AIAA 2013-4386
AIAA Aviation
August 12-14, 2013, Los Angeles, CA
2013 Aviation Technology, Integration, and Operations Conference

Discovery of Abnormal Flight Patterns in Flight Track Data

Bryan Matthews
SGT Inc.
NASA Ames Research Center
Moffett Field CA, USA
bryan.l.matthews@nasa.gov

Ashok N. Srivastava
NASA Ames Research Center
Moffett Field CA, USA
ashok.n.srivastava@nasa.gov

John Schade, Dave Schleicher, Kennis Chan,
Richard Gutterud, Mike Kiniry
ATAC Corporation
Santa Clara, CA, USA
[jes, drs, koc, rag, mdk]@atac.com

***Abstract*— The National Airspace System (NAS) is an ever changing and complex engineering system. As the Next Generation Air Transportation System (NextGen) is developed, there will be an increased emphasis on safety and operational and environmental efficiency. Current operations in the NAS are monitored using a variety of data sources, including data from flight recorders, radar track data, weather data, and other massive data collection systems. Although numerous technologies exist to monitor the frequency of known but undesirable behaviors in the NAS, there are currently few methods that can analyze the large repositories to discover new and previously unknown events in the NAS. Having a tool to discover events that have implications for safety or incidents of operational importance, increases the awareness of such scenarios in the community and helps to broaden the overall safety of the NAS, whereas only monitoring the frequency of known events can only provide mitigations for already established problems. This paper discusses a novel approach for discovering operationally significant events in the NAS that are currently not monitored and have potential safety and/or efficiency implications using radar-track data. This paper will discuss the discovery algorithm and describe in detail some flights of interest with comments from subject matter experts who are familiar with the operations in the airspace that was studied.**

Keywords: PDARS; safety; anomaly detection; data mining; anomaly discovery.

I. INTRODUCTION

The National Airspace System (NAS) is perhaps one of the most complex dynamical systems created by humans. It safely transports millions of people with high reliability and extremely low accident rate. As new technologies and procedures are introduced in the NAS during its evolution to NextGen, a critical requirement is that efficiency rate improves while maintaining or exceeding current safety standards. NASA, in partnership with the FAA and industry, is developing new technologies that can automatically monitor the massive data sets being collected in the NAS to discover precursors to potential safety events. These decision-support systems can be used by analysts to track, identify, and compute the trend of emerging safety issues. We contrast this approach to discovering previously unknown safety events with the more traditional approach of computing the frequency of known safety events. The former approach has the potential to discover emerging safety issues, which are previously unknown, whereas the latter approach can be used to monitor the occurrence of known problems. We believe that to develop a truly comprehensive approach to ensuring safety in operations, both approaches need to be pursued, since the proposed method in this paper can be used to compliment the current state- of-the-art. This paper discusses a novel approach for discovering statistical anomalies that may have operational significance. The ability of this approach to discover anomalies in massive data sets comprised of continuous and discrete data streams from flight recorder data has already been established [1] and [2] . This paper presents the approach adapted to work with data generated from radar tracks. The power of the approach that we have taken is that it can be adapted to different data sets to discover many different types of anomalies. This paper discusses the adaptation we made for radar-track data in detail and shows examples of five operationally significant anomalies.

The paper is organized as follows. We begin with a background on the Performance Data Analysis and Reporting System (PDARS) which is an FAA program that captures much of the data used in this study, and then discuss the current approaches to monitoring the frequency of known anomalies. We then discuss the discovery algorithm and the data used for this study. We next move to a discussion of five examples of operationally significant anomalies. Each example is concluded with a safety analysis. We then discuss the conclusions and address future research.

II. BACKGROUND

A. PDARS Program

The Performance Data Analysis and Reporting System (PDARS) provides Federal Aviation Administration (FAA) decision makers at multiple levels of the Air Traffic Organization (ATO) with a dynamic set of comprehensive tools and methods for monitoring the health, safety, and efficiency of day-to-day Air Traffic Control (ATC) operations. PDARS itself is a product of innovative, collaborative research between NASA and the FAA recognized for its excellence by receiving the NASA Administrator's Turning Goals into Reality (TGIR) award in 2003 and achieving full technology transfer to the FAA in 2005 [3]. The PDARS program is managed by the ATO Office of System Operations Services and is routinely used operationally by over a dozen organizational units within the FAA.

The core of PDARS consists of an ever-evolving aviation data collection, processing, and dissemination platform able to accept nearly any surveillance or otherwise collected positional data and merge that with other geo-referenced or contextual aviation related data (e.g. weather, terrain, schedules, etc.) to produce information that is "actionable" to decision makers at multiple levels in a complex Air Navigation Service Provider (ANSP) organization such as the FAA. The development of PDARS has been from the beginning and continues to be driven by the needs of the user base: those actively involved in direct operation of the NAS and the associated challenging areas such as safety, efficiency, and environmental concerns [4].

PDARS processing creates the best four dimensional trajectories possible from existing sensor data or other inputs with a minimum of smoothing or other techniques in order to preserve as closely as possible the measured trajectory of each flight. The processed flights are stored in a database to allow for immediate access for operational analysis and maintained in a data warehouse to allow for historical and trend analysis. Naturally this large collection of data is an excellent source for the application of advanced data mining technologies such as those presented in the research here.

With approval from the FAA, we applied the anomaly detection algorithms discussed in this paper on a portion of the PDARS data warehouse. The approach currently used focuses on measuring the frequency of occurrence of known events based on previously identified issues. With such a large collection of accurate performance data available, the prospect of discovering previously unknown anomalies in the data set was intriguing to both NASA researchers as well as FAA personnel from a safety and operational efficiency perspective.

