AirQMAS: A Collaborative Multi-agent System for Air Quality Analysis

Mihaela Oprea*, Mădălina Cărbureanu,**
Elia Georgiana Dragomir***

**Department of Automatic Control, Computer Science and Electronics, Petroleum-Gas University of Ploieşti, Ploieşti, 100680, Romania (e-mail: mihaela@ upg-ploiesti.ro).*
*** Department of Automatic Control, Computer Science and Electronics, Petroleum-Gas University of Ploieşti, Ploieşti, 100680, Romania, (e-mail: mcarbureanu@upg-ploiesti.ro)*
**** Department of Information Technology, Mathematics and Physics, Petroleum-Gas University of Ploieşti, Ploieşti, 100680, Romania (e-mail: elia.dragomir@yahoo.com)*

Abstract: The paper presents a multi-agent system, AirQMAS, developed for air quality analysis at different stations from the Romanian national air quality monitoring network. The multi-agent system includes collaborative agents that work together in order to improve the overall efficiency of the system. Two types of agents are used: environmental agents and meteorological agents. The learning capability is provided to the environmental agents, that use rules generated by the C5.0 data mining algorithm for the analysis of air quality index. A first version of the system was implemented in Zeus, and was tested as a simulation for a local network of 6 stations from the Ploiesti town, analyzing the concentrations of some air pollutants.

Keywords: artificial intelligence, agents, collaborative systems, machine learning, environmental engineering, air pollution.

1. INTRODUCTION

The development of efficient environmental quality monitoring, analysis and control systems became one of the main challenges for scientists from various fields of research, environmental sciences, meteorology, chemical and petrochemical engineering, physics, chemistry, electronics, geology, automatic control, computer science, artificial intelligence etc. The technology of intelligent agents and multi-agent systems provides the software infrastructure for the implementation of geographical distributed environmental systems that make tasks like monitoring, analysis, and control of the environmental quality (Weiss, 1999). Moreover, the agent-based environmental systems can use the collaboration ability of the agents in order to solve better different environmental problems.

In the last period, several agent-based collaborative systems had been reported in the literature with application in domains like robotics, planning and scheduling in manufacturing systems, civil constructions, computer-aided design, e-learning, and business process management. A brief review of the literature emphasizes the intensive research work already done in this direction. Most of the industrial collaborative agent-based systems were developed for the manufacturing area (planning, scheduling, product design etc), as can be seen, for example, in Nishioka (2004), Hao et al. (2006), Kong et al. (2006), and Hernández et al. (2011).

The architecture of a collaborative multi-agent system (named CAPPS) for scheduling intensive production management was proposed by Nishioka (2004), for use in Japan. In Kong et al. (2006) it is described a new process planning method for assembly lines based on collaborative systems. An application from the domain of civil constructions is presented in Zhang and Hammad (2007), for the coordination of the construction equipment operation in real time by using collaborative agents. Another industrial application is provided by Hao et al. (2006), which proposed a collaborative e-Engineering environment for mechanical product design, based on software agents. The implementation of the collaborative intelligence ability in agent-based web applications are also reported in the literature. An example, is given by Xu and Wang (2002), which tackled the problem of multi-agent collaboration for B2B workflow monitoring. Some examples of environmental systems based on intelligent agents are given in (Doran, 2001), (Giannetti et al., 2005), (Nichita and Oprea, 2007), (Dragomir, 2011).

Collaborative agent-based systems can have better performance accuracy by using the learning capability. In this case, different machine learning techniques can be used (e.g. data mining, inductive learning, artificial neural networks). In Xiong et al. (2009) it is proposed the use of novel data mining hybrid techniques, combined with a biological approach.

An important issue that is necessary for the implementation of collaborative agent-based systems, is the ontology development. The conceptualization of the expertise domain of the collaborative system is given by the ontology of that domain, including concepts/terms and relations between concepts/terms that are used by agents during communication and social actions (Castelfranchi, 1998), (Mařík et al., 2001).

From the brief overview of some collaborative agent-based systems that were presented in this introduction, we can conclude that there are many research challenges in the development of such performant systems. In this paper, we shall focus on the use of collaborative intelligence and learning ability in the environmental systems for air quality analysis.

The paper is organized as follows. Section 2 discusses the main issues of agent-based environmental monitoring. The AirQMAS system is described in section 3. The architecture of the system, the ontology, the data mining module, and the main characteristics of the system (collaborative intelligence and learning) are presented. A case study of using the AirQMAS system is detailed in section 4. The last section concludes the paper.

