Risk-based Classification of Incidents

William S. Greenwell, John C. Knight, Elisabeth A. Strunk,
Department of Computer Science, University of Virginia; Charlottesville, VA, U.S.A.

Keywords: safety-critical systems, incident investigation, risk assessment

Abstract

As the penetration of software into safety-critical systems progresses, accidents and incidents involving software will inevitably become more frequent. Identifying lessons from these occurrences and applying them to existing and future systems is essential if recurrences are to be prevented. Unfortunately, investigative agencies do not have the resources to fully investigate every incident under their jurisdictions and domains of expertise and thus must prioritize certain occurrences when allocating investigative resources. In the aviation community, most investigative agencies prioritize occurrences based on the severity of their associated losses, allocating more resources to accidents resulting in injury to passengers or extensive aircraft damage. We argue that this scheme is inappropriate because it undervalues incidents whose recurrence could have a high potential for loss while overvaluing fairly straightforward accidents involving accepted risks. We then suggest a new strategy for prioritizing occurrences based on the risk arising from incident recurrence.

Introduction

By their very nature, commercial aviation accidents demand our attention. Major accidents can create spectacular scenes of carnage and destruction that threaten public confidence in commercial air travel. At the very least, accidents remind us that, while very safe, there is still some risk in commercial air travel, and they often force engineers and regulators to rethink their safety analyses and add additional safeguards to the air transit system. It is out of a desire to improve safety and prevent the recurrence of tragedy that society demands investigations into accidents in order to learn as many lessons from them as possible.

Although major accidents receive the most publicity, less severe accidents and even incidents in which no loss is incurred can be equally valuable in their ability to provide lessons [2]. Despite this, incidents rarely command the attention that accidents do, and this is a serious imbalance with possibly serious consequences. This paper presents two commercial aviation events involving safety-critical software systems in which the failure of those systems contributed to the occurrence of the events. The first event resulted in a crash with hundreds of fatalities. Although the second event did not develop into an accident, the failure of the system involved led to a near-collision between two jumbo jets. After summarizing the events, we show that the first event received a much more rigorous investigation than the second, even though the latter could have resulted in almost twice the number of fatalities. We then suggest an alternative incident classification scheme that we claim will more appropriately match investigative resources to events whose recurrence would likely have catastrophic consequences.

Both of the events that we discuss in this paper could have been prevented in many ways. However, the need for change in incident classification is illustrated very clearly by the fact that both events were *preceded* by similar incidents that indicated the possibility of a systemic problem [3]. Our strategy attempts to exploit such leading indicators to prevent future accidents.

Accidents Versus Incidents

Before proceeding, it is useful to distinguish accidents from incidents. Numerous definitions exist for these terms; however because this paper focuses on two commercial aviation events, the definitions we use are those adopted by the Federal Aviation Administration (FAA) and the National Transportation Safety Board (NTSB). Those organizations define the terms as follows:

> *Aircraft accident*—an occurrence associated with the operation of an aircraft which takes place between the time any person boards the aircraft with the intention of flight and all such persons have disembarked, and in which any person suffers death or serious injury, or in which the aircraft receives substantial damage.
>
> *Incident*—an occurrence other than an accident, associated with the operation of an aircraft, which affects or could affect the safety of operations [11].

According to these definitions, for an unsafe occurrence involving an aircraft to be considered an accident, it must take place when persons are aboard the aircraft and result in some form of loss—death, serious injury, or damage to the aircraft. Otherwise, the occurrence is considered to be an incident. Most aircraft accidents and incidents occur while the aircraft is in operation, which implies that persons will be aboard at the time of an occurrence. Thus, loss is the key factor that distinguishes an accident from an incident, which agrees with the distinction made by Leveson [16]. All occurrences affecting safety begin as incidents, and whether they remain incidents or develop into accidents depends upon their outcomes. Incidents and accidents might share similar event sequences, meaning that incidents are sometimes precursors to accidents. Investigating incidents can lead to recommendations that help prevent future accidents.

