An Agent-Based Approach for Data Fusion in Homeland Security

Dr. José Miguel Castillo Chamorro[1], D. Rafael de Solís Montes[2]
[1]Pontifical University of Salamanca, Madrid –Spain-
[2]Superior School of Telecommunications, Polytechnic University in Madrid, Madrid –Spain-

***Abstract* — This article presents an agent-based solution for data fusion in Homeland Security. Communication technology has been developed very fast in the last decades. We can get lots of data in milliseconds. Our current problem is to process such amounts of data in order to provide useful information. We have to focus our effort on developing intelligent information systems able to handle big amounts of data extracting or revealing relations among data and able to produce information easily understandable for the human user. That is the case of data fusion in tactical operations, especially in the field of defense and Homeland security. Our research is focused on obtaining a Multi-agent system able to inference future enemy's actions or behaviors from data received from heterogeneous sensors.**

***Keywords* — MAS, Data fusion, tracks, merge, inference, Homeland Security.**

I. Introduction

In the last decades the concept of HOMELAND security plays an important role in our lives. Everybody is concerned about the necessity of protection against undesired attacks from outside our borders or from internal terrorists' threats.

Nowadays, the concept of military defense is merged with the national security concept. Both are complementary and have a common aim: to preserve the social stability against external or internal threats. Furthermore, this need is spread in most occidental countries which implies the necessity of taking common measures by developing common regulations and promoting technological initiatives that make our community less vulnerable.

This article introduces an agent-based architecture [1] [2] as a solution to solve data fusion problems in a tactical environment. One of the advantages that this solution has is that it can be used also for training; so the operational team can be trained in a virtual environment by using the same architecture.

In tactical operations either in the military or for civil purposes like borders control or coasts surveillance, data fusion is a process that facilitates intelligence tasks and will be an important input for planning operations.

When using different sensors to locate objects or targets we will get the same number of synthetic representations as sensors. Our first problem consists of comparing features of these tracks in order to determine if all of them are representing the same target or there are multiple objects in the scene. After that, we have to know if we are dealing with a single object or a group of organized targets. Through sensors we can know about the current and past location of targets, but it is necessary to extrapolate future positions in order to prevent undesired enemy actions to happen.

We have faced the problem with the idea of giving a general solution for such kind of different problems under a single architecture.

This article is divided into different sections. In Section II we introduce the need of data fusion in tactical operations. An analysis of the State of the Art in that field is made in Section III. The system architecture that gathers the four phases in a data fusion system is depicted in Section IV. Our first approach to a communication solution among agents is explained in Section V. In Section VI, the multi-agent operational approach is introduced. And finally we end the article with our plans to complete the research and the conclusions.

II. Data Fusion in Tactical Operations

Data fusion is one of the main activities within the intelligence process. It is essential to figure out the enemy tactics as well as its plans and purposes.

Any Command and Control system within the field of HOMELAND security requires an easy-to-use data fusion system that would not demand specific management tasks as IT personnel will not be available in the lower levels.

The unbalanced capabilities of command and control currently available require an evolution of the platform-centric to a network-centric capability that allows the use of unique data fusion capabilities.

Defense operations cover a wide spectrum of threats and deployment scenarios that range from conventional war through limited operations, crises response operations, asymmetric conflict, and terrorism.

The successful execution of fast moving operations needs an accelerated decision-action cycle, increased tempo of operations, and the ability to conduct operations.

Unilateral capability is important but most planning is made on the assumption of alliance and coalition operations in scenarios that are difficult to predict and which often arise at short notice. Thus the nature and composition of the force structure to meet defense requirements will be specific to requirement and based upon a general and flexible HOMELAND security capability [8]. Data fusion capabilities must allow the deployment of military forces. Additionally, as tactical forces could need to interact with non-governmental organizations, including international aid organizations, Data fusion should be able to pass information to them through web pages in an off-line and secure control way.