2. AGENT-BASED ENVIRONMENTAL MONITORING

Several critical environmental problems occurred in the last years, with a high direct impact on the population and on the environment (e.g. floods; environmental pollution – air, water soil; tsunami; earthquakes; nuclear accidents etc). One of the main solutions that has to be adopted is to have a real time monitoring network, with stations in the most important locations. For example, in the case of air pollution, the network is composed by the monitoring stations placed in the sites were there are air pollution problems (e.g. sites with known air pollution sources such as chemical and petrochemical plants, power plants, heavy traffic etc). An environmental monitoring network has to check continuously the environmental quality standards (national and international), and if the maximum admissible levels are exceeded, to alert the environmental decision factors, and/or the population. Also, some control measures should be taken, if possible, in real time.

The environmental monitoring network can be modelled as a multi-agent system, that is composed by a group of intelligent agents, geographically distributed, in which each station has associated one or more software agents, that monitor and analyze the state of the environment quality in that location, and which inform the supervisor agents if some critical environmental problems can occur. An example of such environmental monitoring network structure based on intelligent agents is given in Figure 1. We have considered five networks for: air pollution monitoring (Air Stations Network - ASN), water pollution monitoring (Water Stations Network - WSN), soil pollution monitoring (Soil Stations Network - SSN), hydrologic monitoring (Hydrometric Stations Network - HSN), and meteorological monitoring (Meteorological Stations Network - MSN). Each network was modelled as a multi-agent system (MAS). The five multi-agents systems are MAS-1 for MSN, MAS-2 for ASN, MAS-3 for SSN, MAS-4 for WSN, and MAS-5 for HSN. As the environmental problems are very complex, and they need additional information from other parts of the environment, that have associated their specific monitoring and analysis network, the multi-agent system must interact with the agents or their resources (e.g. databases) from other parts of the environmental multi-agent monitoring and analysis systems. This interaction can be solved by collaborative intelligence. The meteorological network is involved in most of the environmental problems and need to interact with all the other four multi-agent systems.

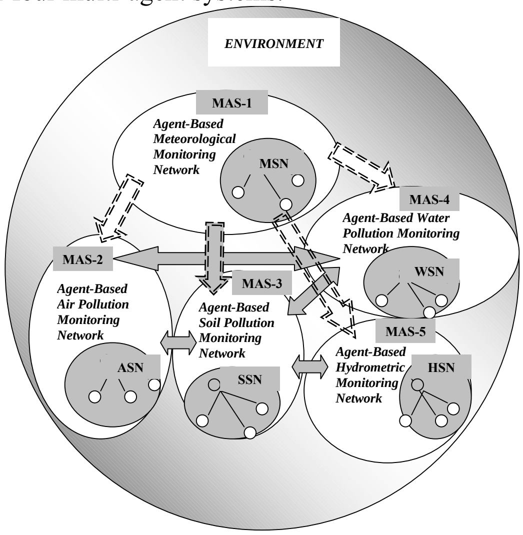


Fig. 1. Agent-based environmental monitoring systems.

Our previous work in this research area include the references: Oprea et. al (2009), a multi-agent system for dam monitoring, and Oprea et al. (2011), which proposes the generic structure of a collaborative agent-based environmental monitoring system, EnvMAS. The main advantages of using an intelligent agents-based solution are given by the reactivity and the pro-activity characteristics of the agents, as well as their social facility, the agent inter-communication in a specific agent language (e.g. FIPA ACL). Moreover, the multi-agent systems are a proper solution for distributed problems solving (Brenner et al., 1998).

3. DESCRIPTION OF THE AIRQMAS SYSTEM

Starting from the generic architecture of EnvMAS, that was introduced in Oprea et al. (2011), we have developed a collaborative multi-agent system for air quality analysis, AirQMAS. As in Romania it was implemented a first version of a National Network for Air Quality Monitoring (www.calitateaer.ro), that provides public measurements of some air pollutants in different locations, we have used the public available data for the AirQMAS system run as simulation.

We are presenting the architecture of the AirQMAS system, the ontology, the data mining module, and the characteristics of the system.

3.1 The architecture of the AirQMAS system

In figure 2 it is given the architecture of the AirQMAS system. As it can be seen two multi-agent systems from the EnvMAS generic architecture were used, MAS-1 (the multi-agent system for air quality monitoring and analysis), and MAS-2 (the multi-agent system for meteorological monitoring), that collaborate and share the same ontology, EnvMAS_Onto. Also, MAS-2 has a data mining module (DM) that is used by the environmental agents for knowledge extraction from the databases with measurements of air pollutants concentrations. Each multi-agent system has a supervisor agent (SA), that can be a global supervisor agent (GSA) or a local supervisor agent (LSA), in the case of MAS-2. The AirQMAS system has two main types of agents: environmental agents and meteorological agents. The environmental agents are the software agents that monitor and analyze the evolution of specific air pollutants concentrations in different locations. Examples of air pollutants that are monitored are the following: carbon monoxide (CO), sulphur dioxide (SO_2), nitrogen monoxide (NO), nitrogen dioxide (NO_2), ozone (O_3), and particulate matters (PM). The meteorological agents are the agents that provide information about all meteorological parameters that can influence the degree of environmental pollution: temperature, speed of wind, humidity etc.

