Prediction and Visual Intelligence Platform for Detection of Irregularities and Abnormal Behaviour

Konstantinos Demestichas[1], Theodoros Alexakis, Nikolaos Peppes, Konstantina Remoundou, Ioannis Loumiotis, Wilmuth Muller, Konstantinos Avgerinakis

[1]Institute of Communication and Computer Systems, 9, Iroon. Polytechniou Str.,
15773 Zografou Athens, Greece
cdemest@cn.ntua.gr

Abstract. Nowadays, (cyber)criminals demonstrate an ever more increasing resolve to exploit new technology so to achieve their unlawful purposes. Therefore, Law Enforcement Agencies (LEAs) should accommodate an approach that surpass the existing limits in policing practices. In this light, the authors introduce an innovative platform that provides near real-time advanced social behavior analytics using irregularities detection based on historical patterns.

Keywords: Big Data, abnormal behaviour detection, crime detection.

1 Introduction

Law Enforcement Agencies (LEAs) have an increasing necessity to combine, prioritize and analyze heterogeneous massive data streams. Another relative challenge is that data generation considering crimes, nowadays, is massive and mostly in semi-structured or unstructured data format. Big Data technology provide powerful tools by means of social networks analysis, semantic technologies as well as the utilization of advanced linguistic models. Therefore, LEAs must develop and integrate future-proof solutions and tools which will empower them with supreme analytical and predictive capabilities against terrorists, Organised Crime Groups (OCGs) and individuals.

The platform which is described integrates future-proof solutions and tools which will empower LEAs with the requested supreme analytical and predictive capabilities against terrorists, Organised Crime Groups (OCGs) and individuals. The platform exploits Visual Intelligence algorithms and representation technologies in order to tackle object detection, activity recognition, etc. using surveillance data [1]. Data Mining techniques are also used in order to apply analytics over gathered information from heterogeneous sources. [2]. Finally, semantic reasoning and Big Data analytics drawing on online and other activities can be used to determine relevant behavioural indicators [3]. The following Figure illustrates the concept of the described platform.

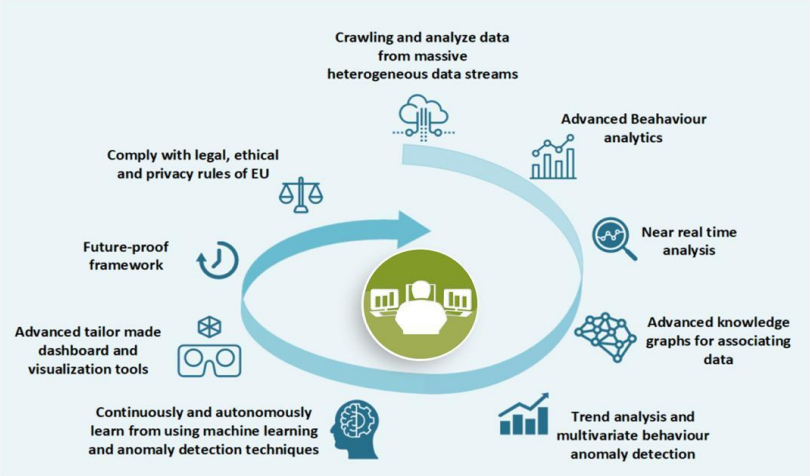

Fig. 1. Platform concept

2 Architecture

The proposed high-level system architecture provides the required tools for LEAs to accelerate their investigations and remain careful in consideration of terrorist and cyber-criminal threats by successfully integrating massive data streams from heterogeneous sources. The system architecture is designed so to meet performance and resiliency requirements at scale. Specifically, in the following paragraphs of this Section they are presented with details all the main functional modules as well as the components that the system generally is consisted of, in a coherent manner. The details supply in a comprehensive way the structure of the system architecture based on the combination of the different components into the common proposed platform which can be deployed into LEAs' and practitioners' facilities. The main effort is being put on standardization of the platform's open architecture and its constituent components, interfaces as well data exchange formats. In order to succeed, extensive monitoring studying and contribution into activities related to standards of ISO/TC 292 (Security and Resilience) in the area of security is foreseen. In short, the system is composed of: i) Data Mining Module for Crime Prevention and Investigation, ii) Visual Intelligence Module, iii) Semantic Information Representation and Fusion Module, iv) Trends Detection and Probability Prediction Module for Organized Terrorism and Criminal Activities, v) Detection Module of Cyber-Criminal Activities and Situation Awareness and vi) HMI Module**.** The next figure illustrates the high-level architecture of the presented platform.

