Iulia MĂRIEŞ, PhD Student
E-mail: iulia.maries@hotmail.com
Diana DEZSI, PhD Student
Department of Economic Cybernetics
The Bucharest Academy of Economic Studies

A GENETIC ALGORITHM FOR COMMUNITY FORMATION IN MULTI-AGENT SYSTEMS

Abstract. *Community formation has certainly gained more and more attention from both the researchers and practitioners in the fields of complex networks and multi-agent systems. Since the number of the possible communities is exponential in the number of agents, an efficient algorithm is needed. Genetic algorithms are very useful tools for obtaining high quality and optimal solutions for optimization problems, due to their self-organization, self-adaptation and high parallelism. The paper proposes a high performance genetic algorithm for community formation, the key concept in our algorithm being a new fitness index, which aims at being a trade-off between intelligence and cooperation, and allows not only community formation but also intelligence to be the driving principle in the community formation process.*
Key Words*: Genetic Algorithms, Collective Intelligence, Intelligence Index, Communities of Practice*

JEL Classification: C45, C63, C92

1. Introduction

Complex systems composed of different interacting subsystems tend to evolve towards more coherence and interdependence as the subsystems mutually adapt. The continuous evolution, adaptation, cooperation and negotiation determine an increasingly diverse, complex and efficient organization [3].

To successfully deal with organizational problems we need to develop collective intelligence as a global civilization. In the context of a global market, collective intelligence enhances competitiveness within organizations and collective performance becomes a critical factor in the organization's development.

Knowledge assets are critical resources that can generate competitive advantage for organizations. A new organizational form that complements existing structures has emerged, and is named the community of practice. Communities of practice are the core of the collective learning and collective intelligence processes and they rely on a constant exchange of knowledge and information between members. These communities can provide a social reservoir for practitioners and knowledge producers to analyze, address and explore new solutions to their problems.

The remainder of this paper is organized as follows: the evolution and the development of community of practice are presented in Section 2. In this framework, collective intelligence is defined in Section 3. Section 4 places the communities of practice in the context of multi-agent systems. In Section 5 we propose a genetic algorithm for community formation based on collective intelligence capacity. Finally, Section 6 comes up with conclusions, relevant implications and directions for future work.

2. The evolution and development of Communities of Practice

The concept "community of practice" was outlined by Lave and Wenger in early 90's, to describe "a group of people who share a concern, a set of problems or a passion about a topic, and who deepen their knowledge and expertise by interacting on an ongoing basis" [19].

In this context, communities of practice are free associations of persons who choose to improve their skills together. Furthermore, they develop new productive collective capabilities, which are sources of value creation in the knowledge-based economies. Through interactions between the members of communities of practice individual knowledge and experiences are shared, new knowledge is developed and problems are solved (Figure 1).

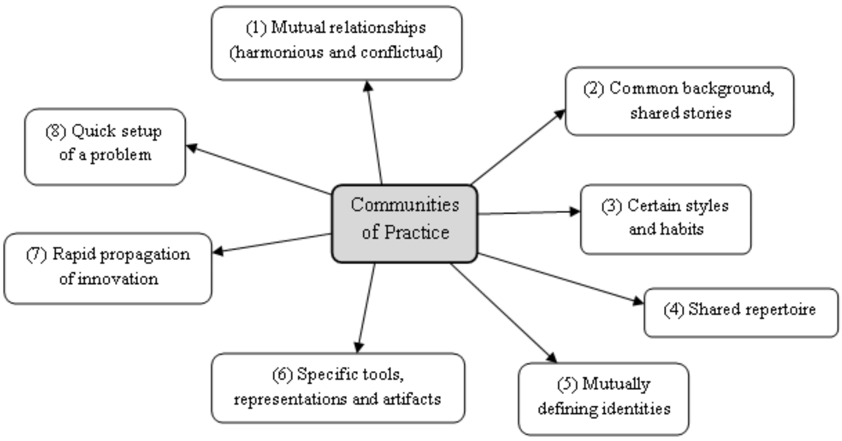

Figure 1. The key characteristics of communities of practice (Source: compiled from [19])

Communities of practice can be defined as "groups of people informally bound together by shared expertise and passion for a joint enterprise" [18]. In this respect, we can say that a community of practice is delimited by three dimensions:

- The shared repertoire of communal resources (routines, habits, artifacts, vocabulary, styles) that members have developed over time;
- The relationships of mutual engagement that bring members together into a social entity;
- The "joint enterprise", as it is understood and continually renegotiated by its members.