Each agent has a list of tasks to perform according to the environmental problem that need to be solved at a given time. Some examples of tasks are: *Monitor_P, Analyze_P*, *Forecast_P*, *Envir_Status_Report*, *Initiate_Collaboration*, *Knowledge_Extraction*, *Control_Measure*.

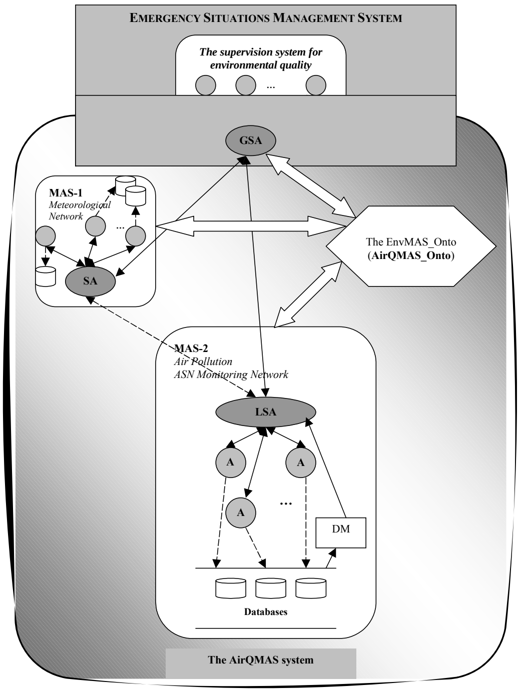


Fig. 2. The architecture of the AirQMAS system.

3.2 The AirQMAS_Onto ontology

The agents of the AirQMAS system share a common environmental ontology, EnvMAS Ontology, with general terms, and specific terms for each environmental problem. In our case, the terms specific to air pollution and meteorology are used. The ontology AirQMAS_Onto included in EnvMAS_Onto, was derived from the generic ontology AIR_POLLUTION_Onto, described in Oprea (2009). Examples of terms used by AirQMAS_Onto are as follows: air_pollutant, site_location, speed_of_wind, humidity, temperature, monitoring_station, station_report, air_quality_index, CO, NO, NO2, SO2, Ozone, PM10, PM, air_pollutant_concentration, control, time_interval, general_report, control_measure. The ontology provides a conceptualization of the domain of air quality monitoring and analysis.

3.3 The data mining module

The environmental agents use knowledge extracted under the form of production rules (if-then rules) from the databases with air pollutants concentrations by using a data mining technique, the C5.0 algorithm, an improved variant of ID3 and C4.5 (Quinlan, 1993). This technique is doing a greedy search through the space of possible decision trees. The details about the rule extraction experiment done for the AirQMAS system are given in the next section.

3.4 The characteristics of the AirQMAS system

The main characteristics of the AirQMAS system are the collaborative intelligence and the learning ability of the environmental agents. The collaborative intelligence is provided by the ability of agents to work jointly in order to solve a specific environmental problem. For example, in the case of air quality analysis problem, the environmental agents (local and supervisor) work jointly with the meteorological agents (local and supervisor), by sharing the same ontology and assisting each other when it is performed a general report on air quality. The learning ability is specific to the environmental agents, and is provided by the data mining module (DM), that allows rule extraction from the databases with air pollutant concentrations measured in a specific period of time, and a specific region. Through data mining, the agents are learning which air pollutants has higher influences on the air quality in a specific region.

In the next section, we will analyze the two characteristics for a particular case of air pollution analysis.

4. CASE STUDY

We have considered the case of a multi-agent system MAS-2 (shown in Figure 3), with six air quality monitoring stations from the Ploiesti town, and a supervisor agent. The AirQMAS system will provide the air quality index and the air quality general report. The air pollutants that were analyzed are CO, SO_2, PM_{10}, H_2S, NO, NO_2, O_3.

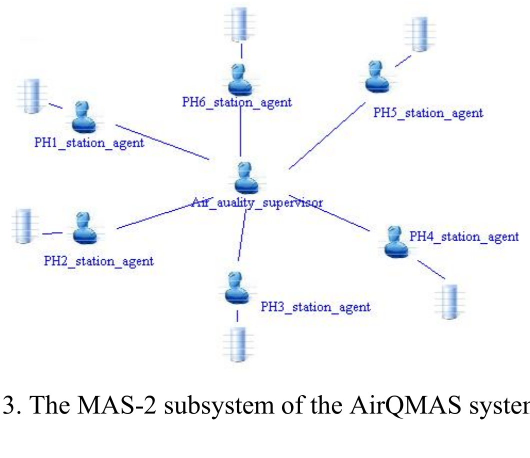


Fig. 3. The MAS-2 subsystem of the AirQMAS system.

