OPTIMAL ORGANIZATIONAL HIERARCHIES IN FIRMS

Sergey Mishin

Russian Academy of Sciences, Institute of Control Sciences, 65 Profsoyuznaya St.,
117997 Moscow, Russia. Tel.: +7-095-334-9051; fax: +7-095-334-9331.
E-mail: smishin@newmail.ru

Abstract. One of the main problems in modern economy is to construct an efficient organizational hierarchy allowing to control the firm with minimal cost. This paper describes the mathematical model of optimal hierarchies in firms. Optimal hierarchies for several classes of cost functions are obtained. Particularly, sufficient conditions for tree optimality, 2hierarchy (any manager has two immediate subordinates) optimality and two-tier hierarchy optimality are defined.

Keywords: optimal hierarchy, control cost, structure, firm.

1. Introduction

Any firm is an organization of economic agents (employees)[1]. In the organization employees conform to some rules (mechanisms) regulating their activity and providing the achievement of the general goal of the firm.

The employees in the organization are specialized. Therefore, they are more efficient than the set of self-employed (non-organized) agents. But the employees with different specialization must be coordinated to achieve the general goal. Coordination is a fundamental problem of any organization because activity of a team must be planned and monitored, individual goals must be coordinated, etc. Some organizational hierarchy[2] is created to fulfill the coordination functions (administrative labor) in the firm.

---

[1] Below we use the terms "organization" and "firm" as synonyms.

[2] The employees on higher tiers of the hierarchy have more authority than the employees on lower tiers. It allows to control the firm even when conflicts between the employees exist.

On the one hand, the hierarchy increases efficiency of the employees' interactions (for example, due to the planning and monitoring informational, material and other flows). On the other hand, the performance of coordination (control) functions is costly. In modern economy organizations become increasingly more complex. As a result, the proportion of managers in organizations may exceed 40 % (see, for instance, Radner, 1992). So, the key factor of firm's efficiency is the optimality of the hierarchy.

Two-tier hierarchy can be optimal for small firms. In this hierarchy workers on the first (lowest) tier are immediately subordinated to a single manager. As the firm grows, the single manager can not control all interactions between the workers. Therefore, one has to hire several managers to the second tier of the hierarchy and to delegate them the responsibility to control business interactions (flows) within the subordinated groups of workers. But interactions between subordinated groups cause interactions between the managers on the second tier. Several managers on the third tier must control these interactions, etc. In such a way multi-tier hierarchy arises. A superior manag-

er in the hierarchy has an authority over his or her subordinates (managers or workers) and a subordinate provides the information to and follows the instructions from his or her superiors.

The design of the hierarchy is one of the aspects of organization design. The process of organization design (and re-engineering) is divided into three phases[3] (see, for instance, Mintzberg, 1979, Williamson, 1975):

I. Technology design: the number of workers, their functions and interaction rules are determined.

II. Hierarchy design: the number of managers and their subordination are determined.

III. Mechanism design: superiors' authorities over their subordinates are determined[4].

Typically, an expert in the appropriate field performs the technology design (phase I).

There are many mathematical models of control mechanisms (phase III). Two-tier hierarchy mechanisms (principal-agent problems) have been researched in detail (see, for instance, Hart and Holmstrom, 1987; Grossman and Hart, 1982 and 1983). There exist the models of control mechanisms in some types of multi-tier hierarchy (e.g. Melumad, Mookherjee and Reichelstein, 1995) explore the delegation mechanism in three-tier hierarchy).

In this paper we concentrate our attention on phase II. Several papers are focused on the hierarchy optimization problem (phase II) or joint optimization of hierarchy and mechanisms (phases II and III). The study of the hierarchic organizations was pioneered by Simon (1957). His model is based on the following assumptions:

1. The employees on the first tier are the only workers performing production labor. All employees on higher tiers are managers performing only administrative labor (control functions).
2. Any employee in the hierarchy has the only immediate superior on the next hierarchical tier. Thus, any hierarchy is a tree. And only employees on adjacent tiers may interact directly.

3. The wage is the same for all employees on one tier. The span of control (the number of manager's immediate subordinates) is the same too. So, employees on one tier are assumed to be identical.
4. The span of control is the same on different tiers of the hierarchy.
5. The wage on the next tier is a constant multiple of the wage on a previous tier. The constant is an exogenous number, which does not depend on the tier and other parameters of the hierarchy.

Williamson (1967) explores a similar model and proves that firm size is limited because of "loss of control" (employees' efficiency decreases from an upper tier to a lower tier). The interlayer efficiency differential is an exogenously given constant. Calvo and Wellisz (1978) explain the wage and the efficiency endogenously. Employee's efficiency depends on his or her wage and the span of control of the immediate superior. The larger the manager's span of control is, the less is his or her subordinates' effectiveness, as individual subordinate is controlled rarely. Using this assumption Calvo and Wellisz (1979) consider the profit maximization model. The profit equals the difference between income (the number of workers multiplied by their effectiveness) and total wages of all employees. In this model both different spans of control and wages on different tiers are possible. Thus, Calvo and Wellisz dispense stringent assumptions 4 and 5 and prove important principles, for example, that in the optimal hierarchy the higher tier, the more employee's efficiency and wage per efficiency unit.

Keren and Levhari (1983) optimize the hierarchy's decision-making time[5] (delay on each tier equals the span of control plus constant). Average cost per employee is calculated for the hierarchy with minimal decision-making time. This cost allows to calculate the limits of the firm's size. Similar information processing models are explored in numerous papers (see, for example, Van Zandt, 1996; Bolton and Dewatripont, 1994; Radner, 1993).

Qian (1994) explores Calvo and Wellisz (1979) model by using optimal control techniques, a method pioneered by Keren and Levhari (1979). Continuous approximation is considered (continuous number of employees on each tier). In this case, the optimiza-

---

[3] In practice these three phases may not be altogether independent. But it is rather difficult to optimize all these phases at once. To simplify the problem each phase is usually considered separately.

[4]   For example, employees' rights and responsibilities are determined.

