Detection and Analysis of Aviation Safety Events using Historic Flight Data

Bae-Seon Park*, Seong-Min Han†, Hak-Tae Lee†, Hyejin Lim‡, Haeyoon Byeon‡ and Hyun-Jin Jung‡
Email: *pearship@inha.ac.kr, †{zx12op@inha.edu, haktae.lee@inha.ac.kr}, ‡{hjlim, hyoon74, hjinjeong}@iiaci.or.kr
*Industrial Science Technology Research Institute, Inha University, Incheon, 21999, Republic of Korea
†Department of Aerospace Engineering, Inha University, Incheon, 21999, Republic of Korea
‡Institute for Aerospace Industry-Academia Collaboration, Incheon, 21999, Republic of Korea

***Abstract*—In this paper, the final approach overshoot, localizer deviation, and glideslope deviation, which frequently occur in the approach phase, were detected and their occurrence was analyzed using ADS-B data of arrivals that landed on Jeju International Airport runway 25 in 2019. The detected events were analyzed based on the defined characteristics. This paper defines and analyzes aviation safety events based on their characteristics. Safety events can be determined by start time, duration, and gap between consecutive events and these characteristics subdivide the event definition into instantaneous, short continuous, and long continuous events. For complex events, they can be defined as simultaneous or sequential events. The thresholds of the event gap and duration were determined using the distributions from the event detection results and the detected events were classified into instantaneous, continuous, and complex events. This paper also visualizes the analysis results using an event tree, which provides information about the events and their propagation in chronological order. The event tree shows the process by which different events are derived from the first event. In addition, it helps to predict the subsequent events and the probability of their occurrence after the specific event has occurred.**

***Index Terms*—Aviation Safety, Safety Event, Event Detection, Event Analysis**

I. Introduction

An aviation safety event is a specific incident that occurs during flight operations. Runway excursions or incursions, abnormal runway contact, and loss of separation are common safety events that are defined and managed as risk factors that can lead to serious accidents. Most organizations manage Key Performance Indicators (KPIs) based on significant safety events. For example, the Commercial Aviation Safety Team (CAST) and the International Civil Aviation Organization (ICAO) created the CAST/ICAO Common Taxonomy Team to manage the major event categories [1].

Several researchers have studied different safety events and their causes. Janakiraman et al, Matthews et al, and Ackley et al proposed methods to detect precursors of safety events based on machine learning techniques [2]–[5]. Other researchers used big data analytics to analyze significant safety events and their contributing factors. Oh et al. analyzed rejected takeoffs and examined the antecedents that contributed to them [6]. Han et al. presented runway excursion analysis to find the related antecedents using QAR data [7]. Sherry et al. analyzed go-arounds and missed approaches and identified the factors from pilot and controller reports [8].

Previous work has focused on identifying precursors to individual safety events and has not presented new approaches to analysis. These studies used different definitions of precursor and event, limiting the definition of precursor to low-level parameters that trigger specific events, such as aircraft state data. They also analyzed significant safety events, most of which are managed as KPIs.

Though significant events have been analyzed, it is necessary to expand the safety event categories to include minor and less serious ones. As shown in Fig. 1, there can be a lot of minor events overshadowed by what gets reported. Some incidents which are unnoticed and not managed such as unstable approach, high or low energy states, and missed approach can be the causes of the severe safety issue. From 2007 to 2016, 48% of severe accidents occurred in the final approach phase, especially eight minutes before landing [9]. Numerous events may occur in this phase and these minor issues come out with various combinations. For example, after a certain event, there are cases where multiple events occur simultaneously, or the same event repeats continuously. Also, several events may start at the same time and are followed by only a single event. It is important to analyze events that have complex causalities and can be the causes of each other.

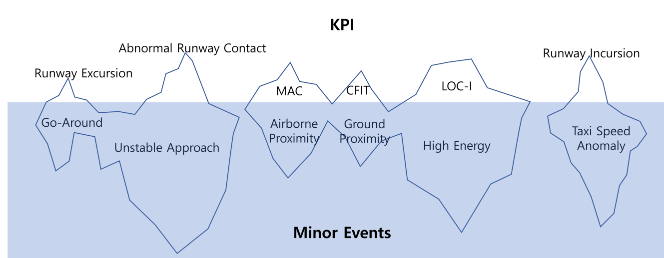


Fig. 1. **Safety event iceberg**

This paper presents detailed definitions of safety event based on their characteristics. Using historical flight data, the final approach overshoot and unstable approaches are detected and analyzed by suggested definitions. The results of the

This work is supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant RS-2020-KA158275 and RS-2021-KA163373).

analysis are presented as tree structures to provide at-a-glance information that is easily understood by subject matter experts.

