

Eine erste Simulation zur Optimierung des Durchsatzes beim Einsatz automatischer Sprengstoffdetektion für die Handgepäckkontrolle an Flughäfen

Yanik STERCHI, Adrian SCHWANINGER

*Institut Mensch in komplexen Systemen (MikS), Hochschule für Angewandte Psychologie, Fachhochschule Nordwestschweiz (FHNW)
Riggenbachstrasse 16, CH-4600 Olten*

Kurzfassung: Die Kontrolle des Handgepäcks von Flugpassagieren leistet einen wichtigen Beitrag zur Flugsicherheit. Sprengstofferkennungssysteme (explosive detection systems, EDS) bieten bezüglich Sicherheit eine wertvolle Ergänzung zur Röntgenbildbeurteilung durch Sicherheitspersonal, können aber zu mehr Fehlalarmen führen, was wiederum den Durchsatz reduzieren kann. Mittels ereignisorientierter Simulation wurde untersucht, wie EDS in Abhängigkeit von Maschineneinstellung, Sicherheitspersonal und Alarmauflösung den Durchsatz bei der Handgepäckkontrolle beeinflusst. Um das Verhalten des Sicherheitspersonals in Verbindung mit EDS zu modellieren, wurden Daten aus einem Experiment verwendet. Die Resultate der Simulation legen nahe, dass trainiertes und erfahrenes Sicherheitspersonal bei Einsatz von EDS mit tiefen Fehlalarmraten Durchsatzeinbussen kompensieren kann, EDS mit höheren Fehlalarmraten jedoch schnellere Verfahren und zusätzliches Personal bei der Alarmauflösung erfordern.

Schlüsselwörter: Luftsicherheit, Sprengstofferkennungssysteme (explosive detection systems, EDS), Mensch-Maschine-Interaktion, ereignisorientierte Simulation, Durchsatz

1. Einführung

Da eine sichere Luftfahrt essenziell für Gesellschaft und Wirtschaft ist, werden üblicherweise alle Passagiere inklusive mitgeführtem Gepäck an Flughafen-Sicherheitskontrollen (FSK) überprüft. Diese Überprüfung soll sicherstellen, dass keine verbotenen Gegenstände (Schusswaffen, Messer, unkonventionelle Spreng- und Brandvorrichtungen sowie andere Gefahrenobjekte) an Bord der Flugzeuge gelangen. Die Leistung der FSK hängt sowohl von menschlichen Faktoren, Maschinenattributen als auch den Prozessen ab, die deren Interaktion definieren (Koller et al. 2009; Mendes et al. 2013; Schwaninger et al. 2010; Schwaninger 2006).

An einer typischen FSK müssen Passagiere ihr Gepäck sowie weitere mitgeführte Gegenstände, z.B. Jacken, Kopfbedeckung und Inhalt der Hosentaschen, kurzzeitig abgeben. Gepäck und Gegenstände werden dann von einem Röntgengerät durchleuchtet und ein Mitglied des Kontrollpersonals (Screener) durchsucht das Röntgenbild nach verbotenen Gegenständen. Wird ein verdächtiges Objekt entdeckt, wird das Bild als "nicht OK" gekennzeichnet und das zugehörige Gepäckstück (oder sonstige Gegenstände) zur weiteren Untersuchung (sekundäre Kontrolle) durch ein weiteres Mitglied des Kontrollpersonals umgeleitet. Diese sekundäre Kontrolle ist

typischerweise eine manuelle Durchsuchung des Gepäcks und/oder Sprengstoffspurendetektion (explosives trace detection, ETD), welche darin besteht, an mehreren Stellen des Gepäckstückes einen Abstrich vorzunehmen und diesen dann von einer Maschine auf Spuren von Sprengstoffrückständen hin zu analysieren (Butler & Poole 2002; Wells & Bradley 2012). Handelt es sich bei dem verdächtigen Objekt tatsächlich um einen verbotenen Gegenstand, spricht man von einem "Treffer". Handelt es sich hingegen um einen harmlosen Gegenstand, bezeichnet man den Alarm als "Fehlalarm".