B. Current State-of-the-art Research

Air Navigation Service Providers (ANSPs) around the world are being challenged to constantly improve their ability to identify and mitigate current and emerging safety risks. Over time there has been an historical transition from a focus on measuring Outcome-based safety metrics focused on analyzing incidents and accidents (through internal governmental incident reporting [5] as well as publicly-available annual safety reports [6]), to measuring Process-based safety metrics (such as those focused on implementing Safety Management System processes [7]) to a more recent focus on measuring Precursor-based safety metrics (with approaches such as that pursued by NASA's National Aviation Operations Monitoring Service (NAOMS) project [8]. This transition has moved from a more reactive safety analysis approach to a more proactive, data-driven approach focused on searching for aviation system risks.

A recent report from the Governmental Accountability Office (GAO) recommends significant improvement in both data quality and analysis capabilities [9]. As new Air Traffic Management systems like US's Next Generation Air Transportation System (NextGen) [10] and Europe's Single European Sky ATM Research (SESAR) [11] get implemented in tandem with increased air traffic demand, there is all the more sensitivity and concern for successfully identifying existing and new safety risks as early as possible and developing and using new analytical techniques.

Recent safety research has focused on the development of new precursor-based safety analysis of historical aviation operations data as opposed to simulation-based risk modeling methods of ATM procedures and technology using tools such as NLR's Traffic Organization and Perturbation AnalyZer (TOPAZ) [12] or real-time simulations such as those exercised in the European En Route Air Traffic Soft Management Ultimate System (ERASMUS) project1. These emerging research approaches include new safety analysis tools, automated safety data analysis techniques, and quantitative analysis of new data sets.

The international ATM community has been developing advanced tools to perform safety analysis off of operational or simulated data. As one recent example, France's ANSP, the Direction des Services de la Navigation Aérienne (DSNA), has developed the termed BISCOT (human Based risk and decision taking Complexity integrated tOolkiT) toolkit that has been focused on performing collision risk modeling from both radar surveillance data and planned airspace changes [13].

In addition, new safety data mining and analysis techniques are being developed such as the automatic safety data gathering and reporting being performed by EUROCONTROL's Automatic Safety Monitoring Tool (ASMT) [14]. Such automated safety monitoring tools support a range of data-driven and operational analysis-driven safety analyses such as baselining, trend analysis, correlation analysis, and propagation analysis.

A multitude of global aviation data sets have been analyzed with a safety operational analysis focus. More recently, some success has been made with aircraft Flight Operational Quality Assurance (FOQA) data [2], regional air traffic data [15], and even text-based incident reports [16].

Taking a step forward, our research had as its goal the data-driven identification of safety precursors in a multi-air traffic facility region's surveillance data over a multi-year time span. The ultimate goal of our research is to have an operational safety risk assessment tool for ANSPs that can use an automated

1. http://www.atm-erasmus.com/

quantitative approach to identifying new operational safety risks in the current aviation system and serve as an "early warning system" upon the introduction of new ATM technology and procedures.

C. Multiple Kernel Anomaly Detection: Proposed Algorithm

The algorithm used in this study is built upon the one class support vector machine (1-class SVM), which is a method developed in the field of machine learning to perform anomaly detection using a kernel function that measures the pairwise similarity between patterns. This algorithm discovers statistical outliers that may have operational significance. The MKAD (Multiple Kernel Anomaly Detection) [2] algorithm was developed at NASA and is a specific example of a 1-class SVM. The key property of the algorithm lies in its ability for information fusion from a variety of sources of varying nature. For example: the algorithm can combine data from flight data recordings, radar tracks, text reports, weather information, etc. To achieve this it constructs a kernel using a similarity function for each source and combines the kernels to make a decision on which patterns are atypical. As the MKAD algorithm name implies multiple kernels are generated for each feature (for this study latitude, longitude, and altitude tracks were used) and combined to form a single kernel for the 1-class SVM optimization. The combination of kernels is done with an equal weight-age average across the 3 feature kernels as shown in equation 1. In this study n is equal to 3.

Combined Kernels:

$$\kappa(\vec{x}_i, \vec{y}_j) = \frac{1}{n}\sum_{m=1}^{n} \kappa_m(\vec{x}_i, \vec{y}_j) \qquad (1)$$

As in all 1-class SVMs a similarity metric (or kernel function) must be defined that is appropriate for the data and domain. In this study the similarity function that was chosen for all features is the cosine similarity function [17] shown in equation 2:

Kernel Function: $$\kappa_n(\vec{x}_i, \vec{y}_j) = 1 - \frac{\vec{x}_i \cdot \vec{y}_j}{\|\vec{x}_i\|\|\vec{y}_j\|} \qquad (2)$$

This kernel was chosen for its similarity properties, which has a straightforward implication in the kernel space for this particular domain. For example: flight tracks that approach from the same direction will have similarity values close to 1 while flight tracks that approach from the opposite direction, i.e. mirror images, will have a similarity values close to 0. This similarity interpretation is consistent for anomaly detection using 1-class SVMs. It is important to note that users are not limited to using the cosine similarity function and may choose from various other similarity functions [17].

While the cosine similarity kernel does not suffer from the added complexity of having to optimize tuning parameters, it is sensitive to bias and scaling affects in the data. To account for these effects each feature was normalized to unit length by using the feature's global maximum and minimums.

Once the kernel is constructed the 1-class SVM can be derived using with the following optimization problem and constraints shown in equation 3 and 4.

