Cognitive Computing and Internet of Things, Vol. 43, 2022, 145–155
https://doi.org/10.54941/ahfe1002703

Digital Twins for the Intelligent Detection of Malicious Activities in Urban Spaces

Zoheir Sabeur, Alessandro Bruno, Liam Johnstone, Marouane Ferjani, Djamel Benaouda, Deniz Cetinkaya, Banafshe Arbab-Zavar, Muntadher Sallal, and Benjamin Hardiman

Bournemouth University, Department of Computing and Informatics, Talbot Campus, Fern Barrow, BH12 5BB, Poole, England, United Kingdom

ABSTRACT

The S4AllCities project has progressed rapidly during the last twelve months since it began in 2020 for the development of three distinct digital twins that collectively augment intelligence concerning cyber and physical security monitoring in smart urban spaces. These respectively specialize on; *a) Distributed Edge Computing IoT (DEC-IoT); b) Malicious Actions Information Detection System (MAIDS); and c) Augmented Context Management System (ACMS)* (S4AllCities, 2020). These three twins are built under a distributed *System of Systems (SoS)* architecture. Further, they each acquire real-time observations of both cyber and physical spaces while processing data for the critical extraction of knowledge at their levels. The extracted knowledge, represented as "events" at each of the respective twins levels, is communicated across the S4AllCities SoS Apache Kafka communication client/ server protocols. These respectively specialize in advancement of situation awareness at their levels. Namely, for the intelligent edge processing of observations in the urban space under the *DEC-IoT*; the detection of unusualness and understanding of cyber and human behavior under the *MAIDS*; while augmenting all awareness for the final release of threat alerts and proposed regulated responses (*ACMS*). In this paper, we will introduce the S4AllCities SoS overall architecture and the three twins high level functions. Then we will focus on describing our development of the *MAIDS* sub-modules and their functions under the De-Facto Joint Director of Laboratory (JDL) data fusion framework. The JDL framework efficiently enables the intelligent monitoring, detection and interpretation of the potential presence of threats and/or attacks in urban spaces. These attacks are either of cyber, physical, or both malicious nature. The well-known Endsley model for the cognitive advancement of situational awareness is mapped into the JDL framework in the context of critical decision support on cyber-physical surveillance in urban spaces. The JDL is much more adaptive for big data processing, machine learning, context knowledge modelling and augmented situational awareness of the cyber-physical space.

Keywords: Digital twins, Artificial intelligence, Data fusion, Behavior detection, Intelligent agents, Big data

INTRODUCTION

Urban spaces which are often attended by citizens do require robust safety and security measures to put in place by public authorities around the globe.

The deployment of scalable intelligent systems for operating security surveillance and human behavior understanding in urban spaces is of paramount importance. Currently, an ever growing trend of smart spaces are being embedded in urban infrastructure for all sort of purposes. Nevertheless, they make it useful to generate real-time spatial, temporal and contextual data and information for supporting security management. With those, one is enabled to extract critical knowledge which can be guided using artificial intelligence based methods for advancing security and situation awareness in modern cities around the world. These can become an integral part of existing surveillance monitoring systems which are themselves compatible to stream information generated by heterogeneous sets of sensor observation platforms. Therefore, smart spaces can efficiently aid into understanding human induced physical and cyber activities using machine intelligence through information fusion and reasoning for advanced situation awareness. These spaces include communication networks, multi-modal physical transport networks, shopping centers, open markets, festival venues, city's tourist attractions and so forth (Sabeur, 2021a). The provision of information and data from multimodal sensing platforms has indeed the potential to generate a rich Common Operational Picture (COP) for under-standing situations in the both the cyber and physical worlds at the same time. These activities are often related to connected since they are likely to be orchestrated by the same rogue offenders in terms of their intended behavior for disrupting the peace in the urban space. These types of cyber-physical attacks have been occurring more often in recent years in cities around the world, which is of serious concern for security management. The aim now is to urgently deploy rich COPs for cities, through the development of so-called Digital Twins (*DT*) which grasp situations through real-time sensing and understand them through validated machine learning and predictions with greater intelligence and confidence. This will indeed bring security communities and first responders to higher preparedness and ahead of the game for tackling urban crimes and beyond.