In den vergangenen Jahren ist die Erkennung von Sprengstoff für die Sicherheit in der zivilen Luftfahrt zunehmend in den Fokus gerückt. Im Zuge dessen wurden neben ETD auch röntgenbasierte Sprengstofferkennungssysteme (explosive detection systems, EDS) für die Handgepäckkontrolle verfügbar. Diese Systeme nutzen die Information von Multienergie-Röntgengeräten über Ordnungszahl und Dichte des geröntgten Materials zur Erkennung von Sprengstoff (Singh & Singh 2003). EDS bringt bei richtigem Einsatz eine grosse Verbesserung der Detektion von Sprengstoff bei der FSK (Mendes et al. 2015). Für den Erfolg der Technologie ist jedoch auch ausschlaggebend, wie sich diese auf den Durchsatz der FSK auswirkt. Dies wurde daher in dieser Studie für einen spezifischen Prozess mittels ereignisorientierter Simulation untersucht. Zusätzlich wurden zwei Massnahmen auf ihr Potential, potentielle negative Effekte des EDS Einsatzes auf den Durchsatz zu reduzieren, untersucht.

2. Prozessbeschreibung und Annahmen

EDS-Geräte analysieren die Röntgenbild-Information auf potentielle Sprengstoffe noch bevor das Bild dem Screener eingeblendet wird. Diese Analyse geschieht sehr schnell und führt nicht zwingend zu einer Verlangsamung der Röntgenanalyse. Der potentielle Einfluss auf den Durchsatz liegt vielmehr in der Auflösung der Alarme des EDS. Für diese Alarmauslösung existieren zumindest zwei Optionen: Entdeckt das EDS eine Region im Röntgenbild, wo sich Sprengstoff befinden könnte, kann diese Region z.B. durch Einrahmen für den Screener markiert werden. Dieser entscheidet dann, ob eine sekundäre Kontrolle (manuelle Suche oder ETD) durchgeführt werden soll. Alternativ kann der Automatisierungsgrad erhöht und ein Gepäckstück, das den Alarm des EDS auslöst, direkt zur sekundären Kontrolle weitergeleitet werden. In dieser Studie wurde diese zweite Option genauer untersucht. Weiter wurde angenommen, dass Alarme des EDS grundsätzlich durch ETD aufgelöst werden, da diese Methode auf die Erkennung von Sprengstoffen spezialisiert ist. Dies bedeutet gleichzeitig, dass auch Bilder, die einen Alarm des EDS auslösen, durch den Screener begutachtet werden müssen, um auszuschliessen, dass weitere Gefahrenobjekte wie Schusswaffen und Messer vorhanden sind. Liegt jedoch ein Verdacht auf Schusswaffen oder Messer vor, wird das betroffene Gepäckstück manuell durchsucht.

Nach der kurzen Beschreibung des Prozesses der sekundären Kontrolle, sollen nun die Modellannahmen zu einigen wichtigen Variablen erläutert werden. Neben der Prozesszeit sind die wichtigsten Attribute eines EDS bezüglich des Durchsatzes die Menge der generierten Alarme, die aufgelöst werden müssen, und indirekt auch dessen Fähigkeit, Sprengstoff zu erkennen. Gleich wie der Alarm eines Screeners kann auch der Alarm eines EDS sowohl ein Treffer oder ein Fehlalarm sein. Die Häufigkeit, mit der bei Vorhandensein von Sprengstoff ein Alarm bzw. bei

Abwesenheit von Sprengstoff ein Fehlalarm ausgelöst wird, nennt sich Trefferrate respektive Fehlalarmrate. Während die Trefferrate eines EDS keinen direkten Einfluss auf den Durchsatz ausübt, da Sprengstoff im Alltag vernachlässigbar selten vorkommt, muss diese gegen die Fehlalarme abgewogen werden: wird ein EDS empfindlicher eingestellt, steigt sowohl die Treffer- als auch die Fehlalarmrate. Letztere hängt also nicht nur von dem Maschinentyp sondern auch von der Maschineneinstellung ab. Von für die Zertifizierung von EDS-Maschinen zuständigen Behörden wurde uns mitgeteilt, dass derzeit Fehlalarmraten zwischen 6% und 18.6% realistisch sind. Noch tiefere Fehlalarmraten sind jedoch möglich, wenn bloss ein Teil der Gepäckstücke mit EDS überprüft werden (z.B. 3%, wenn ein EDS mit einer Fehlalarmrate von 6% die Hälfte der Gepäckstücke überprüft). In dieser Studie wurden Fehlalarmraten von 1-15% berücksichtigt.