Minimize:

$$Q = \frac{1}{2}\sum_{i,j} \alpha_i \kappa(\vec{x}_i, \vec{y}_j)\, \alpha_j \qquad (3)$$

Subject to:

$$0 \le \alpha_i \le \frac{1}{\eta \upsilon}, \sum_i \alpha_i = 1, 0 \le \upsilon \le 1 \qquad (4)$$

The parameter υ is provided by the user and corresponds to the maximum fraction of data assumed to be anomalous (for this study υ was set to 5%.). The SVM algorithm will attempt to separate the anomalous data from the nominal with the data points located along the boundary or hyperplane called support vectors. The support vectors are used to define the separating hyperplane and are given by: $\{x_i | \alpha_i > 0\}$. The threshold ρ can be calculated to ascertain whether a flight is an anomaly or not and is derived in equation 5.

$$\rho = \frac{1}{length(\alpha)}\sum_{j\in\alpha}\sum_{i\in\alpha} \alpha_i \kappa(\vec{x}_i, \vec{y}_j)\,;\ where\!: \rho \ge 0 \qquad (5)$$

The examples that are determined by the algorithm to be anomalous are rank ordered based on their distance to the hyperplane. The anomalous examples are located on the negative side of the hyperplane, whereas the nominal examples are on the positive side. The magnitude determines the severity of the anomalous example. Only negative examples are marked as anomalous. The calculation of the scores is shown in equation 6.

$$Score(y_i) = \sum_{i\in\alpha} \alpha_i \kappa(\vec{x}_i, \vec{y}_j) - \rho \qquad (6)$$

It is important to note that the examples labeled by the algorithm as anomalous are purely statistical anomalies but not necessarily operationally significant. Subject matter experts still must assess the scenarios and determine whether the examples hold some operationally relevance or not. This tool merely identifies unusual patterns in the data and rank orders them by how unusual they are, but it is up to the user to interpret the results.

D. The Data Management Process

The US trajectory data used in this study comes from the Automated Radar Terminal System (ARTS) computer from the Southern California Terminal Radar Approach Control Facility (SCT), the host computer at Los Angeles Air Route Traffic Control Center (ZLA) and the Airport Surface Detection Equipment Model-X (ASDE-X) system at Los Angeles International Airport (LAX). With FAA approval, NASA was given access to PDARS data from October 2009 to September 2011 from these three surveillance data processing systems. Some surveillance data from the FAA facilities contain data from multiple radar sensors. For example: data from SCT contain data from 11 radar sensors. The coverage for those sensors can overlap. The PDARS system picks the best radar hits to use based on many different criteria in order to produce the best quality of four dimensional (latitude, longitude, altitude, and airspeed) trajectories for flights. Runway information that is not stored in the FAA facilities databases is also computed by the PDARS system based on the geometry of the flights.

Since traffic is coming from different directions for each runway, the data mining algorithms were applied to arrival flights per runway. This study focuses on all the busy runways (25R/07L, 25L/07R, 24L/06R and 24R/06L) at LAX. Around 641,000 flights were analyzed to obtain the final results. This data source provides a much better traffic picture than its counterparts, such as the Enhanced Traffic Management system (ETMS) or the ASD Feed for Industry (ASDI) [4]. Generally, PDARS data has higher sampling rates and trajectory resolution than ETMS/ASDI data. The data was then stitched together to provide smooth trajectories between the ZLA airspace boundary and the gates at LAX. Since the study was focused on finding unusual patterns in commercial aircraft, Visual Flight Rules (VFR) flights with beacon codes from 1200 to 1299 and sensitive flights such as military flights were removed from the data. The additional benefit of removing those flights is that military and VFR flights typically will have unusual flights paths as compared to commercial flights, and by removing those flights the tool is expected to yield better results during the data mining process. Figure 1 outlines the process for collecting and merging the flight trajectories together.

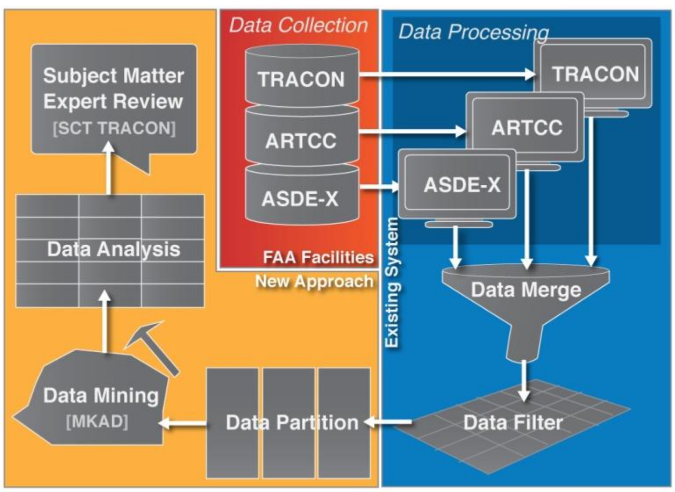

Figure 1: The above process diagram shows the work flow from collection of data at the FAA facilities to the final review of the anomalies by air traffic controllers.

Once the sensitive flights are filtered out the data is partitioned into set of trajectories landing at each runway for each month. Depending on the runway usage the partitions can range from hundreds of flights up to 10,000 flights for some of the busier runways. The flights within the partition are compared against each other as described in Section C and used by the MKAD algorithm. The algorithm is run to compute the outliers, which are reported to data analysts for examination. Trajectories of interest are investigated further with the graphical analysis tool Graphical Airspace Design Environment (GRADE). The overall traffic flow is visualized with the tool to obtain a better understanding of the airspace for each situation. After some scenarios of interest have been identified by the analysts, the key characteristics are summarized with animations and presented to the subject matter experts familiar with the airspace. For this work air traffic controllers from the Southern California TRACON (SCT) facility were consulted with to gather feedback on the anomalous trajectories. Their summaries are presented in Section III.

