Information Fusion and Information Markets in Multi-Agent Systems

Johan L. Perols and Manish Agrawal
Information Systems/Decision Sciences, College of Business Administration, University of South Florida, 4202 East Fowler Avenue, CIS 1040, Tampa, Florida 33620-7800
jperols@coba.usf.edu and magrawal@coba.usf.edu

In distributed problem solving (DPS) the effective and efficient allocation of tasks and resources, and the subsequent fusion of results is of primary importance. Our research objective is to examine the process of information fusion (IF) in multi-agent systems (MAS) and suggest extensions to the MAS paradigm borrowing from IF research. We introduce a novel IF MAS technique (MAS-IF), Information Markets (IM), and show in a simulation that IM outperforms majority (MAJ), average (AVG), and weighted average (WAVG). We additionally argue that IM provides a natural fit with MAS primarily because it allows the various agents to remain autonomous and because it integrates well with a number of MAS coordination techniques, such as contracting, market places, and hierarchical organizations.
(*Multi-Agent System; Software Agents; Information Fusion; Information Markets; Fraud Detection*)

1. Introduction

In DPS, tasks and resources to formulate solutions to sub-problems must be coordinated and the solutions to the sub-problems must be synthesized (Durfee 2001). While MAS are a natural choice for solving distributed problems (Jennings et. al. 1998) MAS researchers have primarily focused on the problem of coordinating tasks and resources, and not on the fusion of data, information and decisions. However, coordination only addresses the first step towards resolving distributed problems, the task results must also be merged before the solution materializes (Durfee 2001).

This prompted us to examine the process of IF in MAS in further detail and suggest extensions to the MAS paradigm borrowing from IF research. We introduce IM as a novel MAS-IF technique that we show is relatively effective compared to AVG, MAJ and WAVG. We furthermore argue that IM provides a natural fit with MAS as it allows the various agents to remain autonomous and also naturally integrates with several MAS coordination techniques such as contracting and market places.

Our research objective is to examine how we can effectively fuse information in MAS. More specifically, our research questions are: (1) how well does different IF decision combiner methods fit in a MAS context; (2) how well does IM perform compared to other IF methods; and (3) does the effectiveness depend on the Event Probability Level and/or Number of Agents?

2. Background

2.1. Multi-Agent Coordination and the Necessity to Fuse Information

In DPS, and more specifically in MAS, coordination of individual autonomous agents is of central importance (Jennings 1996). Due to its importance, coordination has been extensively researched and various types of coordination mechanisms have been proposed and implemented, for example hierarchical organizations, contracting, negotiation, and market places (Nwana et. al. 1996). The coordination mechanisms discussed in the MAS literature are primarily concerned with how to allocate tasks and resource amongst the agents to tackle distributed problems in an effective and efficient manner.

Despite, or perhaps because of the focus on coordination research, the problem of data, information and decision fusion has, to our knowledge, largely been neglected in the MAS literature. Like coordination, IF is a central tenet in DPS and is a research stream in its own right. In a distributed problem the allocation of resources and completion of tasks are only the initial steps towards resolving the problem. The results of the tasks must be merged before the solution materializes (Durfee, et al. 1989; and Durfee 2001). Although MAS literature occasionally recognizes the importance of IF (Sycara, et al. 1996; and Sycara 2002) it does not describe specifically how the fusion is accomplished.

Because of the importance and lack of research of IF in MAS we believe that there is a need to further explore IF concepts in software agent research. We view IF as the flip-side of the coordination problem where coordination is used to accomplish the information generating tasks efficiently and IF is used to effectively combine the generated information. To illustrate what we mean by IF, and to further provide support for why this is a MAS problem we next describe a MAS for employee fraud detection (MAS-EFD) that we are currently developing. See Figure 1 for a conceptual overview of the MAS-EFD. In the MAS-EFD we have sensing agents that sense changes in the information system, i.e. they monitor the system for events such as financial transaction, employee record maintenance, vendor updates, etc. If the sensing agents believe that the event data indicates an anomaly they raise a red flag. The red flags, along with event details, are grouped together by the aggregation agents at various levels of analysis. The data is then analyzed by decision agents that make recommendations based on their perception of the likelihood that the specific object is fraudulent. The assessment agent then fuses these recommendations and makes an overall fraud assessment for the object and reports this assessment to the system user.