Following this introduction, section II presents the safety event definitions based on the event characteristics. In section III, three safety events in the approach phase are detected and analyzed based on the event definitions. The analysis results are explained using an event tree that provides the detailed information of the event propagation. Section IV concludes and summarizes this work.

II. Event Definitions

Event has three characteristics that determine their occurrence. Fig. 2 shows an example of the events in the flight trajectory and the red tracks are the tracks where an event occurred. The time when an event started can be defined as the 'start time' of the event. If an even starts and continues, the time between the start and end of the event is defined as the 'duration' of the event. And the interval between the end of leading event and the start of the trailing event is the 'gap' between consecutive events of same type.

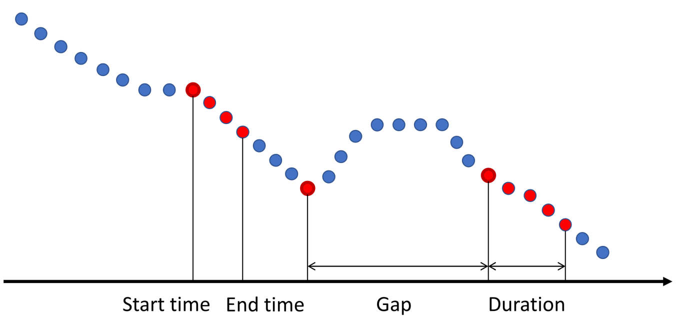


Fig. 2. **Event characteristics**

A. Single event

Three characteristics can be used to define detailed event occurrences. The event gap is a criterion to distinguish individual events. Whether they are the same event or not, they are individual if the time interval between the two events is larger than the gap threshold between the events.

Fig. 3 shows the classification of the single event history. In Fig. 3 (a), the red dots are detected events at 10, 13, and 15 seconds. Unlike reality, the detection results are discrete so the duration of each detection is temporal. Therefore, it is necessary to distinguish whether the detection is temporary or continuous based on certain criteria. If the events are detected every 5 seconds on average, it can be assumed that this event occurs every 5 seconds so the threshold of the gap can be set to 5 seconds. Using a 5-second of gap threshold, each event can be represented as an interval with 5 seconds width as shown in Fig. 3 (b). The intervals that have intersections refer that corresponding events were detected within the gap threshold. The intervals can be merged so the events can be a single 'continuous' event. In Fig. 3 (c), the event at 10 and 13 seconds are merged into a continuous event, and the duration is equal to the actual gap between them. There are no consecutive events after 20 seconds, so the event at 20 seconds has zero duration, which is called an 'instantaneous' event. So, the three events at 10, 13, and 20 seconds can be simplified to the continuous event from 10 to 13, and the instantaneous event at 20 seconds.

The continuous event can be further subdivided into 'short-term' or 'long-term' events depending on whether they continue for less or more than the threshold of the duration. In Fig. 3 (c), the event at 10 seconds can be a short-term event using the 5-second duration threshold. If the threshold is 2 seconds, the event is a long-term continuous event.

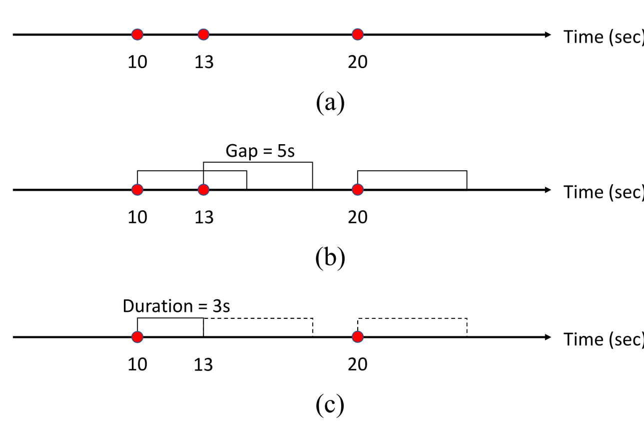


Fig. 3. **Single event classification process**

B. Complex events

However, the event does not always occur individually, so complex events must be considered. The different events at the same time are 'simultaneous' events, whereas when events occur successively after other events, they are 'sequential' events.