The success of a community depends on the way of interaction between community members, depends on communication, cooperation, coordination and knowledge exchange, but also depends on certain characteristics of the setting, characteristics of the individuals, characteristics of the community and even characteristics of the environment (Figure 2). These characteristics are not static; they can change continually, especially in the early phases of a community.

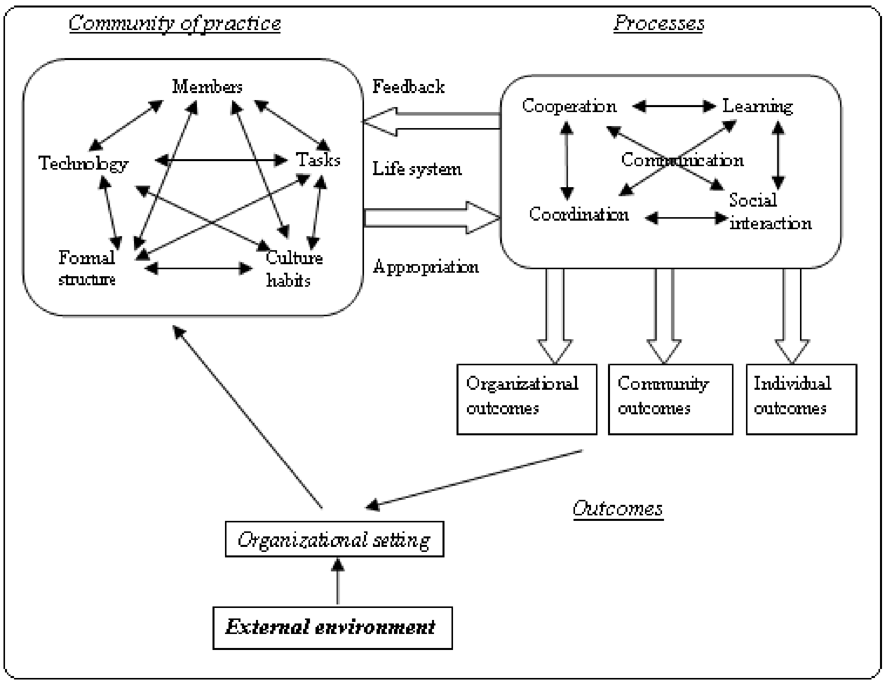

Figure 2. Dynamic Interaction Model (Source: Adapted from [1])

In such a dynamic environment, the previous characteristics represent conditions for the input and output of the community process. Therefore, a successful community depends on the degree to which the processes mutually match each other.

An effective organization includes a set of interconnected communities of practice, each of them dealing with specific aspects of the organization competencies. Knowledge is created and shared, organized and revised, and passed on within and among these communities. In this context, collective intelligence represents the capacity of human communities to enable their members to reach the highest potential and to co-evolve toward more complex integrations through collaboration and innovation in mutually supportive relationships.

The emergent behavior is not attributed to a single individual; it is a global result of coordination of individuals; it is the action of combining simple rules that produces complex results. The interactions between individuals seem to be the most difficult part to understand. A complex dynamic loop is established when individuals induce behaviors which affect other individuals and their behavior with effect on the initial individual. The complex feedback loop shows the only possible solution: an analysis of the emergent phenomena at system level.

As a result, collective intelligence is an emergent phenomenon; it is a synergistic combination of individuals which make the group more capable and intelligent than any individual member.

