CHAPTER 13

Threats and Attack Strategies Used in Past Events: A Review

Konstantinos-Giorgos Thanos, Dimitris M. Kyriazanos, and Stelios C. A. Thomopoulos

13.1 Introduction

The European Union's strategy for integrated border management across all border modalities (air, land, sea) is based on the four-tier access control model "which covers measures in third countries, measures with neighbouring third countries, border control measures at the external borders, risk analysis and measures within the Schengen Area and return" [1]. TRESSPASS EU research project [2] works on assessing the operational benefits and added value from deployment of risk-based border security management concept across all tiers of the access control and all border modalities. An innovative concept and a paradigm shift from current

K.-G. Thanos (✉) · D. M. Kyriazanos · S. C. A. Thomopoulos
Institute of Informatics and Telecommunications, National Centre for Scientific Research "Demokritos", Agia Paraskevi, Greece
e-mail: giorgos.thanos@iit.demokritos.gr; dkyri@iit.demokritos.gr; scat@iit.demokritos.gr

B. Akhgar et al. (eds.), *Technology Development for Security Practitioners*, Security Informatics and Law Enforcement,
https://doi.org/10.1007/978-3-030-69460-9_13

practice, the risk-based approach aims to assist Border Guard Authorities to focus their resources where and when it matters, based on a dynamic and intelligent analysis of risk. The expected impact aims at smarter and more efficient security controls while reducing waiting times and frustration for the increasing number of travellers and passengers across Europe.

Figure 13.1 depicts the TRESSPASS risk-based border security management concept with the TRESSPASS Front End technologies covering Tier 3 Border Control Point (BCP) area, Tier 1 and 2 connected through use of (i) TRESSPASS International Alert System (IAS) and (ii) legacy information systems (e.g. Visa Information System, Schengen Information System, Passenger Name Record, Advance Passenger Information), while Tier 4 is addressed through advanced correlation and analytics, capable of identifying patterns within the Schengen Area that can be connected with higher-risk ranking.

IAS provides a protocol and intelligence sharing component for collaboration with neighbouring and third countries, being critical links in the chain of intelligence. This includes information from all involved authorities, such as border guards, police and customs, about persons of

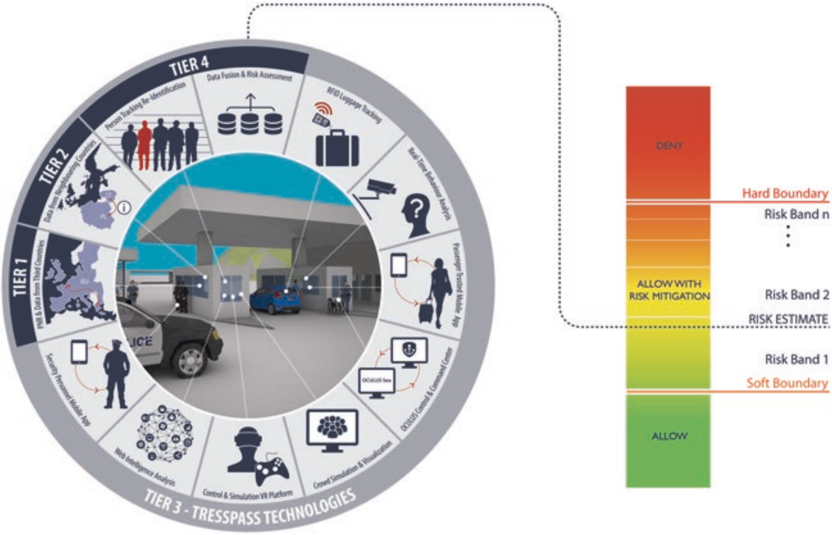
TIER 4
Data Fusion & Risk Assessment
Person Tracking Re-Identification
TIER 2
Data from Neighbouring Countries
TIER 1
PNR & Data from Third Countries
RFID Luggage Tracking
Real-Time Behaviour Analysis
Passenger Trusted Mobile App
OCULUS Control & Command Center
OCULUS Sea
Crowd Simulation & Visualization
Control & Simulation VR Platform
Web Intelligence Analysis
Security Personnel Mobile App
POLICE
TIER 3 - TRESSPASS TECHNOLOGIES
DENY
Hard Boundary
Risk Band n
ALLOW WITH RISK MITIGATION
Risk Band 2
RISK ESTIMATE
Risk Band 1
Soft Boundary
ALLOW