[5] Marschak and Radner (1972) study the effect of delay on the value of decisions. This is one of the first models of hierarchy with managers calculating some "decision" (control action).

tion problem is simpler than discrete problem[6]. If any employee's effort choice is restricted to only zero or maximal effort, then in Calvo and Wellisz model the optimal employee's wage depends only on the span of control of his or her immediate superior[7]. To maximize profit one has to minimize total wages because employees' efficiency (efforts) is maximal. In this case Qian (1994) obtains the optimal hierarchy[8].

Like Qian, in this paper we consider the problem of searching out optimal hierarchy, which minimizes total wage of employees (total cost). However, we differ from Qian and other cited above papers in two important respects. First, we consider manager's wage function depending not only on the span of control, but also on sets of workers controlled by the employees immediately subordinated to the manager. So, manager's wage depends on "specificity" and "complexity" of manager's administrative labor (such wage function is called "*sectional*" in this paper). Thus, we do not assume that employees on one tier of the hierarchy are identical. Second, we consider not only tree-like hierarchies, but also more complex hierarchies with multiple subordination or cross-tier subordination[9]. Therefore, we differ from papers, cited above, because we dispense assumptions 2 and 3 (this paper is based only on assumption 1). The problem of searching out optimal hierarchy considered in this way is much more complicated. To explore this problem we base on the additional assumption: any hierarchy provides the maximal efficiency of employees. In this case to maximize profit we have to find a hierarchy with minimal total wage (total cost). Thus, control mechanisms (phase III) are not considered and manager's wage (cost) function is given exogenously[10]. We suppose that if employee's wage equals to the cost then his or her efficiency is maximal. Par-

ticularly, we entirely abstract from incentive problems[11].

Using an example of sectional cost function one can explore optimality of divisional, functional or matrix hierarchy[12] and prove many dependences considered without formal proof in management science literature[13]. So, the proposed model explains some effects in real firms. Sectional functions are also interesting from the mathematical point of view: any additive (with respect to manager's addition) and anonymous (with respect to manager's permutation) hierarchies' cost function can be represented in sectional form (Mishin and Voronin, 2003).

In this paper we explore optimization methods that can be used to obtain the optimal hierarchy for numerous classes of sectional cost functions regardless of function's specificity and practical interpretations. Particularly, we define sufficient conditions for tree optimality, 2hierarchy (any manager has two immediate subordinates) and two-tier hierarchy optimality. So, the proposed approach allows to construct the theoretical methods, which can be used to solve many problems that have numerous applications in economics. Therefore, the sectional cost function appears to be a useful compromise between detailed description of the real firms and possibility of mathematical modeling.

In the next section we describe the model and consider simple examples. In Section 3 we explore arbitrary sectional cost function and solve the problem

---

[6] Van Zandt (1995) examines the validity of continuous approximation of discrete problem of searching out optimal hierarchy.

[7] Suppose any employee works at full efficiency or shirks. In this case the employee compares expected loss of wage (the wage multiplied by the loss probability) and shirked time utility. To induce the employee to work efficiently one should calculate such wage that expected loss is greater than or equal to the utility. Loss probability inversely depends on span of control of the immediate superior. Therefore, optimal wage linearly depends on the superior's span of control.

[8] Also Qian (1994) explores more complex cases.

[9] It allows to prove insightful optimality conditions for tree, symmetric tree, etc.

[10] In this paper we consider different cost functions. For example, these functions may be defined using technological network (the result of the phase I) and possible controlling mechanisms (the result of the phase III).

[11] It is easy to create incentive mechanism under complete information: costs of maximal efficient employees are compensated and wages of other employees equal to zero (Mishin, 2004).

[12] Comparison of divisional (M-form (multi-divisional form)), functional (U-form (unitary form)) and matrix hierarchies is well-known aspect of hierarchy optimization problem. Advantages and disadvantages of these types of hierarchy are discussed in many papers (see, for example, Mintzberg, 1979). Recently developed models allow to compare mathematically divisional, functional and matrix hierarchies. For example, Maskin, Qian and Xu (2000), Qian, Roland and Xu (1997) explain mathematically advantages of the divisional hierarchy over the functional hierarchy.

[13] For example, Mishin (2005) shows that divisional, functional or matrix hierarchy is optimal for any size of the firm in some circumstances; managers on lower hierarchical tiers must control the most intensive flows because it helps to decrease the strategic managers' costs; the matrix hierarchy is stable with respect to standardization and stability decrease, on the contrary, divisional and functional hierarchies are stable with respect to standardization and stability increase; the divisional hierarchy is stable with respect to horizontal integration and production volume increase, on the contrary, the functional hierarchy is stable with respect to vertical integration and functional links intensity increase.

of searching out optimal hierarchy for several cases. In Section 4 we use these results to analyze cost functions corresponding with different types of interactions between manager and immediate subordinates. Possible extensions of the introduced model are discussed in Section 5. All mathematical proofs can be found in Appendix A.

2. The model of optimal hierarchy

2.1. Workers and managers. Hierarchies

Let $N = \{w_1,...,w_n\}$ be a set of *workers* who can interact with each other. Typically we denote the workers as $w, w', w'' \in N$. In this paper we suppose that the set of workers N is given and fixed.

Let *M* denote a finite set of *managers* who control workers' interactions. Typical managers will be denoted as $m, m', m'', m_1, m_2, ... \in M$. Let $V = N \cup M$ denote a set of all *employees* of the firm (workers and managers). For each manager we need to define his or her subordinates (workers or other managers). Let us define a set of *subordination edges* $E \subseteq V \times M$. Any edge $(v,m) \in E$ means that the employee $v \in V$ is an *immediate subordinate* of the manager $m \in M$. Thus, the edge is directed from the immediate subordinate to the *immediate superior*. An employee $v \in V$ is *a subordinate* of the manager $m \in M$ (manager *m* is *a superior* of the employee *v*), if there exists a path from *v* to *m*.[14] We will say that any superior *controls* his or her subordinates (any subordinate *is controlled* by his or her superiors).