The allocation and execution of the tasks described above can be accomplished using established MAS coordination mechanisms. The process of combining the sensor data, the object information and the recommendation decisions is however not supported by these coordination mechanisms, at least not in the sense of IF. It is this fusion process we focus on in this article and more specifically the fusion of decisions.

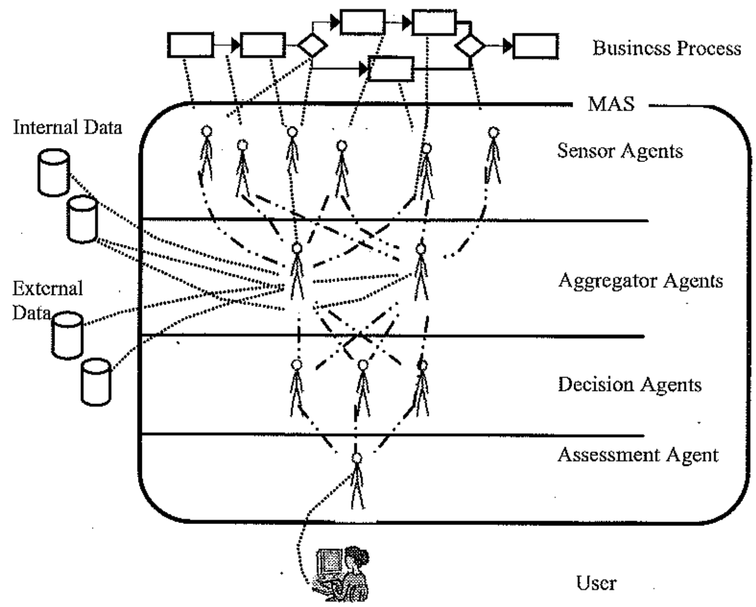

Figure 1: MAS-EFD Conceptualization

2.2. Information Fusion

IF is the process of producing estimates and knowledge about objects and situations based on data coming from various sources (White 1987). The input sources are commonly sensors that capture a certain perspective of objects of interest. The idea is that it is possible to get a more complete picture of the objects and situations if data from many sensors is combined. The most commonly researched problem domains include emergency response (natural catastrophe, terrorist attacks, etc), robotics, network security (intrusion detection), geoscience and defense systems (Valet, et al. 2001).

The fusion process is commonly described using the JDL model, developed by the Joint Directories Research Laboratories of the US Air Force. Figure 2 is a simplified version of the model. At Level 0, sensors capture new data and process this data so that it can be used by the system. In Level 1, the data fusion process synthesizes the sensor data into object information (commonly state information). Information about multiple objects is fused in Level 2. In Level 3 an impact assessment is made and reported to the system user (Steinberg, et al. 1999; and Salerna 2002). In this article we focus specifically on decision fusion.

