COMPARATIVE STUDY OF METROPLEX AIRSPACE AND PROCEDURES USING MACHINE LEARNING TO DISCOVER FLIGHT TRACK ANOMALIES

Bryan Matthews, SGT Inc. NASA Ames Research Center, Moffett Field, CA

David Nielsen, MCT Inc. NASA Ames Research Center, Moffett Field, CA

John Schade, Kennis Chan, and Mike Kiniry

ATAC Corporation, Santa Clara, CA

Abstract

The National Airspace (NAS) is constantly changing and adapting to new and complex challenges, and as a result Next Generation Air Transportation System (NextGen) will need to address these important aspects. These challenges range from increased traffic flow, to reducing environmental impact, to routing efficiency, all while maintaining high safety. The FAA has recently been involved in a number of large scale Metroplex redesigns across the country to enable controllers and pilots to implement more efficient Performance-based Navigation (PBN) procedures on a regional basis. The Houston Metroplex project is one of the first to implement such a large scale change in the NAS, and it significantly changed the traffic flows into the Houston Terminal Radar Approach Control Facility (TRACON) airspace for the two major airports: Houston's George Bush Intercontinental Airport (IAH) and Houston's William P. Hobby International Airport (HOU). This paper addresses an anomaly detection approach that has previously been used to detect operationally significant anomalous flights on approach to Denver International, Newark International, LaGuardia International, and John F. Kennedy International airports. The same method is applied to radar track surveillance data to identify anomalies in the airspace before and after the Metroplex procedure change at Houston. The study covers flights traversing through Houston TRACON (I90) and landing at IAH and HOU over a period of 2 years before and after the significant procedure change. Anomalies identified before and after the procedure change were characterized by their safety risk and operational efficiency to determine whether the types of anomalies that were discovered from before continued to exist or if they were eliminated after the procedure change or if new types of anomalies began to appear.

Introduction

The National Aeronautics and Space Administration (NASA) continues to lead the aviation community in the area of cutting edge research by developing and testing new technologies applicable to NextGen operations. The recently established Airspace Operations and Safety Program (AOSP) [1] builds upon previous somewhat independent efforts focusing on Air Traffic Management (ATM) efficiency based technologies and aviation safety research. With the combination of these research areas, NASA has formulated a strategy to continue the development of advanced ATM efficiency-based technologies that will also be designed with safety implications on the current and future NAS in mind, thus allowing for solutions that are more readily accepted by aviation stakeholders.

One of the thrust areas for AOSP is designated Real-time System-wide Safety Assurance (RSSA). Elements of the vision for this 30+ year effort include the automated monitoring and detection of aviation system hazards and eventually the proactive mitigation of those risks by system operators or intelligent automation enabling a more complex NAS to continue to operate with an exemplary safety record. As part of this effort in the near term, NASA in partnership with the FAA and industry is continuing to develop new technologies and techniques to identify previously undiscovered safety risks through the intense data mining of the large heterogeneous aviation data sets that are collected on a regular basis.

The technology changes to the NAS that were being rolled out as part of NextGen present a ripe opportunity to conduct research on the efficiency and safety aspects of the implementations from a historical data analysis perspective. This is a preliminary step in the overall RSSA vision to move detection of safety risks closer to real-time while complementing the overall efficiency goals of AOSP as well.

This paper presents progress in the development of advanced data mining algorithms as applied to high fidelity surveillance and trajectory data. It builds upon previous research [2] [3] by examining the nature of safety-related events before and after a major implementation of NextGen procedures. It begins to provide an objective look at the safety and efficiency surrounding the Metroplex procedure change at Houston TRACON (I90) where the aim is to discover previously unknown high safety-risk events. In addition, the efficiency of the resulting changes is examined to understand the benefit to stakeholders in the context of safety risk. As with previous efforts, air traffic management subject matter expertise is incorporated into the research to provide specific domain knowledge of operational procedures and to understand the safety and efficiency implications of the discovered events. In addition, the evaluation of algorithm results is made clearer by the inclusion of Air Traffic Control (ATC) voice recordings from communications frequencies involved at the time of the event occurrence.

This paper is organized as follows: First we present an overview of the Performance Data Analysis and Reporting System (PDARS) which delivers several capabilities to enhance this research including serving as the source of the trajectory information. Next we introduce the Metroplex redesign project, describe the Metroplex design process and how data mining can help with the post implementation performance evaluation. We then give an overview of the previous research efforts in assessing safety risk. We then present the primary algorithms used for the discovery of anomalous events and describe the data used for the research. Since the data processing and handling is quite involved, we give an overview of the end-to-end system for algorithm application. We then present the primary results of the research beginning with a discussion of the overall efficiency observed before and after the Metroplex procedures were implemented. Then we review actual traffic scenarios that were identified as operationally significant anomalies along with a brief analysis for each, including subject matter experts' reviews that provided detailed insight into what factors contributed to the anomalous event. Finally we discuss the conclusions and introduce ideas for future research.

Background

PDARS

The PDARS program is managed by the FAA's Air Traffic Organization Office of System Operations Services and is heavily used operationally by over a dozen organizational units within the FAA, many on a daily basis. PDARS provides the FAA organizational managers and decision makers with "actionable" information regarding the efficiency and safety of the NAS [4].

PDARS consists of an ever-evolving data collection, processing, reporting, and dissemination platform able to accept nearly any surveillance or positional data and merge that with other geo-referenced or contextual aviation-related data (e.g. weather, terrain, or schedules). The system routinely produces analysis products including reports and visualizations that provide detailed operational insight to decision makers at virtually any level in a complex Air Navigation Service Provider organization like the FAA.

Key PDARS capabilities used in this research are its routine collection and processing of large surveillance-based trajectory information sets, its categorization of key flight parameters such as runway utilization, its ability to compute additional geospatial measures on aviation data sets on a large scale, and its suitability as a platform for testing new capabilities and validating the results.

Metroplex Background

Similar to some of the other FAA NextGen initiatives, the Metroplex program addresses one of the FAA goals to improve the efficiency of the NAS. The program focuses on implementing a series of horizontal and vertical efficiency improvements and reducing the inter-dependencies among airports in close proximity at metropolitan areas. The Metroplex program takes the "system-of-system" approach to: optimizing traffic flows at the Metroplex system level in order to maximize the overall benefits of the NAS. Although implementation of readily available technologies such as Performance-Based Navigation (PBN) is not the main goal of the Metroplex initiative, it is often considered as one of the main leading technological tools to achieve the goal of Metroplex.

The Metroplex program is a collaborative effort among the FAA lines of business and the users of the airspace. The program life-cycle consists of five main phases [5]:

1. Study Phase
2. Design Phase
3. Evaluation Phase
4. Implementation Phases
5. Post-Implementation Review Phase

Three teams are formed with different responsibilities for these 5 phases. The Study Team during the Study Phase is responsible for identifying and defining issues that limit the FAA's ability to use available airspace in an optimized and efficient way.

The Design and Implementation (D&I) Team is involved in the Design, Evaluation and Implementation Phases. The team takes the concepts and solutions from the Study Team as a starting point to further mature them into deployable solutions. At this point, the team confirms that potential benefits can be delivered, discovers any workload or operational issues, and ensures safety levels are not diminished through human-in-the-loop testing, airline flight simulations, environmental team reviews, and safety management system analyses. The design and evaluation are iterative processes that will only lead to an Implementation Phase once all the assessments are completed and all issues found during assessments are resolved.

