Mathematical Methods of Decision Making for the Problem of Economic Diagnostics of IT Companies

A. E. Kvyatkovskaya, E. V. Chertina
State Federal-Funded Educational Institution of Higher Professional Training Astrakhan State Technical University
Astrakhan, Russia
zima00@list.ru, saprikinae_1912@mail.ru

***Abstract*— The paper considers the creation of mathematical support for carrying out for the task of economic diagnostics of enterprises necessary to evaluate its value. The object of research was selected IT-enterprises, incl. - start-ups that have special characteristics during the growth period. Based on the system analysis, a system of quantitative and qualitative characteristics has been developed to identify the economic state of the IT enterprise and start-ups in the external and internal environment. Scales of indices of various nature are determined. Methods are given that allow us to introduce order and equivalence relations for the peer companies found in order to compare their proximity to the analyzed company. Metrics used to compare companies are considered, taking into account the quantitative and qualitative characteristics.**

Keywords— start-up; business valuation; valuation methods; start-up indicators; economic status; precedents,; peer company; comparative method

I. Introduction

The task of valuations is one of the low-formalized tasks inherent in management in economic systems in conditions of uncertainty. One of the low-formalized tasks inherent in management of economic systems in a context of uncertainty. In our country for over ten years, the results of the valuation of various property objects are integral to the majority of decisions in the private and public sectors. In relation to the study, one of the valuation methods is considered - an analog method based on comparing the enterprise with the most suitable analogical enterprises, choosing the appropriate prototype and transferring its economic properties and trends to the object of research. The solution of this problem is actual both for the growth of the business area and for the growth of the information technology sector in the country, since there are more than 5000 small IT companies and a lot of venture funds and large IT companies interested in buying start-ups in Russia.

At the same time, the intensification of the valuation process through the introduction of information technology is hampered by imperfections or lack of information models and mechanisms that support this process:

- absence of constantly updated knowledge base about the peers required for comparison;
- absence of mechanisms for collecting information from any open source for supplementing the knowledge base;
- absence a frame of reference for peer companies that contain heterogeneous information;
- absence of justification of metrics for calculating the "proximity" of peer companies.

The research objective is to develop and substantiate the mathematical decision-making methods to estimate the value of a business based on an analog approach. The object of research is an IT-company or a start-up, which has during the growth period special values of economic characteristics that are not typical of ordinary companies.

II. Definition of the Characteristics

The set of criteria required for economic diagnostics of an IT company was determined by the example of a startup. A startup evaluation method depends on the stage of: preseed; seed; series A.

At the stage of Preseed, the assessment takes place at a fixed rate of a business angel or an accelerator, the main task of which is to speed up the delivery of early stage projects to the first investor, to refine and help them. It is rather difficult to structure the indicators at this stage, since the start-up does not have formal indicators that allow the construction of a financial model, but only meets the following requirements: an achievable market volume of at least 300 million rubles, deadline – 3-5 years; team of the project - at least two people; the presence of a working MVP (minimum viable product) - minimum viable product.

At the stage of Seed, the objective is to scale the business (increase the number of customers, customer segments, geography, etc.). The evaluation can be viewed from two sides, determining how much investment is needed, based on the team's costs per month and the investor's expectations through a specific time period. It is possible to use the indicators accepted in the international practice for the analysis of investment projects, for example, – NPV (Net present value)

Stage A is the stage of active growth and increasing of the company. At this stage, the following indicators are

The reported study was funded by RFBR according to the research project № 18-37-00130.

978-1-5386-5824-6/18/$31.00 ©2018 IEEE

highlighted: Cash-flow, multiplier, discount rate, scale-out limiters.

Comparing the formation of estimates in three stages, it must be taken into account that the accelerators note that the systematization of the estimation for the Preseed stage is an impossible task, since here the subjective assessment formed after personal communication with the creators is more significant. Therefore, we will consider the Seed & Series A stages. Therefore, we will consider the Seed & Series A stages.