S4AllCITIES PROJECT

The S4AllCities project (S4AllCities, 2020) was launched in 2020. The performance of the developed S4AllCities System of Systems (SoS), which is composed of three major distributed *DT*s, is being tested in three European pilot cities with respective urban smart spaces, which are Trikala Municipality in Greece; the City of Pilsen in the Czech Republic, and the City of Bilbao in Spain. The three S4AllCities SoS *DT*s, respectively include: 1) Distributed Edge Computing IoT (or *DEC-IoT*); 2) Malicious Actions Information Detection System (or *MAIDS*); and c) Augmented Context Management System (or *ACMS*). Collectively, they contribute in real-time, with the flow of relevant critical intelligence for the advancement of situational awareness in the smart space. In this study, our research group activity solely focused on the development of the *MAIDS* Digital Twin detectors in the context of the S4AllCities SoS architecture. The advancement of the *MAIDS* in the last twelve months is described with its early deployment in one of the pilot cities thereafter.

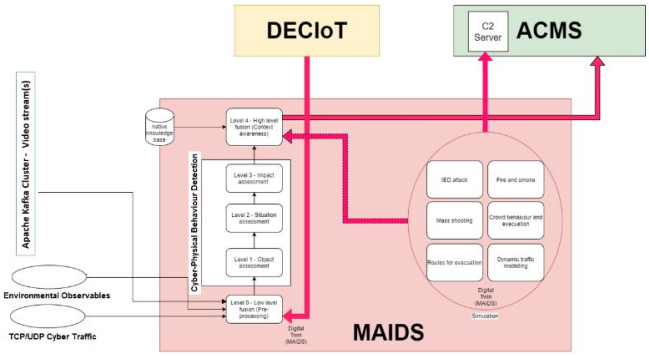

Figure 1: The *MAIDS* overall data flow in context of the *DEC-IoT* and *ACMS DT*s.

The MAIDS Digital Twin

It is important to automate the process of critical knowledge extraction for understanding situations in smart spaces. This is done with the deployment of artificial intelligence methods which enable the building of the so-called digital twin. This is achieved together with the design and conceptualization of independent intelligent agents which can fulfill the tasks of observing the smart environment through their percepts, machine understanding and reasoning, so that they could action their expected tasks in the concerned environment. The intelligent actioning part can be of physical nature, therefore leading to an in-built robotic system or; of information and data flow nature for decision-support to human critical operations. The latter constitute the foundation of our *MAIDS DT* in this sense. Furthermore, it is governed by the overarching Joint of Directors Laboratory (JDL) multi-leveled framework concerning data and information fusion (Lambert, 2009; Sabeur, 2021b). The detection of cyber-physical behavior and understanding in context of urban spaces under the *MAIDS DT* required the deployment of three intelligent core modules under the multi-level JDL framework (Sabeur et al., 2021). These are: *Spatial Fusion for pre-processing multiple environmental sensor observations; Cyber Behavior Detector for understanding cyber traffic activities*; *Physical Behavior Detector for understanding human physical activities*. These core modules are embedded in the *MAIDS DT* for collectively capturing situation awareness in the smart space. As shown in Figure 1 below, the data flow of the *MAIDS*, in context of the S4AllCities SoS with its two other *DT*s is depicted. Flows of data concerning cyber traffic networks, physical environmental sensor networks and digital camera vision are all ingested under level 0 of the JDL framework for pre-processing.

This results into richer generated data, such as environmental data gridded maps which can be fetched on demand, or TCP/UDP time series as well as video stacks at specific time-framing for post-processing and patterns recognition under higher JDL levels 1, 2 and 3. These associate detected behavior through machine learning with correct classification rates at both the physical and cyber domains of the smart space under surveillance. The cyber and physical behavior outputs are pushed to the *ACMS DT* through Kafka messaging in real-time. They are then displayed in the COP user interface of the *ACMS*.