Wie sich die Einführung von EDS in die Handgepäckkontrolle auf deren Durchsatz auswirkt, hängt auch davon ab, wie lange die Auflösung der Alarme des EDS durch ETD dauern. Die Dauer einer solchen sekundären Kontrolle mittels ETD hängt massgeblich davon ab, wo und wie viele Abstriche an einem Gepäckstück vorgenommen werden (Butler & Poole 2002). Um die hier mögliche Variation in der Simulation zu berücksichtigen, wurden drei Bedingungen simuliert, in denen die sekundäre Kontrolle mittels ETD im Durchschnitt 30, 60 bzw. 120 Sekunden dauert.

Ebenfalls wichtig zu berücksichtigen ist die Anzahl der Alarme durch den Screener: Umso mehr Alarme (sowohl Fehlalarme als auch Treffer, z.B. beim Vorhandensein von Flüssigkeiten oder Scheren) ein Screener generiert, desto kleiner ist die Kapazität für die Auflösung der Alarme des EDS. Es gibt mehrere Faktoren, die bekanntermassen die Treffer- und Fehlalarmrate von Screenern beeinflussen. Insbesondere für computer-gestütztes Training konnte schon mehrfach ein starker Effekt auf die Erkennungsleistung (Halbherr et al. 2013; Koller et al. 2008; Schwaninger & Hofer 2004; Schwaninger 2006) und auch spezifisch auf die Fehlalarmrate (Koller et al. 2009) aufgezeigt werden. Arbeitserfahrung scheint erst in Begleitung mit Training einen Effekt zu haben (Schwaninger et al. 2010; Schwaninger et al. 2007), während Alter zwar einen negativen Effekt ausübt, dieser jedoch im Vergleich mit Trainingseffekten eher klein ausfällt (Ghylin et al. 2006).

Um die durchsatzmindernden Effekte der Fehlalarme eines EDS zu reduzieren, stehen verschiedene Massnahmen zu Verfügung. Eine Verkürzung der sekundären Kontrolle durch schnelle ETD-Verfahren könnte bereits als eine Massnahme betrachtet werden, die in der Simulation durch die drei ETD-Bedingungen bereits abgedeckt ist. Als eine weitere Massnahme wurde der Einsatz einer zweiten Person, die sekundäre Kontrollen durchführt, getestet. Auch getestet wurde Unterstützung durch den Screener bei der sekundären Kontrolle, sobald so viele sekundäre Kontrollen anstehen, dass der Röntgenprozess ohnehin unterbrochen werden muss.

3. Methoden

Die Simulation wurde in FlexSim, einer Software für 3D-Modellierung und ereignisorientierte Simulation, implementiert. Layout, Prozesse und Parameter wurden mittels Daten aus einer Arbeitsbeobachtung und Interviews auf die FSK eines europäischen Flughafens abgestimmt. Auf eine Diskussion aller Parameter und Verteilungsannahmen kann hier leider nicht eingegangen werden.

Für die Simulation der Screenerleistung wurden Daten aus einer experimentellen Studie entlehnt. In dieser Studie analysierten 150 zertifizierte Screener Röntgenbilder

entweder mit oder ohne Unterstützung durch EDS (Mendes et al. 2015), welche entweder durch Markierung der verdächtigen Stelle im Röntgenbild oder durch automatisches Weiterleiten zur sekundären Kontrolle bestand. Im zweiten Fall, der dem simulierten Prozess entspricht, zeigte sich weder in der Fehlalarmrate noch in der Beurteilungszeit ein Unterschied zur Röntgenbildanalyse ohne EDS. Bezüglich Fehlalarmraten zeigten sich aber Unterschiede in Abhängigkeit von Erfahrung und Training. Für die Simulation der Röntgenbildanalyse durch die Screener wurden daher die Fehlalarmwerte von drei Referenzgruppen aus der besagten Studie verwendet (siehe Tabelle 1).