Accident & Incident Investigation

Unfortunately, most investigative agencies simply do not have the resources to fully investigate every aviation-related occurrence within their jurisdiction and must prioritize certain occurrences when allocating investigative resources. Agencies typically prioritize occurrences according to the severity of their associated losses. For example, the NTSB classifies an accident as “major” if the accident results in the destruction of a commercial aircraft, multiple fatalities, or one fatality and substantial damage to a commercial aircraft. According to NTSB statistics, 74 major accidents occurred between 1983-2002 compared to 581 accidents receiving less severe designations involving commercial aircraft [5]. NTSB investigators use a special operating manual when investigating major accidents that guides them in collecting evidence, holding public hearings, and preparing final reports [6]. Reports are typically reserved for major accidents; synopses are prepared for less severe accidents and then stored in a database. The NTSB investigates all civil aviation accidents that occur within its jurisdiction. It also selectively investigates aviation incidents but is not required to do so. NTSB incident reports are stored in the board’s accident database and resemble the synopses it prepares for minor accidents.

In addition to the NTSB, the FAA may also investigate civil aviation incidents at its discretion, and informal channels exist for collecting and analyzing incident reports such as the Aviation Safety Reporting System (ASRS), which is administered by the FAA and the National Aeronautics and Space Administration (NASA). Pilots, air traffic controllers, flight attendants, mechanics, and others may voluntarily submit incident reports to the ASRS. These reports are reviewed by ASRS personnel who use them to prepare monthly safety bulletins and to identify immediate safety hazards to report to the FAA [12]. A similar system called CHIRP exists in the United Kingdom. The FAA also maintains partnerships with air carriers and repair stations

known as Aviation Safety Action Programs (ASAPs) that encourage employees to voluntarily report information that might help the FAA identify potential precursors to accidents [15]. Johnson notes, however, that these systems have limitations and in particular tend to focus on direct, short-term fixes to safety problems rather than addressing underlying issues [13].

The NTSB is not unique in prioritizing accident and incident investigations according to loss. Most investigative agencies worldwide distinguish between accidents and "serious incidents," including the U.K. Air Accidents Investigation Branch (AAIB), the French Bureau d'Enquêtes et d'Analyses (BEA), the German Federal Bureau of Aircraft Accidents Investigation (BFU), the Accident Investigation Board of Finland (AIB), the Australian Transport Safety Bureau (ATSB), the Taiwanese Aviation Safety Council (ASC), and others. While the definition of "accident" is typically clear, the term "serious incident" is often not well-defined. The AAIB and BFU offer guidelines that give examples of serious incidents, but admit that these guidelines are not comprehensive. The ATSB uses a five-category system to classify accidents and incidents, but the criteria for categorizing an occurrence are subjective. The Canadian Transport Safety Board (TSB) does not actually distinguish between accidents and incidents but labels both types of events as "occurrences." They classify and investigate occurrences based on "whether the investigation is likely to lead to reduced risk to persons, properly, or the environment" [9]. This is similar to the scheme we propose; however their criteria are still quite subjective.

The effect of allocating resources to accident and incident investigations based on the severity of their associated losses is that less severe accidents might receive only a small amount of attention from investigators, and incidents might not be investigated at all. However, many major *accidents* are preceded by similar *incidents* in which it was only by coincidence that a loss did not occur. This is particularly important in the context of safety-critical software systems because design faults present in such systems can manifest themselves with unpredictable consequences. If the systems control hazardous operations, they might bring direct harm to passengers or crew. Alternatively, if the systems provide advice or warnings to pilots, they might raise false alerts or issue erroneous guidance to pilots, who could inadvertently jeopardize safety by acting on this information.

To illustrate the disparity in the level of attention given to accidents versus that typically given to incidents, we examine the investigations conducted following a major accident and a major incident. The following sections begin with brief descriptions of the occurrences and then present details of their respective investigations.

Korean Air Flight 801

On August 6, 1997 at about 1:42am Guam local time, Korean Air flight 801, a Boeing 747-300, crashed into Nimitz Hill, Guam while attempting a nonprecision approach to runway 6L at A.B. Won Guam International Airport. Of the 254 persons on board, 237 of which were passengers, only 23 passengers and 3 flight attendants survived. The National Transportation Safety Board (NTSB) investigated the accident and classified the crash as a controlled-flight-into-terrain, or CFIT, accident. During its investigation, the NTSB found that a ground-based minimum safe altitude warning system (MSAW), designed to alert air traffic controllers of aircraft flying too low, had been inhibited. In its final report [7], the NTSB concluded that the crash was largely due to pilot error, but also noted:

> "Contributing to the accident was the Federal Aviation Administration's (FAA) intentional inhibition of the minimum safe altitude warning system (MSAW) at Guam and the agency's failure to adequately manage the system."