It is not only the new capabilities and concepts that accompany the Information Age that are different, but also the nature of the technology itself and the ever increasing rate of change that make our times very different. The discomfort associated with the nature of the changes combined with the incredibly rapid pace of change and the very high cost of error associated with decisions that involve our national security create a formidable set of challenges.

Data fusion should actively seek to provide users with both, the most accurate and timely information, and also, that the users can trust the information they have available is exactly the same information to the members involved in an operation.

Data fusion not only helps us detect redundant data regarding targets location but also contributes to enriching our knowledge about the possible enemy actions by guessing an enemy's future manoeuvres.

Specific requirements

The design of the present data fusion project is based on the following requirements:

a) Centralization of data. A standard database model will be the core of the system. There will be a unique database regarding data fusion.

b) Decentralization of inputs. Inputs to the database will be introduced by any acquisition device. Other command levels can also introduce inputs to the database.

c) Data fusion process can be run in high level headquarters on demand or in headquarters of subordinate units.

d) Aggregation and identification of target groups will be done by grouping isolated targets.

The design of the data fusion module has some technical requirements that it are worth highlighting:

- In order to cluster targets and identify undesired groups of targets, it is necessary to have a precise definition of them in the data model. The definition of potential enemy units should be provided beforehand.
- Any time that a headquarters wants to run the data fusion module, the updated data model has to be accessible. This circumstance implies an efficient use of the communications network.

Data fusion is necessary for intelligence phase and supports planning tasks.

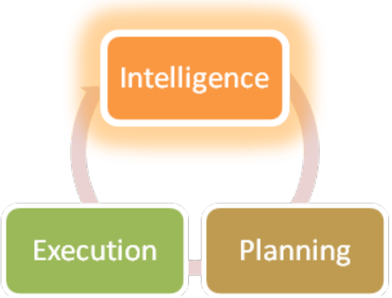

Fig. 1. Intelligence cycle

The phases of the intelligence cycle are illustrated in Figure 1. All these phases are continuous through time. We get information from the scenario that facilitates the planning process of operations. The plan is executed and as a result the scenario changes and we get new information that feeds the intelligence phase that will start a new planning process.

III. State of the Art in Tactical Data Fusion

Data fusion is a process that combines data from multiple sensors, and related information from associated databases, in order to improve accuracy and to make better inferences than could be achieved by the use of a single sensor or data set alone [20].

While the coverage of methodological areas of data fusion systems includes artificial intelligence procedures, pattern recognition, and statistical inference; application areas of data fusion include automated target recognition, guidance for autonomous vehicles, remote sensing, battlefield surveillance, and automated threat recognition systems, such as identification-friend-foe-neutral (IFFN) systems [13].

As explained in [15], the JDL's definition (Joint Director of Labs) of information fusion has four levels, ranging from identifying and tracking targets of interest, to determining whether these targets are real threats.

Other authors divide the fusion process in five, or even six, levels starting in level 0 with the source pre-processing and ending in level 5 with the user refinement.

The JDL model was proposed by the US Joint Directors of Laboratories Data Fusion Sub-Group in 1985 [12].

The model from 1992 is explained by Steinberg, Bowman & White [21] and it proposes five levels for the data fusion process explained as follows:

- Level 0 − Sub-Object Data Assessment: estimation and prediction of signal/object observable states on the basis of pixel/signal level data association and characterization;
- Level 1 − Object Assessment: estimation and prediction of entity states on the basis of observation-to-track association, continuous state estimation (e.g. kinematics) and discrete state estimation (e.g. target type and ID);
- Level 2 − Situation Assessment: estimation and prediction of relations among entities, to include force structure and cross force relations, communications and perceptual influences, physical context, etc.;
- Level 3 − Impact Assessment: estimation and prediction of effects on situations of planned or estimated/predicted actions by the participants; to include interactions between action plans (e.g. assessing susceptibilities and vulnerabilities to estimated/predicted threat actions given one's own planned actions);
- Level 4 − Process Refinement (an element of Resource Management): adaptive data acquisition and processing to support mission objectives.