Für jede Kombination aus den drei Referenzgruppen, den 16 Fehlalarmraten von 0-15% (in 1% Schritten, 0% ist Bedingung ohne EDS) sowie den drei ETD-Bedingungen ($M = 30$ s, 60 s, 100 s) wurden separate Simulationsläufe durchgeführt. Um die Resultate überschaubar zu halten, wurden die zwei Massnahmen bloss in Kombination mit der ersten Referenzgruppe und der mittleren ETD-Bedingung ($M = 60$ s) durchgeführt. Für jede der resultierenden 192 Bedingungen wurde 200-mal eine Stunde des Handgepäckkontrollprozesses simuliert und dann aggregiert. Zur Simulation der Kapazität unter Volllastung wurde ein stets hohes Passagieraufkommen implementiert.

Tabelle 1: Referenzgruppen aus Mendes et al. (2015). Gruppenmittelwerte und Standardabweichungen in Klammern. Trainingszeit schliesst nicht verfügbare Ersttrainings aus. Gruppe 1 stammt von einem anderen Flughafen als 2 und 3.

	Fehlalarmrate	Trainingszeit	Arbeits Erfahrung [J.]	Alter
Referenzgruppe 1	.025 (.039)	101.40 (31.47)	7.68 (4.85)	42.50 (10.52)
Referenzgruppe 2	.040 (.046)	28.56 (12.41)	8.24 (5.78)	36.55 (8.46)
Referenzgruppe 3	.049 (.031)	2.58 (2.00)	< 1 Jahr (-)	30.81 (10.93)

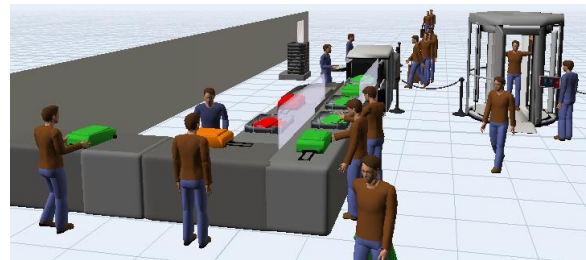


Abbildung 1: Screenshot des 3D Modelles in FlexSim. Passagiere sind in Braun, Sicherheitspersonal in Blau. Als OK eingestufte Gegenstände sind in Grün, für sekundäre Kontrolle umgeleitete Gegenstände in Rot und Gegenstände, die sich im sekundären Kontrollprozess befinden, in Orange.

4. Resultate und Diskussion

Abbildung 2 zeigt den simulierten Durchsatz für eine einzelne FSK-Spur abhängig von der Fehlalarmrate des EDS. Der Durchsatz ging erwartungsgemäss mit zunehmender Fehlalarmrate des EDS zurück, wenn keine Prozessanpassungen vorgenommen wurden. Für das Ausmass des Rückgangs war dabei die mittlere Dauer der sekundären Kontrolle mit ETD entscheidend. Ebenso war der Durchsatz abhängig davon, ob die Screener gut trainiert und erfahren waren. Wie oben erwähnt, ist eine Fehlalarmrate von 6% für aktuell verfügbare EDS unter Analyse aller Gepäckstücke realistisch. Diese Fehlalarmrate führte zu einer Durchsatzreduktion von 40% bei einer mittleren Dauer von 120 s für die EDS-Kontrolle. Reduziert sich diese auf 60 s, nahm der Durchsatz je nach Referenzgruppe (in aufsteigender Reihenfolge) 19%, 23% bzw. 25% ab, bei 30 s