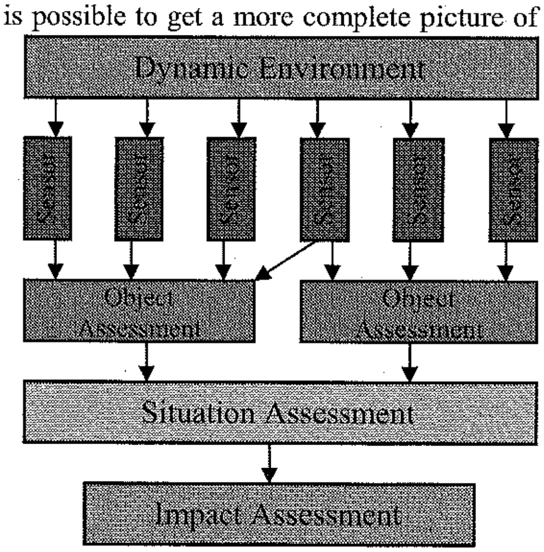

Figure 2: Simplified JDL Model

In MAS-EFD, the decision fusion is actually a combination of multiple classifiers, which is a research stream within the larger IF domain. This research is concerned with determining which decision makers to create and use for a given problem, and how to combine these experts' decisions. The idea is that decision makers have different strengths and weaknesses and that these differences can be exploited by including appropriate decision makers in the ensemble of experts and by using an appropriate decision fusion method to combine the decion makers' outputs (Bahler and Navarro 2000; and Jain, et al., 2000). In this article we specifically examine the method of fusing decisions and not how to select the appropriate decision makers, which we view as a coordination issue.

When reviewing combination methods, it is helpful to categorize the methods according to the homogeneity of the underlying classifiers. Homogenous classifiers ensembles are built using the same learning paradigm, while heterogonous classifiers use different techniques (Bahler and Navarro 2000). We are interested in combining decisions from multiple heterogeneous classifiers and as such we focus only on heterogeneous research. The different heterogeneous classifiers can create their own perception of the state space leading to classifier diversity by focusing on different regions of the state space, utilizing different classification techniques, using different learning algorithms of the specific classification techniques, using different training data, etc. Heterogeneous classifier combination methods are commonly further categorized based on the data type of the classifiers' output and if the method requires training (Shipp and Kuncheva 2002). Table 1 lists common heterogeneous combination methods according to output and learning requirements. Prior research has found, somewhat surprisingly, that MAJ, performs almost at the same level as trained methods that use measurement data, while another basic method, AVG, often outperforms these methods (Bahler and Navarro 2000).

The existing combiner methods, in particular the more complex methods, do not provide a good fit with MAS. They are centralized methods and are not designed to handle changes in the MAS. For example, what happens to methods that require training when decision agents leave or enter the MAS world, decide to focus on a different set of information, or learn? Existing combiner methods are furthermore not designed to integrate with other MAS infrastructure components, for example MAS coordination. We do not believe that existing combination methods provide support for these and similar issues, and as such they do not integrate well with MAS. We therefore propose IM, a novel combination method in both MAS and IF contexts that we argue integrate well with MAS.

		Classifier Output	
		Class Label	Measurment
Training Required	No	Majority Vote (MV)	Max; Min; Sum/Average; Median; Product; Bayesian
	Yes	Weighted MV; Bayesian; Dempster Shafer; Behavior Knowledge Space	Neural Networks; Logistic Regression; Descion Templates

Table 1: Combination Methods

2.3. Information Markets

IM are low volume markets designed specifically for the purpose of information aggregation. Equilibrium prices derived using conventional market mechanisms provide information based on the private and public information maintained by the market participants about a specific situation, future event or object of interest (Plott 2000; and Berg and Rietz 2003). In MAS-EFD the private information is represented by classifier diversity. The power of markets to aggregate information is discussed in prior research. For example, (Roll 1984) shows that juice futures can be used to make weather forecasts, (Berg et al. 2003) discuss markets predicting future presidential election that have outperformed major opinion polls and (Pennock, et al. 2001) present markets predicting Oscar, Emmy and Grammy winners.

In the MAS-EFD the agents trade in the IM using an instrument negotiable within the MAS (i.e. MAS dollars), that can for example be used in coordination. The more successful agents will have more capital and can therefore have a larger influence on resource usage than the less successful agents can. In IM, the agents can maintain their autonomy as without centralized control new agents can be added and removed dynamically from the MAS and the system can dynamically adjust to changes in agent accuracy (changes in an agent's accuracy is reflected in the agent's trading success).

3. Experimental Design

3.1. Experiment overview

We used a 4X3X3 factorial design to evaluate the performance of IM against three other combination methods, AVG, MAJ and WAVG, and to assess whether performance differences are sensitive to the Number of Agents and Event Probability. Rather than using actual fraud data, simulation was used to allow the manipulation of these factors and to control parameters such as diversity and distribution of fraud probability scores, through random assignment.