Fig. 13.1 Overview of TRESSPASS concept for border security risk analysis and management

interest or even high-risk warnings for illegal activities. Front End technologies include all the deployed TRESSPASS sensors and components on the field and in the area of the BCP: surveillance cameras and computer vision algorithms, location-sensing beacons for travellers and carry-on luggage, location-based services offered to travellers and assistive mobile applications for enhanced awareness offered to border guards. Finally, simulation and VR can offer valuable training and "what-if" scenario decision support, feeding also pattern recognition and deep learning algorithms with data that are scarce to find in normal everyday operations. All the aforementioned components provide input to the data fusion and risk assessment procedure, which is responsible of providing risk metrics across the four tiers of access control and most importantly to the right place, time, authority and security officer.

Risk assessment procedure is based on the analysis of risk factors of potential risk for malicious incidents jeopardizing critical infrastructures related to border crossing such as airports and harbours. These risk factors are determined by authorized security authorities and refer to indicators of risk of undesired of illegal activity to take place within the infrastructure. The risk factors can either be determined by reports of security personnel, or by the real-time information stemming from surveillance equipment installed on each infrastructure pre-processed by intelligent components that leverage raw data to meaningful high-level information or by infrastructures visitors profile constructed by travel documents, PNR and other available official documentation. Between risk factors and values and high-level information resulting from surveillance intelligent components or profile data, there is a conceptual gap which is covered by the data fusion component. Data fusion role is to aggregate the available input from the available heterogeneous information sources and approach each risk factor. Although these information sources are defined with the aim of providing evidence about the risk factors, the factor approximation cannot be performed without uncertainty. This uncertainty occurs due to the prediction errors that may result from the raw data pre-processing of the surveillance systems and the statistical uncertainty that result from the probabilistic models used to approach risk factors from the heterogeneous sources input. As a result, the data fusion algorithms were designed accordingly in order to be robust in cases of uncertainty or erroneous input. In the following sections, there will be presented briefly a state-of-the-art overview, various techniques along with corresponding pros and cons and finally the determination of the algorithm realized in the DFA component and the several options and decisions that were needed to be taken. The proof of concept and the corresponding evaluation is presented in the last section.

13.2 Related Literature

13.2.1 Information Fusion

In this regard, Bayes rule (13.1) is exploited in order to link the posterior probability which corresponds to the estimated incident value given the available evidence, with prior beliefs about the expectation of the occurrence and the respective likelihood of the incident.

$$P(z) = \frac{P(z \mid x) \cdot P(x)}{P(z)} \tag{13.1}$$

Although this technique provides more accurate results compared to other method, it is not appropriate for heterogeneous types of sensors and for cases where assignment of probabilities to unknown propositions beforehand is inevitable.

Interval-Based Fusion

Interval-based methods mainly address a crucial weakness of Bayesian methods having to do with uncertainty management. In this approach uncertainty is represented as an interval between upper and lower parameter limits (e.g. $x \in [a, b]$) where no any probabilistic distribution of x over the interval is implied.

Interval-based fusion method has the benefit of providing a good measure of uncertainty in case of lacking probabilistic information. However, these algorithms cannot guarantee convergence. Moreover, these methods are not appropriate for encoding dependencies between variables.

Fuzzy Logic-Based Fusion

Fuzzy logic is a generalization of rule-based reasoning by extending each rule outcome and related fact values from binary (true or false) to real number ranging from 0 to 1. Fuzzy logic inference follows the process below:

1. Fuzzification of input values: Map input variables to membership function.
2. Apply fuzzy rules and compute the output membership functions.
3. Defuzzification of output memberships to specific values, which corresponds to the outcome estimation.

Fuzzy logic inference demands expert knowledge to be provided and is characterized by high complexity in the learning phase of membership functions.

