A Dynamic Risk Assessment (DRA) Methodology for High Impact Low Probability (HILP) Security Risks

Short paper

Carlo Dambra [1], Chanan Graf [2], Jordi Arias [3] and Alex Gralewski [4]

[1] ZenaByte s.r.l., Genova, Italy
carlo.dambra@zenabyte.com
[2] Railsec Ltd., Magshimim, Israel
chanan@railsec.com
[3] ETRA Investigacion y Desarrollo SA. Valencia, Spain
jarias.etraid@grupoetra.com
[4] PROPRS Ltd, Oxford, United Kingdom
alex.gralewski@proprs.com

Abstract. This paper proposes a Dynamic Risk Assessment (DRA) methodology applicable to the so-called High Impact Low Probability (HILP) security risks which, by their very nature, are difficult to identify or occur only infrequently. DRA is based on the processing of Weak Signals (WSs) to protect critical infrastructures and soft targets against HILP security risks before they materialise. DRA allows to rank WSs according to the reliability and credibility of the sources and to correlate them to obtain threat precursors. Experimental results have shown that DRA is effective and helps suppressing irrelevant alerts.

Keywords: Dynamic Risk Assessment, Low Probability High Impact risks.

1 Introduction

This paper proposes a methodology to dynamically assess High Impact Low Probability (HILP) security risks which, by their very nature, are either difficult to identify or occur only infrequently [1]: in this category fall, for example, terrorism, extremism, and lone wolf actions. The dynamic assessment of risks is an essential element of any decision support tool aimed at improving the situational awareness while protecting critical infrastructures and/or soft targets against HILP security risks. Related probabilistic approaches (e.g. [2], [3], [4]) have two major drawbacks: typically present a high number of false positives and, to characterise the problem, require substanial statistical evidence that is not available for HILP security risks that manifest themselves as “black swans”. The proposed Dynamic Risk Assessment (DRA) approach tries to overcome these

drawbacks, by processing Weak Signals[1] (WSs) [5] collected from heterogeneous sources taking inspiration from Intrusion Detection Systems [6]. WSs, once detected and correlated with other WSs, can generate precursor alerts of threats related to HILP security risks to be deeper and further investigated. The paper is organised as follows: Section 2 discusses the DRA approach, Section 3 proposes an application of DRA to a mass gathering event, and finally Section 4 presents conclusions and future work.

2 The proposed DRA approach

The proposed DRA approach bases its reasoning on the processing of WSs, the minimum managed quantum of information. Starting from a static risk assessment, the DRA logic can be summarised in the following steps:

1. Continuously **collect** the WSs potentially representing precursors of threats;
2. **Analyse** each collected WSs and verify if, alone or correlated/grouped with other existing WSs, can represent a more **significant precursor** of a threat;
3. **Present** the potential detected precursor to a security operator for evaluation;
4. **Re-assess the risks** for the considered target accordingly.

Each WS, detected by a given source, contains the following minimal information:

- A unique **ID** that has embedded the reference to the source of the WS;
- The absolute time **t** in which it has been collected;
- The **geolocation** (x, y) - if available;
- A **snapshot** of what has been detected using a pre-defined semantic to help the operator to confirm, discard (false or nuisance alarms) or amend the detection.

Each detected WS is characterised by a **Significance** (**S**) value that is a combination of:

- The **Reliability** (**R**) of the source that characterises the ability of a source to give a true information in a particular context of use;
- The **Credibility** (**C**) of the information generated by the source, that introduces a measure of the degree of confirmation: the more an item of information is confirmed, the higher its credibility and, conversely, the more an item of information is contradicted by others, the less credible it becomes [7].

The Significance of the considered WS detected by the source *m* ranges in the [0, 1] interval and is computed as follows:

$$S(WS_{ID}) = (\alpha \cdot R_m + \beta \cdot C_m) / NF \quad (1)$$

where:

[1] A WS can be defined as "*A seemingly random or disconnected piece of information that at first appears to be background noise but can be recognized as part of a significant pattern by viewing it through a different frame or connecting it with other pieces of information*" [14]

- C_m and R_m are integer in [0, 5] (where 1 is very low and 5 is very high);
- The Normalising Factor NF keeps $S(WS_{ID})$ in [0, 1];
- α and β are correcting factors to tune the role of each factor in the product.

If the proposed methodology is applied to an event (e.g. a concert in a stadium) it is possible to add a further element that consider the Time Distance (TD_t) of the WS from the event date: the closer the detection of the WS to the event date the bigger the TD_t . $S(WS_{ID})$ then becomes:

$$S(WS_{ID}) = (\alpha \cdot R_m + \beta \cdot C_m + \gamma \cdot TD_t) / NF \quad (2)$$

Once received, it is necessary to process WSs to evaluate if they can become, alone or together with other WSs, a **significant precursor** of a threat related to a specific HILP security risk. To this end three structures of **Precursors** are introduced:

- **Suspicious Sign** (**SS**), represents a single WS that has either sufficient significance to become a SS or is related to a high-risk threat. In both cases, S(SS) = S(WS);
- WSs coming from the intelligence services, i.e. **Intelligence Alerts** (**IA**), can be considered a special Suspicious Event with maximum Significance S(IA) = 1;
- **Suspicious Pattern** (**SP**), two or more WSs can create a SP if they have sufficient significance and are linked together according to one of the criteria described below.