The Post-Implementation Review is the last phase of the Metroplex process. In this phase, the Post-Implementation Review team reviews the performance and benefits of the newly deployed procedures and airspace. If unforeseeable issues are identified and/or desired benefits are not delivered, the team works with the appropriate subject matter experts (SMEs) to mitigate each issue found and modify the design to achieve the expected gains from Metroplex.

The Houston Metroplex is a pioneer of the Metroplex program. It also has high public visibility, being on the White House Dashboard of High Priority Infrastructure Projects list [1]. The new Houston Metroplex procedures were implemented in May of 2014. A total of 61 new routes were launched for flights in and out of four Houston area airports[2]. The new design included many PBN and Optimized Profile Descent (OPD) procedures. Although the Metroplex process is a very well-designed process, due to the number of changes to the Metroplex and the complexity of the NAS, operational issues are still discovered in the NAS during the Post-Implementation phase. Some of these issues can only be uncovered over time for the reason that not all the conditions can be simulated due to high simulation costs or predicted ahead of time due to the complex interactions of the NAS. However, with the current process, none of the issues can be detected automatically and may be difficult to objectively measure the severity of the safety risk during the Post Implementation Review phase. The FAA would benefit greatly from having a tool that can automatically identify operationally significant situations, which may be defined as a circumstance that: increases the safety risk to the system, induces higher controller workload, or reduces operational efficiency. Such automation would shorten the discovery period and thus lower the overall risk of Metroplex procedural improvements. This paper proposes certain aspects of such an approach.

Approach

Current State-Of-The-Art

One of the FAA's accepted methods of analyzing the safety of the NAS is to examine Loss of Standard Separation (LoSS) incidents. In recent years the FAA has implemented the Traffic Analysis and Review Program (TARP), which monitors the TRACON airspace operations and reports on LoSS incidents. As of 2012 over 20 facilities have kept TARP running 24 hours a day 7 days a week collecting data[3]. Incidents are reviewed by air traffic control management along with voice recordings, when appropriate, to determine the severity of the situation. After the incidents are reviewed recommendations regarding training are made to help prevent similar incidents in the future from occurring. Controllers also use the voluntary Air

[1] http://www.permits.performance.gov/projects/nextgen-infrastructure-initiative-houston-metroplex-oapm

[2] https://www.faa.gov/nextgen/snapshots/stories/?slide=35

[3] http://www.faa.gov/news/testimony/news_story.cfm?newsid=13572

Traffic Safety Action Program (ATSAP) to report the details of any situation that may have led to a breakdown in safety or may have increased the risk of operations. The ATSAP reports allow for more flexibility in reporting safety risks outside of LoSS, allowing for the controllers to describe the chain of events as well as contributing factors, corrective actions, and perceptions of human performance. In 2013, a report written by the Office of Inspector General [6] indicated that not all LoSS incidents are being identified by TARP as compared to what was reported in ATSAP. This discrepancy signals a need to have additional monitoring that can complement the current tools as well as provide some redundancy in reporting. Outside of its monitoring and reporting tools, the FAA also has evaluation tools, such as the Integrated Safety Assessment Model (ISAM) [7] that assesses the safety of future NextGen operation concepts. While ISAM is still in the research and development phase, these assessment capabilities cover a range of accident categories including: wake turbulence, controlled flight into terrain, taxiway collision, runway collision, and midair collision. The causes are also grouped into the following categories: direct cause, prevention failures, prevention opportunities, and indirect influences. Although this tool is useful for assessing future operations and safety risks in NextGen, it lacks the capability to examine the probability of unknown events or anything outside of predefined incident categories which primarily focus on LoSS with aircraft or controlled flight into terrain. Since these methods do not have the ability to identify unknown events where the safety margins are reduced, they highlight a vulnerability in the system that if left unmonitored without corrective action may result in more significant incidents. This further emphasizes the need for a tool to identify events that may not be reported by the current safety risk monitoring tools.

Anomaly Detection Algorithm

The method, Multiple Kernel Anomaly Detection (MKAD) [8] was selected for this study, and is based on a One-Class Support Vector Machine (OCSVM) algorithm [9]. OCSVMs are widely used in anomaly detection in the data mining community [10] due to its ability to leverage kernel based methods that can transform the data into a non-linear space where the complex relationships between the anomalies and nominal examples can be separated. The principal concept of the method is to separate a percentage of statistically significantly anomalous examples from the greater population using an appropriate distance metric. The distance metric or kernel function is used to calculate a kernel matrix that describes the similarity of all nxn flights in the training set. For this work the kernel function used was the radial basis function shown in Eq 1.

$$\kappa_m(\vec{x}_i, \vec{y}_j) = exp^{-\frac{\|\vec{x}_i - \vec{y}_j\|}{2\sigma^2}} \quad (1)$$

The $\vec{x}_i$ and $\vec{y}_j$ vectors represent feature vectors for two different flights, while σ is a kernel hyperparameter and is required to be tuned by the user. For this study an unsupervised grid search was performed on a subset of the data over the σ values for each feature to determine the σ that produces the best kernel distribution for distinguishing between flights. This process is described in more detail in [3]. Once the kernel matrix is computed the following optimization problem is solved using a quadratic programming solver with the following constraints shown in Eq 2, Eq 3:

Minimize:

$$Q = \frac{1}{2}\sum_{i=0}^{n}\sum_{j=0}^{n}\alpha_i\kappa(\vec{x}_i, \vec{y}_j)\alpha_j \quad (2)$$

Subject to:

$$0 \le \alpha_i \le \frac{1}{nv}, \sum_{i=0}^{n}\alpha_i = 1, 0 \le v \le 1 \quad (3)$$

This yields the weights(α_i)for the training examples, where K represents the kernel matrix, n is the number of examples, and v is a hyperparameter set by the user that represents the expected percentage of examples that are anomalous (for this study v was 15%). Examples with weights above a tolerance define the support vectors and form the resulting hyperplane that separates the normal examples from the anomalous ones.

In classical OCSVM problems a single kernel function is used to compute the kernel matrix. In this work, as in our previous study [3], a separate kernel for each of the features in the data is computed and combined linearly across all M features to form a single kernel as shown in Eq. 4.

$$\kappa(\vec{x}_i, \vec{y}_j) = \sum_{m=1}^{M} W_m * \kappa_m(\vec{x}_i, \vec{y}_j) \quad (4)$$

The kernel parameter: σ was tuned for each kernel separately, using the grid search approach, before combining with equal weights for each feature and therefore leveraging multiple kernels for anomaly detection. The flights corresponding to the support vectors are then compared to test flights using the kernel function and the process for combining the weighted kernels is repeated. The vector of $\alpha_i$weights is multiplied with the kernel matrix resulting in a vector of distances from the origin for each flight. A bias term is computed from the support vectors' mean distance from the origin to determine where the hyperplane lies in one-dimension. Flights below the hyperplane are considered anomalies while flights above the hyperplane are considered nominal. The absolute distance from the hyperplane represents the severity of the anomaly and can be used to rank the flights. It is important to note that the distance ranking by severity of anomaly and statistical significance of that severity is not directly coupled with operational significance. Anomalies that are found still need to be validated by SME review, which can require analyzing additional features that the algorithm currently cannot utilize. This may involve: visually inspecting the original parameter space, viewing animations, and/or listening to voice recordings to help fully define the event in context of increased safety risk.

