Run Length Encoded Dynamic Bayesian Networks for Probabilistic Interaction Modeling

Simone Chiappino, Lucio Marcenaro, Pietro Morerio, Carlo Regazzoni
DITEN – University of Genoa – Via Opera Pia 11A – 16145 Genoa Italy

ABSTRACT

Human behavior analysis for Cognitive Surveillance Systems (CSS) share mainly the concept that it can be time to extend functionalities beyond simple video analytics. In most recent systems addressed by research, automatic support to human decisions based on object detection, tracking and situation assessment tools is integrated as a part of a complete cognitive artificial process. In such cases a CSS needs to represent complex situations that describe alternative possible real time interactions between the dynamic observed situation and operators' actions. To obtain such knowledge, particular types of Event based Dynamic Bayesian Networks E-DBNs are here proposed. In this paper it is shown how, by means of Run Length Encoding (RLE) of off line acquired information, the cognitive system is able to represent and anticipate possible operators' actions within the CSS. Results are shown by considering a crowd monitoring application in a critical infrastructure. A system is presented where a CSS embedding in a structured way RLE E-DBN knowledge can interact with an active visual simulator of crowd situations. Outputs from such a simulator can be easily compared with video signals coming from real cameras and processed by typical Bayesian tracking methods.

Index Terms — Event based DBN, bio-inspired learning, cognitive system, interaction modeling, crowd monitoring

1. INTRODUCTION

Video analytics for surveillance in critical areas is becoming more significant for public security. Several works have been devoted in the last decade to link traditional computer vision tasks to high-level context aware functionalities such as scene understanding, behavior analysis, interaction classification or recognition of possible threats or dangerous situations. Most current automated video surveillance systems can process video sequences and perform almost all key low-level functions, using some principal techniques: visual objects detection - segmentation, object tracking and object classification [1], [2]. For example, in [3] is presented a technique for automatic detection of relevant abnormal events in video surveillance application. Recently, technical interest in video surveillance has moved from such low-level functions to more complex scene analysis to detect human and/or other object behaviors. Several recent studies have proposed the application of smart functionalities to camera and sensor networks in order to move from object recognition paradigm to event/situation recognition one [4]. Among the several disciplines which are involved in the design of next generation security and safety systems, bio-inspired processing approaches [5] represent one of the most promising in terms of capability of provoking improvements with respect to state of the art. In fact studies in cognitive science nowadays allow to explain human reasoning functions at higher semantic level than in previous decades. For example [6] shows how the application of bio-inspired models to safety and security tasks can represent a relevant added value in understanding the dynamic evolution of complex scenes, where multiple patterns interact in accordance with specific dynamic relations feature and time spaces. In addition, to efficiently exploit cognitive capabilities in an intelligent sensor network, the role of data fusion architecture and algorithms is crucial, [7]. In the literature, several works deal with data fusion problem applied to heterogeneous sensors for security [8] and safety tasks [9]. A video analytics automatic system is able to extract visual objects from video frames applying multi levels analysis, in order to model monitored entity behaviors. In [10] experience acquisition importance from the interactions with the environment in order to make decisions, is shown, from intelligent system point of view.
A crowd monitoring scenario is here considered, where a human operator controls an environment and interacts with the people in it, through actions aiming at restoring normality after abnormal crowding event detection. In this section the proposed learning and predictive models are compared. It is shown that operator capability at the decision level can be extended by allowing the system to learn and use more simple operator regulation actions.
The next paper's sections are organized as follows: the proposed model will be discussed in section 2 and 3. Section 4 is focused on the presentation an applicative scenario. In the section 5, results and conclusion are presented.

2. COGNITIVE SURVEILLANCE NODE

In this work, a probabilistic model based on a specific type of event based Dynamic Bayesian Networks (E-DBNs) is discussed that comes from a bio-inspired perspective. The bio-inspired approach and a possible implementation by using the so called Autobiographical Memories (AMs) has been already discussed in [11] [12], based on the work of