3. Collective Intelligence

Collective intelligence is a shared intelligence that emerges from the collaboration of individuals. Collective intelligence explores the collective behavior from the level of quarks to the level of bacterial, plant, animal and human societies. The researchers consider collective intelligence as a subfield of sociology, communication or behavior, computer science or cybernetics.

3.1 Intelligence

Intelligence is related to the complexity of tasks that we are capable of automating and delegating to computers and also to the goal or problem and to the previous knowledge and experience of the solver.

In our context, intelligence is defined as "the ability for attaining goals or for solving problems that puts at work responsiveness of the solver to the situation where the goal or problem arises and use of its previous knowledge and experience." [2]

3.2 Intelligence of Collectivities

"Collective intelligence is neither a new concept nor a discovery. It is what shapes social organizations – groups, tribes, companies, teams, governments, nations, societies, guilds, etc… – where individuals gather together to share and collaborate, and find an individual and collective advantage that is higher than if each participant had remained alone. Collective intelligence is what we term a positive-sum economy.” [13]

Thus, the presence of collective intelligence has been felt for a long time: families, companies and states are groups of individuals that at least sometimes act intelligent. There are groups of insects, such as bee and ant colonies, that are finding food sources acting intelligent. Also, the human brain could be seen as a collection of individual neurons that collectively act intelligent.

Collective intelligence can be defined as a group ability to solve more problems than its individuals [6]. In order to overcome the individual cognitive limits and the difficulties of coordination, a collective mental map can be used. A collective mental map is represented as an external memory with shared access that can be formalized as a directed graph. The efficiency of mental problem-solving depends on the problem representation in the cognitive system. Intelligent agents are characterized by the quality of their mental maps, knowledge and understanding of their environment, capacities for action or goals.

The mathematic measure applied to quantify the collective intelligence is a “collective intelligence quotient”. The elements of collective intelligence, such as displacements, actions of beings or exchange of information, are observed, measured and evaluated. A formal molecular model of computation and mathematical logic for describing the collective intelligence concept has been proposed [17]. The process, random and distributed, is tested in mathematical logics by social structures.

Collective intelligence offers a new perspective to different phenomena. This concept suggests another way of thinking about effectiveness, profitability, teamwork or leadership. The formal hierarchies of traditional organizations need to be replaced by self-organizing communities of practice because “most fields of expertise are now too complex for any one person to master and thus collective intelligence must be brought to bear to solve important problems”.

4. Communities of Practice in Multi-Agent System

Agents and multi-agent systems offer a new possibility for analyzing, modeling and implementing the complex systems. Agent-based vision offers a wide range of tools, techniques and paradigms, with a real potential to improve the use of informational technologies.

In a dictionary, an agent is defined as “someone or something who acts on behalf of another person or group”. This type of definition is too common to be considered operational, so agents have been defined to be “autonomous, problem-solving computational entities capable of effective operation in dynamic and open environments" [9]. Therefore, agents offer a new and appropriate route to the development of complex systems, especially in open and dynamic environments.

To describe an agent-based environment for formal modeling of communities of practice, we have used the methods of computational organization theory. The computational organization theory is a multidisciplinary field that integrates perspectives from artificial intelligence, organization studies, system dynamics and simulations.

In this framework, a comprehensive interaction model has to fulfill the condition:

- to be a dynamic representation, allowing change and development over time;
- to have a strong social dimension, so that members could learn, work and interact with others;
- to recognize the existence of general and particular communities of practice associated with particular occupations and organizations.

Symbolic representation and reasoning techniques from research on artificial intelligence are used to develop computational models of theoretical phenomena. Once formalized through a computational model, the symbolic representation can be developed to simulate the dynamics of members' behavior. The empirical validation, reflecting the dynamic behavior of the organization in communities of practice, determines the results and outcomes of the model to be considered already externally validated and generalized. This approach enables us to integrate qualitative behavior determined by the symbolic models with quantitative dynamics generated through simulations.