Data Processing

The surveillance-based trajectory data used in this research come from both Houston I90 and Houston Air Route Traffic Control Center (ARTCC) (ZHU). The PDARS system collects and processes raw data from these two ATC facilities individually to transform it from individual data records into flight track-based information.

Then, information from the two facilities is merged together. During the merging process, the PDARS system selects the best radar hits to merge together based on many different criteria in order to produce the best quality of four-dimensional (latitude, longitude, altitude, and time) trajectories for flights. The resultant data provides analysis ready trajectories within the ZHU and I90 airspace boundaries. Commercial Instrument Flight Rules (IFR) aircraft are the primary focus of this study, while Visual Flight Rules (VFR) flights with beacon codes from 1200 to 1299 and military flights are removed from the data at this stage. The benefit of removing those flights is that military and non-discrete code VFR flights typically have unusual flights paths as compared to commercial flights, and by removing those flights the algorithm is expected to yield more relevant results. Also at the same stage, additional flight information required for this study such as destination airport and landing runway are computed. Although this study's focus is on Houston TRACON airspace, information such as flight route (field10 of the FAA filed flight plan) is obtained from the ARTCC data which is particularly useful when analyzing the results in the findings section in this paper. Due to the sensitivity of the ATC data, these data first go through a filtering process to remove any sensitive flight information.

After the data is transformed and filtered, it is used to calculate the minimum separations between aircraft for MKAD for use as one of the determining factors for anomaly detection.

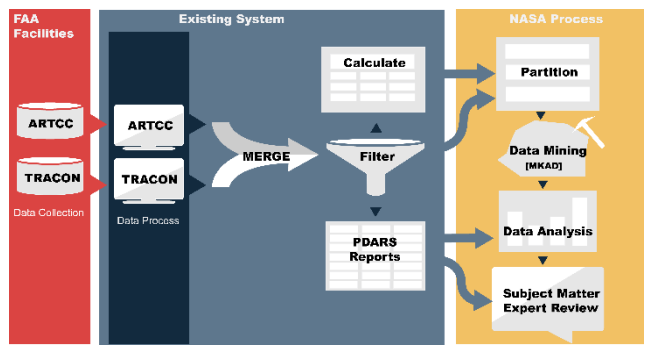

Figure 1. Process flow chart.

To format the data into a structure suitable for the MKAD algorithm, each trajectory and its corresponding features are converted into a uniform vector with common distance traveled. The trajectory starts by finding the TRACON boundary crossing point and using the ground speed. The next 40 flown miles are used to create a four-dimensional trajectory of latitude, longitude, altitude, and distance to the nearest neighboring aircraft, averaged over every nautical mile (NM). The trajectories are then partitioned by unique flow, based upon the arrival fix they use to enter the TRACON airspace, their entry altitude, and landing direction. By partitioning the data this way, populations of similar flights are grouped together and their underlying similarity allows MKAD to create a tighter model of their nominal behavior. For

each partition on each day, MKAD uses the previous 30 days of historical data from that partition as training data against which the test day is compared. Figure 1 shows the data flow process from collection at the facilities down to the analysis SME review steps.

To support data analysis and SME review of the MKAD results, various PDARS metrics such as number of go-arounds and measurements at the time when a flight begins its base-to-final turn into an airport were used for validations. In addition, efficiency metrics such as time flown from 100 NM and 40 NM to IAH (which is similar to Terminal Arrival Efficiency Rate metrics) were referenced and computed in PDARS to investigate the overall TRACON efficiency impacts of the procedure change on arrivals into Houston after Metroplex procedures were implemented.

Findings: Efficiency Routing

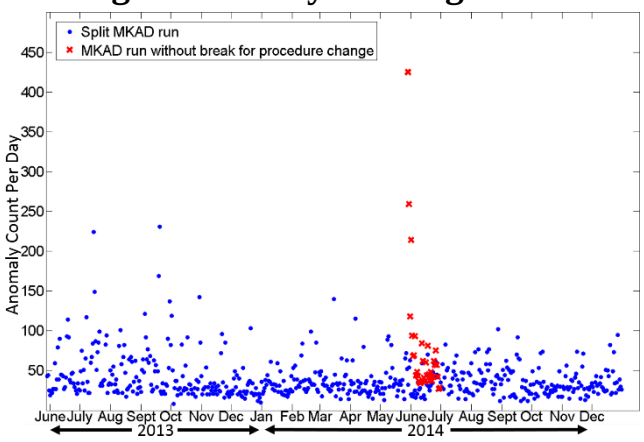

Figure 2. MKAD anomaly count for each day.

Using the approach described in the *Data Processing* section, MKAD evaluated all flights landing at IAH and HOU for the 2014 calendar year. MKAD produces a ranked list of anomalies for each day of the year with the total anomaly count for each day from June 2013 to December 2014 shown in Figure 2. The algorithm uses a sliding window of training examples prior to each test day. When training days were chosen from the month prior to the procedure change on May 29th and tested on days after, a higher anomaly count (shown in red) is observed. When only using training days from the same Metroplex implemented operations, the spike in the anomaly count disappears (shown as blue in the plot). This was expected behavior and validated the ability of the algorithm to detect a significant change in the system.

Days with relatively higher anomaly counts were also investigated to determine if significant changes in routing were identified. One such day in August 2014 (after the Metroplex procedure change) was found to have a significant spike, primarily on the North East (NE) flow. Upon further examination it was determined that a large number of RNAV equipped flights were using the ground based conventional route. This was found to be due to bad weather on the new route paths, forcing the controllers to fall back to using the conventional route. Figure 3 illustrates the flows for this day and the anomalous flights (trajectories colored in pink) labeled by MKAD. The anomalies followed the conventional arrival route due to adverse weather patterns on the RNAV routes. The nominal flights (trajectories colored in blue) labeled by MKAD follow the RNAV routes during good weather periods on the same day.

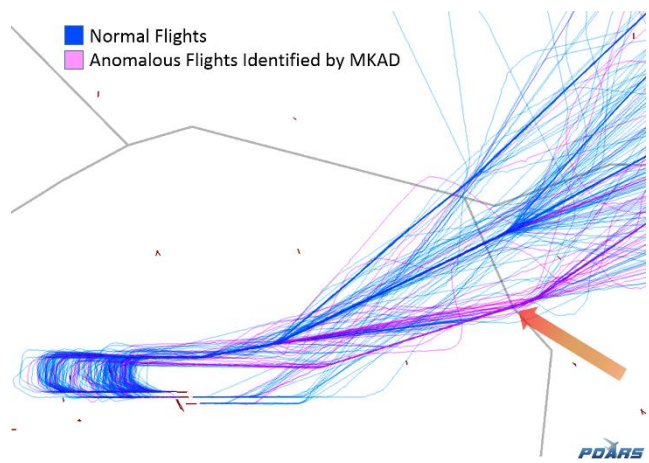

Figure 3. MKAD anomalies for high anomaly day in August 2014.

