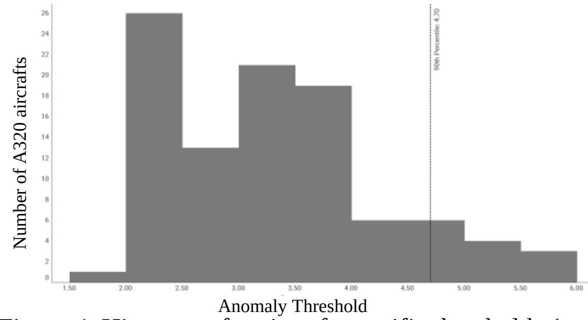


Figure 4. Histogram for aircraft-specific thresholds (μ+2*σ)

4. CONCLUSION

This study represents a significant advancement in the management of in-flight Wi-Fi systems, demonstrating how big data and advanced analytics can be leveraged to enhance connectivity reliability. Despite the complexity of the task, our findings offer promising directions for future research, particularly in further refining the NWiHS metric and exploring its application across different aircraft types and operational contexts. As the aviation industry continues to evolve, the continuous improvement of PHM strategies will play a crucial role in meeting the increasing demands for reliable IFC systems.

ACKNOWLEDGEMENT

We express our profound gratitude to David Cronin, Keith Shackelford, and Chris Maldonado from the Cabin System Engineering Team at United Airlines for their invaluable partnership and contributions that played a pivotal role in the realization of this paper. Additionally, we wish to express our sincere appreciation to the Data Engineering team at United Airlines, especially Missy Tillander, Kevin Hsu, Rajeswari Bhojan, and Kristen Lin, for their unwavering support and expertise in fulfilling our data requirements. Their collective efforts have been instrumental in the success of this work.

REFERENCES

Baumann, E., Hsu, C., Buba, H., & Cox, T. (2023). An Introductory Approach to Time-Series Data Preparation and Analysis. Annual Conference of the PHM Society, 15(1). https://doi.org/10.36001/phmconf.2023.v15i1.3561

Clark, J., Liu, Z., & Japkowicz, N. (2018). Adaptive Threshold for Outlier Detection on Data Streams," *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA)*, Turin, Italy, 2018, pp. 41-49, doi: 10.1109/DSAA.2018.00014

George, G., Haas, M. R., & Pentland, A. (2014). Big data and management. Academy of Management Journal, 57(2), 321-326.

Kordestani, M., Orchard, M., Khorasani, K., & Saif, M. (2023). An Overview of the State of the Art in Aircraft

Prognostic and Health Management Strategies. IEEE Transactions on Instrumentation and Measurement, 72, 1-15.

Rosenbaum, P. R. (2002). Observational Studies (2nd ed.). Springer-Verlag.

Rula, J. P. et al. (2018). “Mile High WiFi: A First Look at In-Flight Internet Connectivity.” Proceedings of the 2018 World Wide Web Conference.

Shumway, R. H., & Stoffer, D. S. (2017). Time Series Analysis and Its Applications: With R Examples (4th ed.). Springer

Zio, E. (2022). Prognostics and Health Management (PHM): Where are we and where do we (need to) go in theory and practice. Reliability Engineering & System Safety, 218, Part A.

BIOGRAPHIES

Ehsan Rahimi, PhD Ehsan serves as a Senior Analyst at United Airlines, where his extensive experience encompasses the fields of Econometrics and Machine Learning modeling. He earned his PhD in Transportation Engineering from the University of Illinois at Chicago, United States. Specializing in the application of data-driven methodologies, Ehsan tackles intricate challenges within the transportation sector, with a particular focus on aviation. He is also an active member of the Travel Survey Methods Standing Committee at The National Academies of Sciences, Engineering, and Medicine.

Shuang Ling, MS Shuang serves as a managerial data scientist with a focus on predictive maintenance, where she leads a dedicated team in the development and implementation of prognostic solutions across crucial systems including flight control, air conditioning, and engine bleed systems. Holding Master of Science degrees in Predictive Analytics with Computational Methods and Marketing Communications, Shuang has accrued substantial expertise in crafting predictive models that enhance the reliability and efficiency of complex systems. Her work, characterized by the integration of advanced analytics and interdisciplinary knowledge, has established her as a key figure in the advancement of maintenance practices within her domain.

Luis Mesen, Director of TechOps Data Analytics Luis Mesen is a results-oriented professional with a distinguished career in data analytics within the aviation sector, currently serving as the Director of TechOps Data Analytics at United Airlines. Known for his precision and analytical acumen, Luis has been instrumental in leveraging data analytics to spearhead strategic initiatives and promote operational excellence at United Airlines. His extensive background encompasses applied statistical analysis, operations research, aircraft performance engineering, air traffic flow management, geospatial analytics, airline operations, and software development. Luis’s contributions to the field are marked by his successful implementation of AI/ML models, completion of comprehensive analytics projects, and the delivery of insightful data analyses that have markedly improved business decision-making processes.