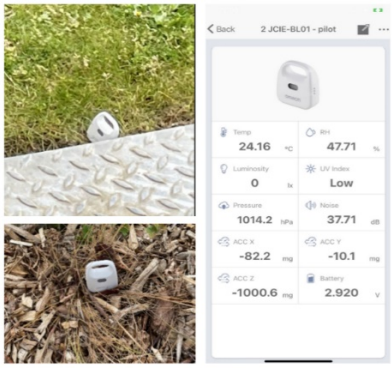

Figure 2: The Omron sensors 2JCIE-BL01 with the smart phone application.

Figure 1 also depicts other modules outside the realms of intelligent detection of behavior, such as the various simulations of crowd dynamics using modelling methods and so forth. These are developed by other collaborating partners in the S4AllCities Consortium. The higher level 4 fusion, concerning reasoning on the output of the intelligent behavior detectors combined with native knowledge on the smart space security operation is developed by other partners in the S4AllCities project. The various output of these mentioned components are each independently connected to both the *DEC-IoT* and *ACMS DT*s.

Spatial Data Fusion With RBF Networks

While the *MAIDS* adopts the JDL data fusion framework for efficient data processing and analytics, it automatically organizes itself for ingesting the right observation data for investigating on either cyber or physical behavior, or both, for subsequent high level reasoning and advanced context awareness. Such awareness concern: 1- Perception; 2- Comprehension; 3- Projection; 4: Decision; and 5- Performance feedback for operating in the smart space and preserving safety and security (Endsley, 1995). The multi-sensing of environmental parameters which are observed at multiple locations in the smart space have been generated. We used a network of "Omron" wireless environmental sensors for covering an experimental space in order to capture multiple environmental parameters. (See Figure 2 below).

These included: *Temperature; Luminosity; Atmospheric pressure, Relative humidity; UV Index; Ambient noise; and 3D Mechanical accelerations*. Among these parameters, our initial interest on observation and measurements in space and time included: "Temperature" and "Ambient noise". These two parameters are indeed very relevant for us to set thresholds on physical behavior (Sabeur, 2021a). These are influential in terms of capturing the physical context when observing crowd behavior. The generated data from the Omron sensors and recorded on Omron's smart phone environmental applications can be pushed to a cloud database for further data analysis under spatial fusion. We opted for the Radial Basis Functions (RBF) networks based spatial fusion of multiple sensor observations (Broomhead, 1988). These are

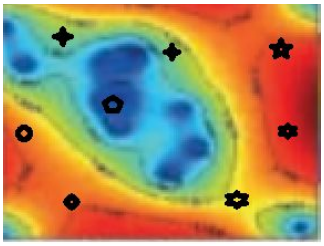

Figure 3: Sensor spatial fusion with RBF networks.

commonly used for building robust spatial function approximations in the following way:

$$y(x) = \sum_{i=1}^{N} \omega_i \varphi(\|x - x_i\|)$$

Where *y(x)* is approximated by the sum of *N* radial basis functions $\phi(x)$, using weights ω_i, each of which correspond to what is called "centre *i* ", where the environmental sensors are respectively located. The influencing element for function *y(x)* to vary spatially depends on how far the value of *x* is with respect to the centres. However, it still guarantees the correct value of the environmental parameter at each centre *xi*. The choice of the basis functions reflect on how tunable the approach is. These functions may include Gaussian, multi-quadratic, inverse quadratic, inverse multi-quadratic, polyharmonic splines and more. With such RBF deployment for spatial sensor data fusion in our study, one can predict the environmental conditions at space locations where sensing is not available. Furthermore, one is still capable to predict such environmental conditions even if some of the sensors in the network breakdown. In this sense, it guarantees the resilience of the spatial fusion module to maintain its delivery of the environmental parameter in space and time, while it also maintains its functionality when new sensors are plugged in into the network. Nevertheless, this all comes with a cost on evaluating estimated spatial uncertainties, while proceeding with data control in the prediction process.