For days throughout the 2014 calendar year a large number of anomalies that were identified were found to be due to flights vectored off of the RNAV routes. These included some of the higher anomaly count days in the summer of 2013 and again in the summer of 2014, when the region typically experiences adverse convective weather patterns (however the summer months in 2014 appear to be less severe based on the MKAD anomaly count). Other anomalous routing was typically found for sequencing on the base leg before joining the final leg during high traffic times. High traffic times were determined by analyzing the arrival counts from the FAA's Aviation

System Performance Metrics (ASPM) data set [11]. The threshold was set to 2- σ above the mean (~18.5 flights per 15 min) to identify high traffic times. Although, the anomalies found on these days were not high safety risks, the observed effects were still statistically and operationally significant and warranted further investigation of the flow dynamics.

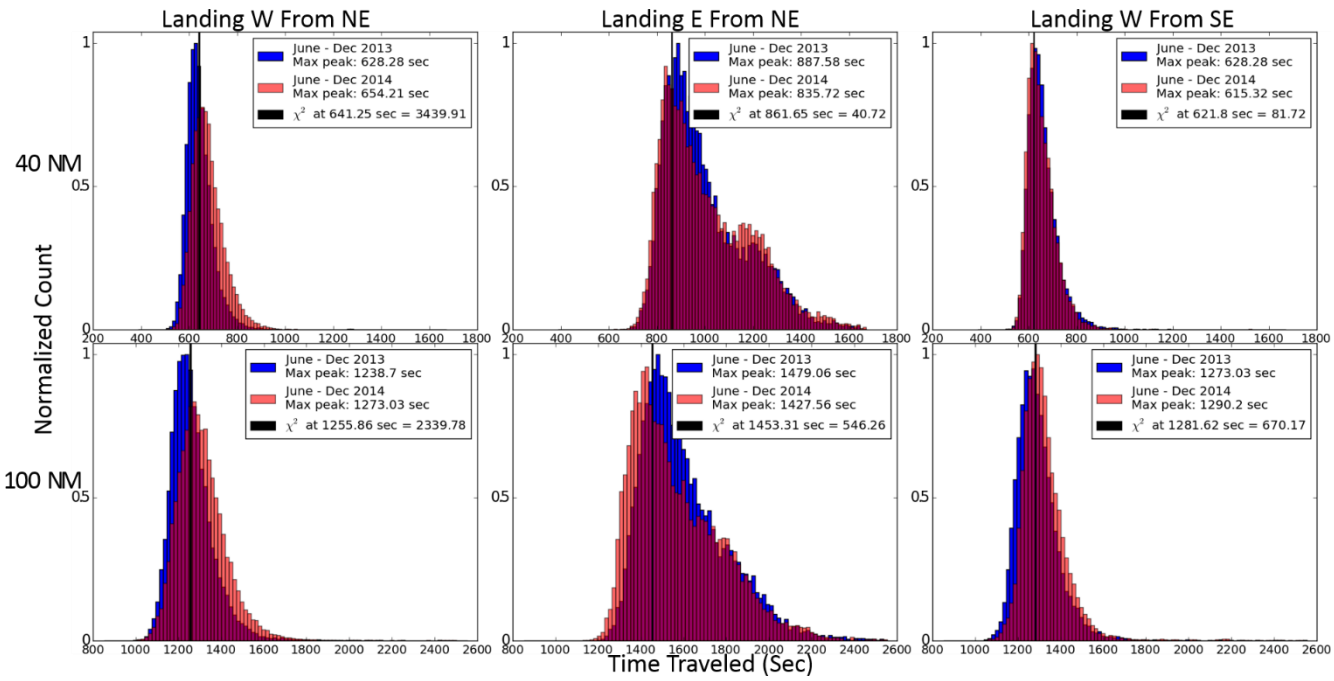

Figure 4. Histograms and χ^2 values for transit time distributions from 40 NM and 100 NM to the airport.

Additional analysis was performed to determine how well the new procedures were handling traffic in terms of time spent from 100 NM and from 40 NM to the airport. Histogram distributions were plotted for each flow from June through December for both 2013 and 2014 giving a comparison of "before" vs "after" implementation, while preserving the same seasonal effects in traffic flows. Many of the flows show no distinguishable changes in transit time efficiency before vs. after implementation; however, three flow distributions (shown in Figure 4) were found to have statistically significant shifts in distributions based on testing the χ^2null hypothesis (all significant χ^2 values are well above 7.879 signifying a null hypothesis probability<0.005 for 1-degree of freedom). Figure 4 shows the three most significant flow distributions and χ^2 values for the threshold halfway between the peak distributions (indicated by the vertical black line) for landing W from the NE, landing E from the NE, as well as landing W from SE. The times spent between both the 100 NM and 40 NM circle distances until landing have a significant shift in the peak distribution times. The distributions in blue are from before the Metroplex implementation while the distributions in red represent post-implementation behavior for the same 7-month time period from June through December for 2013 and 2014 respectively. The flow landing W from the NE shows an increase in the peak distribution's transit time for both the 40 NM (~25 sec.) and 100 NM (~35 sec.) circle distances; however, landing E from the NE flow shows a slight reduced time for 40 NM (~50 sec.) and 100NM (~50 sec.) circle distances. And in looking at the flow landing W from the SE there is a small significant change for the 40 NM (~15 sec) circle distance; however, at the 100 NM circle distance there is a slight increase in the peak distribution transit time (~20 sec).

After looking at all traffic times the analysis was narrowed to high traffic times to determine what impact the traffic had on the distributions. As before many of the flows do not show any distinguishable changes; however, the same 3 flow distributions are shown in Figure 5 for the high-traffic time periods. In keeping with the same color scheme as before the distributions in blue are from before, the Metroplex implementation while the distributions in red represent

post-implementation behavior for the same 7-month time period from June through December for 2013 and 2014 respectively. The χ^2 values in the legend are for the thresholds halfway between the two peak distributions and are indicated by the vertical black line. As before, the flow landing W from the NE shows an increase in both the 40 NM (~35 sec.) and 100 NM (~55 sec.), while the distributions for the flow landing E from the NE shows no significant changes (according to the χ^2test) in the 40 NM or 100 NM circle distances (it is important to note the flight counts for these high-traffic times are in the hundreds, which explains the sparse histogram distributions). The flow landing W from the SE shows similar distributions during all traffic times with no significant changes (according to the χ^2test) at the 40 NM range, but at the 100 NM range there is an increase in the peak distribution transit time (~55 sec.). With the traffic from the NE corner accounting for the largest number of arrivals at IAH with over 32% followed by the NW at 29% and the SE at 19%, any increases or decreases in time for these flows would have a significant impact on the overall airspace efficiency for arrivals. Though these distribution comparisons do not account for total flight time traveled from gate-to-gate, they do give some insight into noticeable shifts in the local traffic efficiency near the Houston Metroplex.