Definition 1. A directed graph $H = (N \cup M, E)$ with a set of managers *M* and a set of subordination edges $E \subseteq (N \cup M) \times M$ is the *hierarchy* controlling the set of workers *N* if *H* is acyclic, any manager has at least one subordinated employee and some manager controls all workers. Let $\Omega(N)$ denote *the set of all hierarchies*.

Definition 1 ex ante excludes graphs with cycles (each manager in a cycle is a superior and subordinate of another managers in the cycle, which contradicts the main point of the term "subordination") and "managers" without subordinates. According to Definition 1 there exists a manager controlling all workers. Therefore, any set of workers has a common superior and any hierarchy is able to control all workers' interactions.

Any nonempty set of workers $s \subseteq N$ will be called a *group of workers*. We can start from a manager *m* and consider his or her immediate subordinates. After that we can consider their immediate subordinates, etc. Acyclicity implies that finally we determine nonempty set of workers subordinated to the manager *m*. This set $s_H(m) \subseteq N$ will be called *manager's m subordinated group of workers* in some hierarchy *H*. In other words, the manager *m controls the group of workers* $s_H(m)$. We will leave out inferior index "*H*" in notation $s_H(m)$ if it is clear what hierarchy we analyze. It will be convenient to think of worker $w \in N$ as having a subordinated "group" $s_H(w)=\{w\}$ which consists of this worker only. In other words, any worker $w \in N$ "controls" the elementary group $s_H(w)=\{w\}$ in any hierarchy $H \in \Omega(N)$.

In Fig 1 a hierarchy is constructed over the horizontal plane that corresponds with workers. In the figure the part of hierarchy subordinated to the manager *m* is shown. This part consists of immediate subordinates of the manager *m* and his or her subordinates not controlled immediately. The subordinated group of workers $s_H(m)$ is outlined by ellipse.
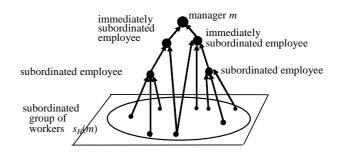

Fig 1. An example of manager and subordinated group of workers

Below we prove two technical lemmas which are necessary only to prove another proposition.

Lemma 1. *For any hierarchy H and any manager* $m \in M$ *the equality* $s_H(m) = s_H(v_1) \cup ... \cup s_H(v_k)$ *holds, where* $v_1,...,v_k$ *are all immediate subordinates of the manager m. For any employee v subordinated to the manager m the inclusion* $s_H(v) \subseteq s_H(m)$ *holds.*

Consider an example. In Fig 2 a manager *m* has two immediate subordinates m_1 and m_2. The group of workers $s(m) = \{w_1, w_2, w_3, w_4\}$ is subordinated to the manager *m*. The groups $s(m_1)=\{w_1,w_2\}$ and $s(m_2) = \{w_3, w_4\}$ are subordinated to the managers m_1 and m_2 respectively. Thus, the group $s(m)$ is divided into the subgroups $s(m_1)$ and $s(m_2)$:

---

[14] There exists such sequence of managers $m_1, m_2, ..., m_k \in M$ that the employee *v* is an immediate subordinate of the manager m_1 ($(v,m_1) \in E$), the manager m_j is an immediate subordinate of the manager m_{j+1} ($(m_j, m_{j+1}) \in E$) for each $1 \le j \le k-1$, $m_k = m$.
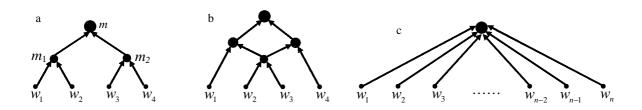
Fig 2. a) An example of 2-tree, b) An example of non-tree hierarchy, c) Two-tier hierarchy

$\{w_1, w_2, w_3, w_4\} = \{w_1, w_2\} \cup \{w_3, w_4\}$. In this example subgroups do not overlap. In general case subgroups can intersect (see Fig 2 b).

Definition 2. A hierarchy H is a *tree*, if only one manager m has no superiors and all other employees have exactly one immediate superior. The manager m will be called the root of the tree.

An example of the tree is shown in Fig 2 a. Hierarchy in Fig 2 b is not a tree because one manager has two immediate superiors.

Lemma 2. *In any tree any manager's immediate subordinates control non-overlapping groups.*

Thus, in the tree the immediate subordinates of any manager do not "duplicate" each other (do not control the same worker).

Definition 3. A hierarchy H is *r-hierarchy* if any manager has no more than r immediate subordinates, where $r > 1$ is some integer number. If *r*-hierarchy H is a tree then H will be called *r-tree*.

The term *span of control* (the maximal number of immediate subordinates, which can be controlled by one manager) is often used in practice. In our terms, if the span of control equals r then the hierarchy is *r*-hierarchy. Lemma 2 implies that immediate subordinates of any manager in a tree control non-overlapping groups. Thus, the maximal number of immediate subordinates equals n (if all immediate subordinates are workers). So, the span of control in any tree does not exceed n. And *two-tier hierarchy* with single manager controlling all workers (see Fig 2 c) has the maximal span of control.

2.2. Sectional cost functions. Optimal hierarchies

Definition 4. Cost function of the manager $m \in M$ in a hierarchy $H = (N \cup M, E) \in \Omega(N)$ is called *sectional* if it is given by:

$$c(s_H(v_1), \ldots, s_H(v_k)), \qquad (1)$$

where $v_1, \ldots, v_k$ are all immediate subordinates of the

manager m, $s_H(v_1), \ldots, s_H(v_k)$ are groups controlled by employees, $v_1, \ldots, v_k$, $c(\cdot)$ is a nonnegative real function of set of groups.[15] Cost of total hierarchy equals to total managers' costs[16]:

$$. \qquad c(H) = \sum_{m \in M} c(s_H(v_1), \ldots, s_H(v_k)) \qquad (2)$$

A hierarchy $H^* \in Arg \min_{H \in \Omega(N)} c(H)$ with minimal cost is called the *optimal hierarchy*.

Several optimal hierarchies may exist. This paper focuses on the problem of searching out *some* optimal hierarchy (we will reference to it as to "*optimal hierarchy problem*"). So, for given cost function[17] we need to search out an optimal hierarchy (the number of managers and their subordination) from $\Omega(N)$, which minimizes the cost of control of the workers.

