Self-managed Crowdsourcing

Vladislav Protasov, Zinaida Potapova, Eugene Melnikov
Center of Computing for Physics and Technology
Moscow, Russia
protvlad@gmail.com, zinaida.potapova@gmail.com, apinae1@gmail.com

***Abstract*—The paper presents a technology of crowdsourcing organization that is considered to provide optimal solutions of complicated intelligent problems from large group of experts. This approach is developed as an extension of the Evolutionary Solutions Coordination method; it supposes dividing entire expert group into small ones and automatically defining best experts on each iteration to form the most competent groups on further iterations. Finally, the group of the most competent experts gives full and the most rational solution. The paper also presents an experiment on the Eysenck's tests solving. It demonstrates significant superiority of self-managed crowdsourcing over individual solution approach.**

Keywords—crowdsourcing; collective intelligence; coordination; generation; estimation; election; social.

I. INTRODUCTION

Growth of the Internet and evolution of web technologies gave birth to a new kind of collective interaction of the web users – crowdsourcing, which means creating needed solutions, ideas and other intellectual products with contribution of large number of people, usually Internet users. This term was first used by the journalist Jeff Howe [1]. In Russia, this technology is developed by a young company Witology performed a set of considerable projects for Russian economy [2]. Unfortunately, according to experience of crowdsourcing practice, the technology is rather spendy, it requires usage of expensive software, large calculation resources and numerous project coordinators, so called facilitators [3], it makes the technology almost unreachable for small companies and certain Internet users, who also may represent interesting social projects.

The technology of Self-managed Crowdsourcing presented in this paper does not have these disadvantages. It utilizes a special algorithm that is able to coordinate collective solution search automatically, without human intervention [4].

The first section of the paper deals with theoretical basics of the presented algorithm. It explains the logic of three-cycle scheme step by step. In the next section the theory of self-managed crowdsourcing is presented. After that some experimental results will be provided. The experiment considers a group of students that solve IQ tests using the provided theory. The last section gives a conclusion upon the described experiment.

II. THREE-CYCLE SCHEME OF CROWDSOURCING

Papers on crowdsourcing state that group solution can significantly overcome individual results if there is good algorithmic platform of collective work organization [5]. Method of Evolutionary Solutions Coordination described by Protasov et al. in [6] can be used as a kind of crowdsourcing approach on large experts number. Therefore, there is a need to develop such a collaboration technology that leads the number of coordination steps to minimum.

According to the procedure of Evolutionary Solutions Coordination, in order to minimize coordination steps number, the following organization of self-managed crowdsourcing can be considered. On the first cycle, each expert fills slots of the projects with his own proposals. On the second cycle, basing on predefined regular graph [7] of connections experts receive others' solutions and fill up empty slots of their own solutions with received variants that seem correct in their view [8]. On the third cycle, for each slot of the project, the value chosen by greater than a half of experts is considered as the group's choice. The group's solution is formed of such slots' values.

Of course, on the first cycle, it is desirable to provide equiprobable slots filling. Regular graph is applicable for this task because there is only iteration of solutions coordination and random genetic-like expert pairs selection is not necessary. It is also noticeable that each cycle is performed by all experts concurrently and the operation time does not depend on their quantity.

The obtained results are also significant for the artificial intelligence construction of large number of uniform modules, e. g., neurons cluster or primitive computers forming homogeneous environment.

The scheme of three-cycle coordination is illustrated on Fig. 1.