Discussion

The spatial fusion of environmental parameters using our Omron's wireless sensor network with the smart phone application has been very useful for us to deploy in our project field experiments as "plug and play" devices. The fetching of a large matrix of environmental parameters, such as ambient temperature, provide good context about the space under surveillance with the likelihood of crowd to frequent it. Specifically, one expects a higher probability of crowd outdoor activities in smart spaces during good weather conditions, than otherwise. Temperature values at requested locations is also pushed to the *ACMS DT* on demand using the S4AllCities Kafka message broker. See Figure 4 below.

Our next study will entail the spatial fusion using RBF networks, for other environmental parameters particularly ambient noise levels. These directly

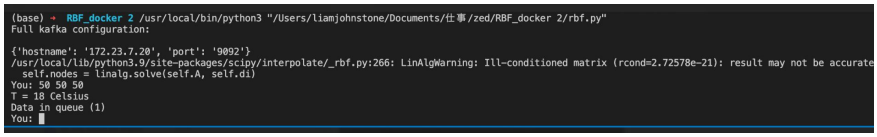

Figure 4: Spatial data fusion module using Kafka messaging / Docker technology.

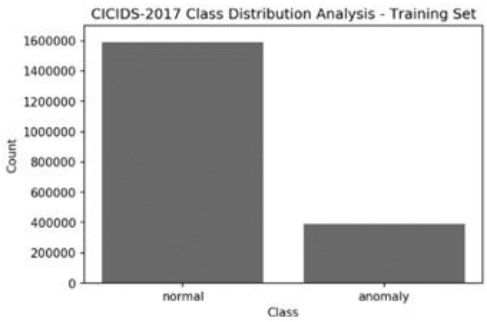

Figure 5: Class distribution of CICIDS-2017 dataset.

influence our assessment of behavior in the smart space. In particular, the presence or less presence of crowd in the space can be implicitly captured through spatial noise levels, with detection of crowd physical motions by the physical behavior detector module.

CYBER BEHAVIOUR DETECTION

The ever-growing number of organized cybercrimes around the world, particularly in the context of smart cities environments has led the scientific community bring in artificial intelligence and embed it into existing scalable systems for big data processing on security surveillance. However, the challenges remain for achieving highly per-forming machine learning (ML) algorithms concerning cyber behavior detections with correct classification rates. In the S4AllCities project, we initially adopted exemplar cyber data for training our various ML algorithms on cyber traffic network behavior (Sabeur, 2021a; Sabeur et al., 2021). These included CICIDS-2017, NSL-KDD and UNSW-NB15 datasets, as shown in Figures 5, 6 and 7 respectively. These are open access datasets which were selected for their balanced range and similarities (usual or normal, and unusual or anomalous conditions).

Following the above, we implemented five different classifiers for pursuing a comparative work and identification of the best performing ones, while use the three open datasets. The ML classifiers included, the Naïve Bayes (NB), Logistic Regression (LR), Decision Tree (DT), Random Forest (RF) and Gradient-Boosted Tree (GBT).

The respective performances, using various metrics, of the five classifiers is summarized in the tables below.

The five classifiers worked mostly well for the detection of unusual behavior when tested on cyber communication traffic data. They can indeed be used to monitor cyber-behavior in context of urban environments, where

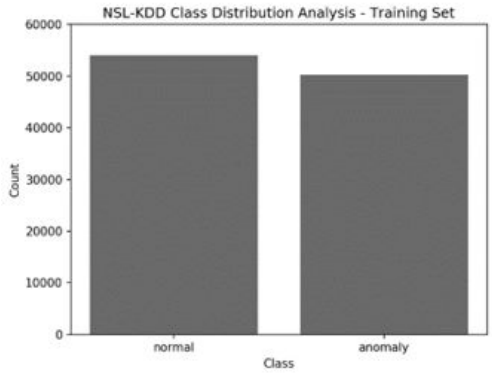

Figure 6: Class distribution of NSL-KDD dataset.

