INCENTIVES IN MULTI-AGENT SYSTEMS

Dmitri Novikov
Institute of Control Sciences
Profsojuznaya St., 65, Moscow,
117806, Russia,
E-mail: nov@ipu.rssi.ru

KEYWORDS
Multi-agent system, intelligent behavior, game theory, decentralized control.

ABSTRACT

The incentive problem for multi-agent active system (AS) is solved on the basis of the agents game decomposition, i.e. by the construction of the incentive system, which implements the optimal (from principal's point of view) strategies of the agents as the dominant strategies equilibrium (DSE). The problem of information aggregation in multi-agent systems is also considered.

INTRODUCTION

In most of the game-theoretical models, explored in the active systems theory (Burkov and Enaleev 1994; Novikov 1995; 1997; 1998) and in the theory of contracts (Hart 1983; Mas-Colell et al. 1995), active systems, which consist of one principal and one agent, are considered. The absence of the general approaches towards the incentive problem solving in the case of multi-agent systems may be explained by the fact that efficient methods of the agents game analysis are unknown, while existing methods (Mookherjee 1984; Novikov 1997) are characterized by a very high complexity and they do not provide the researcher with the analytical solution. In this paper the method, based on the incentive function, which decomposes the game of the agents by implementing the DSE, is proposed.

INCENTIVE PROBLEM

Consider deterministic multi-agent AS, which consists of the principal and n agents. Agent's strategy is the choice of the action; principal's strategy is the choice of the incentive function (control parameter) – a mapping of agents' actions onto the set of their feasible rewards (remuneration).

Denote $y_i \in A_i$ – the action of i-th agent, $i \in I = \{1, 2, ..., n\}$ – the set of agents, $y = (y_1, y_2, ..., y_n) \in A' = \prod_{i=1}^{n} A_i$ - a vector of agents action, $y_{-i} = (y_1, y_2, ..., y_{i-1}, y_{i+1}, ..., y_n) \in A_{-i} = \prod_{j \neq i} A_j$ - a situation of the game for i-th agent.

Preferences of the principal and the agents are reflected by their goal functions. Principal's goal function $\Phi(\sigma, y)$ is the difference between his income $H(y)$ and total rewards of the agents: $\upsilon(y) = \sum_{i=1}^{n} \sigma_i(y)$, where $\sigma_i(y)$ – is the reward of i-th agent, $\sigma(y) = (\sigma_1(y), \sigma_2(y), ..., \sigma_n(y))$. The goal function of i-th agent $f_i(\sigma_i, y)$ is the difference between the reward and his costs $c_i(y)$, i.e.:

$$f_i(\sigma_i, y) = \sigma_i(y) - c_i(y),\ i \in I, \quad (1)$$

$$\Phi(\sigma, y) = H(y) - \sum_{i=1}^{n} \sigma_i(y). \quad (2)$$

Suppose that the principal and the agents at the moment of making their decisions posses the information about all the goal functions and all the sets of feasible strategies (this information is common-knowledge). The principal chooses his strategy first and reveals his choice to the agents, whereupon the agents choose the actions in order to maximize goal functions under the given incentive functions.

Introduce the following assumptions (an introduced assumption is considered to be valid hereafter):

A.1. $\forall\ i \in I\ \ A_i \subseteq \Re_+^{m_{ш}}$ is a closed interval with the left end in zero.

A.2. $\forall\ i \in I$ 1) the cost function $c_i(\cdot)$ is continuous; 2) $\forall y_i \in A_i$ $c_i(y)$ does not decrease in y_i, $i \in I$; 3) $\forall\ y \in A'\ c_i(y) \geq 0$; 4) $\forall\ y_{-i} \in A_{-i}\ \ c_i(0, y_{-i}) = 0$.

A.3. Incentive functions are piece-wise continuous and nonnegative.

A.4. Principal's income function is continuous and achieves its maximum on nonzero actions of the agents.