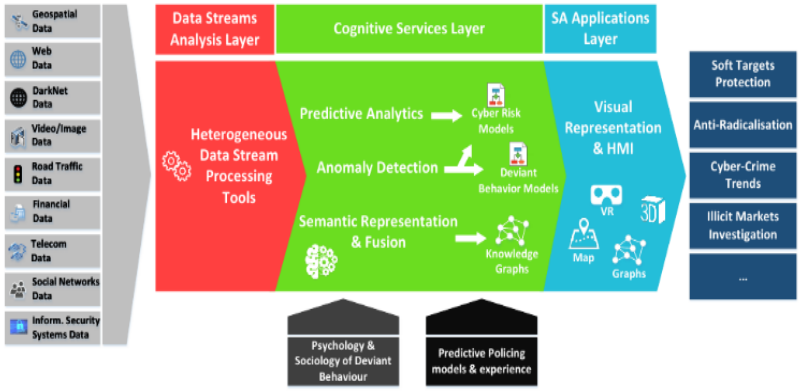

Fig. 2. High-level architecture

2.1 Data Mining Module for Crime Prevention and Investigation

Data mining is used to extract valuable information from the existing data from various sources such as web, dark web, social media, etc. as shown in Figure 2. Crawling and mining take place by using crawl points (data sources) from which posts and references are extracted and stored for analysis purposes. Eventually data from multiple sites are processed and exported into a common format that is available to the other components for further processing and analysis.

To develop and study crime patterns we used existing open-source web and social media mining tools in conjunction with API's provided by social media as well as extensive and scalable open-source web crawlers software projects. However, the difficulty of accessing Dark Web sites as well as the existence of specific rate limits on social media constitute a possible risk. The contingency plan pays attention to key Dark Web Sites and narrow social media access by using specific keywords and phrases.

2.2 Visual Intelligence Module

The visual intelligence module is implemented by using available face recognition and face detection algorithms such as YOLO [4], SSD [5], Faster R-CNN [6], VITAL [7] and other widely used and tested algorithms. This specific module is engaged in order to achieve identifications of persons and objects contained in images and videos crawled from the Web, social media and footage from static or moving cameras as well as surveillance systems. The face and object recognition through visual data is focused on the detection of suspicious objects. Finally, suspicious or abnormal activities are also being tracked by the adoption of crowd analysis and human action recognition with spatio-temporal localization.

The starting point for the prementioned implementations comprises Deep Convolutional Network, Deep faces, SSD coupled with a YOLO architecture, KCF, Goal-based descriptors and Swarm intelligence for crowd analysis and abnormal activity detection

[8]. Alongside with visual intelligence techniques, machine learning algorithms and techniques are used in order to predict and estimate outliers and abnormal person and object activities based on trends and motifs discovered by the gathered visual data.

2.3 Semantic Information Representation and Fusion Module

Another tool of the described platform is the Semantic Information Representation and Fusion Module that gathers data from various and heterogeneous sources like: geospatial data, web data, darknet data, video/image data, road traffic data, financial data, telecom data, social networks data and information and security systems data. Thus, a module dedicated to data and information fusion is mandatory. This module is able to apply information fusion to the heterogeneous data that are collected from the different sources listed above. The development and application of this module results to information transformation into valuable knowledge. The baseline of this module includes Knowledge graphs, JDL model, OWL language and Markov Logic Networks along with the use of appropriate semantic information fusion models. Thus, the module can extract useful patterns and hidden relationships among different datasets that can lead to trends discovery and abnormal behaviour detection.

2.4 Trends Detection and Probability Prediction Module for Organised Terrorism and Criminal Activities

All the aforementioned modules gather data from various sources in order to create various datasets. Generating all these datasets the next step is to predict organised terrorism and criminal activities. In this direction big data analytics techniques are applied over the collected data in order to identify hidden trends inside the datasets. Analytics results lead to predictive models which can be considered as the link between data and decision-making processes.

This module uses machine learning algorithms which are developed by engaging appropriate open source libraries alongside with predictive policing software. More specific, the platform engages many widely used Artificial Intelligence Algorithms such as Artificial Neural Networks (ANN), decision trees, pattern recognition and Lifelong Learning Algorithms (LLA).

The risks that the development engineers spotted are the inaccurate results that were extracted from the prediction model and the false positive alert that a model generated in some cases. These risks may cause an increasing false alarm rate (FAR) in the early stages of deployment which can be tackled by the user feedback. A proposed contingency plan foresees the use of data sources of higher degree of diversity as well the creation of more sophisticated models for explaining deviant behaviours.