On the other hand, the Data Fusion Information Group (DFIG) model is divided into six levels, as explained bellow:

- Level 0: Source Preprocessing/subject Assessment
- Level 1: Object Assessment
- Level 2: Situation Assessment
- Level 3: Impact Assessment (or Threat Refinement)
- Level 4: Process Refinement
- Level 5: User Refinement (or Cognitive Refinement)

Level 1 is in charge of fusing filtered sensors data to determine the identity, position and speed of the entities.
Four functions can be distinguished:

- Data alignment: this task transforms the data received from the heterogeneous sensors into a common spatial and temporal frame.
- Data association: this task performs the correlation observations from the sensors into different groups. Related data is represented as a single distinct entity.
- Tracking: this task aims to determine an estimation of the position and speed of the targets through the multiple observations of positional data coherently for successive instants.
- Identification: this task classifies the objects originated by the measurements of their characteristics.

Feature-level fusion, referred to state-level fusion, involves the extraction of representative features from sensor data. In feature-level fusion, features are extracted from multiple sensor observations, and are combined into a single concatenated feature vector which is used as an input to pattern recognition approaches based on neural networks, fuzzy logic, clustering algorithms or template methods [20][25].

Modeling procedures involved in data fusion consist of association, estimation and identity declaration [20]. First of all, association determines which pairs of observations belong together, by representing observations of the same entity. Commonly used association measures include correlation coefficients, distance measures, association coefficients, and probabilistic similarity measures. Next, parameters of the fusion model are estimated using maximum likelihood estimator, least square estimator, or Kalman filter estimator. Finally, for identity declaration, typically one of the following three level of fusion is used: data-level fusion, feature-level fusion and decision-level fusion level [13].

IV. Four Levels, one architecture

Taking into account the State of the Art in Data Fusion, in this section we are going to illustrate the conceptual architecture on which we are going to develop the multi-agent system. This architecture is based upon the Steinberg approach [21] and it is described in details as follows:

Five are the main characteristics that a data fusion system should include:

- Capacity to detect and merge duplicate tracks
- Capacity to identify a single target by its features
- Capacity to identify groups of targets regarding their behaviours
- Capacity to assess the scenario and extrapolate future targets positions or actions.
- Possibility of indicating mitigation actions that can prevent enemy attacks from happening.

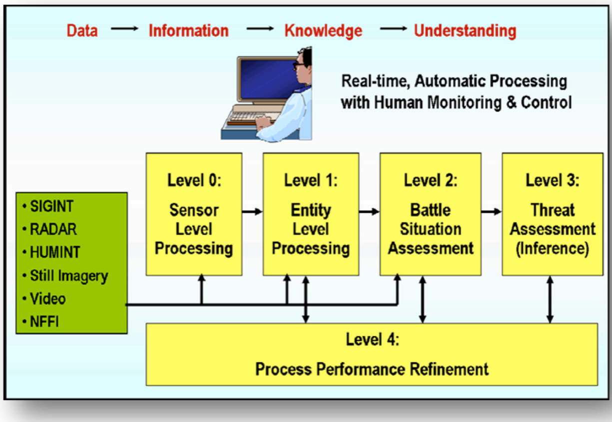

Fig. 2. Four levels, one architecture

The above figure shows the integration of the four different levels in data fusion in a single architecture.

A. Level 1 Data Fusion Submodule

At level 1 the following associations are made (in the case that several sources are available), recognition, identification and tracking of received tracks from different sources or associated tracks.

At this level the association of various data sources available is made. This level receives information about level 0 tracks and user settings and user corrections, this process of data fusion is made up of a collection of functions, these functions are recognition of tracks received from different sources, to obtain information from the merging of several tracks in the data fusion processes, the spatial and temporal alignment of tracks and the application of predictive algorithms to simulate the future position about an identified track

To summarize, Level 1 has functions to discriminate non valid information and corrupted signals, and provides the capability to generate recognized tracks and generate output information for other levels. The information generated from this level will be tracks with entity types information associated.