The experiment was performed using Swarm, in which 10,000 assessment objects were created for all 30 replications in all 36 treatment groups. Each object was randomly classified as either fraud or non-fraud, but the overall fraud probability was controlled using the Event Probability factor (0.01, 0.1 and 1%). Each object was also assigned a raw fraud probability score. While the score was highly skewed to the left (closer to 0) when the object was actually fraudulent it was more likely that the randomly generated raw fraud probability score was higher than when it is was not fraudulent. Measurement level classifier output (fraud probability decisions) was then simulated for each object in the form of fraud probability decisions and the number of decisions per object was manipulated using the Number of Agents factor (5, 15 and 50).

Based on prior research indicating the importance of classifier diversity for overall ensemble performance (Brown et al. 2005), we include two Swarm parameters to create diversity, the knowledge and chance parameters simulating the agent's intrinsic capabilities and random error, respectively. The knowledge parameter was randomly assigned to each decision agent and always improved the classifier's decision. The chance parameter was randomly assigned to each agent for each object and could improve or worsen each decision. The classifier output was then fused using IM, AVG, WAVG and MAJ, and Specificity (SP), Sensitivity (SE), False Negative (FN), False Positive (FP) and Hit-Rate (HR) were measured for each treatment group.

3.2. Fusion Methods

MAJ and AVG can be considered to provide the minimum base-level performance that other combiner methods can be gauged against, although as stated earlier, these methods have performed surprisingly well in prior research. We implemented a WAVG where the weight placed on each classifiers recommendation is determined by the classifier's historic hit-ratio as compared to the other classifiers' hit-ratios. This implementation does not require training as the hit-ratios are dynamically updated, thus allowing for comparison with MAJ, AVG and IM that also do not require training.

We implemented a simple IM based on gambling. In our betting algorithm (figure 3) weights are established in a decentralized fashion by each agent keeping track of their own funds that are based on bets placed and/or won. The size of each bet depends on the agent's current funds and the expected utility of the bet, which increases the more the agent's fraud assessment diverts from the house's assessment. Because weights are assigned dynamically, agents do not require training and the weights are redistributed if the environment changes. Notice that fraud outcomes are only known for events that were investigated and as such the IM and WAVG weights are only updated when investigation the system recommends.

We also tested a more complex IM betting algorithm where the average probability was determined based on the specific odds that created an equilibrium (bets for = bets against event). This more complex IM however proved to be less efficient while at the same time not being more effective.

I. Get p_{ij}, r_i, and f_i
II. If for at least on member of I p_{ij} > cut-off then
- a. Calculate number of agents in gamble: $n_j = \Sigma_i 1$
- b. Calculate average probability: $p_j = (\Sigma_i p_i) / n_j$
- c. Calculate agent i's bet on event j: $q_{ij} = (p_{ij} - p_j) f_i r_i$
- d. Calculate total funds: $F = \Sigma_i f_i$
- e. Calculate adjusted probability decision: $P_j = \Sigma_i (p_{ij} f_i / F)$
- f. If P_j > cut-off then
  - i. Calculate odds for event j: $O_j = 1/P_j - 1$
  - ii. Update holdings for agent i: $f_i = f_i - q_{ij}$
  - iii Update agent i's holdings if j=true: $f_i = f_i + (O_j + 1) q_{ij}$

where, for a specific object let p_{ij} be the probability assigned by agent i that the object is of event j; and r_i be the risk aversion and f_i the current holdings of agent i. A_j is the decision agent's knowledge of the moving average probability of event j.

Figure 3 - IM Betting Algorithm

4. Data Analysis

The ANOVA assumptions were verified using normal probability and residual plots as well as relying on the robustness of ANOVA when the design is balanced and the treatment group size is relatively large.