III. RESULTS

Approximately 40 flights were selected from the list of anomalies identified by the algorithm for further analysis by subject matter experts. Out of these, 15 were deemed to hold some operational significance. These flights were presented to two SCT TRACON controllers familiar with the everyday operations at the center. For brevity five representative occurrences of the 15 will be discussed in this section. Each occurrence discussed will provide a short synopsis of the situation, offer the controller's feedback for the possible explanation(s) of what may have lead up to the flight's unusual behavior and describe each occurrence's relevance to safety. In the following descriptions the aircraft identified with the unusual trajectory is referred to as TGT AC. Other flights in the airspace will be referred to as FLTXX. The north and south complexes at LAX refer to runways 24/06 and 25/07 respectively. The Extended Runway Center (ERC) lines are shown in the figures to give a sense of horizontal alignment with the runways for the aircraft's turn to final.

A. Occurrence 1

A B737 (TGT AC) landing at LAX (west configuration/north complex) is issued a 360 degree maneuver on base leg for runway 24R. Weather at the time was visual meteorological conditions. Figure 2 is a PDARS GRADE graphic showing TGT AC relative to a nominal flight path for intercepting the final approach course for runway 24R.

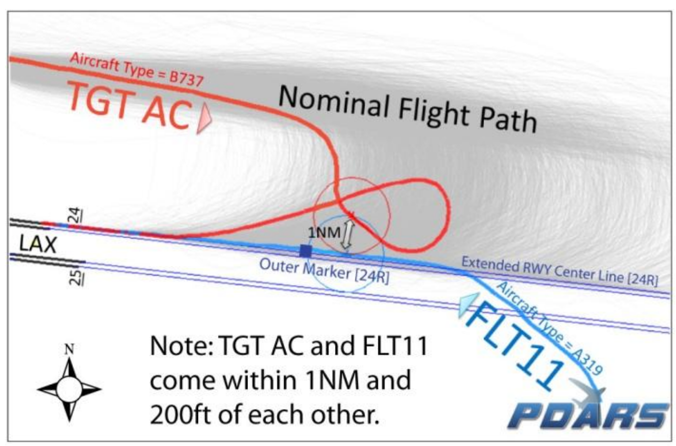

Figure 2: Above shows a 360 maneuver to avoid the traffic approaching from the south resulting in an interception of ERC in close to the airport:

The following are possible explanations/observations for this abnormal approach event. On the downwind leg the TGT AC was in close proximity to another landing flight (FLT11). Visual separation may have been used, or the TGT AC had the wrong aircraft in sight. FLT11 may have been originally cleared for the south complex, but then assigned to the north complex (probably to accommodate pilot's request for the north runway). One possible scenario was that the controller misjudged and thought there was an adequate hole in traffic flow, subsequently the TGT AC was vectored too sharply resulting in the controller having to use the 360 degree turn in order for TGT AC to be re-sequenced back into the traffic flow. Another scenario could have been that the controller realized late that there was a lack of spacing on the TGT AC and ultimately issued a 360 degree turn for spacing. It may have been possible that the TGT AC could have taken another aircraft's instructions and turned too early, and since there was a foreign carrier involved in the landing traffic at this time a communication problem may also have existed.

Safety Review: The action by the controller and or pilot resolved a more serious situation. Although TCAS was probably alerting the pilot, with parallel approaches at LAX the pilots may receive multiple traffic alerts (TA) or resolution alerts (RA), thus creating a more complex situation. This situation displays one course of action that prevented continued loss of separation.

B. Occurrence 2

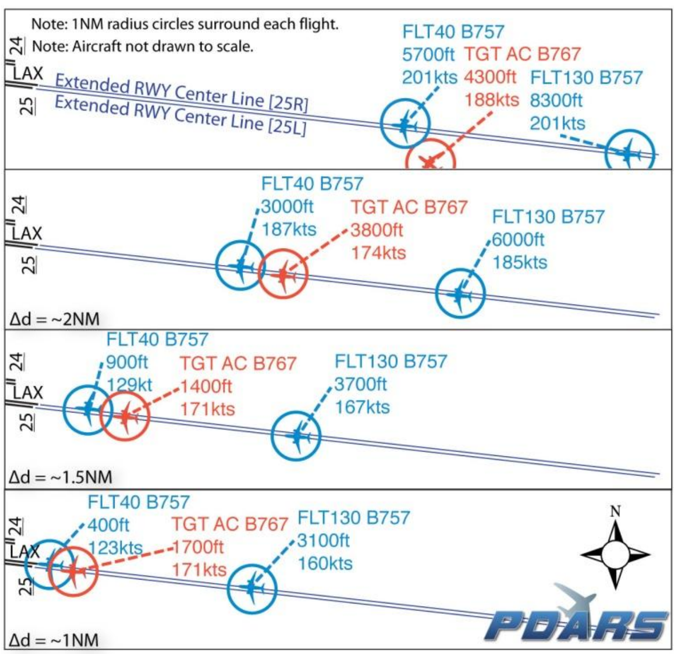

Figure 3: Above shows an impending overtake of the proceeding aircraft on a parallel runway during final approach. This was averted by executing a go-around during final approach.

A B767 landing at LAX (west configuration/south complex) is issued a go-around due to overtake of the proceeding aircraft on final approach to runway 24R. Weather at the time indicated few clouds at 1,500 feet and visibility 10 statute miles. Figure 3 is a GRADE graphic showing a sequential closure of separation distance and speed differential between the TGT AC and FLT40 that resulted in the necessity for executing a go-around.