The papers of B. Payne [1–2] and S. Nasser [3] are the most popular papers in this area of research, which are much talked in online research. They are devoted to the valuation of companies, including start-ups to various stages of investment.

To analyze the selected stages, we use five commonly used assessment methods of startups, summarizing the indicators on which they are based. The methods were determinate after undertaken studies in the largest business incubators in Russia, which mark the feasibility and adaptability of the selected methods to Russian conditions. It should be noted that most methods are based on data from comparable companies or basic estimates: the Berkus method, the method of summation of risk factors, the venture capital method, the discounted cash flow method, the comparison method.

The characteristics that generate the above methods are grouped as qualitative and quantitative, it was done for the subsequent structuring and scaling. In total, 15 quantitative and 14 qualitative indicators were selected, including 9 types of risk.

TABLE I. CHARACTERISTICS OF START-UP

Quantitative characteristics		**Qualitative characteristics**
	Unit	
Customer Acquisition Cost (CAC)	Rub.	Team evaluation
Cash-Flow	Rub	Scaling drivers
Multiplier	–	Scaling limiters
Market capitalisation	Rub	Strategic relationship
Backlog	Rub.	Product introduction or sales start
Operating profit	Rub.	Quality of the prototype
Sensible idea (cost base)	Rub.	Managerial risks
ROI (Return On Investment)	%	Risks at different stages of business development
Discount rate	%	Political risks
Expected growth rate	%	Marketing risks
Regular monthly income	Rub.	Risks related to financing / raising of capital
Number of persons employed	Piece	Litigation risks
EBITDA (Earnings before interest, taxes, depreciation and amortization)	Rub.	International risks
Gross profit	Rub.	Reputational risks
		Risks associated with a potentially profitable exit from a startup

The evaluation system of an IT company under the given set of characteristics will determine a point set in the criteria space that have a formal criterion representation. In order one company to serve as a good analog for other evaluation, it is desirable that they resemble in many characteristics, at the same time it is possible to prioritize, reinforcing the weight significance of a particular characteristic.

III. DETERMINATION OF THE METHOD FOR SELECTING A PEER COMPANY

For the selection of peer companies, we apply one of the decision-making methods – the method of case-based reasoning, using knowledge of known situations or cases (precedents), which in our case are peer companies. We define the set (IT) of IT companies considered in the selection of analogues. The information about a set IT is represented in the form $IT = \{it_i, i = \overline{1,n}\}$. To determine the properties-characteristics of each IT company it_i we compare a set of characteristics $K = \{k_j\}, j = \overline{1,m}$. Then each IT company can be represented in a form $it_i = \{f_1(k_1), f_2(k_2), ..., f_m(k_m)\}$, where $f_j(k_j)$ is a characteristic function that defines a subset $k_j^* \subseteq K$ or the i-th IT company.

Once the it_i peer companies are extracted, you need to select the "similarity" to the it^* precedent, describing the degree of proximity by the formula:

$$R(it_i, it^*) = \frac{\sum_{j=1}^{m} \rho(f_j^i(k_j), f_j^*(k_j)) \cdot w_j}{\sum_{j=1}^{m} w_j},$$

where $p(f_j^i(k_j), f_j^*(k_j))$ – a metric is calculated by m characteristics of analog and precedent $f_j^i(k_j)$ and $f_j^*(k_j)$; w_j - a degree of importance of the *j-th* characteristics.

The choice of the metric is the most difficult problem. The inhomogeneity of the characteristics does not allow us to introduce an algebra of operations on the given set. The most famous is the mathematical method of nearest neighbor [1], which is able to measure the degree of proximity for any characteristic.

$$mnear(it^*) = \underset{it \in IT}{argmax} \sum_{j=1}^{m} [f_j^i(k_j) = f_j^*(k_j)] \cdot w_j,$$

where $[f_j^i(k_j) = f_j^*(k_j)]$ - is an error indicator that takes a logical value to a number by the rule [false] = 0, [true] = 1.

For quantitative characteristics it is also possible to use Euclidean distance or the Manhattan metric, provided that all characteristics are reduced to a single measurement scale or normalized.