5. A Genetic Algorithm for Community Formation based on Collective Intelligence Capacity

A genetic algorithm represents an iterative process that applies genetic operators such as selection, crossover and mutation to a population of elements. The elements, called chromosomes, represent possible solutions to the problem.

Each chromosome has associated a fitness value which quantifies its value as a possible solution. Obviously, a chromosome representing a better solution will have a higher fitness value. The chromosomes compete for survival based on their fitness value. The crossover operator transfers genetic material from one generation to another. The mutation operator introduces new genetic material into the population.

Genetic clustering or partitioning algorithms include selection, crossover and mutation operators, adaptations of these operators and also some totally different operators. Adaptation is essential, well defined fitness functions and suitable operators are needed to encode potential solutions and to induce the evolution towards the optimal clustering.

The community formation problem is a partitioning problem, aiming to find good partitions of a set of agents into disjoint communities. The solution of a problem must satisfy various constraints, otherwise the solution is invalid. The objective of the grouping is to optimize the fitness function.

Classical genetic algorithms cannot be directly applied to partitioning problems. The structure of the chromosomes is item oriented instead of group (community) oriented, so a special encoding scheme is needed to transfer the relevant structures of grouping problems into genes in chromosomes.

5.1 The Encoding Scheme

The encoding scheme focuses on transferring the genes into relevant groups or communities. This encoding scheme ensures both the transmission of the genes from one generation to the next and a better quality estimation of the regions they occupy in the search space.

In this context, a chromosome can be represented as a set of a number of mutually disjoint communities:

$$\left\{a_{i_1}, \dots, a_{i_{k(1)}}\right\} \dots \left\{a_{i_{n-k(C)+1}}, \dots, a_{i_n}\right\} \quad (1)$$

where $k(j)$, $j = 1, \dots, C$, denotes the length of the community j (the number of agents in that community) and C the number of communities encoded in a chromosome.

5.2 The Fitness Index

The most important part is to find a measurement of the suitability of an agent i into the community c based on the execution of task j. We called this measurement *intelligence score* (Table 1). The intelligence score is denoted by μ_{ij}, where $0 \leq \mu_{ij} \leq 100$.

The intelligence scores measure agents' intelligence in executing the current tasks by forming communities, combining agents with the tasks based on their intelligence.

The overall performance of a community c in respect to a task j is defined by an *intelligence index*, denoted μ_j^c.

Table 1. The intelligence scores of *n* agents participating in *p* tasks.

	Task 1	…	*Task j*	…	*Task p*
Agent 1	μ_{11}	…	μ_{1j}	…	μ_{1p}

…	…	…	…	…	…
Agent i	μ_{i1}	…	μ_{ij}	…	μ_{ip}
…	…	…	…	…	…
Agent n	μ_{n1}	…	μ_{nj}	…	μ_{np}

Communities for which the intelligence of completing the task *j* matches exactly or exceeds the necessary capacity are valued to 1. The ones for which the intelligence does not match the necessary capacity are valued to 0. The *intelligence index* has the following form:

$$\begin{cases} 1, if \ \sum_{i \in c} \mu_{ij} \geq 100 \\ 0, if \ \sum_{i \in c} \mu_{ij} < 100 \end{cases} \tag{2}$$

The aggregate *intelligence index* of the community *c* for completing the overall task is calculated as:

$$\mu^c = \prod_{j=1}^{p} \mu_j^c \, , 0 \leq \mu^c \leq 1 \tag{3}$$

The *intelligence index* of the partitioning solution (partition of *n* agents into *C* communities) represents the average μ^c values:

$$f = \frac{\sum_{c=1}^{C} \mu^c}{C} = \frac{\sum_{c=1}^{C} \prod_{j=1}^{p} \mu_j^c}{C}, 0 \leq f \leq 1 \tag{4}$$

where *C* represents the number of communities in the solution.