The following are possible explanations/observations for this abnormal approach event. At 1.3 nm from the runway threshold the TGT AC is exhibiting a greater than 40 knot overtake on FLT40. TRACON or LAX Tower probably instructed TGT AC (B767) "don't overtake FLT40 (a B757) on final for runway 25R" (this is normal procedure involving a heavy jet). If wake turbulence were a potential cause for the go-around, the TGT AC should have been issued a go-around earlier; therefore, it appears that the controller did not think wake turbulence (overtake) was an issue during the approach (inboard traffic on adjacent runway 25R was a B757). Another possibility was that TGT AC may have been too fast to land on runway 25L (an unstable approach), and aircraft could not reduce speed enough to remain behind inboard traffic (B757) on approach to runway 25R.

Safety Review: apparently the trail aircraft was not able to reduce airspeed and configure the aircraft for an approach to remain abreast of or behind the inboard traffic. The Controller resolved the situation with the re-sequencing, thus avoiding a wake turbulence issue at the point of touchdown.

C. Occurrence 3

A B747 initially on left downwind for landing on runway 24R and just abeam the airport is vectored in a teardrop flight profile to the south complex for a landing on runway 25L. Weather at the time is overcast clouds with a low ceiling of 500 feet. As Figure 4 illustrates, after crossing south of the runway 25L ERC, the TGT AC is subsequently turned too early by the controller resulting in an extreme intercept angle of approximately 59 degrees which leads to the overshooting of the runway 25L ERC and crossing into the north complex traffic.

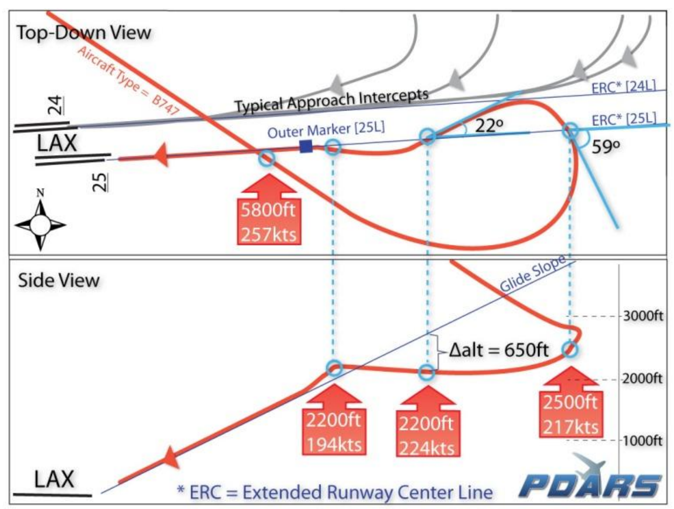

Figure 4: Above shows an aircraft turning too early resulting in overshoot of the ERC and drifting into the north complex.

The following are possible explanations/observations for this abnormal approach event. The controller's clearance to turn and the intercept localizer (LOC) was made too early (for a B747) thus forcing the aircraft into an unstable approach, although pilot accepted the clearance. The aircraft was vectored over the top (a seldom used LAX operation, normally aircraft are vectored on the north downwind) therefore resulting in a 360 degree turn to the downwind, a difficult maneuver for a B747. The B747 flight path should have been extended further southeast to allow for a successful turn-to-final and avoiding an overshoot turn situation. Since there was no traffic to the area south of the B747 decent trajectory the opportunity did exist to extend the approach further to the southeast, which would allow a normal descent and turn. Even during visual conditions the controller should have extended the aircraft further to provide an appropriate turn to final.

Safety Review: A very unusual operation for a B747 and interesting that the pilot accepted the clearance for such a steep turn to final. Fortunately, the north downwind traffic was not a factor, which could have resulted in a wake turbulence situation if separation was not maintained.

D. Occurrence 4

An A320 approaching over the water during midnight configuration operations (2:40 a.m.) for landing on runway 06L at LAX executes a 360 degree turn inside the Outer Marker (OM). Weather indicated overcast conditions with cloud bases at 400 feet with rain, mist, and east winds. The TGT AC speed at the OM on first approach was excessive (205 knots) and when speed did not dissipate enough inside the OM, a 360 degree turn was initiated for re-sequencing (See Figure 5).

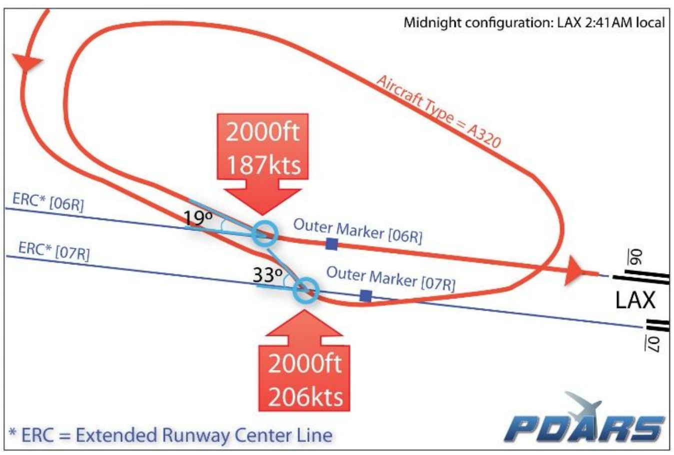

Figure 5: Above shows a high energy descent followed by a 360 degree turn executed inside of the outer marker.

The following are possible explanations/observations for this abnormal approach event. No other LAX traffic was in the vicinity of airport during initial approach. Normally with no traffic, a controller would not issue speeds to aircraft and the pilot would be responsible for the speed of the aircraft on approach. Since the aircraft approached the OM too fast, although at an appropriate altitude, the TGT AC probably requested a go-around and/or a re-sequence back to the airport indicating LAX Tower approved an early turn back over the ocean. It appears that the aircraft was unstable and did not acquire the glide path and was not able to fly the approach, and/or the pilot may have selected the wrong localizer or was not configured properly for final approach, i.e., too fast.