If the exact match of characteristics is not required (or it is not attainable), it is possible to use the Zhuravlev metric:

$$mzhur(it^*) = \sum_{j=1}^{m} if(|f_j^i(k_j) - f_j^*(k_j)| < \varepsilon), then\ 1,\ else\ 0),$$

where ε is a given level of deviation of j-characteristics of the analogue and precedent from each other.

The number of characteristics has an effect on output error, since the curse of dimension may arise: according to the law of averages, the sums of a large number of deviations are very likely to have very close values. This fact subsequently leads up to the need to form a set of informative characteristics, but will require retrospective observations for them to form a sample of data, to reveal the dependence or multicollinearity.

For qualitative characteristics, it is possible to use the measure of Hamming's similarity by determining the maximum number of matching characteristics of a precedent and an analogue. If you cannot enter a metric, various proximity measures are used.

After the database of precedents is formed in any way - manual or automated, it is possible to allocate relationships of order and equivalence for the objects filling it [2]. Using a geometric approach to the solution of this problem, the importance of which was stressed by D.A. Pospelov [3], it is possible to represent analogs and precedents as independent information objects and, in the future, to compare them both by individual characteristics and in general.

Analyzing analogues using the equivalence relation, the original set is divided into equivalence classes $[it^*] \subset IT$ of element $it^* \in IT$ in the form of a subset of elements equivalent to $it^* : [it^*] = \{it \in IT \mid it \sim it^*\}$.

The classes of analogs can represent both nominal and ordinal scales. In the first case, they can be constructed in two ways: by clustering and using expert estimates. In the second case it is possible to use the partitioning of the original set into Pareto classes with subsequent ordering of these classes.

When analogues analyzed using the order relationship, precedents are arranged by rank in the absence of an accurate analog. Let's highlight the following decision-making tasks, using the ranking of analogues along the proximity to the precedent:

- the task of ranking analogs based on knowledge of their states at a given time $t^*(it_i, k_j), i = \overline{1,n}, j = \overline{1,m}$;
- the task of ranking analogues based on knowledge of their states at different times (for example, corresponding to the stages) $it^*(t_g, k_j), \mathrm{g} = \overline{1,\mathrm{s}}, j = \overline{1,m}$;
- the task of ranking analogues according to a given characteristic $k^*(it_i, t_g), i = \overline{1,n}, g = \overline{1,s}$;
- the task of ranking analogues on aggregate characteristics $k(it_i, t_g), i = \overline{1,n}, g = \overline{1,s}$.

In the latter case, the equal importance of characteristics is considered when the decision-maker can or cannot reliably establish priorities between them. In the case of equal characteristics, a set of incommensurable undominated alternatives are formed - the Pareto ITP set. Thus, in the case of the solution is selected not just one but many peers, which ultimately makes the final decision difficult. In this case, apply mathematical methods that narrow the Pareto set, for example, the method of median distributions [4, 5]. The advantage of the method is the combination of qualitative and quantitative assessments.

It is also possible to construct various functions for selecting $C^K(IT)$ and $C^D(IT)$ in case the absence of information about the relative importance of characteristics and the availability of characteristics of both quantitative and qualitative type. They narrow the Pareto set and take into account only the mutual relations between the estimates of the analogs without taking into account the absolute values of the differences in the estimates by characteristics.

For two analogues $it_i, it_l \in IT, i,l = \overline{1,n}$ we define the number of characteristics by which it_l has more proximity to it^* than it_i. For analogs whose maximum is this number, we define on the IT-set a numerical function $q_{i,l} = q(it_i, it_l)$ taking values corresponding to the maximal numbers found, where $q(it_i, it_l)$ is the number of characteristics over which it_l exceeds the variant it_i, in other words, is closer to the precedent.

A choice function for the C^K was constructed, considering the number of dominant characteristics of the analogue, which are close to the precedent, choosing the maximum values of the row of the matrix $Q_{IT} = \{q_{i,l}\}$ and then separating the minimal of them:

$$C^K(IT) = \{it_i \in IT \mid i \in Arg\min_i q_i, i = \overline{1,n}\},$$

where $q_i = \max_l q_{i,l}$.