the neuro-physiologist A. Damasio [5]. Damasio's theories describe the cognitive entities as complex systems capable of incremental learning based on experience of the relationships between themselves and the external world. Two specific brain processes can be defined to formalize the above concept called proto-self and core-self. Such devices are specifically referred to monitor and manage respectively the internal status of an entity (proto-self) and the relationships with the external world (core-self). Thus, a crucial aspect in modeling a cognitive entity following Damasio's model is represented, on the one hand, by the capability of accessing entity's internal status and, to the other hand, by the knowledge and analysis of the surrounding environment. These concepts are the guidelines to design a Cognitive Surveillance Node, in which two types of sensors are mapped into a sensing framework, namely endo (or proto-sensors) and eso-sensors (or core-sensors). Endo-sensors and eso-sensors monitor the internal or external state of the interacting entities, respectively. Dore *et al.* demonstrate AM capability to store the spatio-temporal information about the relationships which occur between two interacting entities. Specifically, it is proposed that proto-core-proto or core-proto-core triplets of events are stored into the AM, defining passive interactions (passive triplets) and active interactions (active triplets) between a reference entitiy and an interacting one. In [6], [11], [12] the states dynamic evolution of each entity is represented by a probabilistic model based on E-DBNs. Furthermore, Coupled E-DBNs (C E-DBNs) are introduced as guideline for interactions representation problem between entities. In this paper a new type of C E-DBNs is proposed, in which the causality relationships, among the events, is highlighted. In the proposed C E-DBNs structure, strong causality constraints are introduced. A new encoding for the events is proposed, based on Run Length Encoding (RLE), in order filter *relevant information* only. Using RLE it is possible to compact a sequence of equal events (e.g. null events considered as less relevant). This paper integrates such a RLE C E-DBNs framework in a Cognitive Surveillance Node (CSN).
Using the above explained concepts it is possible to redefine the AM structure based on C RLE E-DBNs. A sample CSN framework based on the JDL model [13], which is mapped into a Cognitive Cycle (CC) [11] is presented in Figure 1. The task is to establish a bridge between the concepts introduced by Damasio and the effective implementation of the system. In particular, it will be shown how the proposed C RLE E-DBNs can describe an AM structure capable to provide a more efficient knowledge representation, about different interactions between entities, at different levels of the JDL model.
In this work, the proposed architecture is applied and evaluated in the crowd monitoring domain, where the goal of the system is to analyze and classify crowd interactions in order to maintain a proper security level in the monitored area and to put in action effective countermeasures in case of detection of panic or overcrowding situations.

3. PROPOSED COGNITIVE MODEL

In this paper, a new type of Coupled Event based DBNs (C E-DBNs) is presented. Such a structure is able to provide an efficient knowledge representation for what concerns interactions.

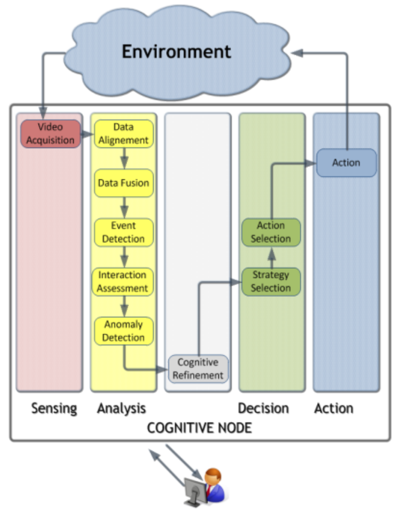

Figure 1: Cognitive Surveillance Node framework

The first step is to introduce a different way to represent dynamic state variables within each single E-DBN that describe proto and core dynamics. The variables that describe the state and time at which occurs are considered as disposed according to a reversed hierarchic priority with respect to classical DBNs as in [11]. Let us define the state label produced by core/proto source tracking and by classifying the result as:

$\bar{S}_j = \left\{S_{1,j}, \dots, S_{N_j,j,}\right\}$ is the set of possible state labels observed at each time by the system, indicated as Y, where each $S_{i,j}$ denotes the i^{th} component associated either with core source (if $j = C$) or with proto source (if $j = P$). Let us define a State Sequence as a temporal series of state variables $\Sigma_j = \{(\sigma_m^j, t_m) : m = 0 \dots M-1, \sigma_m^j \in \bar{S}_j, t_0 = T_0, t_m = T_0 + m\Delta t\}$.