Denote M the set of feasible (satisfying A.3) incentive functions, $P(\sigma)$ – the set of equilibrium (under the given incentive functions σ) strategies of the agents (the type of equilibrium is not specified yet, except the assumption that the agents choose their actions independently and simultaneously and are not allowed to interchange additional information or utility), which is referred to as the set of implementable actions.

The efficiency of management (guaranteed efficiency of incentives) is the minimal value of principal's goal function on the set of actions, implemented by these incentives.:

$$K(\sigma) = \min_{y \in P(\sigma)} \Phi(\sigma, y). \quad (3)$$

The incentive problem is to point out the feasible incentive function σ^*, which has maximal efficiency on the set M:

$$\sigma^* = arg \max_{\sigma \in M} K(\sigma). \quad (4)$$

In that specific case, when agents are independent (the reward and the costs of each agent depend only on his own action), the following compensative incentive function is optimal (correctly - δ-optimal, where $\delta = \sum_{i=1}^{n} \delta_i$):

$$\sigma_{iK}(y_i) = \begin{cases} c_i(y_i^*) + \delta_i, & y_i = y_i^* \\ 0, & y_i \neq y_i^* \end{cases}, \; i \in I, \quad (5)$$

where $\{\delta_i\}$ are arbitrary small strictly positive constants. Optimal action y^*, implemented by the incentive function (5) as the DSE, is defined as the solution of the following optimal incentive compatible planning problem:

$$y^* = arg \max_{y \in A'} \{H(y) - \sum_{i=1}^{n} c_i(y_i)\}.$$

THE SOLUTION OF THE INCENTIVE PROBLEM FOR MULTIAGENT SYSTEM

Introduce several definitions. The set of Nash equilibrium (NE) $E_N(\sigma)$ is:

$$E_N(\sigma) = \{y^N \in A \mid \forall i \in I \; \forall y_i \in A_i \; \sigma_i(y^N) - c_i(y^N) \geq$$
$$\geq \sigma_i(y_i, y_{-i}^N) - c_i(y_i, y_{-i}^N)\}, \quad (6)$$

An action $y_{i_d} \in A_i$ is a dominant strategy of i-th agent iff

$$\forall y_i \in A_i, \; \forall y_{-i} \in A_{-i} \; \sigma_i(y_{i_d}, y_{-i}) - c_i(y_{i_d}, y_{-i}) \geq \sigma_i(y_i, y_{-i}) - c_i(y_i, y_{-i}).$$

If under the given incentive function all the agents have dominant strategies, then the corresponding vector of actions is implemented by this incentive function as DSE.

If the reward of each agent depends only on his own action, then the definitions of NE $E_N(\sigma)$ and DSE $y_d \in A$ take the form of:

$$E_N(\sigma) = \{y^N \in A \mid \forall i \in I \; \forall y_i \in A_i \; \sigma_i(y_i^N) - c_i(y^N) \geq$$
$$\geq \sigma_i(y_i) - c_i(y_i, y_{-i}^N)\}, \quad (7)$$

$y_{i_d} \in A_i$ – is a dominant strategy of i-th agent iff

$$\forall y_i \in A_i, \; \forall y_{-i} \in A_{-i} \; \sigma_i(y_{i_d}) - c_i(y_{i_d}, y_{-i}) \geq \sigma_i(y_i) - c_i(y_i, y_{-i}).$$

Fix some arbitrary vector of actions $y^* \in A'$ and consider the following incentive function:

$$\sigma_i(y^*, y) = \begin{cases} c_i(y_i^*, y_{-i}) + \delta_i, & y_i = y_i^* \\ 0, & y_i \neq y_i^* \end{cases}, \; \delta_i \geq 0, \; i \in I. \quad (8)$$

<u>Theorem 1.</u> If the principal applies incentive function (8), then y^* is a DSE. Moreover, if $\delta_i > 0$, $i \in I$, then this DSE is unique.