Five complete models, one for each dependent variable, were constructed and overall significance tested using a global f-test ($p<0.0001$). At a significance cutoff of 0.0036 ($\alpha=0.05$ allowing us to examine 14 single p-values) the 3-way interaction was not significant (lowest $p=0.9598$). After removing the 3-way interaction, the 2-way interactions involving method were examined. The interaction between method and fraud probability was significant for SP, FP and HR ($p<0.0001$, $p<0.0001$ and $p=0.0003$ respectively), but not for SE and FN (lowest $p=0.3844$). The interaction between method and Number of Agents was not significant in any model (lowest $p=0.0194$). After removing the insignificant 2-way interactions in the SE and FN models the Method main effect was examined and found to be significant ($p<0.0001$) in both models. To investigate the performance of IM relative to the other combiner methods we performed a Tukey-Kramer post-hoc analysis ($\alpha=0.05$) for the significant interactions (three models) and Tukey's post-hoc analysis ($\alpha=0.05$) for the significant main effects (two models). IM significantly outperformed AVG ($p<0.0001$), MAJ ($p<0.0001$) and WAVG ($p=0.0007$) in the main effects comparison of mean SE and FN differences (Figure 4) In the interaction comparison of SP, FP and HR means IM significantly outperformed AVG and MAJ at all fraud probability levels, and WAVG at a fraud probability level of 0.1 ($p<0.0001$). IM also performed better than WAVG at 0.001 and 0.01, but these differences were insignificant. For an example, see the SP means in Figure 5. To determine the overall SP, FP and HR performance differences between WAVG and IM we investigated the main effects and found that IM significantly ($p<0.0001$) outperformed WAVG (Figure 4).

We additionally examined if better performing agents accumulated more funds. Agent funds were positively related to HR and SE ($p<0.0001$), negatively related to FP ($p<0.0001$) and FN ($p=0.0016$) and insignificant for SP ($p=0.3237$).

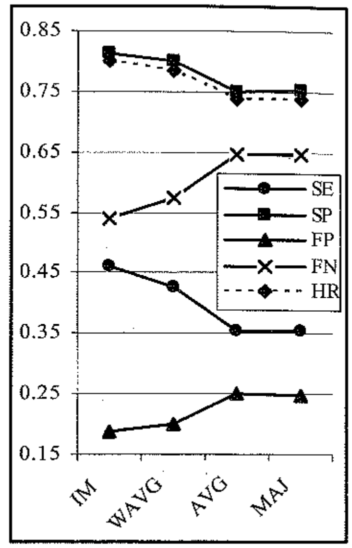

Figure 4: Method Means

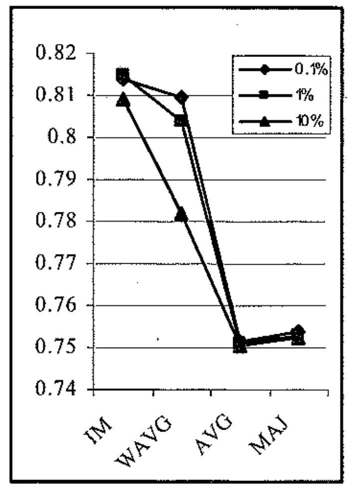

Figure 5: SP Means

5. Discussion and conclusions

The utility of IM as a decision combiner method is supported by the results from the follow-up analysis with parameter estimates showing that IM outperforms AVG and MAJ under all circumstances and either significantly or insignificantly outperforms WAVG. For SP, FP and HR the performance advantage over WAVG depends on the fraud probability level. The advantage is significant overall and at 0.1, but insignificant at 0.001 and 0.01. In regards to the combiner method fit with MAS, the regression results between funds and the performance measures show that better performing decision agents accumulate more funds. As discussed earlier, these additional funds can be used for MAS task and resource coordination allowing better performing agents to decide how to coordinate the MAS.

Through our discussions and experimental results, we make two distinct contributions. (1) We demonstrate the importance of IF in MAS and provide a first step towards bridging the two research streams. (2) We introduce IM, a novel fusion method, that we show is effective and fits well with MAS.

6. Limitations and Future Research

We used a fairly simple IM implementation and it is possible that other more sophisticated market mechanisms can provide additional performance benefits. Future research could attempt to model more complex trading behavior, for example marginal traders, and markets.