In order to define an *event* a *mobile window* $W_m(.,.)$ is defined. Such operator can be applied to a couple of two consecutive elements of Σ_j:

$A_m^j = W_m\left((\sigma_m^j, t_m), (\sigma_{m-1}^j, t_{m-1})\right)$, where A_m^j represents the *event*. It can be fully described as: $A_m^j = (\sigma_m^j, \sigma_{m-1}^j, t_m)$ where $\sigma_m^j, \sigma_{m-1}^j \in \bar{S}_j$. Informative content of an event can be analyzed as follows:

$$\begin{cases} (\alpha_m^j, t_m): \ \alpha_m^j = (\sigma_m^j, \sigma_{m-1}^j) \text{ if } \sigma_m^j \neq \sigma_{m-1}^j \\ \qquad \emptyset^j = (\emptyset, t_m) \text{ if } \sigma_m^j = \sigma_{m-1}^j \end{cases} \quad (1)$$

Where $\emptyset$ is a *null event* (i.e., not relevant event), while σ_m^j defines a relevant proto/core event. Let us now consider two alphabets from which state transitions estimated by Y can take values $\mathcal{E}_j = \{\alpha^{j,1}, \dots, \alpha^{j,V_j}\}$, where $\alpha^{j,i}$ is an event-symbol describing a specific state transition and V_j is the number of events. It is possible to define the event as: $A_m^j = (\alpha_m^j, t_m)$: $\alpha_m^j \in \mathcal{E}_j$. In general the i-th $\alpha^{j,m}$ symbol is used to represent a specific state label change.

An event sequence Ψ_j is defined as follows: $\Psi_j = \{A_m^j, m = 1 \dots M-1\}$.

Given a sequence of events Ψ_j, in order to compact consecutive null events and as consequently underline relevant events, it is possible to define an run length encoding function, as follows: $R(\Psi_j)$. This operator can generate a RLE sequence $\Upsilon_j = R(\Psi_j)$ as:

$$\Upsilon_j = \{x_1^j \alpha_{m_1}^j ! \, x_2^j \, \alpha_{m_2}^j ! \dots \dots ! \, x_N^j \, \alpha_{m_N}^j\} \qquad (2)$$

Where x_k^j is the number of occurrences of $\alpha_{m_k}^j$ event with $k \in \{1, \dots, N\}$ and m_k is the k-th time instant. The mark "!" is separator between $\alpha_{m_k}^j$ and $\alpha_{m_{k+\Delta k}}^j$. Υ_j is a sequence over an *ordered time k*. Δk is a time sampling sufficient to represent relevant and null events in Υ_j sequence.

It is possible to model E-DBNs starting from RLE encoding definition for events sequence. Considering an discrete index k that represents the relative position of the event in a RLE sequence Υ_j then one can fix $Y_j(k) = [Y_1^j(k), \dots, Y_{V_P}^j(k)]$ where $Y_x^j(t)$ are binary random variables such that $Y_x^j(k) = 1$ when $\alpha^{j,x} \in \mathcal{E}_j$ $\alpha^{j,x} \neq \emptyset$, $x \in 1, \dots, V_j$; and $Y_x^j(k) = 0$ otherwise. Given any couple of events in Υ_j $DY_{i,s}^j(k) = \left(Y_i^j(k+\Delta k), Y_s^j(k)\right)$, it is possible to define a conditional probability $p\left(Y_i^j(k+\Delta k) = 1 \middle| Y_j^j(k) = 1\right)$ as the occurrences that a not null event $Y_i^j(k+\Delta k)$ follows a given not null event $Y_j^j(k)$

It is important to note that any couple of events $DY_{i,s}^j(k) = \left(Y_i^j(k+\Delta k), Y_s^j(k)\right)$ is defined in according to a maximum time W_{max}^j. Such temporal window permits to define a vector of random variables:

$$T_{i,s}^j = T_{DY_{i,s}^j(k)}^j = \left[T_1, \dots, T_{W_{max}^j}\right] \qquad (3)$$

such that if $Y_i^j(k+\Delta k) = 1$ then $T_r = 1$ if $r = \Delta k$ with $r \in \{1, \dots, W_{max}^j\}$. The vector defined before represents a second hierarchic level, in which the occurrence time between two events of the same entity is stored. In particular, the probability of $T_{i,s}^j$ comes out to be dependent on which coupe of successive events occurred at time k and $k+\Delta k$, i.e. one can estimate easily from Υ_j sequences conditioned probabilities of the type

$$p\left(T_{i,s}^j = 1 \middle| Y_i^j(k+\Delta k) = 1, Y_s^j(k) = 1\right) \qquad (4)$$

In Figure 2 is shown a hierarchic RLE E-DBN structure.