The Precursors can be generated combining already collected WSs, SSs, IAs or SPs using either the experts' knowledge to define the rules for grouping WSs or data analytics applied to WSs [8] [9] [10] as follows:

- **Group**: a set of precursors without time and geographic constraints independently of the time sequence in which they are detected;
- **Sequence**: a set of precursors that need to be received in the correct sequence;
- **Area**: a set of precursors within the same area and in a given time interval;
- **Distance from Hot Spots**: a set of precursors in a given time interval all at a distance from Hot Spots (e.g. embassies, police offices, etc.) shorter than a given threshold;
- **Simultaneous Group**: the grouping is generated using the strategy of "**simultaneous events**", i.e. three or more WSs detected within a short period of time;
- **Data Analytics**: the grouping of precursors is generated using **data analytics** approaches for example through the generation of new rules on the basis of data collected in the past. A possible approach is described in [11], where Suspicious Activity Reports (SAR) collected by 911 emergency operators are analysed to identify and prioritise cases of interest from the large volume of SARs;
- **Operators Group**: generated by the operator according to his/her experience.

Precursors' Significance value is computed using the Significance values of all the WSs connected to it. The approach to combine Significance values for an SP with two WSs contributing to it, with significance S_1 and S_2 respectively, is derived from Certainty Factors [12] theory using the following formula

$$S_{1\ and\ 2} = S_1 + (1 - S_1) \cdot S_2 \quad (3)$$

Having more than two WSs contributing to the same SP, it is possible to iteratively apply the same formula.

Precursors, when triggered by WSs, can be then classified as either **Non-Critical** or **Critical**, i.e. elements that constitute an immediate threat for a given risk. **Critical Precursors** shall be triggered and brought immediately to the attention of a security operator that should take the necessary mitigation actions.

Using the above methodology, the **Risk Level** can be re-assessed using escalation approaches [13]. An example, when dealing with a mass gathering event, based on an IF-THEN-ELSE approach is given in the following:

- IF (Time Distance is Big) AND (no Critical precursors are triggered) THEN (the Risk Level is Very Low);
- IF (Time Distance is Big) AND (some Non-Critical Precursors are triggered) THEN (the Risk Level is Low);
- IF (Time Distance is Big) AND (at least one Critical Precursor is triggered) AND (Crowd Density is Low) THEN (the Risk Level is Medium);
- IF (Time Distance is Small) AND (at least one Critical Precursor is triggered) AND (Crowd Density is Low) THEN the Risk Level is High;
- IF (Time Distance is Small) AND (at least one Critical Precursor is triggered) AND (Crowd Density is High) THEN the Risk Level is Very High.

Clearly, exact and complete IF-THEN rules and related thresholds need to be defined according to laws, protocols and best practices including also socio-political and environmental conditions.

3 The DRA application to a mass gathering event: an example

3.1 The DRA practical implementation

DRA methodology has been applied to a scenario representing a mass gathering event managed by a Law Enforcement Agency (LEA). The sources of WSs are:

- Normal citizens calling 112 emergency services;
- Stewards recruited to manage the event;
- Human-Centred Computer Vison (HCCV) tools able to semi-automatically recognise car plates, identify vehicles and suspicious behaviours of vehicles and individuals;
- Intelligence services.

The sequence of WS detection, SP generation and DRA is described in **Fig. 1**:

1. On the basis of the received WS, the corresponding values of sensor's credibility and reliability and the time distance from the event are identified.
2. The Significance is then computed using the formulas in Section 2 (with α, β and γ set to 1 for the sake of simplicity) and normalised to get values in the [0; 1] range.