We discuss in detail how the MSAW system at Guam contributed to the accident elsewhere [3]. Essentially, the system had been disabled years before the accident in order to eliminate nuisance low-altitude warnings. Prior to the accident, the FAA had received multiple warnings that MSAW systems were being configured improperly. These included a safety recommendation from the NTSB issued in response to a previous accident urging the FAA to verify the MSAW configurations at each of its air traffic control facilities as well as an evaluation of the Guam facility that noted its MSAW inhibition. After the Korean Air flight 801 accident, the FAA developed a comprehensive program to manage its MSAW installations, but continued to be plagued by accidents in which MSAW configuration errors were cited as contributory factors.

The NTSB began its investigation into the Korean Air flight 801 accident immediately after the crash. The Board adopted its final report, a 212-page document, on January 13, 2000. The report contains 134 pages of factual information pertaining to the accident and 37 pages of analysis. The investigation yielded 36 findings and a set of 15 recommendations mostly addressed to the FAA. During the investigation, the NTSB held a three-day public hearing into the accident in which officials from the FAA, Korean Air, the government of Guam, and other organizations gave testimony. The transcript from this hearing spans approximately 430 pages [8].

British Airways Flight 027

On June 28, 1999, British Airways flight 027, a Boeing 747 carrying 419 passengers and crew members en route to Hong Kong, China, and another Boeing 747 operated by Korean Air Cargo nearly collided in flight over a remote region of Chinese airspace. At their closest point of approach, the two aircraft passed within 600 feet of each other, and the British Airways copilot later recounted that his windshield was consumed by the fuselage of the other jet. No injuries resulted from the incident and both aircraft arrived at their destinations. If the two aircraft had collided, however, it is likely that none of the persons aboard either aircraft would have survived [10].

Prior to the incident, the two aircraft were travelling in opposite directions along the same airway with a safe margin of 2,000 feet of vertical separation. The British Airways passenger flight was flying above the Korean Air Cargo flight. The incident sequence began when a collision avoidance system onboard the Korean Air Cargo flight malfunctioned and mistakenly determined the aircraft's altitude to be 2,400 feet higher than its true altitude. This caused the system to believe that a traffic conflict existed between the two aircraft, which prompted it to erroneously instruct the Korean Air pilot to climb in order to avoid the conflict. Because no air traffic control service was available in the region of airspace in which the aircraft were operating and meteorological conditions prevented the pilots from visually identifying each other's aircraft, the Korean Air pilot had no reason to question the collision avoidance system's instruction and thus complied. This placed the two aircraft on a collision course that neither flight crew detected until moments before the aircraft reached their closest point of approach. British Airways officials later noted that it was only by coincidence that the two aircraft avoided each other and that they would have likely collided had they been using more precise navigation systems such as the Global Positioning System (GPS) navigation systems in widespread use today [1].

With the assistance of Korean Air, the CAA determined that the malfunction in the Korean Air Cargo jet's collision avoidance system was caused by damage inflicted during maintenance to the aircraft's avionics systems. Upon concluding its investigation, the CAA issued an airworthiness directive requiring air carriers using similar systems to periodically conduct inspections to ensure the systems are using correct altitude values. The CAA also notified other European aviation regulatory agencies, the FAA, and equipment manufacturers of the problems it found, and it

issued a recommendation to aircraft operators urging them to consider using more robust schemes for handling altitude data.

The U.K. Civil Aviation Authority (CAA) and British Airways each conducted their own investigations into the incident. The CAA's report does not indicate when its investigation into the incident began; however the report is dated October 28, 1999, suggesting that the investigation lasted at most four months. The report is three pages long and includes eight paragraphs of factual information spanning two pages and a single paragraph of analysis. It contains a single conclusion and three recommendations directed at operators and equipment manufacturers. No public hearing was held in response to this incident. British Airways prepared a more detailed report on the incident, but that report has not been officially released to the public.

Event Comparison

In order to help quantify the difference in rigor for the investigations described earlier, we have summarized data from the events and their investigations in Table 1 below.