The measurements of the air pollutant concentrations were stored in MySQL databases. The data were taken from the Romanian National Network for Air Quality Monitoring. We have collected the datasets from the years 2009, 2010, and 2011.

4.1 Rule extraction by using data mining C5.0 algorithm

According to the Romanian norms and standards for air quality, the air quality index has six values, excellent (1), very good (2), good (3), medium (4), bad (5), and very bad (6). These values are set for the concentrations of each air pollutant. The greatest value of all air pollutant index is set as the air quality index for a given monitoring station. At least three air pollutants index are necessary for each air monitoring station. We have used the See5 data mining tool (demonstration version for Windows), that implements the C5.0 algorithm. This tool extracts patterns from the databases and constructs classifiers under the form of decision trees or rule sets. Figure 4 shows the interface of the See5 system.

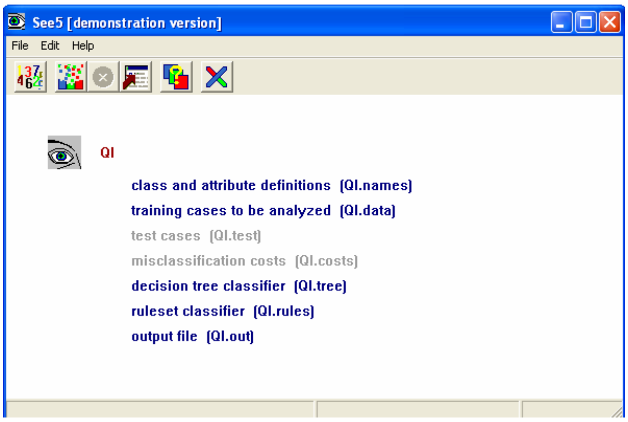
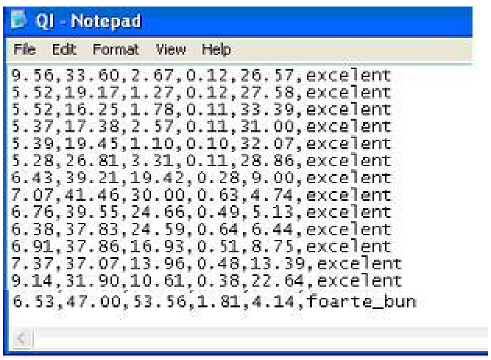


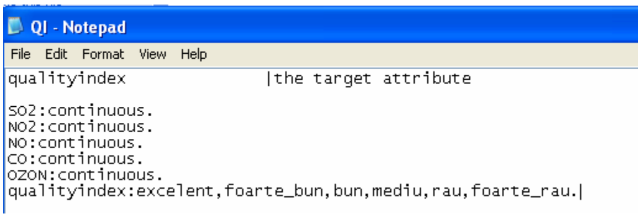
Fig. 4. The interface of the See5 data mining tool.

For the problem of air quality index (QI) analysis, the See5 data mining tool takes as input two files, one with the training data (QI.data), and one with the attributes and classes (QI.names). We have used in our experiment the dataset with air pollutant concentrations hourly measured during January 2011 for five air pollutants (SO_2, NO_2, NO, CO, O_3). The file QI.data contains 400 input data for all five air pollutants, and the air quality index. Figure 5 shows a part of the QI.data file content. Figure 6 shows the content of the QI.names file with the attributes and classes of the case study.



```
9.56,33.60,2.67,0.12,26.57,excelent
5.52,19.17,1.27,0.12,27.58,excelent
5.52,16.25,1.78,0.11,33.39,excelent
5.37,17.38,2.57,0.11,31.00,excelent
5.39,19.45,1.10,0.10,32.07,excelent
5.28,26.81,3.31,0.11,28.86,excelent
6.43,39.21,19.42,0.28,9.00,excelent
7.07,41.46,30.00,0.63,4.74,excelent
6.76,39.55,24.66,0.49,5.13,excelent
6.38,37.83,24.59,0.64,6.44,excelent
6.91,37.86,16.93,0.51,8.75,excelent
7.37,37.07,13.96,0.48,13.39,excelent
9.14,31.90,10.61,0.38,22.64,excelent
6.53,47.00,53.56,1.81,4.14,foarte_bun
```

Fig. 5. The training dataset QI.data (selection).