Fig. 4 (a) shows the detection histories of two different events A and B. To check the simultaneity, each event should be determined to be instantaneous and continuous events based on the gap threshold. In Fig. 4 (b), the gap thresholds of A and B are 2 and 5 seconds. Event A is continuous at 12-13, 19-21, and 23-26 seconds and instantaneous at 15 seconds. Event B is continuous from 10 to 14 seconds and instantaneous at 20 seconds. As shown in Fig. 4 (c), the intervals of events A and B intersect from 12 to 13 and at 20 seconds, so these intersections became the simultaneous events of A and B. All events should not intersect each other except the simultaneous events. Considering the simultaneous events as a new single event, single events and combined events should be separated after the simultaneous events are determined. Fig. 4 (d) shows that events A, B, and A&B were completely separated. For event A, the events at 19-21 seconds are split into two continuous events because A and B occurred simultaneously and instantaneously at 20 seconds.

III. Event analysis

In this study, three events were detected and analyzed in the approach phase. The final approach overshoot and

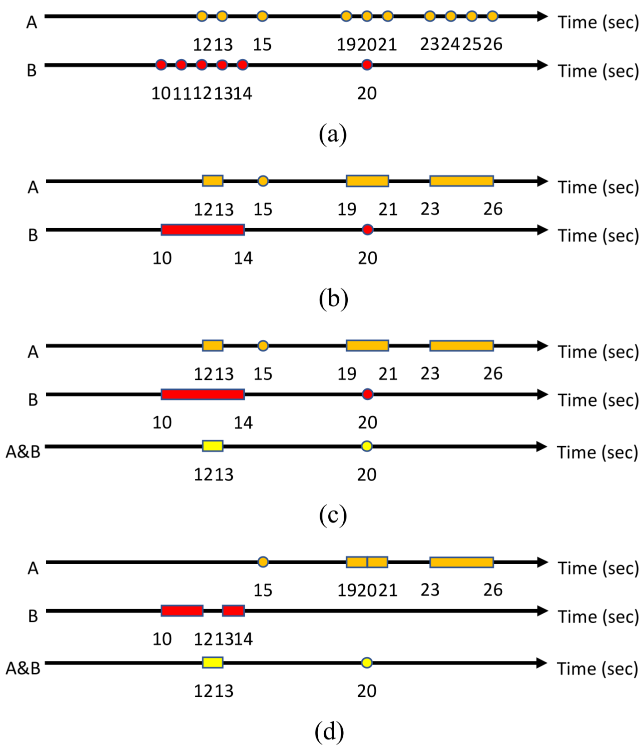


Fig. 4. **Complex events classification process**

localizer/glideslope deviations are common events in the final approach. These are the events that can occur when 8 minutes before landing and trigger the chain of severe incidents or accidents such as runway excursions, ground collisions, or out or courses. They are not included in KPIs, but they can be managed as sub-indicators that can cause major safety indicators.

The events were analyzed as follows: 1. Detect target events using historical flight data. 2. Classify the detected events: (a) Set event thresholds based on the gap and duration of each event. (b) Classify each event into instantaneous and continuous events using their thresholds. Also, check the complex events by combining each event and set the thresholds of the complex events. (c) Classify the complex events using the thresholds. 3. Configure and analyze the event trees based on the classified events.

A. Event Detections

1) Final Approach Overshoot (FO) [10]: The final approach overshoot is an event in which the aircraft trajectory is corrected to realign with the runway after deviating from the extended runway centerline during the final approach, as shown in Fig. 5. If the trajectory and the extended runway centerline do not intersect, it is not an overshoot. The overshoot distance is a perpendicular distance between the aircraft position and the extended runway centerline. The overshoot can be detected by calculating the overshoot distances at each position. In this study, overshoots were detected within 7-14 nmi of the runway threshold, and the occurrence positions were determined by the maximum overshoot distance of each flight.

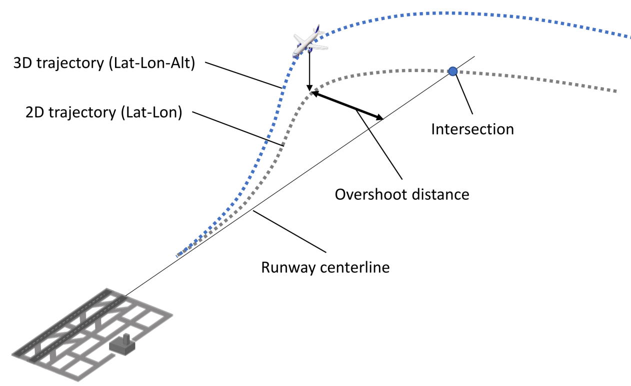


Fig. 5. **Final approach overshoot**

2) Localizer and Glideslope Deviations (LOC, GS): The localizer and glideslope are instruments used to align the airplane with the runway during the final approach. Figs. 6 and 7 show the localizer and glideslope ranges, and the localizer or glideslope deviation occurs when the airplane is out of range. In this study, the deviations were detected within 5.34 nmi of the runway threshold. The localizer deviation was detected when the aircraft was out of the ±1-degree range, and the glideslope deviation was detected when the aircraft was out of the 0.7-degree up and down range based on a 3-degree glideslope.