B. Level 2 Data Fusion Submodule

Level 2 data fusion performs clustering of single entities in order to create higher hierarchy entities. It also performs the recognition, identification and tracking of the higher level entities.

The level 2 sub module accepts the output provided by level 1 sub module and from other sources.

Level 2 data fusion carries out the clustering process of single entities in order to create aggregations. An aggregation is made up of various entities related, that generate a unique element with its own behavior. It performs the recognition, identification and tracking of the aggregations.

The system is able to automatically create clusters. The user can manually create new clusters or uncluster erroneous aggregations.

Level 2 is able to accept changes and new rules generated by the level 4 in order to optimize processes.

C. Level 3 Data Fusion Submodule

At level 3 possible future situations are identified, utility/cost assessment shall be made for each of the proposed actions.

Level 3 looks for a prioritized list of enemy potential targets. These targets are retrieved from tactical database. Once the objectives are listed, the system carries out the targets route prediction by using path finding algorithms.

D. Level 4 Data Fusion Submodule

Level 4 monitors operations from levels 1, 2 and 3 in order to generate statistical data, assess results and optimize processes.

Level 4 permits performing management, control, monitoring, optimization capabilities, statistical functions, management recording, report tools and parameter modifications tools of fusion process, in real time.

Level 4 submodule allows interaction with a user so as to perform queries on historical data, decision making, visualize/edit logs, configure optimization/monitoring, etc.

V. THE MULTI-AGENT COMMUNICATION MODEL

Current specifications and approaches to connectivity in Multi Agent Systems are discussed [18] and related to the needs of a tactical data fusion system. Results from an implementation of a Multi-agent system [10] are used to illustrate how using these theories can enhance the development and improved development of a C2IS (Command and Control Information System) [8].

We are researching AI techniques [17] suitable for providing connectivity among intelligent agents within C2IS and data fusion modules.

We have analyzed the multi-agent systems literature to identify how theories of agent's interaction can help provide a theoretical framework on which to base improvements to our system.

These theories have been successfully applied by Dominique Benech [5] to improve the MAS interaction by using CORBA specifications [22].

KQML [23] is, in fact, a formal specification for an Inter-agent language, to be placed at the application layer in the OSI schema. It makes the communication protocol independent of the contented message semantics.

Application
Presentation
Session
Transport
Network
Link
Physical

Fig. 3. OSI Layers

It is clear that this 'application layer' needs to be divided into other n sub-levels, where the KQML would occupy the lowest one, wrapping the upper levels knowledge in order to transfer it to the lower ones.

Agent
VKB[1]

[1] Virtual knowledge base

Kwowledge language (KIF)
KQML

Fig. 4. Application layer sub levels

KQML [9] doesn't describe a complete MAS architecture, in that sense, is, so to say, a wrapper for the exchanged knowledge, similar to a PDU at the application layer.

KQML head (interaction info)	Body (shared Knowledge)

Fig. 5. KQML Message

KQML establish an inter-agent speech mechanism, regardless of the knowledge that is to be exchanged.

This interaction protocol is not clearly defined, so we will use it for the syntax and semantics but for the interactions we need to use another specification (FIPA) [11], which is an evolution of KQML.

Besides, this specification rests in the upper levels for the ontology definition, which needs to be the same between the agents, so they have to be in the same domain's problem.

On the other hand, they also have to share the syntax to be able to understand that content.

It's interesting to take into account the study that has been accomplished, related to the different distributed management protocols, where CORBA is not simply a protocol but a complete architecture for distributed systems. CORBA [3] works as a middleware to offer an abstract view of the network to application objects. Thus, it also adds an intermediate layer and communications are slower than when using lower-level protocols'.

These extra complexities drive us not to use CORBA as a MAS support platform for multi-agent communications in C2IS.

TCP could be used as a transport protocol for the purpose of distributed management, but a higher-level protocol is necessary in our context to answer to the requirements we defined'.

