

Digital recruitment of innovative project teams' leaders based on genetic algorithms

E E Kovshov¹, V S Kuvshinnikov¹ and L E Osipenko²

¹ JSC"NIKIMT – Atomstroy", Moscow 127410, Russia
² Moscow City University, Moscow 129226, Russia

KovshovEE@atomrus.ru

Abstract. Over the past decade, innovation has become an important part of the development for large business companies. The search for promising topics, as well as the recruitment of potential applicants teams who can implement an innovative project with the maximum benefit for business at the lowest cost, led to the search for approaches to solving these problems using IT. Since the problems of such research are covered separately in different industry areas, we implemented an approach to completing project teams based on genetic algorithms. A wide review of the literature allowed us to identify some important characteristics of innovative projects, as well as to summarize the portfolio of competencies required for potential designers. These attributive characteristics of the project and the applicants’ competence should be taken into account. The developed genetic algorithm allows searching for multiple parameters simultaneously and generating new solutions through the use of different operators. The results open up a new range of opportunities for business companies and government organizations interested in successful implementation of innovations.

1. Introduction

Currently, the economic performance of business companies and their ability to retain competitive industry advantages depend on their innovation potential and innovation activity [1-3]. Innovation management is a rather complex and multi-criteria task, since it includes an analysis of a significant number of such heterogeneous elements as human resources, financial and technological components of an innovation project (figure 1).

One of the central components of innovation management is the workforce. Postulating the thesis "Cadres are all-important", owners of business corporations adequately assess not only the importance of forming a cohesive, efficient project team, but also the search for alternative options of "efficiency of designers ― quality of innovative product" balance. The latter suggests that other experts, subject your material to such powerful checks that they can bring the race of evidence to the point where all your resources may not be enough to win [4].

This kind of "battle" increases the cost of proof race, multiplies the number of tests, changes the shape of new objects, thereby actualizing the need to find solutions that can track and analyze set of indicators and parameters described by big data and processed by complex algorithms.

A wide range of identified problems leads to the formulation of three important tasks: justification of the main attribute characteristics of an innovative project and a team of designers, as well as ways to manage them based on IT solutions.
We will reveal each of these tasks sequentially.

**Figure 1.** Structural components of the innovation management process [4].

2. **Attribute characteristics choice**

Speaking about the search of ideas for innovative projects by business companies, it should be noted that a characteristic feature of modern post-industrial society is the "transformation of science into technoscience". In the symbiosis "fundamental science ↔ technology → wealth of society", science is a generator of new ideas and principles for high-tech developments. Technology, in turn, creates new products and solutions that expand human capabilities [5].

The "applied science" trend has led to an increase in the part of technological developments in total scientific research volume, and as a result — the possibility of their typology based on the subject area nature into research, design and complex.

The commercialization of science and the gradual replacement of fundamentals with the criterion of efficiency and practical utility — determine the range of an innovative project targeted results, which varies from supporting, involving the improvement of existing knowledge and solutions, to universal and breakthrough.

Traditional typology based on the execution duration implies the presence of short-term, medium-term and long-term projects.

These main attribute characteristics of the innovation project led to the following discursive field for discussion of the project team. Formed like theater companies that are created temporarily for the production of a play, such project teams gather "performers" of different roles. The diversity principle assumes that the scientific productivity of each member of the project team will increase with the increase of its functions and specialties, generating nonlinear interaction between them and, as a result, not additive, but a more dramatic increase in the dependence of the collective result on the efforts of individual specialists [6, p. 226].

In this regard, we consider generalizing the main attributes of an innovative project team as a system object is important. The controlling mechanism of such team is the motivation of designers to get the maximum benefit for business while minimizing costs.

K. O. Sokolov identifies external and internal factors that ensure the effective work of the innovative project team. Among other things, he highlights managerial and organizational components, such as the employment of an applicant in other projects as a leader, expert, performer, as well as the assessment of the professional, personal and psychophysiological potential of members of the innovation team [7].

We believe that the functioning of an innovative project team will be successful if it has one or more specialists with an academic degree who are responsible for the formulation of problems, the construction of research design, as well as having sufficient authority and experience to attract the most effective designers [8].

Practice shows that in some cases the success of an innovative project is determined by the participation of a frontman; a designer with special competencies, such as technical or industrial expertise, etc. [9].
Group projects share the need for finding efficient ways to improve the effectiveness of communication between designers, for example, when discussing complex issues related to project design, clarifying and adding incomplete requirements.

For breakthrough innovation projects, it is important to justify the involvement of single inventors sitting in the "ivory tower", such as Edison or Wright.

Since the third-party IT specialists involvement is usually very expensive, and, like any designer, they have both pros and cons, and their performance in a particular innovative project depends on a variety of components, it is important to offer some objective abstract metric of "coolness" for them, too. We have accepted the main programming statuses as such attribute characteristics: Junior, Middle, and Senior.