III. THE TECHNOLOGY OF SELF-MANAGED CROWDSOURCING

Let us consider the Evolutional Coordination Method application based on social web-platform – an Internet site constructed specially for organizing crowdsourcing technology. The number of experts is not defined previously. Collective work can be described as follows. The experts registered on the site form small groups, e. g., 10 people each group. Each expert is provided with a predefined project task and an instruction which contains his groupmate list for sending initial ideas and time interval allotted for ideas generation. After finishing the first stage and receiving other users' solutions for expertise, the experts perform an estimation of the received solutions. All slots' values which seem correct for them are copied to their own solutions. After the received variants estimation and slots filling, all solutions are saved on the site. Then, the special program, Project Moderator, picks out best

 ISBN: 978-1-61208-352-0

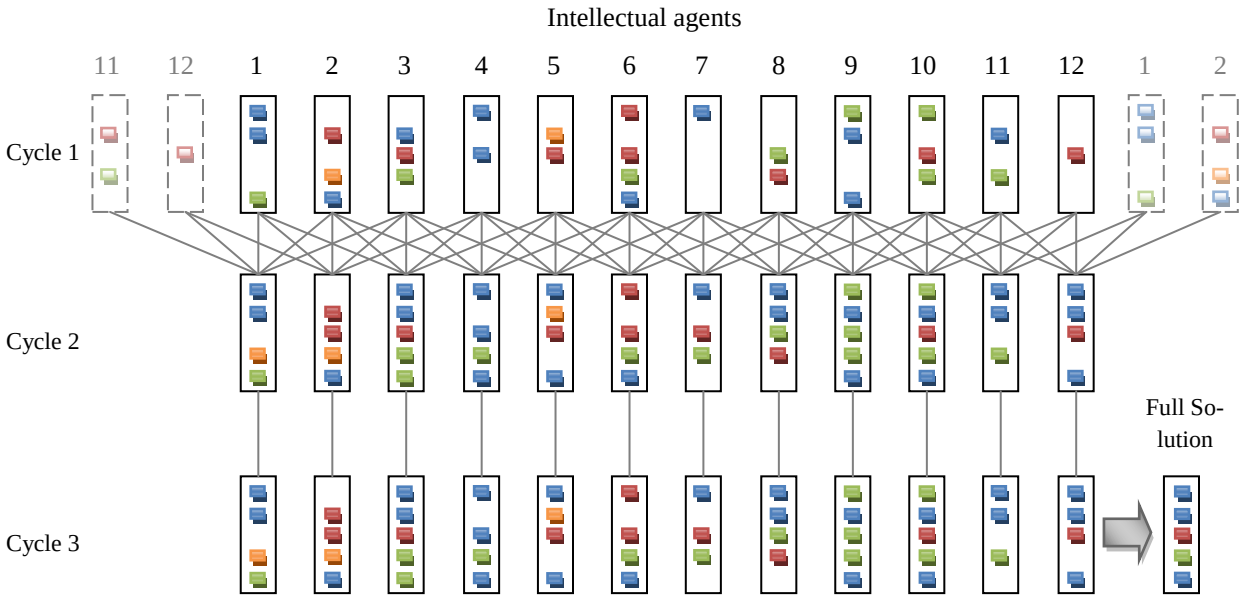

Figure 1. Three-cycle coordination scheme

experts (ones with greatest estimation ability in a group) from all groups and allocates them into new groups. These groups are given the best variants of previous expert generation and estimation cycle is repeated. Further, the best experts are selected again from these groups; the process proceeds until the best final group is selected among experts, and they obtain the last and the most fully defined solution.

Such scheme of crowdsourcing organization induces self-managed processes of the best experts distinguishing and obtaining the best solution with small number of iterations. Existing crowdsourcing schemes widely utilize such technology with inviting large groups of human moderators. This makes projects very expensive. But, the technology described in this paper supposes to entrust the coordination work to a program.

IV. Experimental Testing of Self-Managed Crowdsourcing Technology

In order to test the self-manged crowdsourcing technology, some experiments in Internet were carried out. 7 groups of students consisting of 3 to 7 people participated in the experiments. Eysenck intellect testing with predefined test examples was selected as a test task [9]. It included 50 questions where the answers were single keywords. The experiment was hold according to the crowdsourcing technology described above. Students registered beforehand on the site of the project and were united into groups of 3, 4, 5, 6, 7, 5 and 7 people (see the Table 2). According to the Eysenck methodology each group was given 30 minutes to complete the test, the time intervals allocated for the first and second cycles were defined depending on a group size: 18 minutes for 7 and 24 minutes for 3 people for the first cycle – the time varies because more experts need more time for coordination stage.