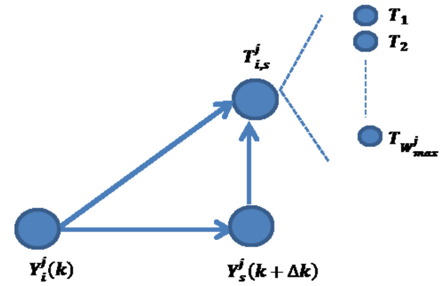

Figure 2: Example of a RLE E-DBN for generic couple of events $\boldsymbol{DY_{i,s}^j(k)}$where it is possible to note the second hierarchic level in which is stored the occurrence time between events.

Let us suppose one has to deal with two RLE sequences, Υ_P and Υ_C. Under the hypothesis that Υ_P and Υ_C are produced by a time aligned couple of proto and core intelligent sensors, sharing the same starting time t_0 and processing data at the same time sampling Δt, it is possible to define their RLE intersection (I-RLE) sequence Υ_{PC} as:

$$\Upsilon_{PC} = \left\{\dots \dots ! \, x_{\xi-1}^C \alpha_{m_{\xi-1}}^C ! \, x_\xi^P \, \alpha_{m_\xi}^P ! \, x_{\xi+1}^C \, \alpha_{m_{\xi+1}}^C ! \dots \dots\right\}, \quad (5)$$

where only one proto event occurs between two core events. It is possible to represent the interactions, between proto and core events, in a I-RLE sequences using Coupled RLE E-DBN. Given any couple of events $DY_{i,s'}^{jj'}(k) = \left(Y_i^j(k+\Delta k), Y_{s'}^{j'}(k)\right)$, in which $j \neq j'$ and $i \in 1, \dots, V_j$ while $s' \in 1, \dots, V_{j'}$, a third hierarchic level can be defined as the vector (of random variables):

$$T_{i,s'}^{jj'} = T_{DY_{i,s'}^{jj'}(k)}^{jj'} = \left[T'_1, \dots, T'_{W_{max}^{jj'}}\right] \qquad (6)$$

given that if $Y_i^j(k + +\Delta k) = 1$ then $T'_r = 1$ if $r = \Delta k$ with $r \in \{1, \dots, W_{max}^{jj'}\}$. In particular, the probability of $T_{i,s'}^{jj'}$ comes out to be dependent on which pair of successive events occurred at time k and $k+\Delta k$, i.e. from $\Upsilon_{jj'}$ sequences it possible to write the conditioned probabilities as follows:

$$p\left(T_{i,s'}^{jj'} = 1 \middle| Y_i^j(k+\Delta k) = 1, Y_{s'}^{j'}(k) = 1\right) \qquad (7)$$

Let us now consider a triplet of events as $TY_{i,s',s}^{jj'j}(k) = \left(Y_i^j(k+\Delta k), Y_{s'}^{j'}(k), Y_s^j(k-\Delta k)\right)$, it is possible to define associated conditional probabilities as follows:

$$p\left(\frac{T_{i,s}^j = 1, T_{i,s'}^{jj'} = 1}{Y_i^j(k+\Delta k) = 1, Y_{s'}^{j'}(k) = 1, Y_s^j(k-\Delta k) = 1}\right) =$$
$$p\left(T_{i,s'}^{jj'} = 1 \middle| Y_i^j(k+\Delta k) = 1, Y_{s'}^{j'}(k) = 1\right) \cdot$$
$$p\left(T_{i,s'}^{j'j} = 1 \middle| Y_{s'}^{j'}(k) = 1, Y_s^j(k-\Delta k) = 1\right) \qquad (8)$$

It is possible to assume the following definitions: passive interactions referred to $TY_{i,s',s}^{PCP}(k)$ triplets (proto-core-proto), active interactions when $TY_{i,s',s}^{CPC}(k)$ (core-proto-core). In Figure 3 is shown a hierarchic C RLE E-DBN structure.