Safety Review: An over the water operation was observed in several reviews at LAX involving re-sequencing or go-arounds due to failure of the aircraft to be configured for landing or the controller turning the aircraft onto the approach high and fast. Most of these operations occur from midnight to 6am, which could be contributed to pilot fatigue, unawareness of surface winds, or unprepared for a tight turn to final. Awareness by flight crews and controllers would resolve this risky operation in the future.

E. Occurrence 5

A B757 on the north downwind is cleared for visual approach to runway 25L at LAX with the initial intercept of the ERC attempted at a location inside of the OM. Radar data indicates the aircraft was high and fast. This is a preferential runway for that airline since parking is located at the south complex. After a couple of attempts to establish the aircraft on the final course, the B757 still appears to be in an unstable landing configuration and the pilot elects to execute a go-around. The B757 was vectored south of the airport below Class B airspace at 3,000 feet (floor of Class B at 5,000 feet). Weather at the time indicated few clouds at 3,000 feet and the visibility was good, the operation occurred between 11pm and midnight (See Figure 6).

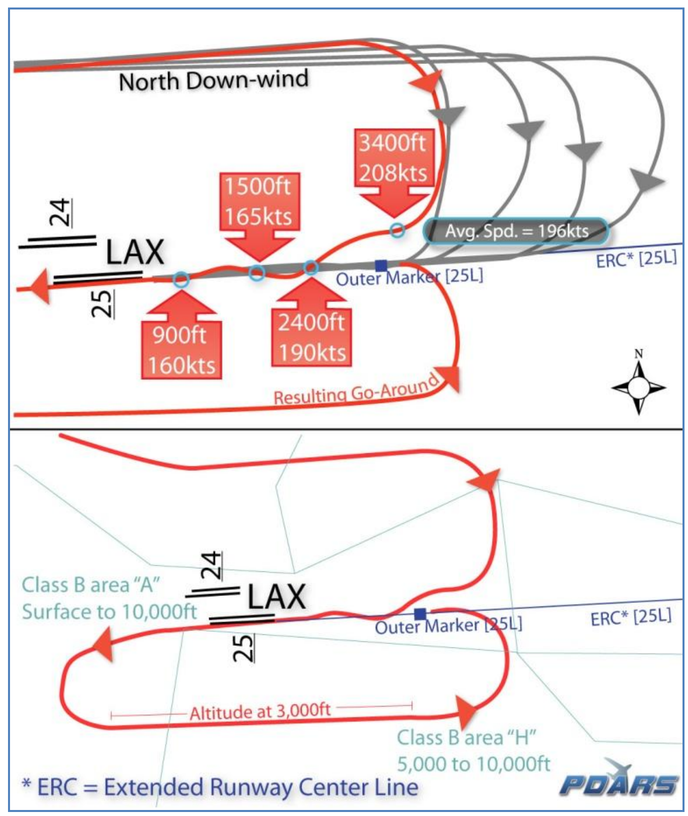

Figure 6: Above shows an aircraft turning too early resulting in an overshoot of the ERC and executing a go-around. During the go-around the aircraft exits class B airspace flying at 3,000 ft and under the 5,000 ft floor of the class B area "H".

Summary/Characteristic Categories	Occurrence 1	Occurrence 2	Occurrence 3	Occurrence 4	Occurrence 5
Aircraft Type	B737	B767	B747	A320	B757
Yearly Season	Spring	Winter	Summer	Summer	Winter
Time of Day	Noon	Afternoon	Night	Late Night	Midnight
Runway Configuration	West	West	West	East	West
Landing Complex	North	South	South	North	South
Complex Switch	X			X	
Go-around		X			X
360 Inside OM				X	
360 Outside OM	X				
Too Fast				X	X
Too High					X
Overtake		X			
S-Turn(s)/Unusual Turns				X	
Intercept ERC Close In	X				
Not Configured to Land				X	
Wrong LOC tuned				X	
Communications	X				
Pilot Initated	X			X	X
Ground Traffic/Vehicle Confliction	X				
Controller Decision/Clearance			X		
Class B Excursion					X
ERC Overshoot			X		
Weather Factor		X	X		

Table 1: Above shows the flight summary characteristic and categories noted for the five selected occurrences.

The following are possible explanations/observations for this abnormal approach event. It appears that the TGT AC is cleared for a visual approach however turned inside the outer marker, at a higher than normal altitude and above normal airspeed. There is no other air traffic in the area and the TGT AC makes a tight turn-on to final. Most-likely the pilot had the airport in sight and accepted clearance for a visual approach before crossing on ERC for the north complex; however the aircraft was not stabilized for landing. TGT AC is a little high on the approach and probably was cleared by the controller to the south complex when the pilot was on the downwind. It appears the pilot made a couple of turns to bleed off speed, since the base leg altitude was not unusually high for a visual approach. It is possible the tower may have changed runways to 25R, although a departure aircraft was observed departing immediately before the B757 executed a go-around. The airspeed was fast for the category of aircraft, although altitude was manageable inside of the OM. Miscommunication between the controller and pilot may have resulted in the aircraft executing a go-around

Safety Review: The B757 apparently was cleared for a visual approach on the north downwind without conflicting aircraft for the airport. The time of the approach was between 11pm and midnight, resulting in a nighttime visual approach to runway 25L. It appeared the pilot was on an unstable descent and unable to configure the aircraft for landing. The pilot attempted a couple of "S" turns for descent and reduction of speed. It does not appear that there were any restrictions to the approach. On the initial go-around the B757 was vectored south of the runway since there was a jet departing runway 25R. The B757 was vectored back to the airport at 3,000 feet (normal return on the downwind is at 5,000 feet to remain in Class B airspace).