Table 1 - Comparison of Korean Air flight 801 and British Airways flight 027

	Korean Air 801	British Airways 027
Classification	Accident	Incident
Persons On Board	254	419
Fatalities	228	0
Injuries, Serious	26	0
Injuries, Minor	0	0
Total Casualties	254	0
Aircraft Damage	Destroyed	None
Investigation Length (months)	30	4
Final Report Length (pages)	212	3
Factual Information (pages)	134	2
Analysis (pages)	37	1
Findings / Conclusions	36	1
Recommendations	15	3

The first seven fields listed in Table 1 assess the loss from each incident and the remaining fields attempt to capture the level of rigor applied in the subsequent investigations. Examining the fields pertaining to loss, the near-collision involving British Airways 027 had no casualties compared to

a 90% fatality rate in the Korean Air 801 accident. In addition, neither of the Boeing 747s involved in the near-collision sustained any damage from the incident, whereas the 747 involved in the Guam accident was destroyed.

Comparing these events solely on the basis of loss is deceiving, however, as the British Airways incident could have easily developed into an accident with almost twice the number of fatalities as the Korean Air flight 801 crash in Guam. As British Airways officials noted, it was entirely by luck that the British Airways passenger flight and the Korean Air Cargo flight did not collide. By the time the Korean Air pilot inadvertently placed his aircraft on a collision course with British Airways flight 027, all of the barriers designed to prevent midair collisions had been defeated, and conditions were sufficient for a collision to occur. Indeed, if the incident were repeated under similar circumstances it is likely that a collision would occur, which suggests that the risk of a recurrence of the British Airways flight 027 incident is at least as severe as that of a recurrence of the Korean Air flight 801 accident if not more so.

Given the risk of a recurrence of the near-collision, one would expect a thorough investigation to be conducted in order to determine what prompted the Korean Air Cargo pilot to suddenly veer toward the aircraft flying above. The remaining fields in Table 1 suggest that this was not the case. While in general criteria such as investigation length, report length, and number of findings or recommendations are not indicative of an investigation's thoroughness, the differences indicated in Table 1 between the two incidents are too extreme to ignore. The factual information, analysis, findings, and recommendations from the CAA's investigation into British Airways flight 027 are only a fraction of those from the NTSB's investigation into Korean Air flight 801. This is not because the former was a simple incident. On the contrary, several factors contributed to the loss of separation and subsequent near-collision, including design faults present in the incident aircraft's collision avoidance systems and the systems they interfaced with, human factors issues concerning the manner in which traffic information was displayed to the flight crews, and broader issues concerning the role that collision avoidance systems play in the overall air traffic system. The CAA's report failed to examine these issues, and consequently missed an opportunity to correct problems that might contribute to future incidents, possibly with more dire outcomes.

Under the loss-based accident classification schemes employed by most investigative agencies, such a catastrophic outcome would be necessary for a major investigation to be undertaken, even though the findings and recommendations would likely be the same as if an equally rigorous investigation had been conducted into the incident alone. This should not be the case. New classification schemes are necessary in order to better allocate investigative resources to incidents whose recurrence could have more severe consequences.

In reviewing this comparison, one might argue that the vast difference between the Korean Air and British Airways events was not necessarily because of their associated losses but rather due to the fact that different agencies investigated each event. Had both events been investigated by the NTSB or CAA, the figures might have matched more closely. The NTSB's incident reports tend to match the CAA's report in length, however, and often do not contain immediate safety recommendations (although incident data is aggregated for use in later recommendations). Similarly, if the Korean Air flight 801 accident had occurred in British airspace, it would have been investigated not by the CAA but by the AAIB, whose formal reports are similar to the NTSB's final reports in structure and length.

Risk-based Classification of Incidents

The term "incident" can be defined in a variety of ways but typically involves the failure of a network of barriers designed to protect a system from one or more hazards. An incident becomes an accident when it is coupled with a loss event such as a crash or collision in which damage or casualties are incurred. It is often the case that luck determines whether an incident develops into an accident and, if so, what the extent of the loss will be.

When investigating accidents, investigators can issue recommendations aimed at preventing the associated incident or at mitigating the severity of the loss, and they usually do both. While attempting to mitigate loss given the occurrence of an incident can help to reduce the severity of accidents, some degree of loss is almost always inevitable. On the other hand, if the incident itself is prevented, it cannot develop into an accident and thus no loss will occur. Therefore, recommendations aimed at preventing incident recurrences are likely to be more effective in preventing future losses. Indeed, 13 of the 15 recommendations issued by the NTSB in response to the Korean Air flight 801 accident were aimed at preventing the recurrence of incidents in which aircraft descend below safe altitudes during final approach. Only two focused on mitigating losses by suggesting improvements to Guam's emergency response units.