While the simulation allowed us to control certain factors and parameters, our results might not generalize to real environments with actual classifier agents making decision based on real data. Future research examining the generalizability of the results presented in this article would be beneficial.

We did not analyze the efficiency of the combiner methods. We however believe that the IM method would have no efficiency disadvantage compared to other complex combiner methods and only a minor disadvantage to the base combiners. We base this on the distributed character of the resource usage, which should improve scalability, the few and fairly simple computations and low memory overhead requirements, that not all events are analyzed, etc. Nevertheless, it is important that the efficiency of the methods is understood in order to make accurate scalability and cost-benefit assessments.

Finally, we only compared IM to three other combiner methods. Although these methods have performed relatively well in prior research, future research should compare the effectiveness and efficiency of IM to additional combiner methods.

7. References

Bahler D., and Navarro, L., *Methods for Combining Heterogeneous Sets of Classifiers*, 17th National Conference on Artificial Intelligence, 2000.

Berg, J., Nelson, F., and Rietz, .T., *Accuracy and forecast standard error of prediction markets*, Working Paper, University of Iowa, 2003.

Berg, J., and Rietz, T., *Prediction Markets as Decision Support Systems*, Inf. Sys. Frontiers, 2003.

Brown, G., Wyatt, J., Harris, R., and Yao, X., *Diversity creation methods: a survey and categorization*, Information Fusion, 6, 2005.

Durfee, H.E., *Distributed problem solving and planning*, Multi-agent systems and applications, 2001.

Durfee, H.E., Lesser, V.R., and Corkill, D.D., *Trends in Cooperative Distributed Problem Solving*, IEEE Transactions on Knowledge and Data Engineering, 1989.

Jain, K.A., Duin, P.W.R., and Mao, J., *Statistical pattern recognition: a review*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000.

Jennings, R.N., *Coordination Techniques for Distributed Artificial Intelligence*, In: Foundations of Distributed Artificial Intelligence, Wiley, 1996.

Jennings, R.N., Sycara, K., and Wooldridge, J.M., *A Roadmap of Agent Research and Development*, Autonomous Agents and Multi-Agent Systems, 1998.

Nwana, S.H., Lee, L., and Jennings, R.N., *Co-ordination in software agent systems*, BT Technology Journal, 14:4, 1996.

Oliven, K. and Rietz, A.T., *Suckers Are Born but Markets Are Made: Individual Rationality, Arbitrage, and Market Efficiency on an Electronic Futures Market*, Management Science, 50:3, 2004.

Pennock, D., Lawrence, S., Giles, L., and Nielsen, F., *The Real Power of Artificial Markets*, Science, 2001.

Plott, R.C., *Markets as Information Gathering Tools*, Southern Economic Journal, 2000.

Roll, R., *A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market*, The Journal of Finance, 1984.

Salerna, J., *Information Fusion: A High-level Architecture Overview*, Proceedings of the Fusion, 2002.

Shipp, A.C., and Kuncheva, I.L., *Relationships between combination methods and measures of diversity in combining classifiers*, Information Fusion, 3, 2002.

Steinberg, A.N., Bowman, C.L., and White, F.E., *Revisions to the JDL Data Fusion Model,* Proceedings of the SPIE - Sensor Fusion: Architectures, Algorithms, and Applications, 1999.

Sycara, K., *Agents supporting humans and organizations in open, dynamic environments*, Proceedings of the First international Joint Conference on Autonomous Agents and Multiagent Systems, 2002.

Sycara, K. et al., *Distributed Intelligent Agents*, IEEE Expert: Intel. Syst. and Their Applications, 1996.

Valet, L., Mauris, G., and Bolon, P., *A statistical overview of recent literature in information fusion*, Aerospace and Electronic Systems Magazine, IEEE, 16:3, 2001.

White, F., *Data fusion lexicon*, Joint Directors of Laboratories, Technical Panel for C 3, Data Fusion Sub-Panel, Naval Ocean Systems Center, San Diego, 1987.