Evidence-Based Fusion

Evidence-based fusion is distinguished from the other probabilistic methods by treating uncertainty not with the strict sense of probability but in a more generalized context where Kolmogorov axioms are not verified. In these methods mass functions are assigned to elements, sets and subsets of elements. Particularly, each incident is represented by a basic probability in the interval [0, 1], which expresses the amount of relevant evidence which is available for supporting this incident. Moreover two uncertainty measures are defined, the belief function *Bel()* and the plausibility function *Pls()*, both ranging from zero to one, which correspond to an upper and lower bound, respectively, of the incident uncertainty.

Evidence-based fusion methods generalize Bayesian inference and additionally have the capability of managing heterogeneous types of information sources, and it is efficient in cases of assignment probabilities to unknown propositions beforehand. However, evidence-based fusion methods are less accurate than Bayesian methods, and they are characterized by higher time complexity.

Summary

The above-mentioned methods are summarized below:

Fusion algorithmic approach	*Pros*	*Cons*
Bayesian fusion	More accurate compared to other fusion methods	Not appropriate for heterogeneous types of sensors A priori assignment of probabilities to unknown propositions

Fusion algorithmic approach	*Pros*	*Cons*
Evidence-based (Dempster-Shafer)	Generalizes Bayesian fusion Capability for heterogeneous sources fusion Assignment of a priori probabilities to unknown propositions is not needed It is closer to human perception and reasoning process	Less accurate than Bayesian approach Higher time complexity
Fuzzy logic-based fusion	Ability of enhancing data quality	Expert knowledge is needed High complexity of membership learning
Interval-based fusion	Intervals provide a good measure of uncertainty in case of lacking probabilistic information	Difficult to get results that converge to specific value Difficult to encode dependencies between variables

13.3 Proposed Solution

The proposed solution regards the intelligent automated real-time surveillance of a critical infrastructure and the corresponding risk detection. In this regard, several sensing components are distributed within the infrastructure which in real time record measurements related to crowd behaviour. These measurements correspond to raw numerical data related to crowd behavioural aspects. To this end, an efficient methodology is needed capable of tolerating the possibility of faulty or missing values, but on the other hand keeping the performance to a satisfying level. Moreover, due to the nature of the sensors input, it is mandatory to the implemented algorithms to be robust with heterogeneous types of sources and have to be flexible with expert knowledge provided in the system. From the operational point of view, the proposed approach consolidates and reconciles usage requirements from all involved security practitioners and parties. This includes requirements, modus operandi and legacy systems within a very fragmented operational ecosystem: border guards, custom authorities, infrastructure and transport operators and neighbouring and third country authorities. Based on these requirements, there is no perfect match by any of the proposed in literature algorithms; thus it is needed an adaptation of some of the proposed approaches which would regularize the weaknesses of one algorithm with the capabilities of another. In this context, a hybrid algorithm is proposed that relies on Dempster-Shafer algorithm and inherits the uncertainty capabilities of the evidence-based algorithms and the robustness in heterogeneous data sources. The system embeds a fuzzy logic rule-based classifier which has the role of regularizing the algorithm's result with any (if any) expert knowledge provided to the system. In this context, the proposed algorithm initially pre-process raw data coming from the sensors in order to distinguish exploitable data from non-exploitable ones and then apply machine learning methods (dimensionality reduction, supervised/unsupervised classification) for extracting high-level meaningful knowledge regarding potential incident factors that impact the risk of undesired behaviours. This information is considered as evidence for potential risk of unforeseen crowd behaviour. To this end this evidence is examined by a fusion system which evaluates the respective information and concludes to potential risk factors related to undesired behaviours and their intense level. Apparently, due the uncontrolled process of measurements acquisition, uncertainty cannot be neglected. As a result, at the information stage, it is proposed a Dempster-Shafer-based fusion algorithm [4, 5], which can tolerate not only uncertainty but lack of a priori knowledge of uncertainty level as well. Dempster-Shafer algorithm however corresponds to a high computational complexity. To this end, evidence sources are cluster based on their relevance to each risk factor. This way each risk factor is estimated taking into account only the most relevant and omits incorporating the ones with least contribution. The fusion process is presented in Fig. 13.2.