This dynamic model describes the relationships within the Autobiographical Memory in terms of cause effect relationships storing the frequencies of occurrences of event triplets $p\left(\frac{x_y^j(t+\Delta t)}{x_{x'}^{j'}(t),x_2^j(t-\Delta t)}\right)$, that can be describe by Coupled DBNs, the occurrence times, between the triplets, are stored into temporal histograms. Now the conditional probabilities are defined by Coupled RLE E-DBNs as follows:

$$p\left(\frac{Y_i^j(k+\Delta k)}{Y_{s'}^{j'}(k),Y_s^j(k-\Delta k)}\right) \equiv p\left(\frac{T_{i,s}^j=1,T_{i,s'}^{jj'}=1}{Y_i^j(k+\Delta k)=1,Y_{s'}^{j'}(k)=1,Y_s^j(k-\Delta k)=1}\right) \quad (9)$$

and the occurrence times are embedded into more compact structure as upper hierarchic levels.

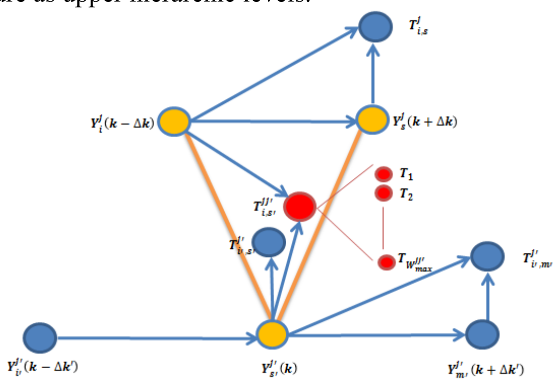

Figure 3: Example of a C RLE E-DBN for generic triplet of events $\boldsymbol{TY_{i,s',s}^{jj'j}(k)}$where it is possible to note three hierarchic levels in which are stored the occurrence time between events

The equation (9) permits to perform the prediction task when an external event $Y_{s'}^{j'}(k)$ is detected by the system, the AM is analyzed to establish which was the previously occurred internal event $Y_s^j(k-\Delta k)$. The Autobiographical Memory is then examined to establish which is the internal event $\hat{Y}_i^j(k+\Delta k)$ that is more likely to occur:

$$\hat{Y}_i^j(k+\Delta k) = max_{\hat{Y}_i^j}\left\{p\left(\frac{x_y^j(t+\Delta t)}{x_{x'}^{j'}(t),x_2^j(t-\Delta t)}\right)\right\} \quad (10)$$

4. CROWD MONITORING

A crowd monitoring scenario is here considered, where a human operator controls an environment through actions aiming at restoring normality after abnormal crowding event detection. In this section the proposed learning and predictive models are compared. Operator reactions are monitored by CSN endo-sensors as changes in the system status that are associated with operator actions when the operator assumes control of action flow. On the other hand, CSN eso-sensors are associated with crowding changes. This approach, according to the proposed bio-inspired learning model, permits a knowledge transfer from human operator to the CSN. The human-environment interactions are stored into the AM and the operator actions are strictly linked to the human interpretations of the scene under control (e.g. dangeros crowding or safety situation). The state labels produced by operator (proto source) are the actions, which correspond to the interpretations of the crowding situations (core source).

A simulation environment has been developed, in order to produce different crowding scenarios. When an operator observes some anomalous crowding situations can produce alarm messages (AlMes), which are viewed on some control panels, to redirect the people flow. Circumstances when doors opens to force quit (FQ) the people, is considered as dangerous situations. Instead, in normal crowding situations an operator can produce non-relevant actions (NULL) or can restore the normal (RtoN) conditions of the environment. Self-Organizing Map (SOM), can be employed in order to reduce the dimensionality of the crowding states. E.g., SOMs can be used in all the case in which multi-camera sensors are monitoring different critical zones. As an example, it is considered the sequence of crowding events (α_m^C, t_m) (core events), shown in Figure 4; for each of them the operator associates specific actions (α_m^P, t_m) (proto events).