2.5 Detection Module of Cyber-Criminal Activities

In addition to trends detection and probability prediction module for organised terrorism and criminal activities that presented in paragraph 2.4 there is the detection module of cyber-criminal activities. This very module focuses not only to identify anomalies

but also on behavioural indicators as well as revealing previously unknown associations and rules that are connected to cyber-criminal activities. For these purposes advanced big data analytics techniques, based on artificial neural networks and classification methods are applied to the collected data. In this light the three widely used machine learning algorithms: K-means clustering, Support Vector Machines and Deep learning algorithms, are being used in this module. The development and integration of these algorithms was held with open source libraries for numerical computation in order to achieve faster results.

As presented above in paragraph 2.4 risks which appeared in module are not only the inaccurate results of the model but also poor-quality model results over time. Thus, again contingency plan gives attention to the selection of more complex model, fit the training frequency as well test and modify the model.

2.6 Situation Awareness and HMI Module

Last but not least we have the situation awareness and HMI Module. This module demonstrates to the end-users the gathered information and analysis results produced by the aforementioned modules in order to increase the situation awareness of the decision makers and practitioners. The baseline comprises open-source libraries for visual analytics in addition to powerful, secure, and flexible end-to-end analytics platforms for data visualization and representation purposes.

Possible problems and risks that may occur and must be overcome could be the inadequate offered visualization tools for some LEAs, the requirement for additional data views in certain use cases and the difficulty in using and handling the visualization environment. In addition, different LEAs use different tools so the adoption of a new tool must be as close as possible to their tools. Thus, the developers of the platform take into account the feedback from LEAs and end users in order to create a common user-friendly HMI. The design of the HMI follows the main design principles of the LEAs' HMIs and tries to simplify the environment in order to attract users to adopt it. It is of outmost importance to receive the feedback from users in the early stage of deployment in order to update and patch the visualization tools so to assure the credibility of the platform's results and increase the LEAs productivity.

3 Conclusion

In conclusion, taking into account the increasing needs of LEAs for future-proof solutions and expertise adoption so to fight crime, we presented a platform and its modules which engage state-of-the-art tools and technologies. The modules presented are interconnected and aim to enhance LEAs with supreme crime prediction and prevention capabilities. Finally, it offers a future-proof framework that is open to the deployment of new situation awareness applications, novel cognitive services and additional data stream analytics tools, both from first- and third- parties, adopting standard and well-documented interfaces.

It is worth noted that the development and the deployment of this very platform acknowledges the legal, privacy, ethical and societal concerns of predictive policing and data science method and integrates independent assessment, while participating into an open dialogue with civil society organisations, security stakeholders, practitioners and policy makers.

Acknowledgement
This work has been performed in the context of the PREVISION project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 833115. The paper reflects only the authors' view and the Commission is not responsible for any use that may be made of the information it contains.

References

1. Hu, P., Ramanan, D.: Finding Tiny Faces. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1522-1530, (2017).
2. Davis, P. K., Walter, L. P., Brown, R. A., Douglas, Y., Parisa, R., Voorhies, P.: Using Behavioral Indicators to Help Detect Potential Violent Acts: A Review of the Science Base, Santa Monica, Calif.: RAND Corporation, RR-215-NAVY, (2013).
3. Jeelani, A., Muqeem, A. Big data and semantic web, challenges and opportunities a survey. International Journal of Engineering & Technology, 7, pp. 631-633, (2018).
4. Hu, P., Ramanan, D. Finding Tiny Faces, In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1522-1530, (2017).
5. Derpanis, K.G., Lecce, M., Daniilidis, K., Wildes, R. Dynamic scene understanding: The role of orientation features in space and time in scene classification, In: Proceedings CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.1306-1313, (2012).
6. Ren, S., He, K., Girshick, R., Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, In: IEEE Transactions on Pattern Analysis and Machine Intelligence 29, 91-99 (2015).
7. Song, Y., Ma, C., Wu, X., Gong, L., Bao, L., Zuo, W., Shen, C., Lau, R.W.H., Yang, M. VITAL: VIsual Tracking via Adversarial Learning, 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8990-8999, (2018).
8. Kaltsa, V., Briassouli, A., Kompatsiaris, Y., Hadjileontiadis, L. I., Strintzis, M. Swarm Intelligence for Detecting Interesting Events in Crowded Environments, IEEE Transactions on Image Processing, 24, pp. 2153-2166, (2015).