When introducing the incentive function (8), the principal applies the following **principle of decomposition**: he offers to i-th agent – "If you choose the action y_i^*, then I'll compensate your costs *regardless the choice of all other agents*. If you choose any other action, then your reward will be zero". Using such a strategy, the principal decomposes the game of the agents.

If the reward of each agent depends only on his action, then, fixing the situation y_{-i}^* of the game as a parameter, the following incentive function is obtained from (8):

$$\sigma_i(y^*, y_i) = \begin{cases} c_i(y_i^*, y_{-i}^*) + \delta_i, & y_i = y_i^* \\ 0, & y_i \neq y_i^* \end{cases}, \; \delta_i \geq 0, \; i \in I. \quad (9)$$

<u>Theorem 2.</u> If the principal applies incentive function (9), then $y^* \in E_N(\sigma)$. Moreover:

a) if the following condition holds:

$$\forall y^1 \neq y^2 \in A' \; \exists i \in I: \; y_i^1 \neq y_i^2 \text{ и } c_i(y^1) + c_i(y^2) >$$
$$> c_i(y_i^1, y_{-i}^2) - \delta_i, \quad (10)$$

then y^* is the unique NE;

б) if the following condition holds:

$$\forall i \in I, \; \forall y^1 \neq y^2 \in A' \; c_i(y^1) + c_i(y^2) \geq c_i(y_i^1, y_{-i}^2) - \delta_i, \quad (11)$$

then y^* is a DSE;

в) if the condition (11) holds and $\delta_i > 0$, $i \in I$, then y^* is the unique DSE.

The optimal value of the vector y^*, is defined as the solution of the following optimal incentive compatible planning problem: $y^* = arg \max_{t \in A'} \{H(t) - \upsilon(t)\}$. The efficiency of the incentive function (9) is:

$$K^* = H(y^*) - \sum_{i=1}^{n} c_i(y^*) - \delta.$$

The following theorem gives the solution of the incentive problem for the class of multi-agent systems, restricted by assumptions A.1-A.4.

<u>Theorem 3.</u> The class (with a parameter y^*) of the incentive functions (8), (9) is δ-optimal.

Results of theorems 1-3 were obtained under the assumption of noncooperative behavior of the agents. But, it is easy to verify, that this results are also valid in the case, when the agents may form coalitions (under transferable and nontransferable utility).

In the incentive problems, considered above, the optimal incentive functions were not continuous: costs of the agents were compensated when the desired vector of actions was chosen, in all other situations the rewards were zeroes. If the continuity of the incentive functions is required, then the optimal solution of the incentive problem is given by the following theorem (intuitively, it is sufficient that the marginal incentives are greater then the marginal costs in the vicinity of the implemented action).

Let the cost functions are continuous, the sets of feasible actions are compact and the hypothesis of benevolence (HB) (the agents choose from the set of implementable actions the action, preferred by the principal) is valid. Define the following incentive functions:

$$\sigma_i(y) = c_i(y)\, q_i(y_i^*, y), \quad (12)$$

where $q_i(y_i^*, y)$ is a continuous function, which satisfies

$$\forall\, i \in I \;\; \forall\, y_i \in A_i \;\; \forall\, y_{-i} \in A_{-i} \;\; q_i(y_i^*, y) \leq 1,\;\; q_i(y_i^*, y_i^*, y_{-i}) = 1. \quad (13)$$

Theorem 4. Under the HB y^* is implemented by (12)-(13) as a DSE.