So the Multi-agent system for Data fusion will require to organize the higher layers to provide the needed services, but keeping the simplicity and low-cost of TCP at the transport layer.

So we have to conclude that for the aim of COBALT, the CORBA specification is suitable as a management and transport protocol, but the mentioned drawbacks make us to decide for using sockets TCP/IP in C2IS, passing to the higher levels the responsibility to manage the rest of the requirements.

This section has focused on the inter-agent communication protocol, in the following section the multi-agent model regarding operational functionalities is faced.

VI. THE MULTI-AGENT OPERATIONAL MODEL

In this section, we illustrate a new approach to data fusion based on a Multi-Agent System. The objective consists of the construction of a model that faces the problem of extracting information from data produced from different sources.

From an operational point of view, agents within a Multi-Agent model have to perform cognitive tasks [4] that facilitate to automate the complexity of the data fusion process [6].

The following figure illustrates the agents that a multi-agent model includes to elaborate data producing information and to make decisions with the supervision of the human operator.

The data fusion system is an open system, which means that it can receive tracks and locations from different sensors, such as radars, human observers, unmanned air vehicles, etc. The data received from heterogeneous resources need to be processed and stored following a specific format. The DB Storing Agent is in charge of collecting data from different sensors and applying the correct format to those data to store them in a database.

Each track and location is produced taking into account spatial and time coordinates regarding the specific location and identification source. All this tracks need to be aligned in terms of location and time. The Alignment Agent is in charge of doing such task.

Once tracks have been pre-processed, the Fusion Agent is responsible for merging different tracks in order to examine if tracks produced from different acquisition sources are representing the same target.

The behavior of individual tracks have to be analyzed to check if they belong to group, so the Classifier agent is in charge of clustering tracks and analyzing their behaviors through time.

The Predictive Agent has the mission of figuring out the location, manoeuvre and intentions of a cluster of tracks in a near future.

Since the Data Fusion System works following the Intelligence Cycle, in every cycle the system will obtain an output that can be compared to the real scenario, consequently the system output can be validated. Possible errors in alignment, fusion, classification or prediction can be treated by the Optimizer agent with the idea of improving the performance of the Agents responsible for such tasks.

Finally, the Intelligent Planning Agent will give mitigation strategies or recommendations to thwart enemy's manoeuvre.

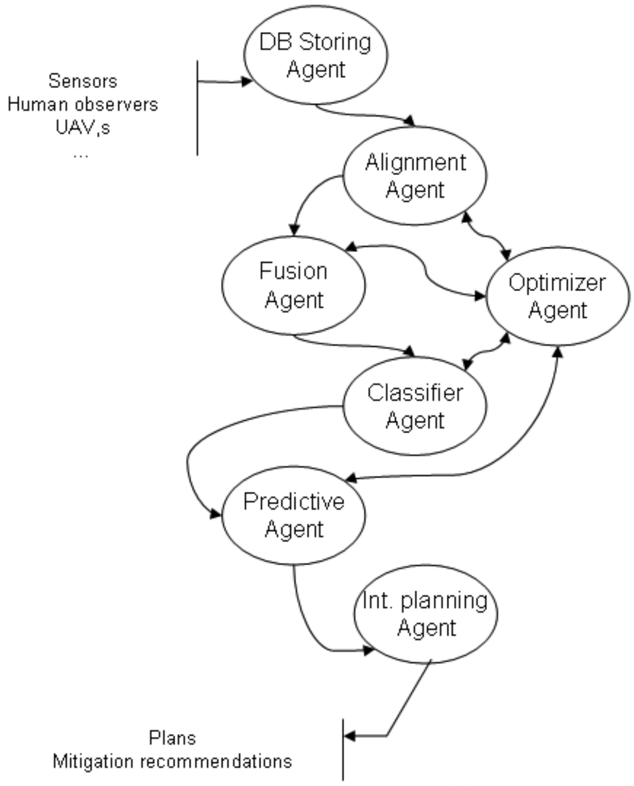

Fig. 6. The Multi-Agent operational model

A. Tools to build the Agents

Each cognitive task to be developed by every agent needs a specific Artificial Intelligence procedure [7] [19]. Some of them are described in the following paragraphs.