The first stage comprises the translation of the high-level information, (a) as results from the pre-processing of the raw measurements by each sensing device and (b) every available profiling knowledge base to linguistic variables each one corresponding to a quantified representation of the concept that the respective high-level information is referring to. These variables are the bases for defining the mass functions of each evidence. To this end these mass functions are defined by expert knowledge given a priori in two ways: (a) either as a direct assignment of an uncertainty value based on the estimated contribution of the respective source to risk evidence (b) or as a rule of combination of various information sources resulting in a specific evidence. Unlike the first case where mass function values are defined directly, in the second case each rule provided by expert knowledge is implemented as a fuzzy rule, and the concluded evidence mass value is calculated based on fuzzy logic inference. The method is

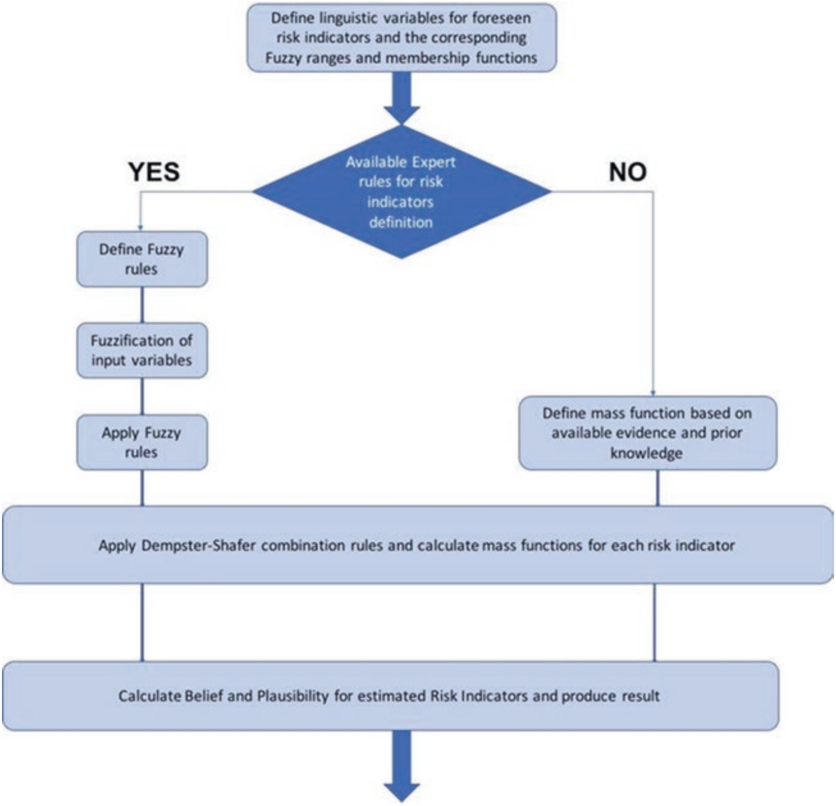
Define linguistic variables for foreseen risk indicators and the corresponding Fuzzy ranges and membership functions
Available Expert rules for risk indicators definition
YES
NO
Define Fuzzy rules
Fuzzification of input variables
Apply Fuzzy rules
Define mass function based on available evidence and prior knowledge
Apply Dempster-Shafer combination rules and calculate mass functions for each risk indicator
Calculate Belief and Plausibility for estimated Risk Indicators and produce result

Fig. 13.2 Algorithmic process flow

based on literature proposed approaches [6, 7] where fuzzy logic inference is used for determine the mass function of evidence that is determined by a specific expert rule that combines high-level information resulting from pre-processing of sensing components data streams. The next stage corresponds to the solution of potential conflicts of evidence where Dempster-Shafer combination rule is applied where m_i, m_j are observations of sensor S_i and S_j, respectively. This rule can be generalized iteratively where the result of each iteration is fed to the following one as sensor measurement:

$$\left(m_1 \oplus m_2\right) = \{0; \quad A = 0 \frac{\sum_{B_i \cap B_j = A} \quad m_1\left(B_i\right) \cdot m_2\left(B_j\right)}{1 - \sum_{B_i \cap B_j \neq 0} \quad m_1\left(B_i\right) \cdot m_2\left(B_j\right)}; \quad A \neq 0 \tag{13.2}$$