Given that accidents begin as incidents and that incident prevention should be the focus of investigations, incidents are opportunities for investigators to identify problems and suggest safety improvements without the losses associated with accidents. Accident classification schemes based on loss alone place a low priority on incidents even though those incidents might be indicative of safety problems that could lead to more catastrophic outcomes should they recur. By itself, loss is a poor indicator of an incident's potential for learning new lessons and preventing future incidents. Therefore, classification schemes based on loss should be de-emphasized in favor of new schemes in which resources are allocated to incident investigations based on the risk associated with the incidents' recurrence. To this end, the fundamentals for such a scheme are presented below.

Risk is defined as the probability that an event will occur multiplied by the anticipated cost derived from the occurrence of the event. When an incident occurs, it suggests the presence of a deficiency in the safety systems involved that, if not corrected, could lead to recurrences of the incident. A useful measure of the importance of an incident, therefore, is the total risk that society faces if nothing is done to prevent recurrences. The total risk of such a recurrence is given in Equation 1 below.

$$\begin{aligned} \text{Total Risk} &= E[\#\text{ Recurrences}] \times E[\text{Cost}] \\ &= P[\text{Incident Recurrence}] \times \text{Exposure} \times E[\text{Cost}] \end{aligned} \tag{1}$$

The term E[# Recurrences] represents the expected number of recurrences of the incident if nothing is done to reduce the likelihood of recurrence and is the product of P[Incident Recurrence], the probability that the incident will happen again, and exposure, the number of opportunities for the incident to recur. The term E[Cost] is the expected cost of the incident given that it has occurred and is defined in Equation 2 below.

$$E[\text{Cost}] = \sum_{i \in S} \text{Cost}(i) \bullet P[i] \tag{2}$$

Equation 2 is simply the expectation of the random variable Cost associated with a particular incident. *S* represents the set of all possible outcomes that might result from the occurrence of the incident. For each possible outcome *i*, the cost of *i*, namely the loss, is multiplied by the probability that *i* occurs. The summation of these products yields the expected value of the random variable Cost, which is the expected cost of the incident.

As defined earlier, exposure is the number of chances for an incident to occur. If a particular system has a chance of contributing to an incident each time it is operated, then the exposure from the system is the number of times the system is operated multiplied by the number of such systems in existence. When the system in question is used widely and frequently, this number can become quite large. For example, consider the in-flight breakup of TWA flight 800 over the Atlantic Ocean in 1996. The NTSB concluded that the probable cause of the accident was an explosion of the aircraft's center wing fuel tank, and the Board identified design issues affecting all Boeing 747 airplanes [14]. Exposure in this case would be the number of Boeing 747s in operation multiplied by the number of flights each aircraft would be expected to make in its lifetime. Given the popularity of the 747 and the near impossibility of surviving a commercial aircraft breakup at cruise altitude, the exposure and *E*[Cost] terms of the Total Risk equation would be very large, stressing the importance of implementing the Board's recommendations and reducing *P*[Incident Recurrence] in order to reduce the risk to an acceptable level.

The terms *P*[Incident Occurrence], exposure, and *E*[Cost] follow one's intuition in prioritizing incidents. Clearly, an incident with a high probability of recurrence with high expected costs warrants significant investigation, particularly if numerous systems are already deployed that might also be susceptible to the incident. Likewise, an incident with a small probability of recurrence, a low expected cost, or for which there are only a handful of susceptible systems that are rarely used might warrant only a minor investigation. Thus, Total Risk can be used as a metric to prioritize incident investigations, to determine where investigative resources would be best spent, and to decide which areas regulators, aircraft operators, and equipment manufacturers should focus on first when following up on investigators' recommendations.

As a second example of the use of Total Risk, consider the incident involving British Airways flight 027. It is very difficult to estimate the probability of recurrence but not impossible. The rates of failure of the relevant hardware components are probably known as is the rate of undetected damage occurring during maintenance. The cost of such an incident were it to result in an accident would be very high since there would be considerable loss of life and equipment. Exposure is also likely to be very high because of the prevalent use of TCAS. Thus, a rough estimate of the total risk could be calculated quickly and used as an indicator of the significance of the incident.