DB Storing Agent: The agent receives data from different sources and formats, and yields records in a formatted database. The core of this agent is based on mapping inputs with a structured format.

Alignment Agent: This agent has been developed by using proved algorithms like Kalman filter.

Fusion Agent: This agent uses logical rules to analyze if the location and time coordinates from different sources are related to the same track.

Classifier agent: Neural networks are used to develop the Classifier agent. A Multilayer Perceptron works efficiently in this cognitive task.

Predictive Agent [24]: In order to extrapolate future locations and manoeuvres, historical data of the same track or cluster and a pathfinder analysis on a GIS are required.

Intelligent planning Agent: To develop the Intelligent planning Agent we are using intelligent searches based on Artificial intelligence techniques as well as Neural networks to classify the importance of targets.

Optimizer agent: The Optimizer Agent is being developed by validating historical data and the user's decisions compared to the real scenario. The system will carry out automatically all Data Fusion Process taking into account former user's decisions.

VII. Future Works

In this article we have presented a solution based on agents to develop an architecture that supports the solution for data fusion problems in tactical operations.

The agents-based architecture is designed not only for being implemented and used in Command and Control centres, but also it can be used in the training field in order to improve operators' skills when dealing with operational data fusion problems. In the near future, we have to work in depth in the training area, by implementing new agents that facilitate such tasks.

Presently we are tackling the study of the application of data fusion on other domains.

In order to validate the architecture and new approach showed in this article, in near future we are going to develop some software prototypes that individually demonstrate the feasibility of this solution in each data fusion phase.

VIII. Conclusion

One of the most important advantages that this work offers is an agent-based approach to solve data fusion problems in tactical operations.

Computers have a larger capacity to process data than humans have. When receiving data in headquarters or HOMELAND security operational centers, the processing time for a person will be larger than for the computer. For that reason, computers play a vital role in this issue.

Humans cannot easily detect redundancies in a large amount of data they are dealing with; a computer is able to do it.

A Data Fusion module is going to receive large amounts of tracks, which have to be aligned, merged and recognized almost in real time in order to proceed with the Intelligence cycle.

The lack of time to process large amount of data and the use of complicated reasoning rules raise the probability of human errors. The use of a multi-agent solution for the data fusion module avoids this kind of problem.

References

[1] A modern approach to Distributed Artificial Intelligence. Gerhard Weiss. The MIT Press. Cambridge, Massachusetts. 1999.

[2] Aarsten, A., Brugali, D. y Vlad, C. (1996). Cooperating Among Autonomous Agents, Proceedings of the 4th International Conference on Control, Automation, Robotics, and Vision. Singapore

[3] ACL Message Representation in Bit-Efficient Encoding Specification XC00069C 29 august 2000. Approved for Experimental

[4] Agre y Rosenschein (1996). Computational Theories of Interaction and Agency, MIT Press,

[5] Benech, D.; (1999). Interaction Frameworks for Distributed and Cooperative Paradigm of Intelligent Systems and Network Management.. University Paul Sabatier, Toulouse

[6] Castillo, J.M. (2009); Una solución a la planificación de operaciones para la defensa basada en agentes inteligentes. Book. ISBN 978-84-9781-473-7

[7] Castillo, J.M., et al., (2006); Strategic Planning: A new approach through MECIMPLAN. Proc. of the IAT Int. Conf. Honk-Kong (China).

[8] David S. Alberts et al, Understanding the Information Age Warfare.

[9] DRAFT Specification del KQML Agent-Communication Language.Tim Finin y otros. 1993

[10] Durfee, E. ; Cox, J. et ali. (2001). Integrating Multiagent Coordination with Reactive Plan Execution. Proceedings of the ACM Conference on Autonomous Agents (Agents-01), pages 149-150, June.