As a result, a subset of analogues is formed, which have a greatest number of characteristics close to the precedent. The resulting subset has less potency than the Pareto set, and $C^K(IT) \subseteq IT^P$.

Consider the second method of generating analogues, closed to precedent using Q_{IT} matrix. The dominant index of the set IT was defined, equal to $\min_{it \in IT} Q_{IT}(it)$. The value of the choice function $C^D(IT)$ is a subset of all variants of $it \in IT$ with a minimum IT dominant index:

$$C^D(IT) = \{it^* \in ALT \mid Q_{IT}(it^*) = \min_{it \in IT} Q_{IT}(it)\}.$$

A circular n-tournament selection function C^T was constructed:

$$C^T(IT) = \{it^* \in IT \mid QM_{IT}(it^*) = \min_{it \in IT} QM_{IT}(it)\},$$

where $QM_{ALT}(it_i) = \sum_{l=1}^{n} q_{i,l}$.

This function also narrows the Pareto set, forming a subset of analogues close to the precedent, with $C^T(IT) \subseteq C^K(IT) \subseteq IT^P$.

IV. Conclusion

The paper studies the matters of economic diagnostics of IT companies and startups in the task of business valuation, based on the method of reasoning on precedents and comparison of peer companies. The characteristics are selected, and the choice of metrics and proximity measures for quantitative and qualitative characteristics of analogs are considered. Mathematical methods are given, that order set of analogues close to the precedent.

The proposed methodology was used as an aid for valuation of IT projects of a business incubator LIFT (the technical park FABRIKA, Astrakhan city). After that, conclusions were drawn about the completeness of the metrics, which describe the process of valuation for start-ups; the adequacy of the proposed methods; the practical relevance of the research results were made.

REFERENCES

[1] Payne B. Methods for Valuation of Seed Stage Startup Companies. Available at: http://www.angelcapitalassociation.org/blog/methods-for-valuation-of-seed-stage-startup-companies/ (accessed: 14.04.2018).

[2] Payne B. Startup Valuations: The Risk Factor Summation Method. Available at: http://billpayne.com/2011/02/27/startup-valuations-the-risk-factor-summation-method-2.html (accessed: 14.04.2018).

[3] Nasser S. Valuation For Startups – 9 Methods Explained. Available at: http://medium.com/parisoma-blog/valuation-for-startups-9-methods-explained-53771c86590e/ (accessed: 14.04.2018).

[4] Anand S.S., Hughes J.G., Bell D.A., Hamilton, P. Utilising Censored Neighbours in Prognostication. Workshop on Prognostic Models in Medicine. Eds. Ameen Abu-Hanna, Peter Lucas. Aalborg, (AIMDM'99). Denmark. 1999.,pp15-20.

[5] Karpov L.E., Yudin V.N. Methods of data mining when building a local metric in output systems based on use cases. Institute of System Programming, Russian Academy of Sciences. Preprint, 2006. (in Russian)

[6] Pospelov D. Modeling of reasoning. Experience in the analysis of mental acts. Moscow. Publ. Radio and Communication. 1989. P.184. (in Russian)

[7] Kosmacheva I. Algorithms of Ranking and Classification of Software Systems Elements. Kosmacheva I., Kvyatkovskaya I.Y., Sibikina I., Lezhnina Y. Knowledge-Based Software Engineering: Proceedings of 11th Joint Conference, JCKBSE 2014; Volgograd State Technical University [etc.]. Springer International Publishing, 2014. Series: Communications in Computer and Information Science; Vol. 466. Pp. 400-409.

[8] Pham Quang Hiep. Methods and Algorithms of Alternatives Ranging in Managing the Telecommunication Services Quality. Pham Quang Hiep, Kvyatkovskaya I.Y., Shurshev V.F., Popov G.A. Journal of Information and Organizational Sciences. 2015. Vol. 39, no. 1, pp. 65–74.