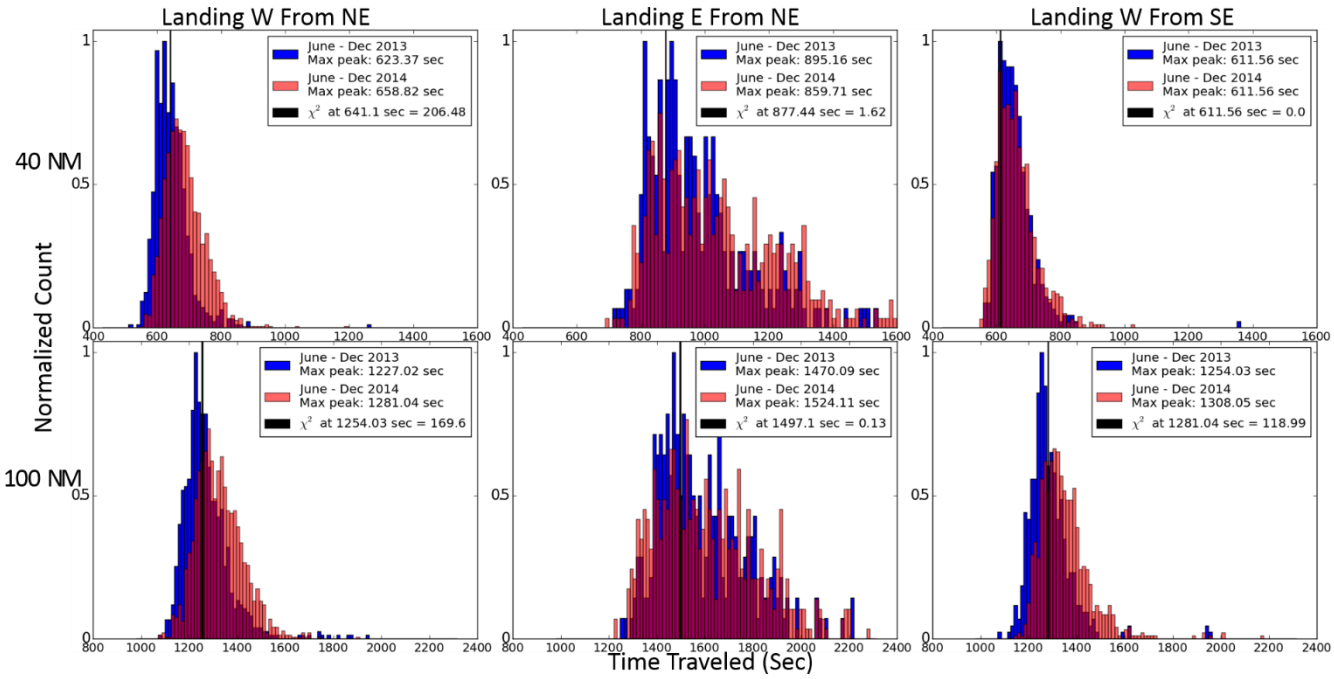

Figure 5. Histograms and χ^2 values for transit time distributions from 40 NM and 100 NM to the airport during high-traffic periods.

Findings: Individual Scenarios

After examining the high level-anomaly patterns per day the ranked list for individual flights was examined for situations that had an increased safety risk. The following 5 scenarios were presented to an active air traffic controller familiar with the I90 operations. Voice[4], when available, were used to help further explain the situations. Each of the following scenarios provides a description of the situation including weather, the aircraft's state (ground speed, altitude, heading, etc.), summaries by the subject matter experts/controller describing the possible explanation(s) of what may have led up to the flight's unusual behavior, and explains each scenario's potential relevance to safety risk. The Outer Marker (OM) is used as reference for arrivals on final approach. Target Aircraft (TGT) is used to denote the

[4] Voice source: http://www.liveatc.net/

flight of interest found by the algorithm. The Extended Runway Centerline (ERC) is used as reference for flights turning onto the final approach.

Configured to Land Long

Description: A Boeing 737 (TGT AC) arriving from a northwesterly direction is cleared for an approach to runway 8 Right (8R) at IAH. The TGT AC intercepts the 8R ERC 9 NM outside the OM at 9,000 feet with a ground speed of 314 knots. TGT AC then tracks an inbound flight arriving 1.1 NM outside OM at 4,100 feet and 256 knots at which time an S-turn is initiated to the south. Upon completing S-turn, TGT AC crosses ERC at intercept angle of 45 degrees, 2,100 feet, 252 knots and 2.2 NM from runway threshold. TGT AC does not turn to re-intercept the localizer but continues on its current heading, further descending to 1,900 feet. Thirty seconds later, while executing 360 turn TGT AC climbs back to 2,000 feet and proceeds to complete 360 turn re-intercepting localizer and landing on runway 8R. Approximately 5 minutes after TGT AC lands on 8R, a Boeing 737 (FLT49) from the west is also cleared for an approach to 8R and proceeds to intercept ERC at 8,500 feet, 318 knots, and 10.8 NM outside OM. FLT49 proceeds inbound crossing OM at 3,700 feet and 256 knots. Fifty seconds later FLT49 executes a go-around and 11 minutes later it is vectored back to the localizer and lands 8R (see Figure 6).

Explanation: The TGT AC and FLT49 are overnight "red-eye" (arriving in the early morning hours) flights from the US West Coast and were the second and third arrivals at IAH during the early morning time period. IAH was operating in a West configuration with the first arrival flight from the Southeast landing under the West operation on runway 27. The two arrivals from the West (upon entering the TRACON (I90)) were at typical altitudes one would expect to affect a normal descent to the downwind leg for a landing to the west, in this case on runway 26 Left (26L). Upon entering the TRACON, the TGT AC was advised by the IAH controller to expect 8R for landing and was questioned as to whether the pilot was able to execute the approach to which pilot responded that the pilot "may need S-turns" to comply. The pilot then made an S-turn just inside the OM in order to dissipate more altitude and reduce ground speed. After switching to the tower controller and crossing ERC, the pilot then requested a 360 turn and during next four minutes maneuvered to re-intercept the ERC at 900 feet, 175 knots, and 2.2 NM from runway threshold. In the case of FLT49, the aircraft was high and fast on approach resulting in the pilot informing the IAH tower that the pilot was "unable to make landing due to tailwind condition" leading to subsequent go-around. Wind conditions at the time were 7 MPH from the South with rain reported in the approach area (but the voice recording did not indicate rain affected either flight's approach).

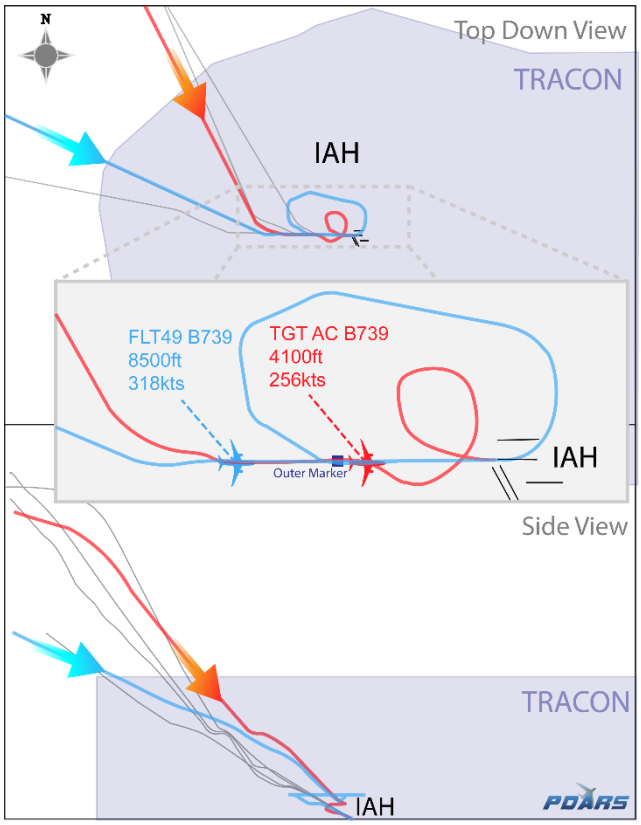

Figure 6. Configured to land long scenario.