Follow-up Actions: A second important use of the concept of Total Risk is to guide the actions taken following an investigation. If Total Risk is high, then the follow-up actions should have a high probability of reducing it to an acceptable level. Many options are available to investigative and regulatory agencies and they need to be used carefully. At one extreme is the option of grounding the fleet and at the other there is the option of no action. In between, there are a variety of possibilities including required inspections, required equipment replacement, required equipment redesign, and so on. There are also options about how quickly any action should occur. Selection among options is a difficult activity if there is no effective mechanism for rating the seriousness of an incident.

Using British Airways flight 027 as an example once more, the actions taken following the incident were insufficient and fragmented despite the fact that Total Risk by the estimation above

was very high. Upon concluding its investigation, the CAA issued an airworthiness directive requiring air carriers using similar equipment to check and periodically inspect the equipment to ensure that it is functioning properly and notified other aviation regulatory agencies as well as equipment manufacturers of the problems it found. It also issued a recommendation to aircraft operators urging them to consider using other encoding schemes for transmitting altitude data since that was part of the problem. The CAA's recommendations did not require mandatory changes and the probability that they would reduce total risk to an acceptable level was small. More importantly, the report by British Airways contains useful insights about the incident yet it has not been made public nor led to appropriate general recommendations.

Iterative Reclassification: As an incident investigation proceeds, new details will emerge that affect the risk of future recurrence. The terms comprising the Total Risk equation will change as the breadth of possible event sequences is narrowed, faults are identified, and remedies are enacted. Consequently, new Total Risk assessments will periodically need to be made, and an investigation's priority relative to others will rise and fall as it is reclassified. After developing an initial set of recommendations, investigators might find that the risk associated with an incident has been reduced to the extent that their efforts would be better spent investigating other incidents with higher Total Risk assessments. Moreover, each reassessment will presumably lower the error in the estimate. Relying only on the initial Total Risk estimate is insufficient because this estimate is based on preliminary information and probably will not have a high degree of confidence associated with it. Therefore, in addition to the Total Risk metric for classifying incidents, a process is necessary to reassess incidents periodically in order to improve the confidence associated with Total Risk estimates.

Until an incident has been categorized, the initial Total Risk assessment cannot be performed, and the investigation into the incident should be given a high priority. Once assessed, the incident can be investigated according to its relative priority among other incidents. Investigators might then choose to reassess the incident on a strictly periodic basis (i.e. monthly or quarterly) or in light of major revelations concerning the investigation that might affect Total Risk, such as when a significant piece of evidence is discovered, when a defect is revealed, when a public inquiry is concluded, when recommendations are issued, or when remedies are implemented. Each reassessment will narrow the confidence interval on Total Risk. If reassessing an incident causes its Total Risk to increase, the investigation should be intensified until the risk is mitigated; if Total Risk decreases, resources can be diverted to more urgent investigations. The investigation may be concluded when investigators are confident that Total Risk has fallen below a predetermined acceptable level, which may depend on the incident's categorization, the type of operation (commercial vs. general aviation, scheduled vs. unscheduled), the flight rules in effect, the type of aircraft, and possibly other factors.

The goal of investigating incidents is to learn lessons that help to prevent the incidents from recurring. Some incidents might be symptomatic of severe defects that could lead to future casualties if not corrected; others could be fairly straightforward and involve accepted risks. By employing the risk-based metric and process proposed above, investigators might be able to determine more accurately which incidents have greater potential for teaching important lessons. Doing so would enable them to allocate resources first to those investigations that would likely have the greatest impact on safety. As a result, investigative agencies could begin to shift from a reactionary role in which loss motivates change to a proactive one focused on risk reduction.

Initial Total Risk Estimates: The Total Risk analysis as we have described above cannot be applied at the outset of an investigation because the data needed to estimate the parameters of the Total Risk equation will not yet be available; however in order to direct the allocation of

resources during an investigation's initial stages, it would be useful to have an estimate of Total Risk, albeit a very crude one. Although investigators will initially know little about an incident, they will have certain information from which a preliminary Total Risk assessment might be developed. This information includes the incident aircraft's flight plan, the type of aircraft, the stage of flight at which the incident occurred, the approximate time and location of the incident, prevailing meteorological conditions, the aircraft's last communication with air traffic control, and possibly preliminary statements from witnesses. These factors might be assembled into an Initial Total Risk Table containing precomputed standard Total Risk estimates compiled from historic statistical data. For a given incident category and set of circumstances, the table would provide estimates of the probability of incident recurrence, exposure, and the expected cost of the incident. Investigators could select which aspects of Total Risk to read from the table and which to estimate directly based on presently available information.