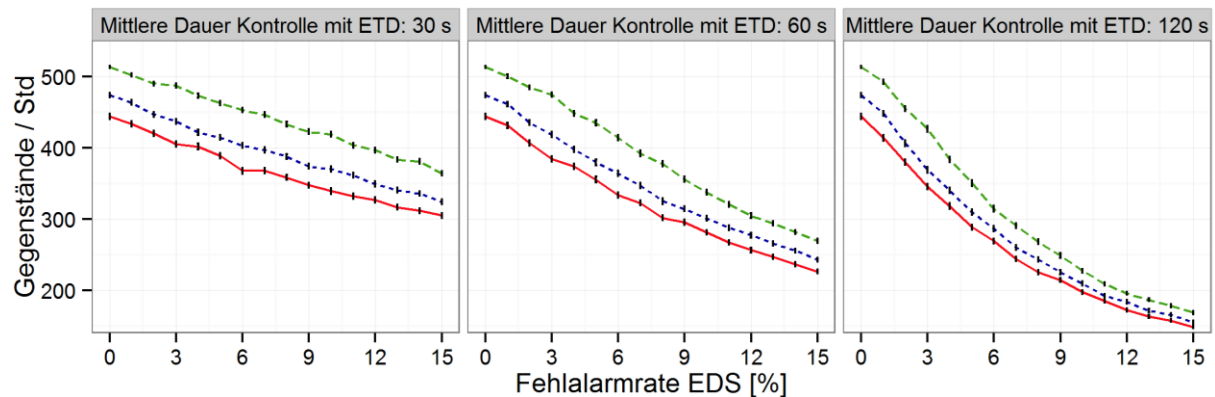


Abbildung 2: Mittelwert und Standardfehler des Durchsatzes in Gegenstände pro Stunde, abhängig von der Fehlarmsrate des EDS (Null entspricht dem Ausgangswert ohne EDS) und Referenzgruppe, grün gestrichelt: Referenzgruppe 1 (Flughafen 1, Arbeitserfahrung > 2 Jahre), blau gepunktet: Referenzgruppe 2 (Flughafen 2, Arbeitserfahrung > 2 Jahre), rot durchgezogen: Referenzgruppe 3 (Flughafen 2, Arbeitserfahrung < 1 Jahr), und mittlerer Dauer des ETD von 30, 60 und 120 s.

12%, 15% bzw. 17%. In diesem letzten Fall ist bemerkenswert, dass die erste Referenzgruppe mit 6% Fehlarmsraten des EDS immer noch einen höheren Durchsatz erzielte als die dritte Referenzgruppe ohne EDS. Die Resultate legen nahe, dass ein hoher Durchsatz trotz EDS möglich ist, wenn die sekundäre Kontrolle kurzgehalten werden kann und die Screener gut trainiert sind.

Abbildung 2 zeigt den Zusammenhang zwischen Durchsatz und Fehlarmsrate des EDS für die Standardkonfiguration sowie die beiden vorgängig beschriebenen Massnahmen. Wie zu erwarten, reduzierte sich der Effekt der Fehlarmsrate des EDS auf den Durchsatz durch den Einsatz einer zweiten Person für die sekundäre Kontrolle stark. Unterstützung durch den Screener bei der sekundären Kontrolle im Falle eines Unterbruchs des Röntgenprozesses zeigte erst bei hohen Fehlarmsraten einen positiven Effekt auf den Durchsatz. In der Praxis könnte diese Unterstützung jedoch positiver ausfallen, falls der Screener nur kurze sekundäre Kontrollen oder Teilprozesse durchführt und somit den Unterbruch des Röntgenprozesses reduziert.

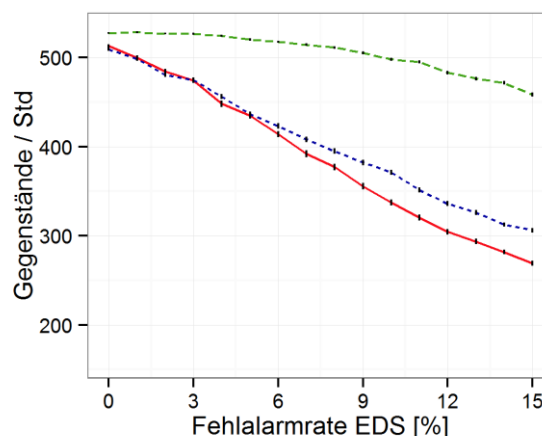


Abbildung 3: Mittelwert und Standardfehler des Durchsatzes in Gegenstände pro Stunde, abhängig von der Fehlarmsrate des EDS (Null entspricht dem Ausgangswert ohne EDS), entweder mit: rot durchgezogen: einzelne Person für sekundäre Kontrolle, blau gepunktet: Screener hilft bei sekundärer Kontrolle bei Unterbruch des Screeningprozesses, grün gestrichelt: zweite Person für sekundäre Kontrolle.