Let us explain Definition 4 using an example (see hierarchy in Fig 2 a). The manager m controls the group $\{w_1, w_2, w_3, w_4\}$ with the help of two subordinated managers m_1 and m_2. Managers m_1 and m_2 control the groups $\{w_1, w_2\}$ and $\{w_3, w_4\}$ respectively. Suppose managers m_1 and m_2 cope with controlling of the subordinated employees. In this case the cost of the manager m does not depend on controlling methods inside the groups $\{w_1, w_2\}$ and $\{w_3, w_4\}$. For example, the managers m_1 and m_2 can control subordinated workers immediately or with the help of some subordinated managers. It is of no importance

---

[15] The function $c(\cdot)$ depends on the set $\{s_H(v_1), \ldots, s_H(v_k)\}$ of groups and does not depend on order of these groups. So, the manager's cost does not depend on numeration of his or her immediate subordinates $v_1, \ldots, v_k$. Some of groups $s_H(v_1), \ldots, s_H(v_k)$ may be the same. In this case the "set" $\{s_H(v_1), \ldots, s_H(v_k)\}$ contains repeated elements.

[16] In expression (2) and below the symbol $c(\cdot)$ denote both manager's and hierarchy's cost.

[17] Cost function may be determined directly (for example, using accounting information about manager's cost). Moreover, some "typical" cost functions may be considered (see examples below).

for manager's *m* cost because direct interactions between *m* and workers are not necessary. Definition 4 implies that the cost of the manager depends only on division of subordinated group of workers between immediately subordinated employees. In the example the group $\{w_1, w_2, w_3, w_4\}$ is divided into subgroups: $\{w_1, w_2, w_3, w_4\} = \{w_1, w_2\}\{w_3, w_4\}$. So, the cost of the manager *m* equals $c(\{w_1, w_2\}, \{w_3, w_4\})$. Thus, we suppose that the cost of a manager depends only on the "*section*"[18] controlled by the manager immediately. In Fig 2 a the "section" of the manager *m* consists of *m* and subordinated managers m_1 and m_2. The cost of the manager does not depend on other parts of the hierarchy, on individual efficiency of managers. Generally the cost of the manager can depend on individual efficiency, hierarchical tier, superiors or the whole hierarchy. Such cost functions are not sectional and are not considered in this paper.

In Definition 4 some of the groups $s_H(v_1)$, ..., $s_H(v_k)$ can be the same or nested one into another. Suppose $s_H(v_1) \subseteq s_H(v_2)$. So, the employee v_1 controls part of the group subordinated to the employee v_2. Thus, one immediate subordinate of the manager *m* duplicates part of the labor of another immediate subordinate. Below we consider only sectional functions satisfying the following assumption:

$$
\begin{aligned}
c(s_H(v_2), ..., s_H(v_k)) \le \\
c(s_H(v_1), s_H(v_2), ..., s_H(v_k))
\end{aligned} \tag{3}
$$

for any groups $s_H(v_1) \subseteq s_H(v_2)$. For example, "auxiliary" immediate subordinate v_1 can waste manager's *m* time discussing some problems inside the group $s_H(v_2)$ (such problems are completely controlled by the manager v_2). So, we can remove subordination edge (v_1, m) with no increase of manager's *m* cost. After removal costs of other managers do not change because groups controlled by all managers do not change.

In some cases in this paper the sectional cost function is given by simplified notation $c(s_1, ..., s_k)$ instead of $c(s_H(v_1), ..., s_H(v_k))$. The value of the function $c(s_1, ..., s_k)$ corresponds with the cost of some manager with immediate subordinates controlling the groups $s_1, ..., s_k$.

2.3. Examples of hierarchies controlling technological interactions

Let us consider several examples of sectional cost function depending on technological flows between workers. Consider a *flow function*, which is given by:

$$
f : (N \cup \{w_{env}\}) \times (N \cup \{w_{env}\}) \to R_+^p \tag{4}
$$

where w_{env} is an *environment* interacting with the workers. Thus, for any pair of workers $w', w'' \in N$ vector $f(w', w'')$ means the *flow intensity* between w' and w'' (*p*-dimensional vector with nonnegative real components). Each component is an intensity of one type of workers interactions or one type of flow (e.g., material, informational or other type of flow). Flow function *f* will be called weighted *technological network f*.[19] For any $w \in N$ the value $f(w_{env}, w)$ is a flow intensity between the worker *w* and the environment. We suppose that the technological network is undirected ($f(w', w'') = f(w'', w')$ for any $w', w'' \in N \cup \{w_{env}\}$). There is no *link* between w' and w'' if and only if $f(w', w'') = 0$ (w' and w'' are *linked* if and only if there are some flows between w' and w'').

Consider an example of managers' administrative labor. Suppose a conflict causes violation of interaction between workers w_2 and w_3 (see the hierarchy in Fig 2 a). The worker w_2 informs the immediate superior m_1 about this interaction problem. The manager m_1 can not solve the problem because the worker w_3 is not subordinated to m_1. Similarly the manager m_2 can not solve the problem after reception of worker's w_3 information. As a result, managers m_1 and m_2 inform their common immediate superior *m* about the problem. The manager *m* makes some decision. Managers m_1 and m_2 pass this decision to the workers w_2 and w_3. In such a way the interaction problems are eliminated. Therefore, the manager *m* controls the flow $f(w_2, w_3)$ and managers m_1 and m_2 participate in this flow control. This example shows that a manager fulfills "obligations" of two following types:

1. The manager controls such flows within subordinated group that are not controlled by subordinated managers. For example, in Fig 2 a the manager *m* controls the flow $f(w_2, w_3)$. The sum of such flows will be called the *internal flow* of the manager *m* and denoted $F_H^{int}(m)$.

---

2. The manager participates in control of the flows between the subordinated group and all other workers, the flows between the subordinated group and the environment. For example, in Fig 2 a the manager m_1 participates in flows $f(w_{env}, w_1)$ and $f(w_2, w_3)$ control. The sum of such flows will be called the *external flow* of the manager m and denoted $F_H^{ext}(m)$.