F. Contributing Factors

Table 1 provides flight summary and characteristic categories for the five occurrences discussed above. For Occurrences 1, 4, and 5 (the most common of the characteristic categories observed by the SMEs), could possibly have resulted from a preemptive action initiated by pilot. Five sets of the identified occurrences had two characteristic categories in common. These categories were a potential switch of the landing complex, a go-around for the flight, approach was too high, approach was too fast, and weather conditions that may have potentially impacted the identified flights' actions. Examples of other flight characteristics investigated by the SMEs as potential reasons for the flight anomalies that were identified by algorithm include: a high energy approach letdown, the pilot unable to see another aircraft, a pilot not having the airport in sight, a flight's excursion into Class B airspace, a possible fly through of the ERC, and/or the pilot originally expecting clearance to the other complex for landing.

IV. CONCLUSION

After discussions with the TRACON controllers there appears to be considerable interest in the anomalies identified using this approach, and possible that the tool could have a significant impact in daily operations for safety analysis. By pre-identifying risk behaviors, the Safety Office can evaluate repetitive anomalies and provide proactive guidance to eliminate risky behaviors. Once the tool's technology readiness level has moved on from a proof of concept and into a working prototype safety analysts can receive immediate feedback on the unexpected anomalies in the airspace and be able to react more quickly in mitigating any undesired effects. The tool also has the potential for new safety metrics to be derived from the flights identified. Discovery algorithms are not new to the data mining community, however, in this domain; the concept has not yet found a foothold. The realization that previously unidentified behaviors are often not monitored or simply classified as singular occurrences can now be better handled with this approach. As evidence by a simple research of Los Angeles Traffic, of over a half a million flights, the group readily identified repetitive risk anomalies and is able to show that the method is scalable to large volumes of data. Through stitching together and identifying high risk operations, the system can react rapidly to resolve these anomalies. Still, this approach represents a ground breaking step in a new direction of aviation safety and has the potential to provide a fresh insight into an already closely monitored complex system.

V. FUTURE WORK

As this tool moves from a proof of concept into a working prototype more advanced automation can be implemented to give a better understanding of the anomalies identified. In addition, having further validation from subject matter experts from various points of view in the NAS will give more justification for using this tool in more regular safety reporting systems. Another validation step is to apply this tool to additional airports in future tests, which will introduce new challenges such as; regional specific weather conditions and neighboring airspace interactions. At this time the tool is partially automated, with only the flights identified as anomalous, but with little or no context reported. For this study analysts needed to examine the interactions with other aircraft to determine the possible contributing factors leading up to the unusual flight profiles, which typically can involve many man hours. With the addition of automated post processing of the flight characteristic into a well designed reporting scheme, analysts may find it easier to recognize the anomalies identified by having the tool highlight the key unusual characteristics of the more unexpected scenarios. New features may also be generated to better characterize the interactions of neighboring aircraft and can easily be incorporated into the algorithm's multiple kernel model. Additionally, linking of pilot/controller audio recordings may also give the analyst better context to the scenarios and help more effectively determine the level of significance in regards to safety, efficiency, and environmental impact.

ACKNOWLEDGMENT *(HEADING 5)*

We would like to acknowledge the following: funding for this work was provided by NASA Aeronautics Mission Directorate, Aviation Safety Program, System-Wide Safety and Assurance Technologies project. Approval to analyze the data was granted by the FAA. We would like to thank Ron Cagle, Staff Manager Southern, California TRACON and Tom Roche, Front Line Manager, Southern California TRACON for their valuable insight regarding the occurrences. And we would also like to thank Peset Tan for his graphical expertise.

References

[1] K. Bhaduri, Q. Zhu, N. Oza, A. Srivastava. "Fast and Flexible Multivariate Time Series Subsequence Search", IEEE International Conference on Data Mining, Sydney, Australia. Pp. 48-57. 2010

[2] Das, Santanu, Matthews, Bryan L., Srivastava, Ashok N., and Nikunj C. Oza, "Multiple Kernel Learning for Heterogeneous Anomaly Detection: Algorithm and Aviation Safety Case Study," 16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington DC, July 25-28, 2010

[3] Statler, Irving C., "The Aviation System Monitoring and Modeling (ASMM) Project: A Documentation of its History and Accomplishments: 1999-2005," NASA/TP-2007-214556, NASA Ames Research Center, Moffett Field, June 2007

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

[5] FAA Ref: FAA Order JO 7210.632, "Air Traffic Organization Policy: Air Traffic Organization Occurrence Reporting, November 2011] and analysis [Murphy, D., and G. Shapiro, "A Quantitative Model for En Route Error Rate Analysis," 7th USA/Europe ATM R&D Seminar (ATM2007), Barcelona, Spain, July 2-5, 2007

[6] Eurocontrol Safety Regulation Commission, "Annual Safety Report 2011," 2011

[7] SMS Ref: FAA, Aviation Safety Analytical Services, "System Safety Management Transformation Program Plan," Version 1.0, Washington, D.C., December 24, 2009

[8] NAOMS Ref: GAO, "Aviation Safety: NASA's National Aviation Operations Monitoring Service Project Was Designed Appropriately, but Sampling and Other Issues Complicate Data Analysis," GAO-09-112, Washington, D.C.: Mar. 13, 2009

[9] GAO-10-414: United States Government Accountability Office, "Aviation Safety: Improved Data Quality and Analysis Capabilities Are Needed as FAA Plans a Risk-Based Approach to Safety Oversight," GAO-10-414, May 2010

[10] NextGen Ref: FAA, "NextGen Implementation Plan," March 2012

[11] SESAR Ref: SESAR, "The Roadmap for Sustainable Air Traffic Management: European ATM Master Plan," Edition 2, October 2012.