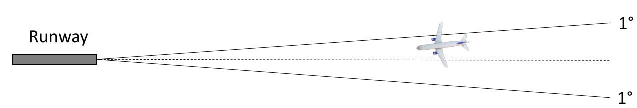


Fig. 6. **Localizer deviation**

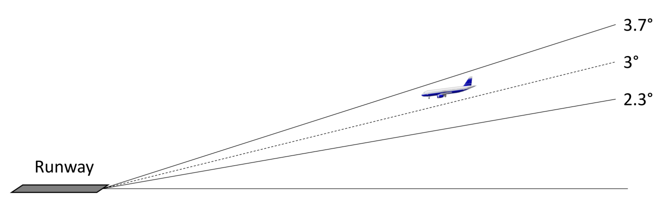
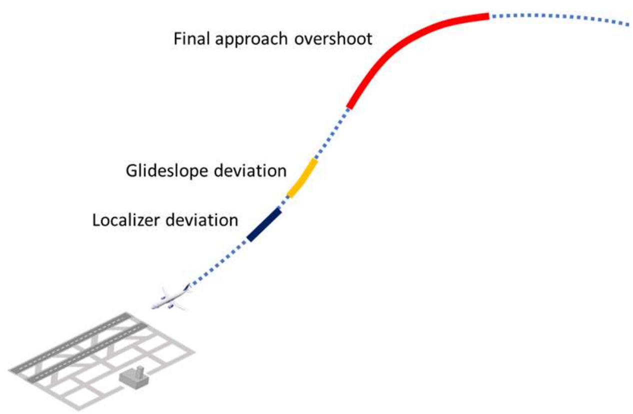
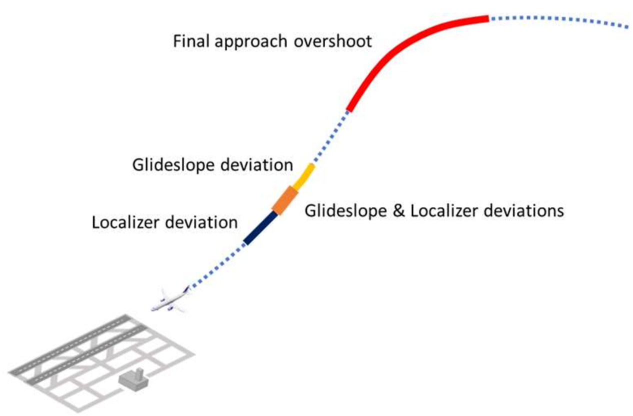


Fig. 7. **Glideslope deviation**

Fig. 8 are examples that three events that are detected in the approach phase. Because of the detection range, the final approach overshoot is detected before the localizer and glideslope deviations during the approach. The localizer or glideslope deviations are detected when the flight is closing to the runway for landing. Two events have the same detection range, so the localizer or glideslope deviations can be detected simultaneously, or detected sequentially.



(a)



(b)

Fig. 8. **Examples of events in approach phase**

B. Event Classification

The events were detected using ADS-B data of flights arrived at Jeju International Airport (CJU) runway 25 in 2019. To detect FO, GS, and LOC events during the approach, the trajectory was resampled at 1-second intervals because the ADS-B data is not uniform. The approach phase is assumed to begin at an altitude of 10,000 ft and at the boundary of the Jeju TMA.

A total of 36,244 flights landed on runway 25 and 20,471 flights had events. The detected events were analyzed to find the thresholds of their gap and duration, and every single event and their combinations were classified by using thresholds.

1) Setting single event thresholds: First, each detected event was analyzed. To determine the event gap and duration threshold for classifying events as instantaneous or continuous, their distributions were analyzed. As the events were detected at 1-second intervals, the start and end times of a detected single event are the same so the minimum gap of the detected event is 1 second. The gap threshold is a reference that determines the temporality and persistence of events detected per second, so it is not recommended to set it too densely or widely. In this study, the threshold was set at 10 seconds, which covers the top 50% of the gap distribution.

Fig. 9 shows the distributions of the individual event durations. The duration thresholds were set to cover the top 50% of the distributions as well as the gap threshold. The final approach overshoot, glideslope deviation, and localizer deviation thresholds were set at 30, 50, and 10 seconds, respectively.