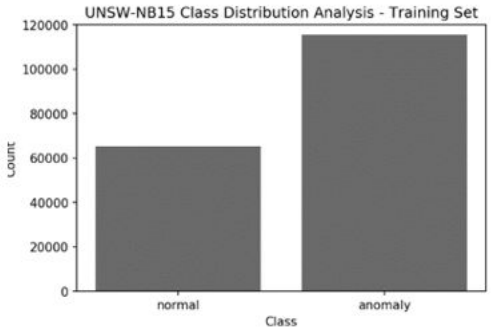

Figure 7: Class distribution of CICIDS-2017 dataset.

Table 1. Performance for the CICIDS-2017 dataset.

Model	Accuracy (%)	Precision (%)	Recall (%)	MCC	ROC AUC
NB	80.3	50.0	0.0	0.0012	0.5000
GBT	98.6	95.5	97.7	0.9572	0.9975
DT	98.2	93.5	97.7	0.9449	0.9752
RF	97.5	93.9	93.5	0.9218	0.9920

Table 2. Performance for the NSL-KDD dataset.

Model	Accuracy (%)	Precision (%)	Recall (%)	MCC	ROC AUC
NB	51.8	45.9	3.5	−0.0079	0.7289
LR	95.4	96.7	93.6	0.9090	0.9857
GBT	98.4	97.9	98.7	0.9676	0.9981
DT	96.3	983	94.0	0.9274	0.8805
RF	95.9	97.9	93.4	0.9176	0.9913

diverse IoT systems are deployed for information communication in smart spaces. Across all three datasets, the NB classifier performed poorly. LR and RF performed satisfactorily across the three datasets, while DT and GBT showed the strongest and most consistent performances. Both DT and GBT

Table 3. Performance for the UNSW-NB15 dataset.

Model	Accuracy (%)	Precision (%)	Recall (%)	MCC	ROC AUC
NB	74.6	59.8	90.6	0.5413	0.4018
LR	88.1	90.3	75.1	0.7385	0.9448
GBT	92.9	89.7	90.9	0.8470	0.9836
DT	91.3	85.2	91.8	0.8150	0.8848
RF	89.3	99.2	71.0	0.7751	0.9459

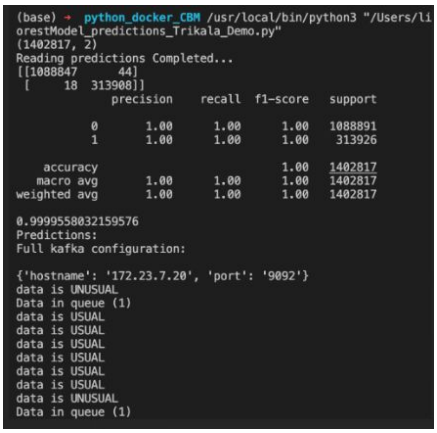

Figure 8: Cyber behavior detection with Kafka messaging / Docker technology.

classifiers strongly performed, nonetheless the GBT comes first since it scored the highest recall and MCC values.

Discussion

The cyber behavior detection module development has been completed and integrated into the *MAIDS DT*, together in context of the S4AllCities SoS for pushing its output towards the *ACMS DT* COP interface. As shown in Figure 8, Dockers technology and Kafka messaging are illustrated for pushing the cyber behavior detector module output to *ACMS* .

Further research on cyber behavior detection continues. We are now investigating the classification of various attack types using labelled cyber communication data. Data is harvested and labelled in real terms at one of our partners pilot sites in the S4AllCities project. The performing DT and GBT classifiers will be trained and tested on the classification of various types of cyber attacks under the *MAIDS DT*.