As an example of how the various ideas we have presented might be used, consider the British Airways flight 027 incident described above. Upon learning of the incident and categorizing it as a loss of separation between heavy aircraft, investigators would consult the Initial Total Risk Table using the categorization and other factors mentioned in the previous paragraph to obtain the initial Total Risk estimate. As the investigation progressed, investigators would rely less on the table and transition to estimating the Total Risk parameters directly, improving the accuracy of the Total Risk estimate in accordance with the objectives of Iterative Reclassification.

Remaining Work: The notion of Total Risk is a starting point for a metric that will allow investigators to assess the importance of incidents more accurately and allocate investigative resources accordingly. By assessing incidents based on the risks of future losses from their recurrence rather than their immediate losses, investigators can be more proactive in detecting safety problems before they contribute to accidents involving casualties or damage to aircraft.

Much work remains to be done before this metric can be put into practice. Because incidents are rare occurrences, estimating their probabilities is difficult. A model of cost will be needed to assess the expected loss associated with an incident that takes into account fatalities, serious and minor injures, and damage to aircraft and other property. Moreover, the estimation techniques and reassessment process presented here are intended to serve as examples and are quite preliminary. Before they can be applied to any investigation, they must first be developed more fully and tested on sample incidents to determine their precision. Statistics concerning incident rates and casualties decomposed according to incident type must be computed in order to estimate the parameters comprising the Total Risk equation. While similar statistics already exist, it is unclear whether they are in a form suitable for this purpose. Perhaps most importantly, investigators will need to set acceptable risk levels and establish criteria for determining which level would apply to a given incident.

Once these challenges are overcome, the estimation and assessment procedures would need to be refined so that they could be employed in the field quickly. Total Risk assessment is an overhead exercise and should not significantly detract from investigators' tasks of analyzing incidents and developing recommendations. While high precision cannot be expected from early estimates, they must be accurate enough to provide a rough indication of the worth of investigating an incident. Likewise, later assessments should help guide investigators in determining which aspects of the investigation to pursue next or whether to table the investigation and turn their attention elsewhere.

Conclusions

Commercial aviation accidents are serious occurrences that demand public investigations in order to correct safety problems and prevent future losses. Incidents are also important, however, since they often present the same opportunities to identify new lessons without the losses associated with accidents. Current accident classification schemes used by investigative agencies to allocate resources to investigations place too great an emphasis on the immediate loss from an accident and undervalue the importance of incidents with no loss. Consequently, incidents suggesting the presence of serious safety problems in onboard and ground-based systems are often ignored or not investigated with sufficient rigor to uncover these problems, which if left uncorrected could contribute to future incidents with more tragic outcomes. This dilemma was illustrated by the large disparity in the investigations conducted into the Korean Air flight 801 and British Airways flight 027 incidents. The latter received a much less rigorous investigation even though both incidents carried a high risk of recurrence.

Precedent existed for both of the incidents described in this paper. The Korean Air flight 801 accident followed a similar incident in 1994 that also involved a mis-configured MSAW system in which a Transportes Aereos Ejecutivos, S.A. Learjet crashed on final approach to runway 1R at Dulles International Airport approximately 0.8 nm short of the runway. The British Airways flight 027 incident followed a similar incident that also involved TCAS processing incorrect altitude data that occurred between two aircraft in January 1998 over Hawaii [3]. These prior incidents indicated the presence of serious problems with the manner in which the affected systems were designed or maintained; however the investigations either failed to address these problems or the follow-up actions were insufficient to correct them. As a result, opportunity remained for similar incidents to recur, *and they did*.

To mitigate this problem, investigators should reconsider the practice of classifying incidents based on their losses, and instead classify them based on the risk of future losses. Adopting risk-based schemes will allow investigators to be more proactive and address safety problems before they contribute to accidents with extensive casualties. For risk-based classification schemes to be useful, techniques will have to be developed for investigators to quickly assess the risk level of incidents early in the investigative process so that they can allocate resources accordingly.