In this context, the problem of community formation can be formulated as follows:

$$\max_C \frac{\sum_{c=1}^{C} \mu^c}{C} \tag{5}$$

Our approach proposes a new fitness index, which aims at being a trade-off between intelligence and cooperation. This “intelligence index” allows community formation, partitioning *n* agents into *C* communities, and also collective intelligence, emerging from the collaboration of agents, to be driving principle in the community formation process.

5.3 The Mutation Operator

The mutation operator should work with groups or communities of agents rather than with individual agents. Thus, three general strategies could be followed: shifting a small number of agents between communities, merging two existing communities or dividing a community in two new communities.

Once the P_{k+1} generation reached the size of the P_k generation, the partitions are sorted in descending order of their fitness index. Then they are subject to mutations based on their fitness index. The fitter the partition, the fewer mutations it undergoes. Each mutation occurs only if predefined mutation probabilities are met. One of the n agents is randomly selected with the predefined mutation probability p_m, being removed from the community and transferred to another community chosen randomly. In the same time, the replaced agent is transferred to the community from which the first agent was removed.

The mutation probability may vary along the population with the fitness index. Only a small partition of the population (approximately 5%) could be considered as elite survivors. The mutation probability for the elite survivors is very low. For the other partitions, when the fitness index gradually decreases, the mutation probability gradually increases.

Moreover, the mutation operator must be able to introduce new communities into a chromosome or to remove existing communities. Such operators may choose to divide the communities with high level of intelligence into two new communities or to merge two communities into an aggregate community.

5.4 The Crossover Operator

The crossover operator should transfer communities by inheritance, for instance to transfer communities from parents to offspring, so that the inherited communities remain valid, exhaustive and mutually disjoint.

At the first step, half of the best distinct parental communities are collected and transferred to the children. The communities in both parents are grouped and sorted in descending order of the fitness index. Half of these communities, the most accurate ones, are transferred to the child. The resulting communities do not necessarily form a partition of the n original agents.

Some of the communities may overlap and some of the n agents may not appear in any child community. Since each parent is a partition, any of the n agents belongs to just one subset of the first parent and to just one subset of the second parent. Thus, each of the n agents belongs to either two or one or zero of the child communities.

The second step is required in order to establish the validity of the child communities, as follows:

- If an agent belongs to two child communities, then it is removed from the community with the lowest fitness index (or removed randomly if the two communities have the same fitness index).
- If an agent belongs to one child community, then no adjustment is necessary.
- If an agent does not belong to any child community, a heuristic approach has to be used in order to integrate it into a community. This approach places the agent in the community with the lowest fitness index μ^c.

5.5 Numerical Example

We have run our experiments on synthetic data. The experiments have shown up that genetic algorithms provide the capability of producing optimal partitions of a set of agents.

To illustrate the measurements and indexes mentioned above we will present a numerical example.

Table 2. The intelligence scores of 7 agents participating in *3* tasks

	Task 1	*Task 2*	*Task 3*
Agent 1	*60*	*44*	*83*
Agent 2	*62*	*22*	*79*
Agent 3	*37*	*89*	*75*
Agent 4	*84*	*90*	*67*
Agent 5	*79*	*48*	*70*
Agent 6	*58*	*64*	*42*
Agent 7	*72*	*87*	*55*

Table 3. The intelligence scores of *3* communities with respect to *3* tasks

	Task 1	*Task 2*	*Task 3*
{1,4,5}	*60+84+79=223*	*44+90+48=182*	*83+67+70=220*
{2,3}	*62+37=99*	*22+89=111*	*79+75=154*
{6,7}	*58+72=130*	*64+87=151*	*42+55=97*

Obviously, the first community has a better performance than the second and the third one. The intelligence scores of each task in the first community are greater than the ones in the second and the third community. We have obtained the following intelligence index for each community:

$\mu_1^1 = 1, \mu_2^1 = 1, \mu_3^1 = 1 \text{ -> } \mu^1 = 1 \cdot 1 \cdot 1 = 1$