AGGREGATION OF INFORMATION IN THE INCENTIVE PROBLEMS

Let the output $z \in A_0 = Q(A)$ of the AS is the function of agents actions: $z = Q(y)$. Principal's goal function $\Phi(\sigma, z)$ is the difference between his income H(z), which depends on the output, and the total remuneration of the agents (incentive costs) υ(z): $\upsilon(z) = \sum_{i=1}^{n} \sigma_i(z)$, where $\sigma_i(z)$ is the reward of i-th agent, $\sigma(z) = (\sigma_1(z), \sigma_2(z), \ldots, \sigma_n(z))$, i.e.

$$\Phi(\sigma, z) = H(\mathrm{z}) - \sum_{i=1}^{n} \sigma_i(z). \quad (14)$$

Agent's goal function $f_i(\sigma_i, y)$ is the difference between the reward and the costs $c_i(y)$, i.e.:

$$f_i(\sigma_i, y) = \sigma_i(z) - c_i(y),\; i \in I. \quad (15)$$

If the actions of the agents are observed by the principal (or when the principal, observing the output, is able to recalculate uniquely their actions), the incentive function $\tilde{\sigma}$, based on the actions, may be used: $\tilde{\sigma}_i(y) = \sigma_i(Q(y)),\; i \in I$. The solution of the corresponding incentive problem is described above. So, consider the case when the principal observes the output and he is not able to recalculate uniquely the actions of the agents, i.e. the aggregation of information (Novikov 1999) takes place.

Introduce the following assumption.

A.5. $Q: A' \to A_0 \subseteq \Re^m$ – is a single-valued continuous mapping ($1 \leq m < n$).

A Nash equilibrium y^N is defined in the following way:

$$\forall\, i \in I \;\; \forall\, y_i \in A_i \;\; \sigma_i(Q(y^N)) - c_i(y^N) \geq \sigma_i(Q(y_i, y_{-i}^N)) - c_i(y_i, y_{-i}^N).$$

Define the set of agents actions, which lead to the given output: $Y(z) = \{y \in A' \mid Q(y) = z\} \subseteq A',\; z \in A_0$. If the principal compensates agents costs, his minimal incentive costs to implement the output $z \in A_0$ are:

$$\vartheta(z) = \min_{y \in Y(z)} \sum_{i=1}^{n} c_i(y).$$

Define the set of agent actions, which lead to the given output and require minimal rewards to be implemented: $Y^*(z) = Arg \min_{y \in Y(z)} \sum_{i=1}^{n} c_i(y)$ and fix some vector $y^*(z) \in Y^*(z) \subseteq Y(z)$.

Introduce the following assumption.

A.6. $\forall\, x \in A_0,\;\; \forall\, y' \in Y(x),\;\; \forall\, i \in I,\;\; \forall\, y_i \in Proj_i\, Y(x)$ $c_j(y_i, y'_{-i})$ does not decrease in y_i, $j \in I$.

Theorem 5. If the principal applies the incentive function

$$\sigma_i^*(x, z) = \begin{cases} c_i(y^*(x)), & z = x \\ 0, & z \neq x \end{cases},\; i \in I, \quad (16)$$

where $x \in A_0$ is a planed output (parameter), this output is implemented as a NE with minimal incentive costs.

The drawback of the incentive function (16) is the presence (besides the NE, defined in theorem 5) of the DSE – the vector of zero (least-cost) actions of the agents. To implement $Y^*(x)$ uniquely, the principal has to pay additionally arbitrary small, but strictly positive, constant, i.e. to apply the following incentive function:

$$\sigma_i^*(x, z) = \begin{cases} c_i(y^*(x)) + \delta_i, & z = x \\ 0, & z \neq x \end{cases},\; i \in I,$$

which is, obviously, δ-optimal.

The optimal value of the output $x^* \in A_0$ is defined as a solution of the following optimal incentive compatible planning problem: $x^* = arg \max_{z \in A_0} [H(z) - \vartheta(z)]$.

Consider the following two cases. The first one is when the principal observes the actions of the agents and he is able to base the incentives both on the actions and on the output. The second case is the case of unobserved actions, when the incentives may depend on the observed output only. Compare the efficiencies of management for this two cases.