$(\boldsymbol{\sigma_{m-1}^C, t_{m-1}})$	$(\boldsymbol{\sigma_m^C, t_m})$	$(\boldsymbol{\alpha_m^C, t_m})$	$(\boldsymbol{\alpha_m^P, t_m})$
(4,0)	(4,3)	(4 → 4,3)	$(NULL, 3{,}1)$
(4,3)	(4,6)	(4 → 4,6)	$(NULL, 6.1)$
(4,6)	(55,9)	(4 → 55,9)	$(NULL, 9.1)$
(55,9)	(55,12)	(55 → 55,12)	$(NULL, 12.1)$
(55,12)	(57,15)	(55 → 57,15)	$(NULL, 15.1)$
(57,15)	(57,18)	(57 → 57,18)	$(NULL, 18.1)$
		…	
(57,36)	(57,39)	(57 → 57,39)	$(NULL, 39.1)$
(30,42)	(29,45)	(30 → 29,45)	$(NULL \rightarrow$ AlMess, 47.5)
(29,45)	(29,48)	(29 → 29,48)	(AlMes → AlMes, 48.1)
(29,48)	(25,51)	(29 → 25,3)	(AlMes → FQ, 52)
		…	
(25,60)	(4,64)	(25 → 4,64)	$(FQ \rightarrow$ RtoN, 66)

Figure 4 Crowding state labels evolutions, definition of core/proto events.

The core/proto events presented in Table 1 correspond to following core/proto events RLE encoding sequences:

$$Y_C = \{2\ \emptyset!\ 4 \rightarrow 55!\ \emptyset!\ 55 \rightarrow 57!\ 8\ \emptyset!\ 57 \rightarrow 30!\ 30 \rightarrow 29!\ \emptyset!\ 29 \rightarrow 25!\ 3\emptyset!\ 25 \rightarrow 4\}$$

$$Y_P = \{14\ \emptyset!\ NULL \rightarrow AlMes!\ \emptyset!\ AlMes \rightarrow FQ!\ 3\ \emptyset\ !\ FQ \rightarrow RtoN\}$$

It is possible to represent the interactions, between proto and core events, in I-RLE sequences using Coupled RLE E-DBNs.

$$Y_{CP} = \{2\ \emptyset!\ 4 \rightarrow 55!\ \emptyset!\ 55 \rightarrow 57!\ 8\ \emptyset!\ 57 \rightarrow 30!\ 30 \rightarrow 29!\ NULL \rightarrow AlMes!\ 29 \rightarrow 25!\ AlMess \rightarrow FQ!\ 2\ \emptyset!\ 25 \rightarrow 4!\ FQ \rightarrow RtoN$$

From Y_{CP} a passive and active AMs can be estimated on the basis of relative frequencies of occcurences of three alternated proto/core events. It is possible to demonstrate that the CSN, using proposed learning approach, is able to detect abnormal crowding events determining appropriate reactions, in according with normal crowding maintaining tasks.

5. RESULTS AND CONCLUSION

This section describes the experiments on several synthetic crowding sequences provided by the crowd simulator [14] [15]. The simulated environment consists of six rooms, monitored by cameras (virtual sensors). These try to reproduce (processed) sensor data coming from different cameras looking at different subsets (rooms) of the monitored scene. A virtual people estimation algorithm outputs the number of people by simply adding some noise to the mere number of people framed by the virtual camera.

The impact of the RLE based C E-DBNs in terms of maintaining a proper security level within the monitored area has been instigated. For this reason we try to evaluate the performances of the system, by defining a crowding index I_i^C for each room i of the monitored environment:

$$I_i^C = \frac{N_i}{N_i^{max}} \quad (11)$$

Where N_i is the number of individual measured in room i and N_i^{max} is the maximum number of people allowed in that room (N_i^{max} is proportional to room areas). System performance has been studied for different values of the temporal window $W_{max}^{jj'}$. Figure 5 presents a crowding index trend using a (heuristically tuned) causality window $W_{max}^{jj'} = 10[s]$: it is possible to note that after a transition period, the crowding indexes tend to stabilize, meaning that a regular stream of people flows through the environment. In this case the CN is able to find and to put in action the best strategy (i.e. doors configuration) in order to regulate the people flow to avoid overcrowding situations.