Safety Risk Review: Both aircraft were cleared to land on the East runway but most likely were expecting to land to the West. In order to facilitate the landings on runway 8R both aircraft required steep rates of descent. By the time each aircraft reached 8R ERC, both were high, fast, and still attempting to stabilize approaches. TGT AC, after executing an S-turn and 360 turn, was able to re-intercept the localizer at 3.6 NM inside the OM and made a successful landing. On the other hand, FLT49 executed a typical go-around flight pattern and a preferred approach intercepting the localizer 2.6 NM outside the OM and crossing the OM at 2,000 feet and 210 knots. Both situations are accepted methods for bleeding off energy, but come with a trade-off of increased safety

risk. The MKAD rankings for these flights were 13 (TGT AC) and 40 (FLT49) out of 51 for the day. An explanation for why the 2nd flight was not ranked higher may be because the path was more of a typical go-around as compared to the 360 maneuver performed by TGT AC. With the implementation of the new procedures, adjusting the decent from cruise altitudes to the TRACON boundary requires more coordination with the ARTCC in relaying the expected runways to the pilots. This level of coordination is now on the order of 100 NM from the airport to give the pilots enough time to be able to hit the desired TRACON boundary altitudes for OPDs, whereas before, altitude level-off periods were built into the descents and allowed for more flexibility within the TRACON.

Long-Haul International Off From New Route

Description: A Boeing 747 (TGT AC) showed a flight path from the Southeast arrival fix that deviated from flight paths observed for the majority of arrivals to IAH on this approach (see Figure 7). Instead of tracking consistent with the normally observed approach the TGT AC tracks as far as 2.5 NM Southwest of initial portion of "backbone" routing while traversing inbound for right downwind leg for runway 8R. Once turning downwind TGT AC also tracked slightly north of normal approach procedure. An evaluation of the vertical profile shows TGT AC descending at a shallower descent rate (starting at FL230 and 12 NM before passing in the vicinity of the Southeast arrival fix). TGT AC then leveled at 17,000 feet for 2 minutes before reestablishing descent and remains above leveling altitude (12,000 feet) of a typical approach. TGT AC continued to maintain 1,300-2,000 feet above normal approach profile until reaching initial portion of downwind leg. Over the next 12 NM, TGT AC paralleled the normal downwind leg until reaching 6,000 feet, finally turning onto the base leg, and landing on 8R.

Explanation: Potential reasons for the non-typical approach were provided by IAH controllers suggesting that possibly the flight may not have been equipped properly and that taxiway restrictions were in effect for certain ground movements during the time period. A check of other days of traffic revealed that the international carrier (same call sign and aircraft type) executed the normal approach on other days and no abnormal Southeast arrival profiles were noted. Similarly, traffic evaluations verified that numerous Boeing 747 landings were made to all three East flow runways during the period of interest. Another explanation for TGT AC's trajectory abnormality may be that the pilot may have experienced some inflight issue requiring the need to deviate from normal flow. No voice communications were available to possibly clarify the reason an atypical approach was utilized.

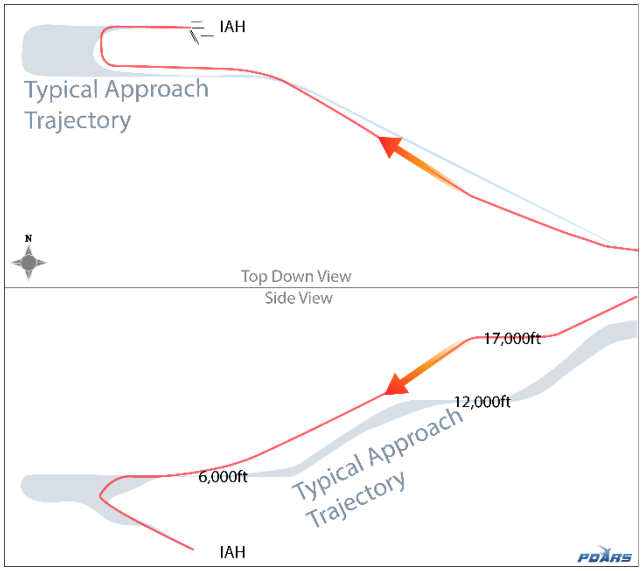

Figure 7. Long-haul international, off from new route.

Safety Risk Review: TGT AC did not fly the expected flight trajectory that the other Southeast arriving flights flew on the day evaluated nor on other days examined. The MKAD ranking for this flight was 1 out of 25 anomalies for the day since the track accounted for a significant deviation from the remaining flights within the flow. It is always best practice to fly the published approach to help reduce communication time and controller workload, which is one of the objectives of the Metroplex project. There was not enough available information pertaining to this occurrence to determine the cause for the atypical approach.

Dual Go-Around Crossing Paths

Description: Figure 8 depicts an Airbus 320 (TGT AC) on localizer to IAH runway 26 Left (26L) at the same time an Embraer E170 (FLT32) is established on the adjacent localizer inbound to runway 26 Right (26R). TGT AC continued inbound from the OM for 71 seconds to 900 feet and 2.2 NM from the runway threshold, at which time a go-around is initiated. Less than a minute later, FLT32, at 400 feet, also initiated a

climb at a distance of 0.6 NM from the runway 26R threshold. Both aircraft maintained runway headings during the missed approaches before initiating climbing right turns, only after crossing the far end of their assigned runways. The TGT AC commenced a right turn to the North reaching 1800 feet, followed 53 seconds later by FLT32 turning right at 1700 feet (TGT AC reached 3,000 feet by this point, see Figure 8). Loss of separation did not occur between aircraft during the tower controller's resequencing instructions to both aircraft. FLT32 was vectored inside of TGT AC track and landed 5 minutes before TGT AC landed on 8L.

Explanation: The TGT AC and FLT32 were not the only flights that executed missed approaches during this time. The flight on approach behind TGT AC for 26L also needed to execute a go-around upon reaching 900 feet. One minute later the airport initiated a change to an East-flow operation with the flight proceeding FLT32 continuing its approach and successfully landing on 26R. Also continuing its approach on an adjacent runway, a second flight trailing TGT AC on the localizer had to execute a missed approach after descending to 600 feet on the 26L localizer. Weather conditions during the time period of these go-arounds were significant factors leading to the landing difficulties experienced. In the time frame of these aborted approaches, the visibility decreased to .1 mile, and weather conditions consisted of .08-1.0 inches of rain, dropping atmospheric pressure, and wind shifts from North to SSW to ESE varying in speeds from 4.6 to 13.8 knots.