[11] FIPA Abstract Architecture Specification. Code XC00001I. 29 january 2001. Approved for Experimental

[12] Gad, A., & Farooq, M. (2002). Data Fusion Architecture for Maritime Surveillance. ISIF

[13] Hall, D. L., Llinnas, J. (1997). An introduction to multisensory data fusion. Proceedings of IEEE 85 (1), 6-23.

[14] Haykin, Simon, (1999). Neural networks. A comprehensive foundation. Prentice Hall.

[15] Keller, J. (2008, November 1). Military & Aerospace Electronics. Retrieved January 23, 2012, from Joining sensors through data fusion

[16] Much, Richard et ali (1999). Intelligent Software Agents. Prentice Hall

[17] Nilsson, Nils J., (1998). Artificial Intelligence: A new synthesis. Mc-Graw Hill.

[18] Riecken, D., (1994). An architecture of integrated agents. Communications of the ACM, 37(7):107-116

[19] Russell N., (2003). Artificial Intelligence: A modern approach. Prentice- Hall.

[20] So Young Sohn, S. H. (2003). Data fusion, ensemble and clustering to improve the classification accuracy for the severity of road traffic accidents in Korea. Safety Science (41), pp. 1-14.

[21] Steinberg, A. N., Bowman, C. L., White, F. E. (1999). Revisions to the JDL Data Fusion Model. Arlington, VA, USA: ERIM International, Inc.

[22] The CORBA Reference Guide. Alan Pope. Addison Wesley. 1998.

[23] TR CS-97-03 A proposal for a new KQML Specification de Yannis Labrou and Tim Finin. Maryland, Baltimore. 1997.

[24] Castillo JM. A crystal ball made of agents. International Journal of Artificial Intelligence and Interactive Multimedia. 2010; Vol. 1, Num. 3, pp:13-17; http://dx.doi.org/10.9781/ijimai.2010.132

[25] Medina-García, V.H.,Bolaños-Castro, S., González-Crespo, R., (2012) Process Management Software as a Script, and the Script as a Pattern, International Journal of Computer and Communication Engineering, Vol. 1, Num. 2, pp. 147-150

José Miguel Castillo Chamorro is the CEO of SoftCast Consulting. From 2011 to 2013 he was the Director of the Area of Information Systems, training and decision making support in 'Tecnalia Research and Innovation'. From 2007 to 2011 he was Director of Division in European Virtual Engineering (Tech. Center). In 2001 he earned his PhD in Telecommunications from the Universidad Politécnica in Madrid. The same year he was awarded with the 'General Fernández Chicarro' prize by the Spanish Ministry of Defence for his work on Operations Research.

He has a large experience leading projects in which simulation, artificial intelligence and project management are involved. He is Associate Professor at the Universidad Pontificia de Salamanca in Madrid. In 2007 he obtained his PhD in the Statistics and Operational Research department from the Universidad Rey Juan Carlos in Madrid. In June 2007 he was awarded with the prize on Research by the Spanish Ministry of Defence. In the same year he was awarded with the prize on Research on the field of Security by the Directorate of Civil Protection and Emergencies (Spanish Ministry of Interior).

Rafael de Solís Montes is the CEO of Proytecsa (www.proytecsa.net). He is Lt.Col SP Army (Ret) with a large experience n C4ISR system in Spain and NATO. His degrees include Master of Science and "Research Aptitude" degree both in Artificial Intelligence and Computational Mathematics, he also possesses a Master of Science in Computer Science. He received post-graduate diplomas in CIS, High Logistics and Cryptology, and has 3 yrs of studies in Economics Science. Between his military courses, parachutist, night combat, PIO can be highlighted. He is completing his PhD thesis on Artificial Intelligence and Computational Mathematics about "Multiagent Systems and Data Fusion techniques in support of the interoperability between information systems" at the UPM University of Madrid. His research interests include Multi-agents Systems, Data Warehousing and co-operation techniques.