PHYSICAL BEHAVIOUR DETECTION

The physical behavior detection module under the *MAIDS DT* specializes in the detection of crowd behavior using camera vision. In the same way as the spatial sensor fusion, physical behavior detection heavily relies on the analysis of image data processing and understanding using computer vision principles (Sabeur et al., 2015; Arbab-Zavar and Sabeur 2020, 2021). We

Figure 9: Individual detection and tracking strategies.

developed methodologies in our previous work which are based on statistical mechanics principles, involving concepts of entropy, kinetic and potential energies of crowds for behavior analyses. These were measured successfully using computer vision and features extraction. In addition, we adopted existing libraries such as YOLO v4 and v5 (Bochkovskiy et al., 2020), to be trained to track individuals within crowds and measure their velocities, to speed up our understanding of crowd behavior in this project. We used open access data in this occasion. In this paper, we will discuss YOLO v4 performances only, while further work using YOLO v5 combined with our original approach using statistical mechanics will be discussed in subsequent papers in the near future.

Physical Behaviour Detection Using YOLO v4

In this subsection, the detection and tracking of crowd individuals using YOLO v4 are discussed. Each individual within the crowd is strategically associated with an ID and bounding box coordinates for tracking them in a scene of interest. With the ID in place, we could track physical motions of each individual over video image sequencing at the micro-scale of the crowd "ensemble" (See Figure 9). The local environmental conditions under which the individual is tracked is also known from the spatial fusion module using the RBF network.

The use of YOLO v4 led to associating ID numbers to each detected individual in scenes of interest. Challenges of occlusions were not present, and we have successfully completed the re-identification of each individual in the scene, as shown in Figure 10. Then, each individual ID was also associated with an instant micro-scale velocity in pixel per second, to construct trajectories and understand the way they evolve in space and time in terms of speed and direction, as shown in Figure 11.

The tracked individuals in the scene of interest combined with the environmental conditions can lead us to associate them with specific behavior. Hence it is important to investigate more on the lines of environmental parameters such as noise levels and not just ambient temperatures while tracking individuals. The output concerning the behavior of individuals is therefore delivered through Dockers and Kafka messaging to the *ACMS DT*, as illustrated in Figure 12 below. Strings of messages are released in time, while tagged as usual, or unusual, behavior depending on the measured velocities of each tracked individual.

Discussion

YOLO v4 presents a good basis for inclusion in understanding crowd behavior. With it we can track individuals and measure their velocities while

Figure 10: Associated ID numbers of individuals within the crowd space.

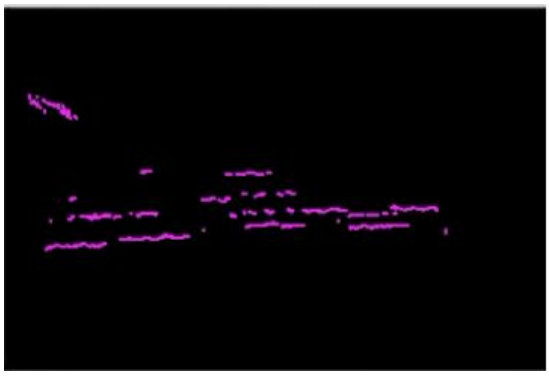

Figure 11: Associated motion trajectories for individuals in a scene of interest.

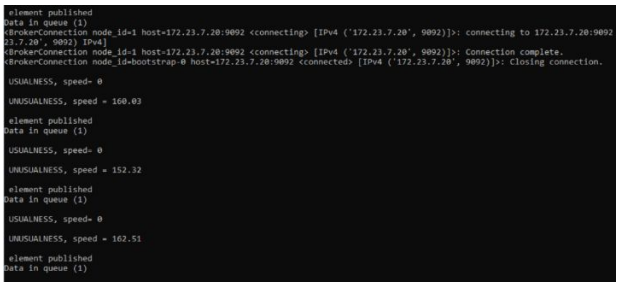

Figure 12: Physical behavior detection with Kafka messaging / Docker technology.

generating their spatial and temporal trajectories. The aim here is to discover trends in the trajectories themselves. The way they form in time and space and the way they could be assembled under specific classes if of great importance to our studies. With the study of such trends, one may be able to extract valuable knowledge on how crowds are formed and evolve in space and time. With the generated trajectories we aim at synthesizing their functions and be able to forecast their future trends while we could validate them using an adaptive predictor-corrector approach.