[12] TOPAZ Ref: H.A.P Blom, K. M. Corker and S. H. Stroeve, "On the Integration of Human Performance and Collision Risk Simulation Models of Runway Operation", 6th ATM Seminar, June 2005, Baltimore, USA

[13] BISCOT Ref: ERASMUS, "Security, Safety, Efficiency and CBA Assessment – Env. Baseline, 2007," Version 1.0, January 10, 2009

[14] Ref: Pozzi, S., Valbonesi, C., Beato, V., Volpini, R., Giustizieri, F.M., Lieutaud, F., and A. Licu, "Safety Monitoring in the Age of Big Data," 9th USA/Europe ATM R&D Seminar (ATM2011), Berlin, Germany, 14th June 2011

[15] Gariel, Maxime, Srivastava, Ashok N., and Eric Feron, "Trajectory Clustering and an Application to Airspace Monitoring," eprint arXiv: 1001.5007, 27 January 2010

[16] Marshall, Patrick, "NASA Applies Deep-diving Text Analytics to Airline Safety," Government Computer News, October 26, 2012

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

AUTHOR BIOGRAPHY

Bryan L. Matthews: holds a B.S. in Electrical Engineering from Santa Clara University. He has worked at NASA Ames Research Center for over 10 years in the fields of data mining and machine learning. Bryan has been involved in NASA's aviation safety program since 2005. During this time his primary research focus has been anomaly detection in numeric data and has helped develop and test a number of open sourced scalable parallel algorithms to address this work.

Kennis Y. Chan: is currently a Senior Aviation Systems Engineer and Planning Analyst at ATAC Corporation. She received a bachelor's degree in Industrial Engineering and Operations Research from University of California, Berkeley, CA, USA and a master's degree in Civil & Environmental Engineering focusing on Transportation Engineering from University of Maryland, College Park, MD, USA. She has 10 years of air traffic analytical/research experience in the fields of air traffic control, aviation data analysis, aviation safety and airspace optimization.

Michael D. Kiniry: is currently a Senior Airspace Modeling Specialist at ATAC Corporation. He received a bachelor's degree in Mathematics from University of California, Irvine, CA, USA and a master's degree in Operations Research from Naval Postgraduate School, Monterey, CA, USA. He was a Navy Pilot and a holds a Multi-engine Air Transport Rating from the FAA. He has over 34 years' experience in leading airspace simulation and analysis projects involving some of the busiest airspace in the United States.

Richard A. Gutterud: severed in the U.S. Army from 1971 to 1975, as an Air Traffic Controller. He was employed by the Federal Aviation Administration for 31 years, as an Air Traffic Controller, Supervisor, Operations Manager, and Traffic Management Officer, with 10 years Traffic Management analytical experience in collaborative development, procedural evolvement, and participation in airspace and air traffic automation. He has participated and tested the implementation of PDARS (at Southern California TRACON) as well as participating member of the CDM (Collaborative Decision Management) work group since 1999. For the past six years he has been employed by ATAC Corporation as Director, Aviation Data Analysis Programs, in Washington D.C. and San Diego CA.

Ashok N. Srivastava, Ph.D: is the Principal Scientist for Data Sciences at NASA Ames and leads the System-Wide Safety and Assurance Technologies project at NASA. He is the Editor-in-Chief of the AIAA Journal of Aerospace Computing, Information, and Communication. He has over 100 papers in peer-reviewed conferences and journals and has edited three books. He is a Fellow of the American Association for the Advancement of Science and an Associate Fellow of the American Institute of Aeronautics and Astronautics. He has won numerous awards including the IEEE Computer Society Technical Achievement Award for "pioneering contributions to intelligent information systems."

John Schade: has over fifteen years of experience in the program management, technical direction, and analysis of aviation related systems concentrating in the areas of air traffic control operations, real-time simulation/animation, and accident/incident investigation. John has extensive experience in metric development, data mining, requirements development, and software implementation. John has specific expertise in the analysis and development of technology to support the analysis of air traffic control radar data, digital flight data recorder data, human-in-the-loop simulation experiments, and performance measures for aviation systems. John has been with ATAC Corporation for 12 years and prior to that was engaged as an Aerospace Engineer at the National Transportation Safety Board (NTSB) for 5 years where he participated in numerous major aviation accident investigations. Currently John is the program manager for the Performance Data Analysis and Reporting System (PDARS) program encompassing activities including task order development, security standardization, resource planning, roadmap implementation, training requirements, documentation generation, and process standards adoption. In addition, John manages key technology development efforts in the areas of application development, system engineering, web design, database development and implementation, algorithmic development, and IT infrastructure procurement and deployment for the PDARS system. John holds two bachelors of science degrees (Aerospace Engineering, Business Finance) from the University of Maryland.

David Schleicher: serves currently as the Chief Operating Officer at the ATAC Corporation with over 22 years of professional experience in aviation research, aviation system and aircraft design, systems and operations analysis, economic analysis, and program management. Mr. Schleicher specializes in the systems and simulation design and engineering analysis of current and future air traffic management (ATM) and cockpit systems and has held previous technical and management positions in the aviation industry and NASA. His more recent activities have included the development and assessment of future NextGen concepts of operations, ATM safety assessments, and development of ATM performance analysis systems. He graduated summa cum laude with a B.S.E. in Mechanical and Aerospace Engineering from Princeton University and received a Master's degree in Aeronautics and Astronautics from Stanford University