This work was funded in part by NASA Langley Research Center under grants numbered NAG-1-2290 and NAG-1-02103.

References

1. Carley, William M. “Wires Crossed: Flawed Safety Device In Jets Gets Blamed For a Near Catastrophe.” *Wall Street Journal.* 12 October 1999, eastern ed.: A1.
2. Federal Aviation Administration, “Aircraft Accident and Incident Notification, Investigation, and Reporting”, Order 8020.11B. 16 August 2000. Washington, D.C.
3. Greenwell, William S. and Knight, John C. “What Should Aviation Safety Incidents Teach Us?” Technical report. CS-2003-12. Department of Computer Science, University of Virginia. 20 March 2003.
4. National Transportation Safety Board. “Accidents, Fatalities, and Rates, 2002 Preliminary Statistics, U.S. Aviation.” <http://www.ntsb.gov/aviation/Table1.htm>
5. National Transportation Safety Board. “Accidents and Accident Rates by NTSB Classification, 1983 through 2002, for U.S. Air Carriers Operating Under 14 CFR 121.” <http://www.ntsb.gov/aviation/Table2.htm>

6. National Transportation Safety Board. *Aviation Investigation Manual: Major Team Investigations*.
7. National Transportation Safety Board. *Controlled Flight Into Terrain, Korean Air Flight 801, Boeing 747-300, HL7486, Nimitz Hill, Guam, August 6, 1997*. Aircraft Accident Report NTSB/AAR-00/01. Washington, DC.
8. National Transportation Safety Board. *Public Hearing in Connection With the Investigation of Aircraft Accident, Korean Air Flight 801, B-747-300, Agana, Guam, August 6, 1997*. 24 March 1998. Honolulu, Hawaii.
9. Transportation Safety Board of Canada. "Investigation Process." (18 September 2002). <http://www.tsb.gc.ca/en/investigation_process/what_we_do.asp>
10. U. K. Civil Aviation Authority. "Hazardous Loss of Separation Between Two Aircraft Over Chinese Airspace." Doc Ref KMH/Pap/059, issue 1. 28 October 1999. London, U. K.
11. "Notification and Reporting of Aircraft Accidents or Incidents and Overdue Aircraft, and Preservation of Aircraft Wreckage, Mail, Cargo, and Records." *Code of Federal Regulations*. 2002 ed. Title 49, Pt. 802, p. 1195.
12. Aviation Safety Reporting System. "Program Overview." <http://asrs.arc.nasa.gov/overview.htm>
13. Johnson, Chris. "The Limitations of Aviation Incident Reporting." <http://www.dcs.gla.ac.uk/~johnson/papers/reminders/>
14. National Transportation Safety Board. *In-flight Breakup Over The Atlantic Ocean, Trans World Airlines Flight 800, Boeing 747-131, N93119, Near East Moriches, New York, July 17, 1996*. Aircraft Accident Report NTSB/AAR-00/03. Washington, D.C.
15. Federal Aviation Administration. "Advisory Circular: Aviation Safety Action Program (ASAP)." Advisory Circular 120-66B. 15 November 2002. Washington, D.C.
16. Leveson, Nancy G. *Safeware: System Safety and Computers*. Reading: Addison-Wesley. 1995.

Biographies

William S. Greenwell, Department of Computer Science, University of Virginia; Charlottesville, Virginia, U.S.A.; telephone - +1.434.982.2298; fax - +1.434.982.2214; e-mail - greenwell@cs.virginia.edu

William Greenwell is a graduate student at the University of Virginia. His primary interests include software dependability, incident analysis, aviation, and Japanese animation.

John C. Knight, Department of Computer Science, University of Virginia; Charlottesville, Virginia, U.S.A.; telephone - +1.434.982.2216; fax - +1.434.982.2214; e-mail - knight@cs.virginia.edu

John Knight is a professor of Computer Science at the University of Virginia. His primary interest is software dependability for safety-critical applications.

Elisabeth A. Strunk, Department of Computer Science, University of Virginia; Charlottesville, Virginia, U.S.A.; telephone - +1.434.982.2292; fax - +1.434.982.2214; e-mail - strunk@cs.virginia.edu

Elisabeth Strunk is a graduate student at the University of Virginia. Her primary interests are software systems dependability and practicing aikido.