In the first case the minimal incentive costs $\vartheta_1(y)$ to implement the vector $y \in A'$ of agents actions are: $\vartheta_1(y) = \sum_{i=1}^{n} c_i(y)$, and the efficiency of management is: $K_1 = \max_{y \in A'} \{H(Q(y)) - \vartheta_1(y)\}$. In the second case the minimal incentive costs $\vartheta_2(z)$ to implement the output $z \in A_0$ are (see theorem 5): $\vartheta_2(z) = \min_{y \in Y(z)} \sum_{i \in I} c_i(y)$, and the efficiency of management is: $K_2 = \max_{z \in A_0} \{H(z) - \vartheta_2(z)\}$.

Theorem 6. $K_2 = K_1$.

Theorem 6 (which may be called the "theorem of ideal aggregation of information") plays an essentially significant methodological role. It states, that, if the principal's income depends on the aggregated output, then the efficiencies of management are the same when the incentives are based on the observable actions or when the incentives are based on the aggregated output (which is (by assumptions A.5 and A.6) less informative).

In other words, the presence of information aggregation does not decrease the efficiency of management. This fact looks like a paradox, as in (Novikov 1997, 1998) was proved that the presence of uncertainty does not increase the efficiency of management. In the model, considered above, the ideal aggregation (Novikov, 1999) takes place due to the principal's indifference between the agents actions, leading to the same output with minimal costs. Conditions A.5 and A.6 are sufficient for the principal's ability to impose the problem of equilibrium search on the agents. It allows to decrease the information, processed by the principal, without loss of efficiency.

CONCLUSION

This paper contains the results of consideration of game-theoretical models of incentive problems for multi-agent active systems. Two basic principles may be formulated: *the principle of costs compensation* and *the principle of agents game decomposition*.

The costs compensation principle states that the optimal incentive function, which implements certain vector of agents actions must exactly compensate costs of the agents. This principle is valid for management problems in both single-agent and multi-agent systems. Agents game decomposition principle is specific for multi-agent systems and concludes in inducing the agents to choose the actions, preferred by the principal, by the appropriate management (see (8) and (9)), which is optimal (see theorem 3) and allows each agent to make decisions independently from another agents (see theorems 1 and 2).

When exploring the information aggregation problem, *the generalization of the costs compensation principle* plays key role: minimal incentive costs of the principal to implement certain output are defined as the minimum of the total agents costs to undertake actions, which lead to this output.

REFERENCES

Burkov V.N., Enaleev A.K. Stimulation and decision-making in the active systems theory: review of problems and new results. *Mathematical Social Sciences*. 1994. Vol. 27. 271 – 291.

Hart O.D. Optimal labor contracts under asymmetric information: an introduction. Review of Economic Studies. 1983. Vol. 50. N 1. 3 - 35.

Mas-Colell A., Whinston M.D., Green J.R. Microeconomic theory. N.Y.: Oxford Univ. Press, 1995. - 981 p.

Mookherjee D. Optimal incentive schemes with many agents. *Review of Economic Studies*. 1984. Vol. 51. № 2. 433 - 446.

Myerson R.B. Game theory: analysis of conflict. London: Harvard Univ. Press, 1991.

Novikov D.A. Project management and incentive problems. *Proceedings of XII International Conference on Systems Science*. Wroclaw, 1995. 2. 387-391.

Novikov D.A. Incentives in organizations: uncertainty and efficiency. *Proceedings of 6-th IFAC Symposium on Automated Systems Based on Human Skills.* Kranska Gora. 1997. 274-277.

Novikov D.A. Incentive mechanisms in multi-agent and dynamic socio-economic systems. *Automation and Remote Control*. 1997. № 6. 3 - 26.

Novikov D.A. Incentives in socio-economic systems. Moscow: Institute of Control Sciences, 1998. (in Russian). 218 p.

Novikov D.A. Mechanisms of multilevel organizational systems functioning. Moscow: "Control problems" Foundation, 1999. 150 p. (in Russian).