The main professional requirement for a junior is the ability to independently and quickly perform the technical tasks assigned to him or her. In addition to energy and determination, it is important for juniors to have a desire to learn and develop, as well as the ability to calmly treat criticism.

The main requirement for a middle developer is an understanding of business requirements and the ability to translate them into specific technical solutions.

The highest-status IT specialist is senior. They should not only have extensive experience, a deep understanding of the libraries and frameworks structure, but also be able to assess the technical risks of the project. Senior has developed communication skills that allow him to recommend constructive ways to refactor code, as well as successfully convince the customer and the team. An example of an attributes set of an innovative project and its potential performers is shown in table 1.

Table 1. Example of an innovative project and its potential performers attributes set with values.

| Attribute                  | Possible attribute values | Numerical value | Relative attribute weight |
|----------------------------|---------------------------|----------------|--------------------------|
| **Innovative project**     |                           |                |                          |
| Classification            | Research                  | 1              |                          |
|                           | Project                   | 2              | 1                        |
|                           | Complex                   | 3              |                          |
| Term of execution         | Short-term                | 1              |                          |
|                           | Medium-term               | 2              | 2                        |
|                           | Long-term                 | 3              |                          |
| Target result             | Supporting                | 1              |                          |
|                           | Universal                 | 2              | 3                        |
|                           | Groundbreaking            | 3              |                          |
| **Potential candidate**    |                           |                |                          |
| Education                 | Secondary                 | 1              |                          |
|                           | Higher                    | 2              | 1                        |
|                           | Academic degree           | 3              |                          |
| Current employment in projects | 3 and more     | 1              |                          |
|                           | 2                         | 2              | 2                        |
|                           | 1                         | 3              |                          |
| Special innovative competences | No                   | 1              |                          |
|                           | Yes (common)              | 2              | 2                        |
|                           | Yes (project specific)    | 3              |                          |
| Experience in innovative development | Separated     | 1              |                          |
|                           | Single (1–2 projects)     | 2              | 3                        |
|                           | Significant (3 and more projects) | 3 | |
| «Teamwork»                | No (individual contributor) | 1           |                          |
|                           | Equally                   | 2              | 3                        |
|                           | Always                    | 3              |                          |
| Programmer status         | Junior                    | 1              |                          |
|                           | Middle                    | 2              | 3                        |
|                           | Senior                    | 3              |                          |
3. **Genetic algorithm building methodology**

As noted by a researcher at the Oxford Internet Institute and The Institute of Turing's Sandra Wachter, the so-called human resources analytics (people analytics) plays an increasingly important role in companies’ decisions on hiring, firing, evaluating performance and improving employees. [10]. However, it should be noted that the process of finding an equation that describes the optimal behavior of any natural system is not a simple task. According to Hoda Lipson, previously, it could have taken a scientist a lifetime to create one such optimization model [11].

Currently, new paradigms are used to solve optimization problems, which search for a balance between efficiency and quality of solutions due to the "survival of the strongest alternative solutions” in uncertain and fuzzy conditions [12–15] et al. One of these paradigms is genetic programming. M. Schmidt expressed its possibilities very vividly, noting that physicists (Newton or Kepler) could run a genetic algorithm on a computer and, after just a few hours of calculations, get laws explaining the fall of an apple or the movement of planets [16].

Genetic programming repeats the principles of biological evolution. The genetic algorithm (hereinafter referred to as GA) starts composing equations by randomly matching different building blocks consisting of mathematical expressions, and then checks how well the resulting equations describe the data. Equations that do not pass verification are rejected, and those that demonstrate a certain potential are recomposed in such a way that they can eventually turn out to be an accurate mathematical model [17].

GA is much more effective than traditional "regression" in some cases. For example, in a regression search, the form of the equation (dependency) is set in advance, and the parameters of the equation are optimized to match the data. However, in some cases it is necessary to search for the equations themselves, including arithmetic operators, trigonometric and logarithmic functions, constants, etc.

The effectiveness of GA for improving the quality of digital recruitment management of innovative project teams is determined by such features as: working with a coded set of parameters; simultaneous search among several alternatives on a given set of solutions; using the target function (hereinafter referred to as TF) to evaluate the quality of solutions; and a probabilistic approach to optimization problems analyzing.

Traditionally, the algorithm for evaluating of innovative project team members’ potential is a step-by-step procedure described below. It starts with the selection of characteristics that define the innovative potential of teams; selecting the appropriate methods of assessment; development of evaluation scenario; evaluation processing; interpretation of results; wording of conclusions; developing recommendations on the team acquisition and development (figure 2).