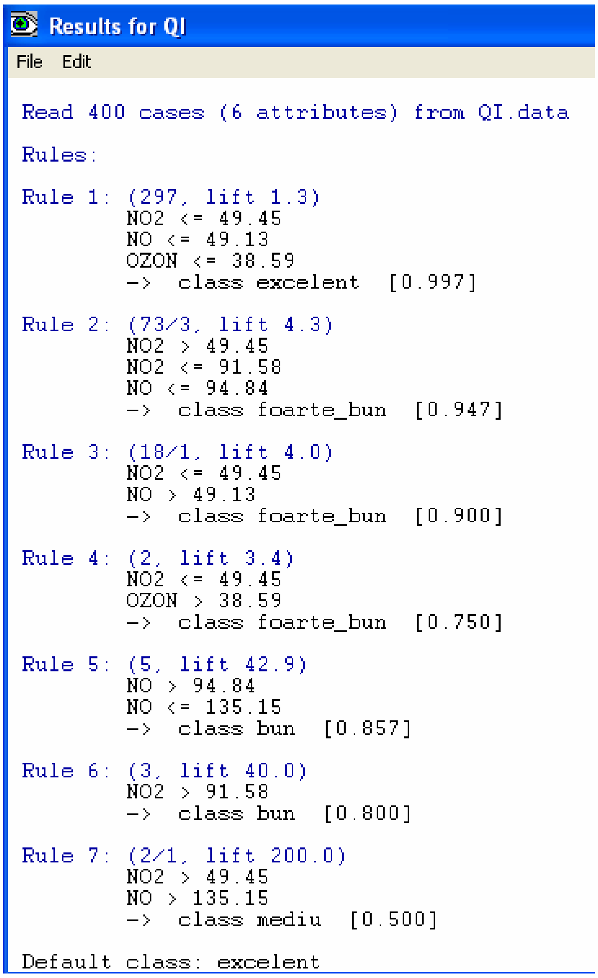


```
qualityindex               |the target attribute

SO2:continuous.
NO2:continuous.
NO:continuous.
CO:continuous.
OZON:continuous.
qualityindex:excelent,foarte_bun,bun,mediu,rau,foarte_rau.
```

Fig. 6. The QI.names file.

We have generated the patterns under the form of rules, which are better predictors, from the viewpoint of prediction accuracy. Figure 7 shows the rules set with seven rules, generated with See5.



```
Read 400 cases (6 attributes) from QI.data

Rules:

Rule 1: (297, lift 1.3)
	NO2 <= 49.45
	NO <= 49.13
	OZON <= 38.59
	-> class excelent  [0.997]

Rule 2: (73/3, lift 4.3)
	NO2 > 49.45
	NO2 <= 91.58
	NO <= 94.84
	-> class foarte_bun  [0.947]

Rule 3: (18/1, lift 4.0)
	NO2 <= 49.45
	NO > 49.13
	-> class foarte_bun  [0.900]

Rule 4: (2, lift 3.4)
	NO2 <= 49.45
	OZON > 38.59
	-> class foarte_bun  [0.750]

Rule 5: (5, lift 42.9)
	NO > 94.84
	NO <= 135.15
	-> class bun  [0.857]

Rule 6: (3, lift 40.0)
	NO2 > 91.58
	-> class bun  [0.800]

Rule 7: (2/1, lift 200.0)
	NO2 > 49.45
	NO > 135.15
	-> class mediu  [0.500]

Default class: excelent
```

Fig. 7. The rules set generated by See5 data mining tool.

For example, rule 3 can be written under the if-then form as follows:

IF NO2<=49.45 **AND** NO>49.13 **THEN**
qualityindex=foarte_bun; //very_good

Some statistical data (missclassified rules, associated errors, confusion matrix, attributes usage etc) are also provided by the See5 tool, as shown in figure 8.

The main conclusions of the statistical analysis are the following:

1) Five rules were missclassified and the associated error is 1.3%.

2) From the confusion matrix it is highlighted the C5.0 algorithm performance over the training data. For example, the C5.0 algorithm missclassified three of the excellent cases as being very good.

3) When we have a great number of attributes it is useful to know the contribution of each attribut to the classification. In our experiment, the greatest contribution to classification is provided by the attributes NO2 and NO (99%), and Ozone (75%).

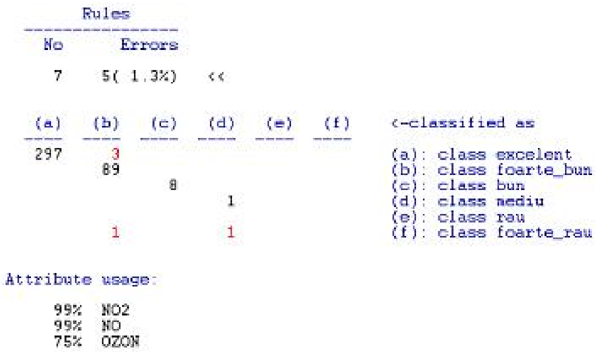


Fig. 8. The statistical report generated by See5 tool.