$\mu_1^2 = 0, \mu_2^2 = 1, \mu_3^2 = 1 \text{ -> } \mu^2 = 0 \cdot 1 \cdot 1 = 0$

$\mu_1^3 = 1, \mu_2^3 = 1, \mu_3^3 = 0 \text{ -> } \mu^3 = 1 \cdot 1 \cdot 0 = 0$

The intelligence index of the partitioning solution (partition of 7 agents into 3 communities) is:

$f = (1 + 0 + 0) / 3 = 0.33$

In order to further develop our genetic algorithm and to maximize the intelligence index of the partitioning solution, the mutation operator should be applied. The strategy we have chosen for optimizing the community formation is to shift a small number of agents between communities. Thus, we have switched *Agent 3* in the second community with *Agent 7* in the third community and obtained a new set of intelligence scores:

Table 4. The intelligence scores of *3* communities with respect to *3* tasks after applying the mutation operator

	Task 1	*Task 2*	*Task 3*
{1,4,5}	*60+84+79=223*	*44+90+48=182*	*83+67+70=220*
{2,7}	*62+72=134*	*22+87=109*	*79+55=134*
{6,3}	*58+37=95*	*64+89=153*	*42+75=117*

Applying the mutation operator, we have obtained the following intelligence index for each community:

$\mu_1^1 = 1, \mu_2^1 = 1, \mu_3^1 = 1 \rightarrow \mu^1 = 1 \cdot 1 \cdot 1 = 1$

$\mu_1^2 = 1, \mu_2^2 = 1, \mu_3^2 = 1 \rightarrow \mu^2 = 1 \cdot 1 \cdot 1 = 1$

$\mu_1^3 = 0, \mu_2^3 = 1, \mu_3^3 = 1 \rightarrow \mu^3 = 0 \cdot 1 \cdot 1 = 0$

The intelligence index of the partitioning solution (partition of 7 agents into 3 communities) is:

$f = (1 + 1 + 0) / 3 = 0.66$

The results we have obtained after applying the mutation operator are significantly better than the previous results. The intelligence index of the partitioning solution has increased from 0.33 to 0.66.

The numerical example seems to be enlightening. Even so, collective intelligence should not be perceived as the sum of individual intelligences, it represents the ability of the community to complete more tasks than its single individuals.

6. Conclusions and Future Work

Communities of practice represent social structures suitable for creating, developing and sharing knowledge in organizations and provide an efficient organizational framework for achieving the creative and learning functions of organizations. Communities of practice can provide a social reservoir for practitioners and knowledge producers to analyze, address and explore new solutions to their problems.

The organizational behavior field is interested in studying organizations as complex social systems. Most of the theories in this domain explore individual and collective human behavior within organizations. Managing collective intelligence within an organization implies combining all tools, methods and processes that can lead to connection and cooperation among individual intelligences.

Our research is based on the theoretical approaches presented in the literature, with emphasis on genetic algorithms applications. The paper proposes a genetic algorithm for community formation based on collective intelligence capacity. This approach introduces the concept of intelligence index, aiming to optimal partitions of a set of agents. The mechanism highlights the relevance of intelligence in community formation and reveals the need for such mechanisms that allow large group of professionals to make decisions better than single individuals.

Our future work in this direction will focus on extending the experiments of genetic algorithms in community formation and applying the algorithm on real data.

Acknowledgement

This article is one of the results of the research activity carried out by the authors under the frame of following projects:
1. “Doctoral Program and PhD Students in the education, research and innovation triangle”. This project is co-funded by European Social Fund through the Sectorial Operational Programme for Human Resources Development 2007-2013, coordinated by the Academy of Economic Studies, Bucharest.
2. “Ph.D. for a career in interdisciplinary economic research at the European standards” (DOCCENT), the European Social Fund through Sectorial Operational Programme for Human Resources Development 2007-2013, project number POSDRU/107/1.5/S/77213.