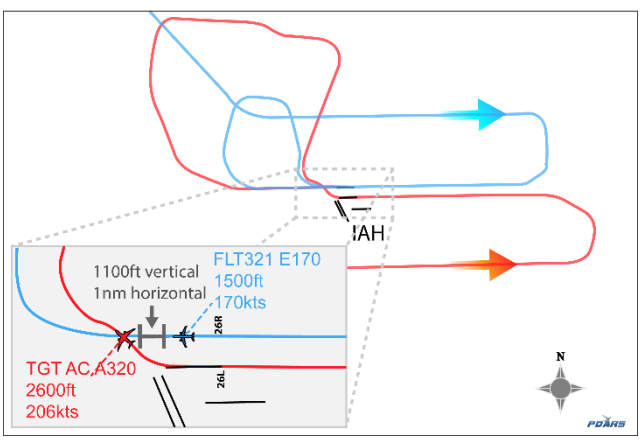

Figure 8. Dual go-around crossing.

Safety Risk Review: The published missed approach procedure for 26L is to maintain a westerly heading, climbing to 3,000 feet to hold at a designated fix, which TGT AC did not follow. The published missed approach procedure for 26R is to climb to 600 feet then make right turn climbing to 3,000 feet direct to a designated fix and hold. The MKAD ranking for TGT AC was 32 out of 71 for the day. No voice recordings were available for this occurrence. The significant weather at the airport and a switch in landing runways required the IAH controller to successfully manage uncommon traffic movements resulting from the operational complexities and increased workload which in turn raises the safety risk.

Climb Via Arrival Conflicts

Description: A Piper Cheyenne turboprop aircraft (TGT AC) approaching IAH from the Northwest and level at 6,000 feet made a southerly deviation from the typical right downwind pattern normally used by aircraft expecting a landing to the west. Just prior to the TGT AC's southerly maneuver, a McDonnell Douglas MD82 (FLT337) departing North crossed under the normal downwind pattern level at 5,000 feet. The two aircraft were vertically separated by 1,000 feet at the moment of the FLT337's tunneling, and were laterally separated by 4.8 NM.

Explanation: A discussion with an IAH controller provided the following insight into the most-likely explanation as to what happened in the above occurrence. In the months prior to the implementation of the Metroplex plan, an interim procedure called "climb via" was implemented to evaluate the workability of interactions between the arriving and departing procedures proposed. The procedure called for flights departing to the North to be restricted on their initial climb up to 4,000 feet until cleared to resume the climb. During the evaluation of the procedure, the IAH controller stated there were substantial occurrences of departures not complying with the climb restriction. In this case, it is possible that the controller may have continued the TGT AC's current heading to ensure separation before turning aircraft back to the right downwind pattern. Shortly after the time of this occurrence, changes were made to the test procedure to mitigate the "climb via" restriction situation. See Figure 9 for an illustration of the event.

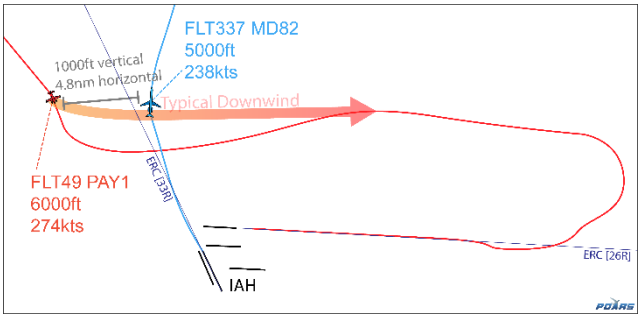

Figure 9. Climb via arrival conflicts.

Safety Risk Review: It is a possibility that the TGT AC overshot the turn for the desired right downwind pattern; however, this same southerly deviation on this arrival stream was also observed in several of the other traffic days evaluated for this study. MKAD ranked this flight 3 out of 37 for the day. Other such events were identified as well by MKAD with the following rankings 7 out of 37, 9 out of 38, and 9 out of 63 for their respective days. Due to the number of "busted clearances" in the past, the controller may have been cautious since FLT337 leveled at 5,000 feet, instead of the specified 4,000 feet climb restriction. This issue continued for several months before an incident occurred in July when a Singapore Airline and Delta Airline flights experienced a LoSS that was reported in the New York Times[5].

Unexpected Runway Closure Due to Weather

Description: An Embraer E190 (TGT AC) arriving from the East was sequenced to a left downwind between two Boeing 737s (FLT306 and FLT210). The event occurred during midday at a time when an extensive weather front 25 miles west of HOU was progressing in a Northwest direction moving away from the airport. In addition, pockets of isolated rain, high and gusting winds, and thunderstorm activities were present but scattered sporadically throughout the I90 area. The leading aircraft FLT306 approached HOU, established on runway 12R localizer. At 700 feet and 1.8 NM from the runway threshold, FLT306 made a right climbing turn to an initial heading of 210. Eighteen seconds later, the trailing TGT AC approaching the OM also initiated a right turn to 280. FLT210, the third landing aircraft, by this point established on the 12R localizer continued its approach for another 2.5 minutes before reaching the same location as TGT AC and then turned right to heading 180 maintaining 2,000 feet (1 NM outside the OM). Shortly after turning right, FLT306 continued its turn to the West as TGT AC maintains its initial heading until both reached 2,000 feet and were separated laterally by 3 NM, where upon FLT306 turned to the South, thus creating a diverging situation. The three go-around aircraft were then climbed by the HOU controller to orbiting patterns of 3,000, 4,000, and 5000 feet. FLT210 the third arrival in the sequence, orbited once and re-intercepted the ERC, landing 10 minutes after aborting the approach. TGT AC and FLT306 landed 15 and 20 minutes after initiating go-arounds, respectively. While the three aircraft were holding west of the airport, another arrival from the Southwest made a 360 spacing turn before proceeding to HOU for landing. Three other aircraft were placed in holding patterns at the TRACON boundary during the airport shutdown and after one holding pattern turn were able to make uneventful landings on 12R. Figure 10 illustrates the event.

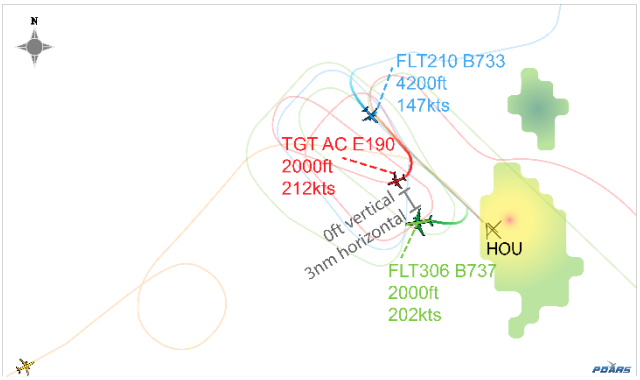

Figure 10. Unexpected runway closure due to weather.