For the GA to work, select a natural parameters set of the optimization problem and encode them into a sequence of finite length in a certain alphabet. The algorithm cycle is repeated until the specified number of algorithm iterations is completed or a required quality solution is obtained on some generation, or if a local optimum is found and premature convergence occurs, when the algorithm cannot find a way out of the local extremum influence area.

In the team selection management optimizing process, the parameter space under study is quite large, and the task does not require finding of the exact global optimum. It is enough to find an acceptable solution from a set of the most suitable ones in a given short time. In this regard, it is most appropriate to use GA at stage 3 (figure 2). The first step is to determine the initial data of the specific innovation project designer selecting process (figure 2).

Each object (any vector element from table 1) can be represented as a set of attributes that characterize this object numerically. Attributes are defined on a limited set of positive values. Selection vector: \( SV = \{F_1, F_2, ..., F_l\} \), where \( F_i \) — frontman number, selected for i-th innovative project, innovative projects \( I = \{1, ..., N_i\} \), frontman candidates \( F = \{1, ..., N_F\} \).
Figure 2. Selection sequence of potential designers for an innovative project implementation.

Table 2. Source data for the genetic algorithm.

|   |   |
|---|---|
| 1. | Potential frontman $F = \{1, \ldots, N_F\}$, $N_F$ — number of potential front-runners |
| 2. | Innovative project $I = \{1, \ldots, N_I\}$, $N_I$ — number of innovative projects |

Thus, the problem of selecting a potential candidate for a particular innovative project performance reduces to the equivalent problem of finding such an SV of candidates which will best contribute to achieving the goal considering given limitations and criteria.

The number of innovative projects should not exceed the number of potential designers. One frontman should be assigned to a specific innovation project. The degree to which the frontman-candidate's attributes (FA) match the project's attributes (PA) is estimated by converting them to
universal attributes (UA) using a matching table. The table is formed depending on the specific case and allows many-to-many relationships. The attribute value is reduced to a scale of universal attributes that contains integers from 0 to 9. Depending on the difference between the corresponding attributes of the candidate and the project, a specific value is taken from a separate table of relative efficiency constants (REC). The value measures the potential compatibility of the candidate and the project.

For a more accurate assessment of the assignment effectiveness, the amount of work (laboriousness) for each of the project attributes (PL) is taken into account separately. The approximate target satisfactory amount of costs for the completion of the i-th project is defined by the formula (1):

$$SC_i = \sum_{k=1}^{N_i} \left( \frac{PL_{i,k}}{REC(RUAF_{j,k} - UAP_{j,k})} \right) + CA_i,$$

where $i$ — project number from $\{1, ..., N_i\}$; $j$ — number of a frontman candidate; $N_i$ — amount of attributes, $k$ — number of a universal attribute; $PL_{i,k}$ — complexity of i-th object's k-th attribute; $RUAF_{j,k}$ — level of the j-th candidate's k-th universal attribute, required for the i-th project's implementation; $UAP_{j,k}$ — the i-th project's k-th universal attribute; $REC()$ — function of relative effectiveness, which values for each attribute are pre-recorded; $CA_i$ — custom individual coefficient for the i-th project, allowing specific implementation requirements consideration. The $CA$ coefficient allows you to adjust the project's requirements for the frontman's abilities, due to increased urgency, execution quality requirements, or other requirements that increase the amount of resources required.

The expected project cost of the i-th project is defined by the formula (2):

$$PC_i = \sum_{k=1}^{N_i} \left( \frac{PL_{i,k}}{REC(UAF_{j,k} - UAP_{j,k})} \right) + CW_j,$$

where $i$ — project number from $\{1, ..., N_i\}$; $j$ — number of a frontman candidate; $N_i$ — amount of attributes, $k$ — number of a universal attribute; $PL_{i,k}$ — complexity of i-th object's k-th attribute; $UAF_{j,k}$ — k-th universal attribute number of j-th frontman candidate; $UAP_{j,k}$ — k-th universal attribute number of i-th innovative project; $REC()$ — function of relative effectiveness; $CW_j$ — estimated current workload of j-th candidate. The $CW$ factor represents the additional resource costs associated with assigning an already employed candidate to the project.

Each specie will be represented by a chromosome when initial population of GA is created. A chromosome is a selection vector $SV$, containing $N_i$ randomly ordered numbers of applicants according to the number of projects. Since a candidate can only be assigned to one project at a time, each gene (candidate's number) must occur no more than once in any relevant chromosome. This type of "enumerable chromosomes with unique genes" is often used in combinatorial problems. The standard crossing operator is incorrect for this type of chromosome, so a more complex two-point crossing scheme is used.