2) Single event classifications and setting complex event thresholds: Since the thresholds of events were set, it is possible to classify each single event in detail. In this section, the single events were classified into instantaneous and continuous events. Also, the complex events that the single events are simultaneously combined were configured and their thresholds were set.

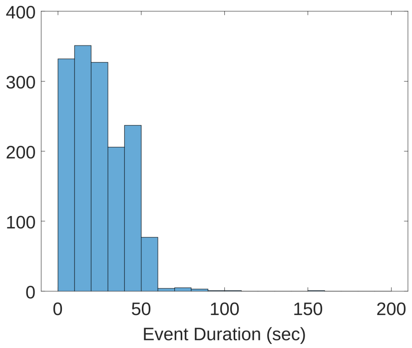
Fig. 10 is an example of the classification of events detected in a flight using gap and duration thresholds. In Fig. 10 (a), the final approach overshoot continued for 64 seconds from 10:00:32 to 10:04:09. The glideslope deviation continued for 26 seconds from 10:04:24 to 10:04:50, and the localizer deviation continued for 19 seconds from 10:04:09 to 10:04:28. Using the duration thresholds of each event, three continuous events can be classified as shown in Fig. 10 (b). Because its duration is longer than a 30-second threshold, the final approach overshoot becomes a long continuous event. Similarly, the glideslope deviation is a short continuous event and the localizer deviation is a long continuous event.

Since the single events were classified into instantaneous and continuous events, the complex events can be determined. In this paper, the glideslope deviation and localizer deviation were detected simultaneously because the detection range of the final approach overshoot was different. The duration threshold of glideslope and localizer deviations was set to 10 seconds including the top 50% interval based on the duration distribution.

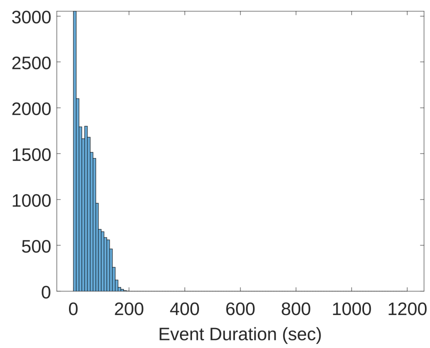
3) Total event classifications: Using the determined thresholds of every single and complex event, all events were classified. Fig. 11 shows the reclassified result of the events in Fig. 10 considering complex events. Fig. 11 (a) shows that the glideslope deviation and the localizer deviation occurred simultaneously for 4 seconds from 10:04:24 to 10:04:28. Since the duration threshold of the glideslope and localizer deviations was 10 seconds, it becomes a short continuous event as shown in Fig. 11 (b). It can be seen that the simultaneous events are separated from the localizer deviation and the glideslope deviation because the complex event is considered a new event type.

C. Event Tree

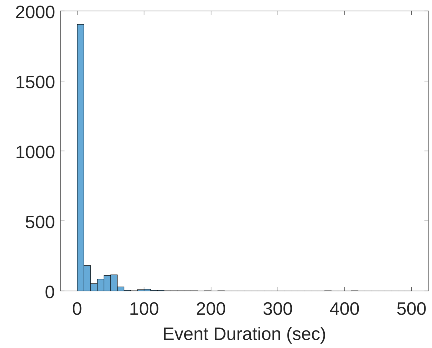
An event tree is a method to analyze events in multiple flights, allowing an intuitive understanding of the context of events and checking the rate of each event frequency. Events in a single flight can be represented in a time-based sequence. Also, the event sequences can be combined into an event tree using the common events of them. A common event means an event with the same occurrence level and type.