Obviously the values

$$F_H^{int}(m) = \sum_{\substack{\{w',w''\} \subseteq s_H(m), \\ \{w',w''\} \not\subset s_H(v_j) \text{ for each } 1 \le j \le k}} f(w',w'')$$

and

$$F_H^{ext}(m) = \sum_{\substack{w' \in s_H(m), \\ w'' \in (N \setminus s_H(m)) \cup \{w_{env}\}}} f(w',w'')$$

depend only on groups $s_H(v_1), \ldots, s_H(v_k)$ controlled by all immediate subordinates of the manager m in the hierarchy H. Therefore, we can consider the following sectional cost function depending only on total manager's flow:

$$c(s_H(v_1), \ldots, s_H(v_k)) = \varphi(F_H^{int}(m) + F_H^{ext}(m)),$$

where $\varphi: R_+^p \to R_+$ is non-decreasing function from R_+^p to R_+. The fact that the function $\varphi(\cdot)$ is non-decreasing means that manager's cost does not decrease when the "volume" of labor increases.

2.3.1. The expediency of multiple-subordination

Consider technological network with four workers and the following intensities of flows: $f(w_{env}, w_1) = 3$, $f(w_1, w_2) = 1$, $f(w_2, w_3) = 5$, $f(w_3, w_4) = 1$, $f(w_4, w_{env}) = 3$ (see Fig 3). This network may be interpreted as a *process line* ("business process"): the worker w_1 gets raw materials from the vendors and passes it to the worker w_2. The worker w_2 executes some production operation and passes the results to the next worker w_3, etc. The last worker w_4 dispatches finished products to the customers. Intensity changes (3,1,5,1,3) may be caused by the specific nature of interactions
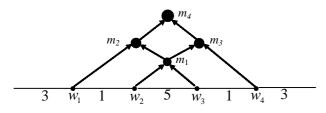

Fig 3. Optimal hierarchy controlling asymmetric process line

between different workers. If all intensities are the same then the process line will be called *symmetric*.

Consider manager's cost function $\varphi(x) = x^3$, where x is the value of the manager's flow. It is easy to see (Mishin, 2005) that the optimal hierarchy H looks like the hierarchy in Fig 3 and all trees are non-optimal[20]. The manager m_1 has two immediate superiors. So, there exists optimal hierarchy with multiple-subordination.

One of interesting questions is the optimality of trees. Tree is a typical hierarchy for many real firms. Considered example shows that in some cases there does not exist optimal hierarchy among the trees. Mishin (2005) proves the optimality of the tree for any symmetric process line and any non-decreasing function $\varphi(\cdot)$. Below we consider a more general sufficient condition for optimality of the tree.

2.3.2. Firm growth with control cost decrease

Consider the asymmetric process line with four workers, the flows $f(w_{env}, w_1) = 1$, $f(w_1, w_2) = 5$, $f(w_2, w_3) = 1$, $f(w_3, w_4) = 5$, $f(w_4, w_{env}) = 1$ and the manager's cost function $\varphi(x) = x^2$, where x is the value of the manager's flow. To start with we suppose that the technological network $N = \{w_2, w_3\}$ consists only of workers w_2 and w_3. So, workers w_1 and w_4 are not part of the firm (for example, the vendor and customer). Then the optimal hierarchy with cost $11^2 = 121$ is shown in Fig 4 a.

Assume we can extend the firm by adding workers w_1 and w_4. For example, this extension can be interpreted as follows. Large wholesale company buys the production firm (the "worker" w_1) and the chain of shops (the "worker" w_4) to control all the stages from production to the ultimate consumer. Large flow $f(w_1, w_2) = 5$ may be caused by purchasing problems, e.g. large quantity of defective goods. Similarly the large flow $f(w_3, w_4) = 5$ may be caused by some selling problems, e.g. customers often return defective goods.

Thus, after the extension the firm controls the whole technological network $N = \{w_1, w_2, w_3, w_4\}$. So, we

---

[20] Managers' m_1, m_2, m_3, m_4 flows equal 7, 5, 5, 6 respectively. So, $c(H) = 7^3 + 5^3 + 5^3 + 6^3 = 809$. In any optimal hierarchy any manager on the lowest (second) tier controls exactly two workers (otherwise his or her cost is greater than or equal to 1000>809). And it is easy to see that in any optimal hierarchy there is a single manager m_1 on the second tier and m_1 immediately controls workers w_2 and w_3 (controls the maximal flow $f(w_2, w_3)=5$). We can calculate costs of all trees controlling employees w_1, m, w_4. It allows to prove that all trees are non-optimal. And it is easy to see that the cost $c(H) = 809$ can not be further diminished. Therefore, H is optimal hierarchy.
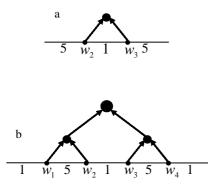
Fig 4. Firm growth with control cost decrease

can reconstruct the hierarchy as shown in Fig 4 b. We can hire two managers on the second tier and give them the responsibility to control the greatest flows $f(w_1, w_2) = 5$ and $f(w_3,w_4) = 5$. The cost of the reconstructed hierarchy equals $7^2 + 7^2 + 3^2 = 107$.

So, control cost can decrease with the technological network growth (including new workers, who were part of the environment). It could be a reason to buy some unprofitable business because it can reduce cost of control of the main business. Such facts often occur in practice[21].

3. Classes of sectional cost functions and corresponding optimal hierarchies

3.1. General form of optimal hierarchy

Proposition 1. *There exists such optimal hierarchy that the following conditions are satisfied*:
(i) all employees control different groups of workers;
(ii) only one manager has no superiors. All other managers and all workers are subordinated (maybe non-immediately) to this manager;
(iii) immediate subordinate of a manager does not control any other immediate subordinate of this manager.

The condition (i) means that there is no pair of managers fully duplicating each other's administrative labor. In other words, there are no managers controlling the same group of workers. In Fig 5 a the example of such duplication is shown. Particularly, the condition (i) leads to the fact that any manager has

at least two immediate subordinates (otherwise Lemma 1 implies that the manager and his or her only immediate subordinate control the same group of workers).