The table contains competence data of all experts: slots number filled on first and second cycles, competence in generation and estimation, forecast of correct answers number (N), IQ level calculated by the Eysenck's formula $IQ = 90 + 2.5N$ and the Intellectual Potential (IP), where IP is defined by the formula

$$IP = \frac{1000}{M} \quad (1)$$

where *M*, for certain expert, is the number of experts with the same competence (both generative and estimative) that provides full solution with probability 0.999.

Analyzing the table, it's easy to descry that IP distribution has considerably greater dispersion then that of IQ level. In this example it vary from 1.5 to 144.7 while IQs of corresponding experts equal 104 and 165. Since IP shows more correctly the certain expert's contribution to overall work, it can be deduced that estimation of experts only by IQ gives undeserved advantage to less smart ones. Table 1 illustrates full results of expert groups work. It contains a number of slots filled on the second cycle, portion of filled slots after first and second cycles and IQ level for each group.

Table 1. Expert groups results

Group N	Slots filled on 2st cycle	Relative slots filling on 1st cycle	Relative slots filling on 2st cycle	IQ
1	19	0,06	0,38	138
2	24	0,18	0,48	150
3	28	0,06	0,56	160
4	25	0,02	0,5	152
5	14	0,02	0,28	125
6	21	0,06	0,42	142
7	23	0	0,46	148
3rd cycle	50	0,52	1	215

 ISBN: 978-1-61208-352-0

TABLE 2. EXPERTS RATINGS

Group *N*	Expert *N*	Slots filled on 1st cycle	Slots filled on 2st cycle	Generation competence	Estimation competence	Correct answers forecast (N)	IQ	IP
1	1	3	26	0,075	0,74	3,75	99	11,3
	2	18	19	0,45	0,0625	22,5	146	30,2
	3	19	21	0,475	0,133	23,75	149	53,2
2	1	6	23	0,156	0,654	7,82	110	23,8
	2	6	0	0,156	0,01	7,82	110	1,9
	3	14	20	0,365	0,333	18,3	136	54,4
	4	14	22	0,365	0,555	18,3	136	63,1
3	1	11	37	0,3	1	15	128	55,1
	2	8	28	0,218	0,69	10,9	117	35,1
	3	1	16	0,02	0,432	1	93	2,9
	4	3	27	0,082	0,706	4,09	100	12,3
	5	13	28	0,355	0,625	17,7	134	62,5
4	1	6	25	0,171	0,543	8,57	111	25,1
	2	1	29	0,02	0,7	1	93	3,2
	3	14	30	0,4	0,222	20	140	52,7
	4	5	6	0,143	0,028	7,14	108	3,9
	5	4	0	0,115	0,01	5,71	104	1,5
	6	11	38	0,314	0,11	15,7	129	26,1
5	1	4	12	0,12	0,348	6	105	14,9
	2	6	11	0,18	0,238	9	113	19,9
	3	11	0	0,33	0,01	16,5	131	4,3
	4	11	13	0,33	0,125	16,5	131	30,1
	5	2	11	0,06	0,36	3	98	7,5
	6	11	21	0,33	0,625	16,5	131	56,8
	7	7	20	0,21	0,65	10,5	116	33,1
6	1	5	19	0,136	0,634	6,82	107	20,3
	2	7	25	0,191	0,9	9,54	114	31,8
	3	9	19	0,245	0,556	12,3	121	38,2
	4	11	14	0,3	0,187	15	128	32,9
	5	10	24	0,273	0,823	13,6	124	47,4
7	1	3	4	0,094	0,025	4,74	102	2,4
	2	4	0	0,126	0,01	6,31	106	1,6
	3	10	13	0,316	0,091	15,8	130	23,3
	4	19	41	0,6	0,917	30	165	144,7
	5	6	28	0,189	0,594	9,47	114	28,8
	6	6	11	0,189	0,135	9,47	114	15,8
	7	12	14	0,379	0,064	18,9	137	23,6

According to the technology of Self-managed Crowdsourcing, on final cycle the group of experts with the best competence in solutions estimation was selected from the whole set of experts. They interchanged their solutions and complemented each other. After the coordination cycle, for the final solution, the slots presented in most experts' solutions were distinguished to form the final solution. The results are presented in the last row of the Table 1. On the last cycle of the experiment selected group of the best experts obtained full solution – all questions of the test were answered.