Explanation: Although pilot voice communications were not obtained, a recording of voice transmissions for the HOU controller were available for analysis. Pertinent content from the recording is summarized: the HOU controller, while sequencing the three arrivals for landing 12R, informed the pilots of a 6 NM weather cell about to move over the airport. Three minutes later the HOU controller informed the approaching traffic that the HOU Tower had just

[5] http://www.nytimes.com/aponline/2014/07/05/business/ap-us-airplanes-near-collision.html?_r=1

declared IFR operations for Runway 12R and subsequently cleared FLT306 for landing. Two and a half minutes later, the HOU controller cancelled FLT306's approach and instructed aircraft to "turn right to 210, climb maintain 3,000 feet." TGT AC meanwhile was inbound to the OM and the HOU controller questioned the pilot as to whether the Runway Visual Range is less than the pilot needed to execute a landing. HOU controller "rogers" unknown pilot's response but immediately cancels the approach clearance, instructing TGT AC to turn right heading 280 and to maintain 2,000 feet. In the interim, FLT306 has turned further right from originally assigned heading 210 and was approaching 2,000 feet (same altitude as TGT AC) and closing towards a lateral separation of 3 NM. The HOU controller issued an immediate left turn to heading 180 with FLT306 complying. These three aircraft were next climbed by the HOU controller to separate altitudes and enter vectoring orbits until the airport weather cleared allowing re-sequencing for landing.

Safety Risk Review: FLT306 was originally given a heading of 210 with climb instructions to 3,000 feet and the TGT AC seconds later issued a heading of 280 maintaining 2,000 feet. FLT306 was further turned to a more westerly heading and while in the turn was approaching a potential loss of separation with TGT AC when the HOU controller immediately turned FLT306 to a diverging 180 heading, mitigating the situation. The MKAD algorithm identified the aircraft arriving from the SW, which was ranked 1 out of 30 due to its unusual maneuver closer to the TRACON boundary. The HOU controller's voice recording implies that TGT AC was maintaining visual contact on FLT306 during this time and that the HOU controller acknowledges TGT AC back-up assistance (since pilot may have called out visual on traffic before the controller had advised).

Conclusions

The analysis of the Houston Metroplex implementation highlighted some significant differences in flow times between the same periods in the previous year as compared to after the procedural change. Although these differences only account for the local airspace from a radius of 100 NM and may not tell the complete story, the findings do give some insight into the airspace operations and could be used as future benchmarking metrics. Further analysis of complete gate-to-gate travel times may provide additional metrics that could capture a more system level view and help better understand the relative importance of the transit time differences.

Along with the traffic flow analysis, some operationally significant events were discovered using the MKAD tool. Some events such as the situation where earlier runway change coordination is needed between TRACON and the ARTCC in early morning operations, as well as the "climb via" procedure with arrivals, uncovered some unintended outcomes from the new procedure implementations. In the "climb via" scenario, a number of ATSAP reports could have been filed to report the issue; however, having an unbiased data mining algorithm such as MKAD automatically identify a set of similar events may have helped highlight this issue earlier on and expedited a corrective action before a more significant safety risk occurred, such as the one in July of 2015. With the current plan to increase traffic flow while maintaining high safety standards, an automated data mining tool such as MKAD can help proactively identify previously unknown risks and enable the alleviation of some of the growing pains of the future NAS.

Future Work

MKAD is currently undergoing further evaluation to determine ways to improve its results. These evaluations include: exploring ways to more intelligently combine the kernels in an optimal way, as well as including additional reports such as the PDARS Turn to Final report, which has been found to increase the performance of the algorithm in detecting operationally significant events instead of just purely statistically significant ones. Other activities to be pursued involve performing active learning experiments where SME feedback is incorporated by automatically generating additional features from the feedback to help better rank the operationally significant events. The active learning technique involves a feedback loop with SMEs that can update the model by incorporating the labels and feedback, which can then be applied to the next day's ranked list of anomalies and better improve the performance. Automatic grouping of anomalies that have similar features or occur in the same vicinity would also help to identify trending safety risks and "hot spots" in the airspace. In addition to the algorithmic improvements, other Metroplex implementations could be targeted for future evaluations by the algorithm. These include: Atlanta, Charlotte, Denver, Florida, Northern

California, North Texas, Southern California, and Washington DC. The FAA has expressed interest in evaluating the tool on additional data using FAA servers, with the possible intent of maturing the technology after evaluation. This effort would have the long-term goal of developing a system that facilities could use to characterize and monitor the increased safety risks of flights within each facility's airspace. With the ability to detect previously unknown safety risks, proactive measures may be taken to prevent more sever safety events. Ultimately this tool would be a valuable addition to the current safety risk assessment methods that the FAA currently uses and, if fully matured, can automatically discover situations that may ultimately help in increasing the overall safety of the NextGen NAS and complement current methods.

References

[1] Cavolowsky, J., and A. Sultan, "NASA Aeronautics Airspace Operations and Safety Program - Discussion with the ARTR," presentation NASA HQ Washington DC, September 26, 2014.

[2] Matthews, B., Srivastava, A., Schade, J., Schleicher, D., et al., "Discovery of Abnormal Flight Patterns in Flight Track Data," AIAA Paper No. DOI: 6.2013-4386 AIAA ATIO Conference, Los Angeles, CA, August 12-14, 2013.

[3] Matthews, B., Nielsen, D., Schade, J., Chan, K., Kiniry, M. "Automated Discovery of Flight Track Anomalies," DOI: 10.1109/DASC.2014. 6979469 Digital Avionics Systems Conference, Colorado Springs, CO, October 5-9, 2014.

[4] den Braven, W., and J. Schade, "Concept and Operation of the Performance Data Analysis and Reporting System," paper number 2003-01-2976, SAE Advances in Aviation Safety Conference (ACE), Montrèal, Canada, September 8-12, 2003

[5] S. Smith, "Metroplex Phased Approach", Managing the Skies, pp. 19-23 July/August 2014.

[6] Office of the Inspector General Report No. AV 2013 - 046, "FAA's Efforts to Track and Mitigate Air Traffic Losses of Separation are Limited by Data Collection and Implementation Challenges," Report No. AV-2013 – 046, February 27, 2013.

[7] Harkleroad, E., A. Vela, J. Kuchar, B. Barnett, and R. Merchant-Bennett. "Risk-based Modeling to Support NextGen Concept Assessment and Validation." MIT Lincoln Laboratory March 2013.

[8] S. Das, B. L. Matthews, A. N. Srivastava, and N. C. Oza. Multiple Kernel Learning for Heterogeneous Anomaly Detection: Algorithm and Aviation Safety Case Study. In Proceedings of SIGKDD'10, pages 47-56, 2010.

[9] Schölkopf, B., and A. Smola. "Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond," MIT Press, 2001.

[10] Jensen, L. J.; Bateman, A. "The Rise and Fall of Supervised Machine Learning Techniques". Bioinformatics 2011, 27 (24), 3331–3332

[11] Ellen A. Cherniavsky and Thor R. Abrahamsen "Aviation System Performance Metrics: Airport Utilization," MITRE Center for Advanced Aviation System Development, McLean Virginia, Technical Document: MP 00W0000090, April 2000

Acknowledgements

We would like to acknowledge the NASA Aeronautics Mission Directorate, Aviation Safety Program, System-Wide Safety and Assurance Technologies project and Shadow Mode Assessment using Realistic Technologies for the NAS project for funding this work. We would like to thank the FAA PDARS Program Office for their approval to analyze the PDARS operational NAS data. We also would like to thank the subject matter experts: Steve Wyloge, and the on-site controller for their insightful comments and perspective on the identified events. Finally we would also like to thank Peset Tan for his graphical expertise that helped clarify the anomalies we discovered.