The condition (ii) means that exactly one manager m has no superiors. This manager controls all workers and all other managers in the hierarchy[22]. The manager m will be called *top manager*. In Fig 5 b there are two managers with no superiors. So, the condition (ii) is violated. Obviously the "redundant" manager can be removed with no cost increase.

The condition (iii) can be interpreted as follows. Assume the manager m_1 is immediately subordinated to the manager m. Then m does not immediately control the subordinates of the manager m_1. The condition corresponds with "normal" activity of the firm, when any manager controls subordinated employees only by means of his or her immediate subordinates, but not directly. In Fig 5 c the top manager m directly controls the workers w_2 and w_3, although these workers are also controlled by subordinated managers m_1 and m_2.

Proposition 1 simplifies optimal hierarchy problem because we can ignore hierarchies that violate condition (i), (ii) or (iii).
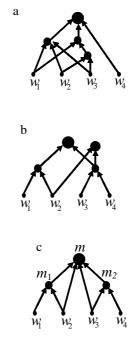

Fig 5. Hierarchies a-c violate conditions (i)-(iii) respectively

---

[21] For example, in'90 s of the 20th century many Russian food plants were transformed into vertically integrated companies by acquisition of farms in the corresponding region. These farms were unprofitable but provided regular supplies of cheap raw materials.

[22] The condition (ii) corresponds to the practice of organization design: there is one and only one top manager whose decisions must be implemented by all other managers and workers (for example, the top manager can eliminate a conflict between any set of employees in the firm).

3.2. Tree optimality condition

It is important to obtain the conditions when some tree is the optimal hierarchy. Below we consider the sufficient condition for tree optimality. This is the so-called group-monotony condition.

Definition 5. A sectional cost function is called *group-monotonic* if the manager's cost does not decrease with the expansion of the groups controlled by the immediate subordinates and with the addition of new immediate subordinates. So, for any groups $s_1,...,s_k$ the following inequalities hold:

$c(s_1,s_2,...,s_k) \le c(s,s_2,...,s_k)$, where s contains s_1 ($s_1 \subset s$); $c(s_1,s_2,...,s_k) \le c(s,s_1,...,s_k)$, where s is any group.

Let us explain Definition 5 by the example. Any manager m communicates with his or her immediate subordinates to solve their interaction problems (corresponding cost may be determined by some non-decreasing function $\chi(\cdot)$ depending on the number of immediate subordinates). Also the manager m may solve part of problems inside each of the groups controlled by immediate subordinates[23] (corresponding cost may be determined by some non-decreasing function $\varsigma(|s_1 \cup ... \cup s_k|)$ depending on the size of the group $s_1 \cup ... \cup s_k$, where $s_1,...,s_k$ are the groups controlled by all immediate subordinates of the manager). Thus, we can consider the following example of cost function:

$$c(s_1,...,s_k) = \chi(k) + \varsigma(|s_1 \cup ... \cup s_k|). \qquad (5)$$

Obviously function (5) does not decrease with any expansion of the groups $s_1,...,s_k$ and with any addition of new immediate subordinates. Therefore, (5) is an example of group-monotonic function.

In some practical situations a manager can decrease his or her cost by increasing the number of immediately subordinated managers ("assistants"). In such cases the function is not group-monotonic. However, if the "assistants" coordination cost is sufficiently high then it is reasonable to model the firm with the help of group-monotonic function.

Theorem 1. *If sectional cost function is group-monotonic then there exists optimal tree.*

According to this theorem if cost function is group-monotonic then optimal hierarchy can be found among the trees[24]. Therefore, to find optimal hierarchy we

can verify the inequalities of Definition 5. If these inequalities hold then optimal hierarchy problem is much simpler because we only need to find the minimal cost tree. Such tree can be found using the algorithms developed by Mishin and Voronin (2002b, 2003). For an arbitrary sectional cost function the exact algorithm's complexity is too high (the minimal cost tree can be found only for 15-20 workers[25]). Consider the cost function given by the expression $c(|s_1|,...,|s_k|)$ (for example, function (5) can be given by $c(|s_1|,...,|s_k|)$).[26] In this case the exact algorithm finds the minimal cost tree for 70-100 workers[27]. If the cost function is group-monotonic then optimal hierarchy problem can be solved using these algorithms. For other cost functions the tree obtained by the algorithms may be non-optimal hierarchy. Still this tree is useful, for example, to compare the best tree with actual hierarchy in the firm.

3.3. Hierarchy and two-tier hierarchy optimality conditions

Definition 6. Sectional cost function is *narrowing* if for any manager m with immediately subordinated employees $v_1,...,v_k$, $k \ge 3$ it is possible to resubordinate several employees from $v_1,...,v_k$ to new manager m_1 and immediately to subordinate m_1 to the manager m with no hierarchy cost increase. Sectional cost function is *widening* if any such resubordination does not decrease cost of hierarchy.

Let us explain Definition 6. In Fig 6 a manager m has three or more immediate subordinates $v_1,...,v_k$. Consider a narrowing cost function. With no hierarchy cost increase we can hire new immediate superior m_1 for j ($1 < j < k$) employees from $v_1,...,v_k$. After the hire the manager m controls these employees with the help of new manager m_1 but not immediately. For example, the result of employees $v_1,...,v_j$ resubordination is shown in Fig 6 b.

Generally any j employees can be resubordinated. So, there exists such permutation $(i_1,...,i_k)$ of numbers

---

[23] For example, the manager can perform some administrative labor when any subordinated worker is dismissed (interview with a new worker, signature of some documents, etc.).

[24] Mishin (2005) proves that the group-monotony is sufficient condition but not requirement for the tree optimality

[25] Using personal computer in several minutes.

[26] Manager's cost depends only on the span of control k (number of all immediate subordinates) and on the numbers $|s_1|,...,|s_k|$ of workers in the groups controlled by the immediate subordinates (but not on individual workers in these groups!).