(a) Final approach overshoot

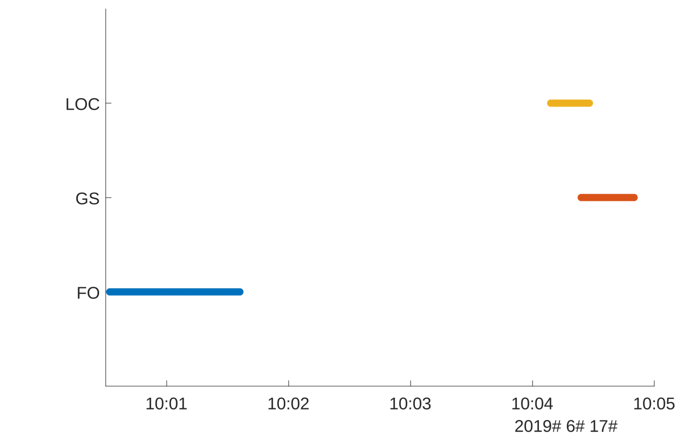


(b) Glideslope deviation

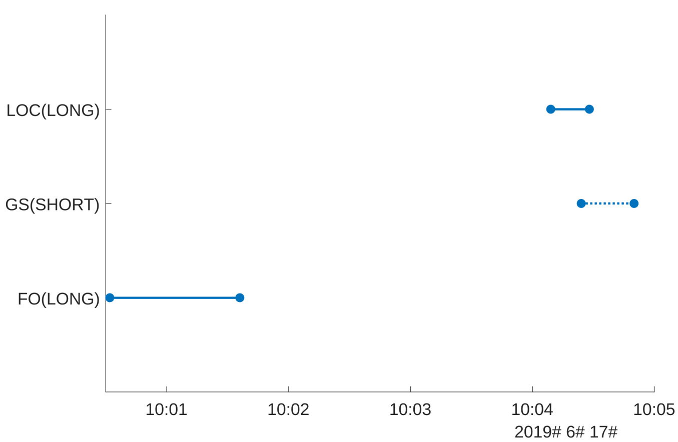


(c) Localizer deviation

Fig. 9. **Distributions of three single event durations**

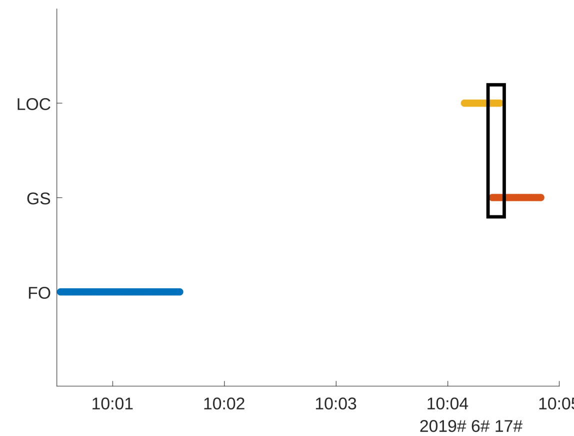


(a) Detected events

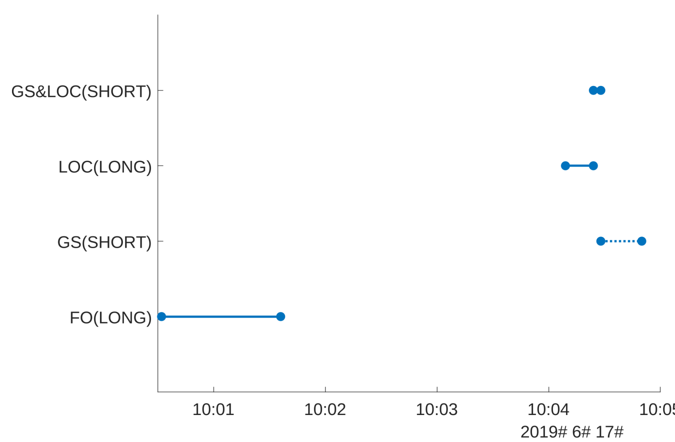


(b) Continuous events

Fig. 10. **Event classification example**



(a) Detected events



(b) Continuous events

Fig. 11. **Total event classification example**

For example, the events in Fig. 11 (b) can be expressed as a sequence FO(LONG) – LOC(LONG) – GS&LOC(SHORT) – GS(SHORT). If there is another sequence FO(SHORT) – LOC(LONG) – GS(SHORT) as shown in Fig. 12 (a), a common event of two sequences is LOC(LONG), which has the same level and type. Using the common event as a junction node, two sequences can be combined and form a tree as shown in Fig. 12 (b). The frequency of the junction event node is updated by summing frequencies of common events of each sequence.