4.2 The AirQMAS system run

We have implemented the first version of the AirQMAS system in Zeus (Collins and Ndumu, 1999), a Java based toolkit for multi-agent systems development, that allow a quick realisation of a prototype multi-agent system, as it has incorporated an automated agent code generation tool. The system tackles the air pollution monitoring and analysis problem as a simulation, and makes air quality analysis in the Ploiesti town, based on the recorded data of concentrations for major air polutants (e.g. carbon monoxide, sulphur dioxide, hydrogen sulfide, particulate matter etc), that were taken from public available databases.

We have experimented four scenarios of the system run. In the first scenario (S1) there are six environmental agents, corresponding to six air quality monitoring stations, a meteorological agent (*Meteo_agent*) which delivers the weather conditions, a local supervisor agent, (*Air_agent*), that centralises the data from the stations agents, and a supervisor agent (*Supervisor_agent*) that receives the information needed for the experiment and analyzes the whole region air quality. The six environmental agents are scanning a database specific to each monitoring station in order to find the record from a certain day (which is an input from the user). These data are sent to the *Air_agent* to be analyzed. The *Air_agent* receives also the weather parameters for that day from the *Meteo_agent*. Based on those data, the *AirQualityStatus* is generated and sent to the *Supervisor_agent*. In the second scenario (S2), there are also six environmental agents for the stations, an agent that centralizes the data and the one which is the supervisor. The difference between the first two scenarios is that the *Meteo_agent* does not exist anymore and the stations agents are searching by themselves in the weather conditions databases for the needed informations. The third scenario (S3) is similar to the first one, except the fact that instead of six stations agents there are twelve stations agents involved in the system. The fourth scenario (S4) is the result of the combination between the second and third scenarios: there are twelve stations agents but there is no meteo agent. The environmental agents are doing the tasks *Monitor_P* and *Analyze_P*, by reading the data from databases and making multiple comparisions between the data suplied and the upper limit of the air pollutants' concentration, according to the air quality standard. The output is given as a general characterization of the environment (*no_problem*, if all the recorded values are normal, *weak_air_pollution*, if there are some possible exceedances, *severe_air_pollution*, if there are recorded values that may have an impact on the human health). If the upper limit is exceeded for a certain pollutant, an alert windows pops up. In the scenarios S1 and S3 the agents are collaborative (C), while in the scenarios S2 and S4 the agents are non-collaborative (NC). The response time of the system during each scenario run was registered. The results are presented in Table 1. As expected, the response time is smaller in the cases when collaborative agents were used.

Table 1. The AirQMAS system response time [seconds]

Experiment	S1 (C)	S3 (C)	S2 (NC)	S4 (NC)
First	**10.23**	**11.24**	13.44	15.20
Second	**10.57**	**11.50**	14.01	15.40

Figure 9 shows a screenshot from the AirQMAS system run for air quality index analysis at all six monitoring stations.

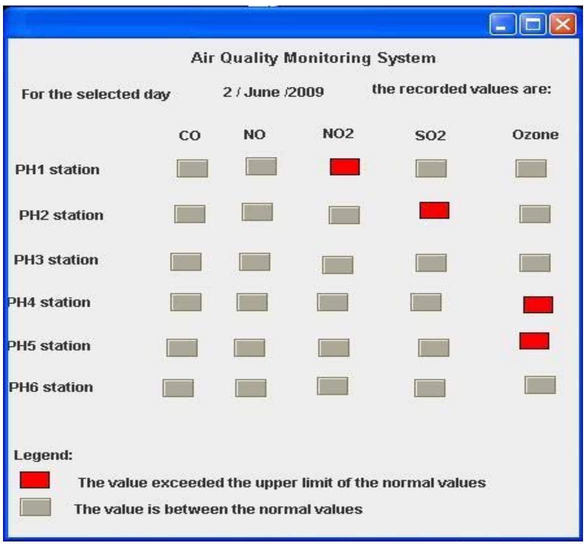


Fig. 9. Screenshot from the AirQMAS system run for air quality index analysis at each monitoring station.

Figure 10 shows a screenshot of the AirQMAS system run with the air quality general report. Other reports provided by the Visualiser utility agent from Zeus are given in Figure 11 (the Agents Society with the messages exchanged between agents) and Figure 12 (the task graph for the general report on air quality).