[27] Mishin and Voronin (2003) also prove that it is impossible to sufficiently reduce exact algorithm's complexity. In the paper noted above some heuristic algorithms are developed. These algorithms have much less complexity and find trees with approximately minimal cost. For arbitrary function given by $c(|s_1|,...,|s_k|)$ two heuristic algorithms are developed. Their complexities grow as n^2 and n^3.
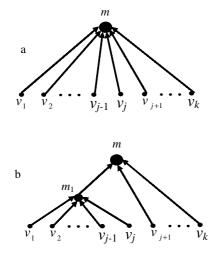
Fig 6. Resubordination for narrowing or widening cost function

$(1, …, k)$ that employees $v_{i_1}, …, v_{i_j}$ are resubordinated. If cost function is narrowing then for any groups $s_1 = s_H(v_1), …, s_k = s_H(v_k)$ controlled by employees $v_1, …, v_k$ some of them can be resubordinated with no hierarchy cost increase. Thus, definition of narrowing cost function can be written as follows. For any groups $s_1, …, s_k$, $k \geq 3$ there exist such number $1 < j < k$ and permutation $(i_1, …, i_k)$ that the following inequality holds:

$$c(s_1, …, s_k) \geq c(s_{i_1}, …, s_{i_j}) +$$
$$c(s_{i_1} \cup … \cup s_{i_j}, s_{i_{j+1}}, …, s_{i_k}) . \qquad (6)$$

Left-hand member of the inequality is the cost of the manager m before resubordination (see Fig 6 a). Right-hand member of the inequality equals to sum of manager's m_1 cost $c(s_{i_1}, …, s_{i_j})$ and manager's m cost $c(s_{i_1} \cup … \cup s_{i_j}, s_{i_{j+1}}, …, s_{i_k})$ after resubordination (see the example in Fig 6 b). Other managers' costs do not change. So, inequality (6) holds if and only if cost of total hierarchy does not increase. Thus, for narrowing cost function we can hire manager's m "assistant" m_1 undertaking part of administrative labor. After that the number of manager's m immediate subordinates decreases. So, the hierarchy becomes "narrower" (the span of control decreases).

Consider a widening cost function. Definition 6 leads to the fact that any described above resubordination does not decrease the cost of a hierarchy. So, for any groups $s_1, …, s_k$, $k \geq 3$, any number $1 < j < k$ and any permutation $(i_1, …, i_k)$ the following inequality holds:

$$c(s_1, …, s_k) \leq c(s_{i_1}, …, s_{i_j}) +$$
$$c(s_{i_1} \cup … \cup s_{i_j}, s_{i_{j+1}}, …, s_{i_k}) . \qquad (7)$$

Thus, for widening cost function it is impossible to decrease the cost of a hierarchy with the help of hiring "assistants".

If inequality (6) or (7) is violated on some overlapping groups $s_1, …, s_k$ but held on any non-overlapping groups (i.e. $s_i \cap s_j = \varnothing$ for all $i \neq j$) then cost function will be called *narrowing on non-overlapping groups* or *widening on non-overlapping groups* respectively.

Theorem 2. *If sectional cost function is narrowing then there exists optimal 2-hierarchy.*

Corollary (from Theorems 1 and 2). *If sectional cost function is narrowing on non-overlapping groups and group-monotonic then there exists optimal 2-tree.*

Therefore, to find optimal hierarchy we can verify inequality (6). If this inequality holds then the cost function is narrowing and we can consider only 2-hierarchies with each manager having two immediate subordinates (minimal span of control) because there exists optimal 2-hierarchy. In this case optimal hierarchy problem is much more easier. If cost function is group-monotonic then we have to verify inequality (6) only on non-overlapping groups $s_1, …, s_k$. If the inequality holds then the corollary leads to the fact that there exists optimal 2-tree[28].

Theorem 3. *If sectional cost function is widening then two-tier hierarchy is optimal.*

Corollary (from Theorems 1 and 3). *If sectional cost function is widening on non-overlapping groups and group-monotonic then two-tier hierarchy is optimal.*

Therefore, to find optimal hierarchy we can verify inequality (7). If this inequality holds then the cost function is widening and two-tier hierarchy with single manager is optimal (singe manager controls all workers immediately, span of control is maximal). If cost function is group-monotonic then we have to verify inequality (7) only on non-overlapping groups $s_1, …, s_k$. If the inequality holds then the corollary leads to the fact that two-tier hierarchy is optimal.

Theorems 2 and 3 show that narrowing functions contrast with widening functions[29]. Narrowing con-

---

[28] Minimal cost 2-tree can be found using the algorithms developed by Mishin and Voronin (2003).

[29] Consider manager's cost function $\varphi(\cdot)$ depending only on his or her internal flow in technological network. If $\varphi(\cdot)$ is superadditive then cost function is narrowing and if $\varphi(\cdot)$ is subadditive then cost function is widening (Mishin (2005)). So, narrowing/widening conditions generalize superadditivity/subadditivity conditions or convexity/concavity conditions (these conditions are equivalent for one-dimensional flows and $\varphi(0) = 0$).

dition implies optimality of 2-hierarchy, which contains the most number of managers. Each manager performs minimal quantity of administrative labor (controls only two immediate subordinates). On the contrary, widening condition implies optimality of two-tier hierarchy, which contains single manager performing all administrative labor (controls all n workers immediately).

It is easy to see[30] that both group-monotonic and non group-monotonic cost function may be narrowing, widening or neither narrowing, nor widening. Moreover, in extreme cases a sectional function may be both narrowing and widening. Interrelationship between classes of functions is shown in Fig 7. Types of optimal hierarchies are shown in the figure too (a tree is optimal for group-monotonic functions, two-tier hierarchy is optimal for widening functions, a 2-hierarchy is optimal for narrowing functions, a 2-tree is optimal for group-monotonic and narrowing functions).