V. CONCLUSION AND FUTURE WORK

Finally, the first cycle of the Self-managed Crowdsourcing includes the creation of population of correctly filled slots. After second cycle, many experts experience considerable growth of correctly filled slots, but most experts correct answers portion is still low and the group has not yet filled all the slots of the projects. Only small group of leaders has estimative ability greater than 0.5, they actually have overcome the Condorcet's border [10]. These very leaders formed the final group that obtained full solution on the last cycle. Taking into account that Eysenck's tests are constructed with consideration of impossibility to solve them fully in given time it can be concluded that expert groups under Self-managed Crowdsourcing technology have done an impossible.

ACKNOWLEDGEMENTS

 ISBN: 978-1-61208-352-0

This work was supported by the Russian Foundation for Basic Research, grant #13-07-00958 "Development of the theory and experimental research of a new information technology of self-managed crowdsourcing".

REFERENCES

[1] Howe J., The Rise of Crowdsourcing, www.wired.com, [retrieved: 04, 2014], p. 1.

[2] Witology, the first Russian crowdsourcing company, www.witology.com/company, [retrieved: 04, 2014].

[3] Shih-Wen Huang, Wai-Tat Fu, "Motivating Crowds Using Social Facilitation and Social Transparency", CHI 2011, Vancouver BC Canada, May 2011, p. 2.

[4] Protasov V. I. and Potapova Z. E., "Self-Governing Crowdsourcing. Theory, Technology and Practice", The 3rd International Conference on Control Automation, Intelligent Systems and Environments, Makhachkala, 2012, RAS Kabardino-Balkar Science Center Press, vol. 1, pp. 91-99.

[5] Protasov V., Vitiska N., and Shelhkova L., "Use of collective intelligence of group of robots for acceleration of acceptance expedient decisions", Conference "Intellectual robotic systems", Russia, Gelendjik, Oct. 2001, pp. 187— 189.

[6] Protasov V., Potapova Z., and Melnikov E., "Overcoming the Condorcet's Border in Collective Intelligence Systems", The Second International Conference on Intelligent Systems and Applications INTELLI, 2013 April 21 - 26, 2013 - Venice, Italy, pp. 2-4.

[7] Regular Graph, en.wikipedia.org/wiki/Regular_graph [retrieved: 04, 2014], p. 1

[8] Protasov V., Vitiska N., and Shustov E., "Solving complex problems with a genetic consilium method", Conference "Intellectual robotic systems", Russia, Gelendjik, Oct. 2001, pp. 53-55.

[9] Eysenck H. J., "The Structure and Measurement of Intelligence. Tests IQ", University of London. Institute of Psychiatry, 2001, p. 128.

[10] Le Marquis de Condorcet, "Essay on the Application of Analysis to the Probability of Majority Decisions", 1785, Les Archives de la Revolution Française, Pergamon Press, pp. 9-23.

[11] Turchin V., "The phenomenon of science. A cybernetic approach to human evolution", Columbia University Press, New York, 1977, pp. 245-250.

[12] Protasov V. I., "Design of Metasystem Tansitions", Physical and Technical Informatics Institute Press, Moscow, 2009; p. 186.

[13] Protasov V., "Generation of New Knowledge by Network Human-machine Intelligence. Statement of a problem.", J. Neurocomputers. Development and application, vol.7-8, Moscow, 2001, pp. 94 – 103.

 ISBN: 978-1-61208-352-0