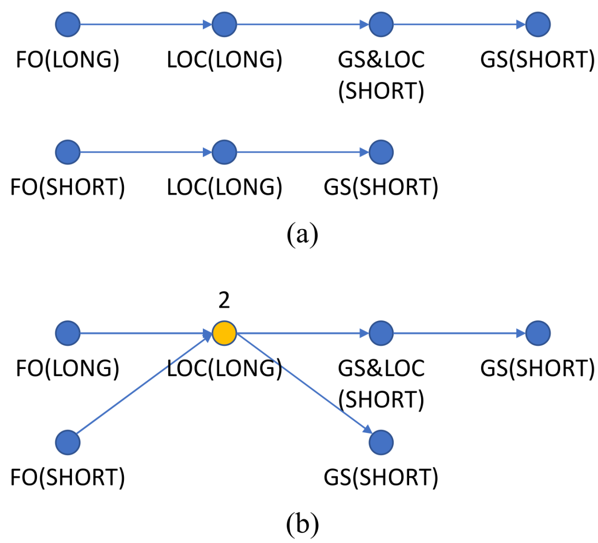


Fig. 12. Event tree example

The events of all detected flights were converted into a sequence, and this was configured as an event tree. In this study, the event that occurred first was defined as the root event, and a total of 12 event trees were constructed by combining sequences with a common root event. Table I shows all the root events, the frequencies of the root events, and the total frequencies and their ratios from the detection result. FO(Instant) is an instantaneous final approach overshoot, which was detected 56 times in total, but was the first detected 51 times. It can be explained that the instantaneous final approach overshoot can occur as a root event with a probability of 91.1%. Glideslope deviation and localizer deviation also have more than a 90% probability of being root events, but the long localizer deviation has a relatively low probability of 72.2%. On the other hand, GS&LOC was detected as a root event with less than 2% probability, which means that a complex event follows the root events of the other types.

Fig. 13 is an event tree that integrates all event sequences starting from the long final approach overshoot. About 50% of a total of 499 long final approach overshoots lead to continuous glideslope deviations. After level 2, most of the glideslope deviations and localizer deviations occurred simultaneously, and the frequency of events gradually decreases as the level progresses.

The event tree shows the probability of an event considering

TABLE I
ROOT EVENTS AND THEIR FREQUENCIES

Root event	Frequency	Total frequency	Root/Total (%)
FO(Instant)	51	56	91.1
FO(Short)	985	1029	95.7
FO(Long)	499	507	98.4
GS(Instant)	344	360	95.6
GS(Short)	10033	10517	95.4
GS(Long)	8201	8601	95.3
LOC(Instant)	14	15	93.3
LOC(Short)	95	99	96.0
LOC(Long)	164	227	72.2
GS&LOC(Instant)	8	1635	0.5
GS&LOC(Short)	13	1924	0.7
GS&LOC(Long)	10	545	1.8

the context of events. Fig. 13 can be described that when a long final approach overshoot occurs first in the approach phase, it may lead to a continuous glideslope deviation with about 50% of probability, and glideslope and localizer deviations may follow simultaneously with a probability of about 15%.

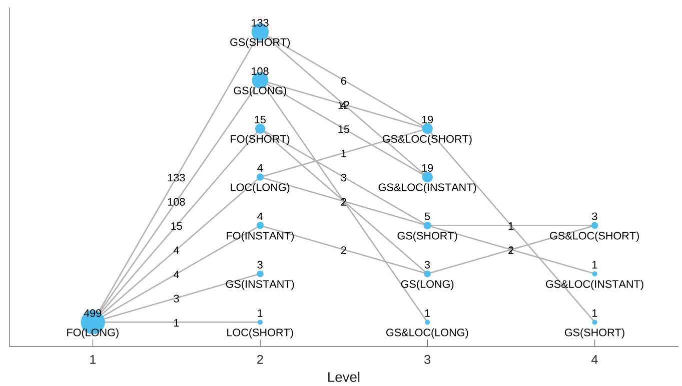


Fig. 13. Event tree from long final approach overshoot

Figs. 14, 15 show event trees starting with the short and long glideslope deviations. In Figs. 14, 15, most of the events at level 2 are glideslope and localizer deviations. Events follow continuously after level 2, but their frequency decreases significantly.

Fig. 14 shows that about 10% of the first short glideslope deviations lead to simultaneous events with localizer deviations at the next level. And 50% of them are short glideslope and localizer deviations that continue for less than 10 seconds. After level 3, the frequency of events significantly decreases, however, the events keep occurring for a long period up to level 6 after the short glideslope deviations.

Fig. 15 is an event tree derived from the long glideslope deviations. The main events at level 2 are the same as in Fig. 14, but the proportion derived from the root event is about 31%, higher than in Fig. 14. This indicates that the longer the initial glideslope deviation continues, the more likely it is that the localizer and glideslope deviations will occur simultaneously.

Mostly, there are few final approach overshoots after glideslope or localizer deviations because the glideslope and localizer deviations are detected after the final approach overshoot. However, if the flight overshoots after the glideslope or localizer deviation as shown in Figs. 14, 15, it indicates that the overshoot was detected during the re-approach after the flight operated go-around after failing to establish the glideslope or localizer.