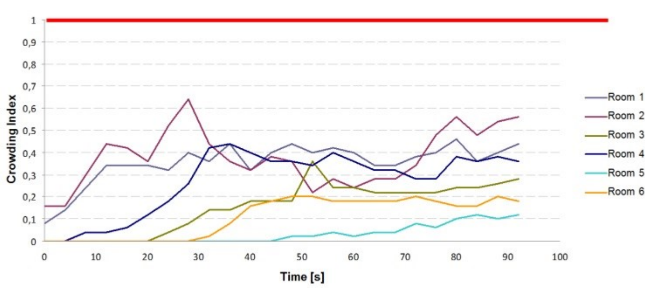

Figure 5 Crowding index with $\boldsymbol{W_{max}^{jj'}}$ = 10[s]

The described work has the purpose to show how coupled a specific type of RLE E-DBNs are a promising tool to represent efficiently several surveillance and security maintenance applications. In particular, the experimental tests show promising results of proposed AM structure capability of the learning the interactions and of predicting their evolution in a crowd control application within a critical infrastructures The presented work shows how the proposed learning model is able to provide a more enriched description of the events for a more accurate interactions analysis while addressing the task of designing next generation surveillance systems.

6. REFERENCES

[1] L. Marcenaro, L. Marchesotti, and C. S. Regazzoni, "Self-organizing shape description for tracking and classifying multiple interacting objects," Image Vision Comput., vol. 24, no. 11, pp. 1179–1191, 2006.

[2] K. N. Plataniotis and C. S. Regazzoni, "Visual-centric surveillance networks and services," IEEE Signal Processing Magazine, vol. 22, no. 2, pp. 12–15, March 2005.

[3] M. Spirito, C. S. Regazzoni, and L. Marcenaro, "Automatic detection of dangerous events for underground surveillance," in IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS 2005). Como, Italy: IEEE Computer Society, September 2005, pp. 195–200.

[4] M. Valera, S. Velastin, "Intelligent distributed surveillance systems: a review," Vision, Image and Signal Processing, IEEE Proceedings 52(2), 192–204 (2005.

[5] A.R. Damasio, "The Feeling of What Happens-Body, Emotion and the Making of Consciousness," Harvest Books (2000).

[6] C. S. Regazzoni, "Emphatic human interaction analysis for cognitive environments," in Proceedings of the first ACM international workshop on Analysis and retrieval of tracked events and motion in imagery streams (ARTEMIS '10). ACM, New York, NY, USA, 1-2.

[7] G.L. Foresti, C.S. Regazzoni, P.K. Varshney, "Multisensor Surveillance Systems: The Fusion Perspective", Kluwer Academic, Boston (2003).

[8] D. Smith, S. Singh, "Approaches to multisensor data fusion in target tracking: A survey," IEEE Transactions on Knowledge and Data Engineering 18(12), 1696–1710 (2006).

[9] B.R. Chang, H.F. Tsai, C.P. Young, "Intelligent data fusion system for predict-ing vehicle collision warning using vision/gps sensing". Expert Systems with Applications 37(3), 2439 – 2450 (2010).

[10] L. Bixio, A.F. Cattoni, C.S. Regazzoni, P.K. Varshney, "Embodied cognition-based distributed spectrum sensing for autonomic wireless systems," in: Y. Zhang, L.T. Yang, M.K. Denko (eds.) Autonomic Computing and Networking, pp. 179–203. Springer US (2009).

[11] A. Dore, A. F. Cattoni, and C. S. Regazzoni. 2010. Interaction modeling and prediction in smart spaces: a bio-inspired approach based on autobiographical memory. *Trans. Sys. Man Cyber. Part A* 40, 6 (November 2010), 1191-1205.

[12] A. Dore and C. S. Regazzoni. 2010. Interaction Analysis with a Bayesian Trajectory Model. *IEEE Intelligent Systems* 25, 3 (May 2010), 32-40.

[13] A. N. Steinberg, C. L. Bowman, and F. E. White, Revisions to the JDL data fusion model, Proc. SPIE 3719, 430 (1999).

[14] S. Chiappino, P. Morerio, L. Marcenaro, E. Fuiano, G. Repetto, and C. Regazzoni, "A multi-sensor cognitive approach for active security monitoring of abnormal overcrowding situations in critical infrastructure," 15th International Conference on Information Fusion, 2012.

[15] S. Chiappino, P. Morerio, L. Marcenaro and C. S. Regazzoni, "A bio-inspired knowledge representation method for anomaly detection in cognitive video surveillance systems," 16th International Conference on Information Fusion , 2013.