Fig. 1. DRA applied to a mass gathering event

Time	Signal/Pattern	Sensor	Description	Reliability R	TD	Significance	Norm. Significance	Risk Level
T01	WS01	Citizen	Suspicious Vehicle	4	1	8	0,06	
T02								
T03	WS02	HCCV	Suspicious Behaviour	3	1	6	0,05	
T04	WS03	Citizen	Suspicious Vehicle	4	1	8	0,06	
T05	**SP01**	DRA Rule	Suspicious Vehicle	**WS01 & WS02 & WS03**			0,17	
T06	Patrol sent to check							
T07	SP01 deleted after operator's check							**1**
T08	IA01	Intelligence	Possible terrorist attack	5	2	50	0,40	**3**
T09	WS04	HCCV	Red truck	5	2	40	0,32	
T10	IA02	Intelligence	Stolen yellow van	5	2	50	0,40	
T11	WS05	HCCV	Suspicious plate detected	5	2	50	0,40	
T12	**SP02**	DRA Rule	Suspicious Vehicle	**IA02 & WS05**			0,64	
T13	IA03	Intelligence	Terrorist presence	5	3	75	0,60	
T14	Reaction due to IA03							
T15	WS06	HCCV	Brown truck	5	3	60	0,48	
T16	WS07	HCCV	Red van	5	3	60	0,48	
T17	WS08	HCCV	Suspicious plate detected	5	3	75	0,60	
T18	WS09	HCCV	Blue car	5	3	60	0,48	
T19	WS10	HCCV	Suspicious plate detected	5	3	75	0,60	
T20	WS11	Steward	Suspicious person	5	3	75	0,60	
T21	WS12	HCCV	Suspicious plate detected	5	3	75	0,60	
T22	**SP03**	DRA Rule	Suspicious Vehicle	**WS08 & WS10 & WS12**			0,94	
T23	Reaction due to SP03							
T24	WS13	Steward	Suspicious person	5	3	75	0,60	
T25	**SP04**	DRA Rule	Probing security	**WS11 & WS 13**			0,84	**4**
T26	WS14	HCCV	Quite dense crowd	4	5	80	0,64	
T27	WS15	HCCV	Yellow van	5	5	100	0,80	
T28	**SP05**	DRA Rule	Ramming vehicle	**SP02 & WS15**			0,93	**5**
T29	Reaction due to SP05							

Through the application of the DRA rules the Precursors are created and, if necessary, Risk Level is modified.

3.2 A possible architectural approach for DRA implementation

DRA has been implemented in the framework of the H2020 LETSCROWD project[2] in a Web-server GIS-based architecture receiving WSs from CCTV-based crowd density estimators, Web-crawling and semantic intelligence on social media, crowd behaviour modelling and humans-as-sensors. SPs above a selected Significance threshold are brought to the attention of an operator to allow a risk-aware decision-making process.

3.3 First experimental results

First experimental results have confirmed the validity and effectiveness of the approach, as confirmed by the involved LEAs and that DRA helps distinguishing irrelevant alerts, thereby reporting only significant threats to operators. The proposed DRA approach is going to be further validated on real scenarios (mass gathering events) from Law Enforcement Agencies (LEAs). The main problem of the DRA application lies in the identification of sources of WSs apart from human-as-a-sensors and (semantic) intelligence: most of the CCTV-based tools are either not sufficiently reliable or facing serious privacy issues.

[2] https://www.letscrowd.eu

4 Conclusions

The proposed DRA methodology has the following advantages over more traditional approaches: it searches for out-of-the-ordinary behaviours, reduces the number of false alarms, does not require large statistical samples and is sufficiently simple to run in real-time. Further research should confirm the first promising experimental results focusing on identifying suitable WSs sources, characterising them in terms of reliability and credibility and evaluating the feedback from LEAs' operators.

5 Acknowledgements

This paper is based on the work carried out in the LETSCROWD project that has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement number 740466.

References

1. UK Government Office for Science, "Blackett Review of High Impact Low Probability Risks," London, 2011.
2. B. Ezell, S. Bennet, D. von Winterfeldt, J. Sokolowski and A. Collins, "Probabilistic Risk Analysis and Terrorism Risk," Risk Analysis, vol. 30, no. 4, pp. 575-589, 2010.
3. J. Brynielsson, A. Horndahl, F. Johansson, L. Kaati, C. Martenson and P. Svenson, "Analysis of Weak Signals for Detecting Lone Wolf Terrorists," in 2012 IEEE European Intelligence and Security Informatics Conference, 2012.
4. M. L. Paté-Cornell, "Fusion of Intelligence Information: A Bayesian Approach," Risk Analysis, vol. 22, no. 3, pp. 445-454, 2002.
5. M. Holopainen and M. Toivonen, "Weak signals: Ansoff today," Futures, vol. 44, 2012.
6. E. Chakir, M. Moughit and Y. Khamlichi, "A Real-Time Risk Assessment Model for Intrusion Detection Systems," in 2017 IEEE International Symposium on Networks, Computers and Communications (ISNCC), 2017.
7. North Atlantic Treaty Organization (NATO) Information Handling Services, "Annex to STANAG 2022 (Edition 8)," 1992.
8. H. Vu, "Deep Abnormality Detection in Video Data," Melbourne, 2017.
9. D. Xu, E. Ricci, Y. Yan, J. Song and N. Sebe, "Learning deep representations of appearance and motion for anomalous event detection," 2015.
10. M. Hasan, J. Choi, J. Neumann, A. K. Roy-Chowdhury and L. S. Davis, "Learning temporal regularity in video sequences," Las Vegas, 2016.
11. K. J. Strom and J. P. M. Hollywood, "Using 911 Calls to Detect Terrorism Threats," June 2009. [Online]. Available: https://www.nij.gov/journals/263/pages/911-calls.aspx.
12. P. J. F. Lucas, "Certainty-factor-like structures in Bayesian belief networks," Knowledge-Based Systems, vol. 14, 2001.
13. UK HM Treasury, " Orange Book: Management of risk - Principles and Concepts," London, 2004.
14. P. J. H. Schoemaker and G. S. Day, "How to Make Sense of Weak Signals," MIT Sloan Management Review, vol. 50, no. 3, 2009.