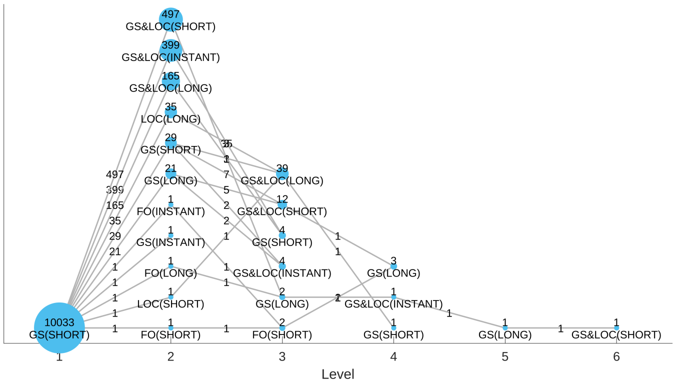


Fig. 14. Event tree from short glideslope deviation

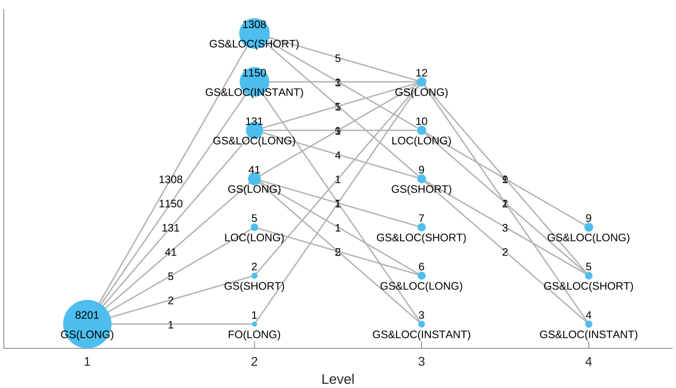


Fig. 15. Event tree from long glideslope deviation

IV. Conclusion

This paper extends the definitions of aviation safety events based on their characteristics. In addition, the event occurrence is analyzed and simplified by transforming the tree structure. Three safety events that occur frequently during the approach phase, but are not included in the main indicators, were detected using historical flight data that landed on runway 25 of CJU in 2019. The results showed that these events occur in combination and have complex patterns. The event tree simplifies these complex event relations and makes them easier to understand. It also provides the insights such as the event probability considering the context of the events.

Future investigations will consider the time-based event tree and advance the event probabilistic model. Events at the same level happen at different times, so the time-based distributions will be applied to the event node. In addition, if a probabilistic model based on the event tree is constructed using large safety event histories, it can help to predict future events after the first event appears.

References

[1] ICAO CICTT, “Aviation Occurrence Categories – Definitions and Usage Notes,” Release 4.8, 2021.

[2] V. M. Janakiraman, B.Matthews, and N. Oza, “Discovery of precursors to adverse events using time series data,” Proceedings of the 2016 SIAM International Conference on Data Mining, Society for Industrial and Applied Mathematics, 2016.

[3] V. M. Janakiraman, “Explaining aviation safety incidents using deep temporal multiple instance learning.” Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2018.

[4] B. Matthews, S. Das, K. Bhaduri, K. Das, R. Martin, and N. Oza, “Discovering anomalous aviation safety events using scalable data mining algorithms,” Journal of Aerospace Information Systems, Vol. 10, No.10, 2013, pp.467-475.

[5] J. L. Ackley, T. G. Puranik, and D. N. Mavris, “A supervised learning approach for safety event precursor identification in commercial aviation,” AIAA Aviation 2020 Forum, 2020.

[6] S. Oh, S. Kim, and Y. Yoon, “An Analysis of Runway Accident Precursors Based on Latent Class Model,” KSCE Journal of Civil Engineering Vol. 24, No. 9, 2020, pp. 2784-2793.

[7] J. Han, Y. Jiao, C. Zhang, C. Wang, B. Shen, “Analysis of the Key Precursor Indicator Parameters of Runway Excursion Incidents Based on Density Clustering Algorithm,” 2022 IEEE 4th International Conference on Civil Aviation Safety and Information Technology (ICCASIT), IEEE, 2022.

[8] L. Sherry, Z. Wang, H. K. Kourdali, and J. Shortle, “Big data analysis of irregular operations: Aborted approaches and their underlying factors,” 2013 Integrated Communications, Navigation and Surveillance Conference (ICNS), IEEE, 2013.

[9] Boeing, “Statistical Summary of Commercial Jet Airplane Accidents: Worldwide Operations 1959-2021,” 2022.

[10] S. Han, B. Park, and H. Lee, “Development of Final Approach Overshoot Calculation Algorithm,” The Korean Society of Aeronautical and Space Sciences (KSAS) Fall Conference, Jeju, Republic of Korea, 2022.