Finally, based on the mass functions values of the available evidence by determining the upper and lower level, the risk factors uncertainty values are estimated by determining the lower and upper values corresponding to the calculated plausibility (13.3) and Belief (13.4) functions respectively.

$$\text{Pls}\left(A\right) = \sum_{B \cap A \neq \emptyset} \quad m\left(B\right) \tag{13.3}$$

$$\text{Bel}\left(A\right) = \sum_{B \subseteq A} \quad m\left(B\right) \tag{13.4}$$

where A = risk factor and B = evidence related to risk factors.

Finally, the result corresponds to a set of risk factors along with their uncertainty interval.

13.4 Conclusion

Real-time risk estimation exploits various types of heterogeneous sources in order to calculate the risk level of a potential malicious behaviour of persons at border crossing points. Risk estimation is based on the realization of a risk assessment model incorporating high-level factors that may be considered as evidence to potential future malicious actions. These high-level factors are approached via an information fusion model which processes input from heterogeneous sources and calculates each high-level factor along with the respective confidence level. The information fusion algorithm relies on the Dempster-Shafer theory where heterogeneous sources values are considered as evidence for the pre-defined risk factors. Additionally, fuzzy logic is incorporated wherever expert knowledge is applicable, in order to assess evidence mass functions with higher certainty. Finally, the algorithm concludes to risk indicators values that feed the risk estimation model, with the aim of assessing potential risks for malicious or suspicious behaviour. The future work will be directed in two ways: (a) evaluation of the proposed method in an experimental realistic use case scenario, where the system will be tested against several behaviour types, normal and abnormal where some will correspond to malicious intensions and some not. Moreover, (b) there will be research of the capabilities of imposing legal/ethics framework as domain knowledge in the system with the aim of constraining the system's response within predefined ethics/legal limits.

Acknowledgement

Part of the work presented in this chapter is supported by and in the context of EU Research Project TRESSPASS – robusT Risk basEd Screening and alert System for Passengers and luggage – funded by the Horizon H2020 Framework Programme of the European Union for Research and Innovation under Grant Agreement number 787120.

Bibliography

1. EU Integrated Border Management scheme. Retrieved on January 2020 and available online at: https://ec.europa.eu/home-affairs/content/european-integrated-border-management_en.
2. TRESSPASS project website. Accessed on January 2020 at: https://www.tresspass.eu/.
3. Durrant-Whyte, H., & Henderson, T. C. (2016). Multisensor data fusion. In *Springer handbook of robotics* (pp. 867–896). Berlin, Heidelberg: Springer.
4. Zheng, Y. (2015). Methodologies for cross-domain data fusion: An overview. *IEEE Transactions on Big Data, 1*(1), 16–34.
5. Wu, H., Siegel, M., Stiefelhagen, R., & Yang, J. (2002). Sensor fusion using Dempster-Shafer theory. In *IEEE Instrumentation and Measurement, Technology Conference*, Anchorage, AK, USA, 21–23 May 2002. Author, F. (2010). Contribution title. In *9th International Proceedings on Proceedings* (pp. 1–2). Location: Publisher.
6. Challa, S., & Koks, D. (2004). Bayesian and Dempster-Shafer fusion. *Sadhana, 29*(2), 145–176.
7. Maseleno, A., Hasan, M. M., Tuah, N., Fauzi, & Muslihudin, M. (2015). Fuzzy Logic and Dempster-Shafer belief theory to detect the risk of disease spreading of African Trypanosomiasis. In *Fifth International Conference on Digital Information Processing and Communication (ICDIPC)*, IEEE, 7–9 October 2015.
8. Yen, J. (2018). Generalizing the Dempster-Shafer theory to fuzzy sets. In *Classic works of the Dempster-Shafer theory of belief functions* (pp. 529–554). Berlin, Heidelberg: Springer.