The effect of frontman candidates assigning is estimated for each selection vector separately using this equation (3):

$$AE_g = \sum_{i=1}^{N_i} LOSS(SC_i - PC_i),$$

where $g$ — number of selection in population; $N_i$ — amount of projects; $k$ — universal attribute number; $LOSS()$ — loss function, which defines the amount of resource losses caused either by a lack
or an excess of resources assigned to the project. Usually, the lack of resources is more dangerous than excess, because it marks the lack of specialist's experience, leads to additional delays, decrease in quality, and may cause failure to meet deadlines. However, assigning an overly valuable specialist to a simple project leads to indirect losses associated with overpayments for professionalism or an unreasonable risk of a temporary experienced personnel shortage if more complex projects are to appear.

To determine and implement the GA we define the following concepts.

**Target function** — numerically represents the result of frontmen selection. Corresponds to the effect of assigning AE.

**Population** — a set of $SV$ vectors.

**Population size** — total number of $SV$ vectors. The population size is set before the GA starts. It remains constant during the entire calculation process.

**Specie** — single $PV$ vector.

**Gene** — part of $PV$ vector.

**GA stop condition** — receiving a solution of the required quality; GA getting into the deep local optimum of the target function; specified search time expiration.

The next step involves defining characteristics and values of attributes (table 2) that determine candidate's and project's unified attributes' weights.

The GA operation sequence used for finding an optimal solution to the frontmen recruiting problem is shown in figure 3.

![Figure 3. The genetic algorithm flowchart.](image)

The given GA model is software implemented in the Jupiter Lab v2 programming environment using freely distributed libraries, modules and software components, mainly open source.

4. **Conclusions**

Thus, the interaction of pedagogy and IT can enrich and expand both sciences in the field of innovative project teams recruitment. Their pedagogical analysis as system objects, supplemented with the capabilities of genetic programming, together allows us to achieve the emergent properties of
innovative project teams, maximizing business benefits while minimizing costs. Clearly aware of the complex difficulties and limitations associated with picking innovative project teams using genetic algorithms, it is necessary to emphasize the possibility of extrapolation to any level of projects implementation, as well as to the rational and emotional spheres of applicants in digital recruitment.

References
[1] Davidson N, Mariev O and Pushkarev A 2018 A Regional factors of innovation activity of Russian enterprises *Foresight* 12 (3) 62–72
[2] Kashparov D and Kochin M 2017 From startup training as a training tool for developing innovations to the innovative economy of the region and country *Bulletin of Immanuel Kant Baltic Federal University* 4 52–9
[3] Kilov D 2018 Improving the efficiency of innovative activity of companies *Foresight* 12 (3) 30–3
[4] Latour Bruno 2013 *Science in action: How to Follow Scientists and Engineers through Society* (St. Petersburg: publishing house of the European University in St. Petersburg) p 414
[5] Mamchur E 2010 Fundamental science and technology: search for mechanisms of interaction *Modern Technologies: Philosophical and Methodological Problems* https://iphras.ru/uplfile/natsc/articals/mamchur/modern-tehnology.pdf
[6] Yablonsky A 2001 *Models and methods of research of science* (Moscow: Editorial URSS) p 400
[7] Sokolov K 2011 Evaluating the potential of an innovative team *Digest-Finance* 8 https://cyberleninka.ru/article/n/otsenka-potentsiala-innovatsionnoy-komandy
[8] Peltz D and Andrews M 1973 *Scientists in organizations. About optimal conditions for research and development* (Moscow: Progress Publishers)
[9] Rasskazova V 2019 Teambuilding in innovative startup projects *Scientific review Pedagogical science* 2-3 66–8 https://science-pedagogy.ru/ru/article/view?id=1898
[10] Ford M 2019 *Rise of the Robots: Technology and the Threat of a Jobless* (Future Alpina non-fiction) p 430
[11] Stevenson M 2011 *An Optimist's Tour of the Future: One Curious Man Sets Out to Answer What's Next?* (Avery No Edition Stated edition) p 384
[12] Krechetov I and Romanenko V 2020 Implementation of adaptive learning methods *The issue of education* 2 252–77
[13] Kasatkina E, Sairanov A, Nefedov D and Rusyak I 2019 In the Application of genetic algorithms for managing organizational systems in the event of emergency situations *Computer research and modeling* 11 (3) 533–56
[14] Chekanin V, Kovshov E and Hue N 2009 Increase of efficiency of evolutionary algorithms at the decision optimization tasks of objects packing *Control systems and information technologies* 37 (3) 63–7
[15] Koza J 2010 Human-competitive results produced by genetic programming *Genetic Programming and Evolvable Machines* 11 (3-4) 251–84
[16] National science Foundation Press release (National Science Foundation) 2009 *Maybe Robots Dream of Electric Sheep, But Can They Do Science?*
[17] Schmidt M and Lipson H 2009 Distilling Free-Form Natural Laws from Experimental Data *Science* p 324