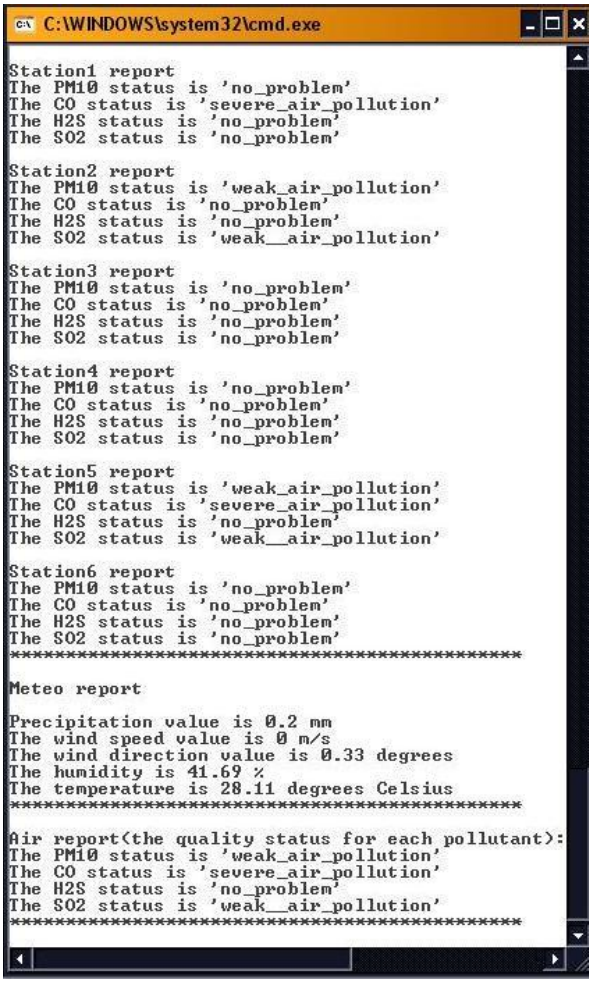


Fig. 10. Screenshot from the AirQMAS system run with the air quality general report.

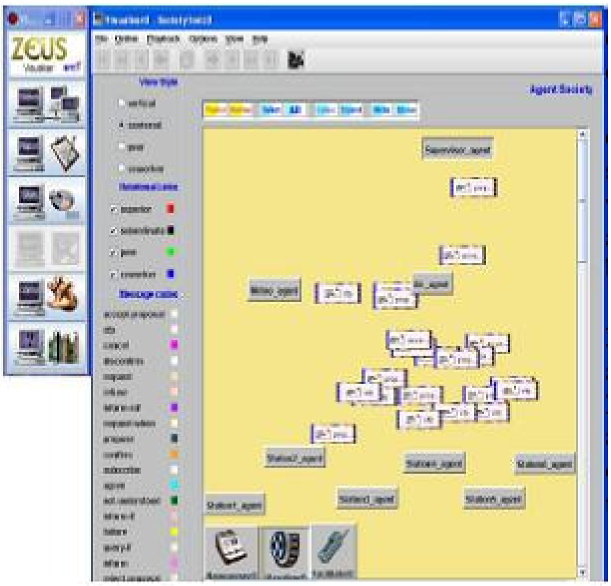


Fig. 11. The Agents Society of AirQMAS system in Zeus.

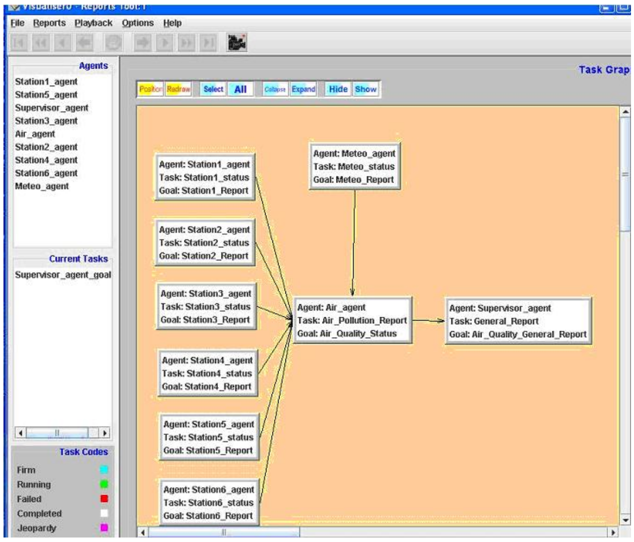


Fig. 12. Screenshot of the task graph obtained with the Visualiser agent during the AirQMAS system run.

In figure 12 it is presented the task graph for Scenario 1, with collaborative agents, when the *Meteo_agent*, the *Air_agent* and all the environmental agents associated to the air monitoring stations are collaborating in order to generate the Air_Pollution_Report that will be sent to the *Supervisor_agent*.

The learning ability of the environmental agents was used in all four scenarios that were experimented.