Zusammenfassend weisen die Resultate der Simulation darauf hin, dass der Durchsatz einer FSK ohne geeignete Massnahmen durch den Einsatz von EDS erheblich reduziert wird. Dieser Effekt kommt hauptsächlich durch die sekundäre Kontrolle zustande, was die Wichtigkeit kurzer ETD-Verfahren betont. Der Einsatz einer zweiten Person für die sekundäre Kontrolle konnte erwartungsgemäss den negativen Effekt auf den Durchsatz stark abfedern. Die Fehlalarmrate der Screener erwies sich als ein sehr wichtiger Faktor. Training und die damit verbundene Reduktion der Fehlalarmrate (Koller et al. 2009) könnte somit den negativen Effekt von EDS auf den Durchsatz zu einem grösstenteils kompensieren. EDS bringt bei richtigem Einsatz eine grosse Verbesserung der Detektion von Sprengstoff bei der FSK (Mendes et al. 2015). Durch die in dieser Studie aufgezeigten Massnahmen lassen sich negative Effekte auf den Durchsatz an einer FSK kompensieren.

5. Literatur

- Butler, V., Poole, R.W. (2002) Rethinking Checked-Baggage Screening. Reason Public Policy Institute, Policy Studies, 297:1–25.
- Ghylin, K.M., Drury, C.G., Schwaninger, A. (2006) Two-component model of security inspection: application and findings. 16th World Congress of Ergonomics, IEA 2006. Maastrich.
- Halbherr, T., Schwaninger, A., Budgell, G.R., Wales, A. (2013) Airport Security Screener Competency: A Cross-Sectional and Longitudinal Analysis. *The International Journal of Aviation Psychology*, 23 (2):113–129.
- Koller, S.M., Hardmeier, D., Michel, S., Schwaninger, A. (2008) Investigating training, transfer and viewpoint effects resulting from recurrent CBT of X-Ray image interpretation. *Journal of Transportation Security*, 1 (2):81–106.
- Koller, S.M., Drury, C.G., Schwaninger, A. (2009) Change of search time and non-search time in X-ray baggage screening due to training. *Ergonomics*, 52 (6):644–56.
- Mendes, M., Schwaninger, A., Michel, S. (2013) Can laptops be left inside passenger bags if motion imaging is used in X-ray security screening? *Frontiers in Human Neuroscience*, 7 (October), 1–10.
- Mendes, M., Hättenschwiler, N., Sterchi, Y., Schwaninger, A. (2015) Advanced cabin baggage (ACBS) study on human-machine performance and automation. Paper presented at the 61st Meeting of the ECAC Technical Task Force, Paris, June 4-5.
- Schwaninger, A., Hofer, F. (2004) Evaluation of CBT for increasing threat detection performance in X-ray screening. In K. Morgan, M.J. Spector (Hrsg.), *The Internet Society 2004, Advances in Learning, Commerce and Security*.
- Schwaninger, A. (2006) Airport security human factors: From the weakest to the strongest link in airport security screening. *Proceedings of the 4th International Aviation Security Technology Symposium*, 265–270. Washington D.C.
- Schwaninger, A., Hofer, F., Wetter, O. (2007) Adaptive computer-based training increases on the job performance of x-ray screeners. *Proceedings of the 41st Carnahan Conference on Security Technology*, 117–124. Ottawa.
- Schwaninger, A., Hardmeier, D., Riegelnic, J., Martin, M. (2010) Use It and Still Lose It? *GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry*, 23 (3):169–175.
- Singh, S., Singh, M. (2003) Explosives detection systems (EDS) for aviation security. *Signal Processing*, 83 (1):31–55.
- Wells, K., Bradley, D.A. (2012) A review of X-ray explosives detection techniques for checked baggage. *Applied Radiation and Isotopes*, 70 (8):1729–1746.