CONCLUSION AND FUTURE WORK

The development of three core modules for the detection of cyber and physical behavior in urban smart spaces and in context of environmental conditions has been developed for the *MAIDS DT* in the S4AllCities project. Their respective output is channeled using Kafka messaging for advancing situation awareness at higher levels of the JDL framework. At level 4 of the JDL, processes of intelligent reasoning are expected to be factored with the detected cyber and physical behavior, while combining them with native context knowledge of the related space. We have recently deployed our *MAIDS DT* technology successfully for demonstration in Trikala, Greece and Pilsen, Czech Republic. The final pilot demonstration of our digital twin capabilities will be tested and demonstrated in Bilbao, Spain in October 2022.

ACKNOWLEDGEMENT

The authors are grateful to the European Commission for partly funding our research work in the S4AllCities project, under the H2020 Programme, Grant Agreement No. 883522. Our early work on cyber behavior detection in the S4AllCities project, which was performed by Liam Collick is acknowledged.

REFERENCES

Arbab-Zavar, B., Sabeur, Z. (2020) Multi-scale crowd feature detection using vision sensing and statistical mechanics principles. Machine Vision and Applications. https://doi.org/10.1007/s00138-020-01075-4

Bochkovskiy, A., Chien-Yao, W., Hong-Yuan, L. (2020). "Yolov4: Optimal speed and accuracy of object detection." arXiv preprint arXiv:2004.10934.

Broomhead, S., Lowe, D. (1988). Radial basis functions, multi-variable functional interpolation and adaptive networks. Royal Signals and Radar Establishment Malvern (United Kingdom).

Endsley, M.R. (1995). Toward a Theory of Situation Awareness in Dynamic Systems. The Journal of Human Factors and Ergonomics Society, 37(1), 32-64

Lambert, D.A (2009) A Blueprint for higher-level fusion systems. Information Fusion, Vol. 10, Issue 1. Pp 6–24.

Sabeur, Z. (2021a) AI3SD Video: Artificial Intelligence for Safer Urban Space. Frey, J. G., Kanza, S. and Niranjan, M. (eds.) AI3SD Autumn Seminar Series 2021. 13 Oct - 15 Dec 2021. doi:10.5258/SOTON/AI3SD0173

Sabeur. Z. (2021b). Lectures Series on Cyber Situation Awareness 2019-2021, Department of Computing and Informatics. Bournemouth University, UK.

Sabeur, Z., Arbab-Zavar, B. (2021). Crowd Behaviour Understanding Using Computer Vision and Statistical Mechanics Principles. In: Bellomo, N., Gibelli, L. (eds) Crowd Dynamics, Volume 3. Modeling and Simulation in Science, Engineering and Technology. Birkhäuser, Cham. https://doi.org/10.1007/978-3-030-91646-6_3

Sabeur, Z., Doulamis, N., Middleton, L., Arbab-Zavar, B., Correndo, G., Amditis,A. (2015) Multi-modal computer vision for the detection of multi-scale crowd physical motions and behavior in confined spaces. In: Advances in Visual Computing, pp. 162–173. Springer, New York.

Sabeur Z., Angelopoulos C.M., Collick L., Chechina N., Cetinkaya D., Bruno A. (2021) Advanced Cyber and Physical Situation Awareness in Urban Smart Spaces. In: Ayaz H., Asgher U., Paletta L. (eds) Advances in Neuroergonomics and Cognitive Engineering. AHFE 2021. Lecture Notes in Networks and Systems, vol 259. pp. 428–441. Springer, Cham. https://doi.org/10.1007/978-3-030-80285-1_50

S4AllCities (2020). Safe and Secure Smart Spaces for all Cities H2020 project ID number 883522. https://www.s4allcities.eu/project