5. CONCLUSIONS

The paper presented a collaborative multi-agent system, AirQMAS, for air quality analysis in a specific geographical region (at local or national level). The agents are solving the air pollution problems by using collaborative intelligence and a learning capability. So far, the system was tested as a simulation. The main benefits of using such a multi-agent environmental system is the fast information provided to the population and to the environmental decision factors when an air pollution episode arise. The AirQMAS system can be integrated in the local or national emergency situation management system associated to the local/national environmental monitoring network.

As a future work we are interested to study the collaborative intelligence between different multi-agent systems associated to various environmental networks (air, water, soil, meteo, hydrology) in order to solve complex environmental problems, such as the analysis of the impact of a severe air pollution to surface water pollution and soil pollution in critical meteorological conditions, for a specific region.

REFERENCES

Brenner, W., Zarnekow, R. and Wittig, H. (1998). *Intelligent software agents, Foundations and Applications*, Springer-Verlag, Berlin.

Castelfranchi, C. (1998). Modelling social action for AI agents, *Artificial Intelligence*, 103(1-2), 157-182.

Collins, J. and Ndumu, D. (1999). *The Application Realisation Guide, The Zeus Agent Building Toolkit*, Intelligent Systems Research Group, BT Labs.

Doran, J. (2001). Agent-Based Modelling of Ecosystems for Sustainable Resource Management. In M. Luck et al. (eds), *ACAI 2001*, LNAI 2086, Springer, 383-403.

Dragomir, E. (2011). Development of a multi-agent system for air pollution analysis, *Doctoral Research Report*, Petroleum-Gas University of Ploiesti.

Giannetti, L., Maturana, F.P., and Discenzo, F.M. (2005). Agent-Based Control of a Municipal Water System. In M. Pěchoucěk, P. Petta, L. Z. Varga (eds), *Multi-Agent Systems and Applications IV - CEEMAS 2005*, LNAI 3690, Springer, 500-510.

Hao, Q., Shan, W., Zhang, Z., Park, S.-W. and Lee, J.,-K. (2006). Agent-based collaborative product design engineering: An industrial case study, *Computers in Industry*, 57(1), 26-38.

Hernández, J.E., Mula, J., Poler, R. and Pavón, J. (2011). A Multiagent Negotiation-based Model to Support the Collaborative Supply Chain Planning, *Studies in Informatics and Control*, 20(1), 43-54.

Kong, S.H., Noh, S.D., Han, Y.-G., Kim, G. and Lee, K.E. (2006). An agent-based collaborative assembly process planning system, *International Journal Advance Manufacturing Technology*, 28, 176-173.

Mařík, V., Pěchouček, M. and Štěpánková, O. (2001). Social knowledge in multiagent systems, *Multi-agent Systems and Applications - ACAI*, LNAI 2086, Springer, 211-245.

Nichita, C. and Oprea, M. (2007). An agent-based model for water quality control, *Proceedings of the Int. Conf. ESCAPE17*, Elsevier, 1217-1222.

Nishioka, J. (2004). Collaborative agents for production planning and scheduling (CAPPS): a challenge to develop a new software system architecture for manufacturing management in Japan, *International Journal of Production Research*, 42(17), 3355-3368.

Oprea, M, Matei, A. and Petre, E. (2009). Agent-based modeling of a dam monitoring system, *Proceedings of the 17th Int. Conf. on Control Systems and Computer Science – CSCS 17*, Politehnica Press, 509-514.

Oprea, M. (2009). AIR_POLLUTION_Onto: an ontology for air pollution analysis and control, *Artificial Intelligence Applications and Innovations III*, IFIP Advances in Information and Communication Technology, volume 296, 135-143.

Oprea, M., Dragomir, E., Cărbureanu, M. (2011). On the use of collaborative intelligence in an agent-based environmental monitoring and analysis system, *Proceedings of the ICSTCC 2011*, 450-455.

Quinlan, J.R. (1993). *C4.5: Programs for machine learning*, Morgan Kaufmann, San Mateo.

See5: www.rulequest.com/see5-win.html, accessed on December 2, 2011.

The Romanian National Air Quality Monitoring Network, www.calitateaer.ro

Weiss, G. (1999). *Multiagent systems – A modern approach to distributed artificial intelligence*, The MIT Press, Cambridge.

Xiong, Y., Zheng, G., Yang, Q. and Zhu, Y. (2009). A Collaborative Multiagent System for Mining Transcriptional Regulatory Elements, *IEEE Intelligent Systems*, 24(3), 26-37.

Xu, D. and Wang, H. (2002). Multi-agent collaboration for B2B workflow monitoring, *Knowledge Based Systems*, 15, 485-491.

Zhang, C. and Hammad, A. (2007). Toward Supporting Construction Equipment Operation Using Collaborative Agent-Based Systems, *Proceedings of the ASCE Int. Workshop on Computing in Civil Engineering*.