REFERENCES

[1] **Andriessen, E., Huis, M. (2001),** ***Group Dynamics and CoPs*** **.** Position paper for ECSCW 2001 Workshop 6, Actions and Identities in Virtual Communities of Practice;
[2] **Garrido, P. (2009),** ***Business Sustainability and Collective Intelligence***; The Learning Organization, Volume 16, Number 3 2009, pp. 208-222;
[3] **Georgescu, V. (2007),** ***Evolving Coalitions of Task-Oriented Agents via Genetic Algorithms to Foster Self-Organization in Digital Business Ecosystems*** . Proceedings of the International Conference on Modeling Decision for Artificial Intelligence;
[4] **Golberg, D. E. (1989),** ***Genetic Algorithms in Search, Optimization, and Machine Learning*** . Addison-Wesley Professional;

[5] **Gruszczyk, W., Kwasnicka, H. (2008),** ***Coalition Formation in Multi-Agent Systems – An Evolutionary Approach*** ; Proceedings of the International Multiconference on Computer Science and Information Technology, pp. 125-130;
[6] **Heylighen, F. (1999),** ***Collective Intelligence and its Implementation on the Web: Algorithms to Develop a Collective Mental Map***; Springer Netherlands;
[7] **Juriado, R., Gustafsson, N. (2007),** ***Emergent Communities of Practice in Temporary Inter-organizational Partnerships*** ;The Learning Organization: The International Journal of Knowledge and Organizational Learning Management, Volume 14, Number 1 2007, pp. 50-61;
[8] **Lave, J., Wenger, E. (1991),** ***Situated Learning. Legitimate Peripheral Participation*** ; Cambridge University Press, Cambridge;
[9**] Luck, M., McBurney, P., Preist, C. (2003),** ***Agent Technology: Enabling Next Generation Computing – A Roadmap for Agent Based Computing*** ; AgentLink;
[10] **Maracine, V., Delcea, C.(2009),** ***How We Can Diagnose the Firms' Diseases Using Grey System Theory***. Economic Computation and Economic Cybernetics Studies and Research Journal, Volume 43, Issue 3;
[11] **Maries, I., Scarlat, E., (2010),** ***Modeling Trust and Reputation within Communities of Practice***; Proceedings of the 2010 IEEE International Conference on Systems, Man, and Cybernetics, pp. 2192-2199;
[12] **Muller, P. (2006),** ***Reputation, Trust and the Dynamics of Leadership in Communities of Practice***; Journal of Management and Governance, Springer;
[13] **Noubel, J., F. (2004),** ***Collective Intelligence, the Invisible Revolution*** ; The Transitioner.org http://www.thetransitioner.org/wen/tiki-list_file_gallery.php?galleryId=1;
[14] **Por, G., van Bukkum, E. (2004),** ***Liberating the Innovation Value of Communities of Practice***; Amsterdam;
[15] **Scarlat, E., Maries, I. (2009),** ***Increasing Collective Intelligence within Organizations Based on Trust and Reputation Models*** ; Economic Computation and Economic Cybernetics Studies and Research Journal, Volume 43, Issue 2;
[16] **Scarlat, E., Maries, I. (2010),** ***Simulating Collective Intelligence of the Communities of Practice using Agent-based Methods*** ; Proceedings of the 4th International KES Symposium on Agents and Multi-Agent Systems, pp. 305-314;
[17] **Szuba, T., Almulla, M. (2000),** ***Was Collective Intelligence before Life on Earth?*** IPDPS Workshops on Parallel and Distributed Processing;
[18] **Wenger, E., C., Snyder, W., M. (2000),** ***Communities of Practice: the Organizational Frontier*** ; Harvard Business Review, Volume 78, Number 1, pp. 139-145;
[19] **Wenger, E., McDermott, R., A., Snyder, W. (2002),** ***Cultivating Communities of Practice: A Guide to Managing Knowledge***. Harvard Business School Press, Boston.