4. Examples of cost function for different types of interaction

Suppose for each worker $w \in N$ some *worker's complexity* $\mu(w) > 0$ (positive real number) is given. Complexity may correspond with "work content" of the worker, his or her professional skills, etc. *Complexity of arbitrary group* of workers $s \subseteq N$ may be defined as total complexity of all workers in s: $\mu(s) = \sum_{w \in s} \mu(w)$. For example, complexity of the group may correspond with total "work content" of

all workers in the group. Sectional cost function depends only on groups $s_1, \ldots, s_k$ controlled by all immediate subordinates of the manager. Let us consider several examples of such sectional cost function that manager's cost depends only on complexities:

$$c(s_1, \ldots, s_k) = [\mu(s_1)^\alpha + \ldots + \mu(s_k)^\alpha - \max(\mu(s_1)^\alpha, \ldots, \mu(s_k)^\alpha)]^\beta, \quad \text{(I)}$$

$$c(s_1, \ldots, s_k) = [\mu(s_1)^\alpha + \ldots + \mu(s_k)^\alpha]^\beta, \quad \text{(II)}$$

$$c(s_1, \ldots, s_k) = [\mu(s)^\alpha / \max(\mu(s_1)^\alpha, \ldots, \mu(s_k)^\alpha) - 1]^\beta, \quad \text{(III)}$$

$$c(s_1, \ldots, s_k) = [\sum_{i=\overline{1,k}} (\mu(s)^\alpha - \mu(s_i)^\alpha)]^\beta, \quad \text{(IV)}$$

$$c(s_1, \ldots, s_k) = \mu(s)^\alpha / \min(\mu(s_1)^\beta, \ldots, \mu(s_k)^\beta), \quad \text{(V)}$$

where $s = s_1 \cup \ldots \cup s_k$ is the group controlled by the manager, $\mu(s_1), \ldots, \mu(s_k), \mu(s)$ are complexities of corresponding groups, $\alpha, \beta > 0$ are some positive real numbers (parameters of the function). Functions (I)-(V) depend on complexities ("work content") of employees of the "section" controlled by the manager immediately. In different firms section may be controlled using different mechanisms. Thus, interaction between the manager and his or her immediate subordinates (inside the section) may be organized in many ways. Below we interpret functions (I)-(V) as manager's cost for different ways of interaction inside the section. In different papers many such ways are considered quali-
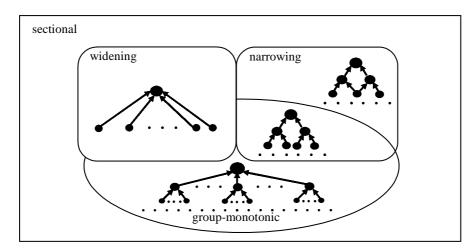
---

[30] See examples below and examples in Mishin (2005).

Fig 7. Interrelationship between classes of group-monotonic, narrowing and widening cost functions

tatively without mathematical modeling (see, for instance, Davies, Smith and Twigger, 1991; Peters, 1987; Jago and Vroom, 1975). We attempt to describe it mathematically.

Suppose among immediate subordinates there exists a "*semi-leader*" that copes with his or her tasks completely even with no superiors' control. Function (I) may correspond with this way of interaction. Manager's cost (I) depends on complexities of groups controlled by all immediate subordinates except the semi-leader. We mean that the immediate subordinate with maximal complexity is a semi-leader.

Suppose there *does not exist a "leader"*. Thus, the manager spends some efforts to control each of his or her immediate subordinates. Manager's cost may depend on complexities of all groups controlled by immediate subordinates. Function (II) may correspond with this way of interaction.

Suppose among immediate subordinates there exists a "*leader*" that helps to solve problems of other immediate subordinates (for example, using his or her experience or authority). Therefore, the cost of immediate superior of the leader decreases. Function (III) may correspond with this way of interaction. Manager's cost (III) depends on complexity of the whole group controlled by the manager and complexity of the group controlled by the leader, which is immediately subordinated to the manager. The greater is complexity of the leader, the greater is his or her importance among other immediate subordinates and the less is the cost of immediate superior.

Function (IV) corresponds with cost of *individual interactions between the manager and all his or her immediate subordinates*. The cost depends on differences between complexity of the group controlled by the manager and complexities of groups controlled by immediate subordinates[31].

Suppose, among immediate subordinates there exists an employee, that controls the group with small complexity. This employee may have *little qualification*. Low-qualified immediate subordinate may increase manager's cost. To control this subordinate the manager may spend much effort. So, manager's cost may

increase because he or she is diverted from solving more complex problems (just such problems must be solved by this manager). Function (V) may correspond with this way of interaction. Manager's cost (V) depends on complexity of the whole group controlled by the manager and complexity of the group controlled by the low-qualified employee, that is immediately subordinated to the manager. The less is the minimal qualification of subordinated employees the greater is the cost of immediate superior.

Let us solve optimal hierarchy problem for functions (I)-(V). Obviously functions (I) and (II) are group-monotonic and functions (III), (IV) and (V) are not group-monotonic. Let us examine narrowing and widening conditions. Below we use the following inequalities (particular cases of the Minkovski inequality, see, for instance, Hardy, Littlewood and Polya, 1934):

$$(x_1 + ... + x_k)^\gamma \geq x_1^\gamma + ... + x_k^\gamma \text{ for any}$$
$$x_1 \geq 0, ..., x_k \geq 0 \text{ and } \gamma \geq 1, \tag{8}$$

$$(x_1 + ... + x_k)^\gamma \leq x_1^\gamma + ... + x_k^\gamma \text{ for any}$$
$$x_1 \geq 0, ..., x_k \geq 0 \text{ and } \gamma \leq 1. \tag{9}$$

Proposition 2. *Function* (I) *is widening for* $\beta \leq 1$ *and narrowing for* $\beta \geq 1$.

Proposition 2 allows to obtain optimal hierarchy for function (I). If $\beta \leq 1$ then two-tier hierarchy is optimal (see Theorem 3). If $\beta \geq 1$ then 2-tree with minimal cost is optimal (see corollary from Theorems 1 and 2). Fig 8 a illustrates optimal hierarchies for function (I).

Proposition 3. *Function* (II) *is widening for* $\beta \leq 1$, *is widening on non-overlapping groups for* $\beta > 1$ *and* $\alpha \geq 1$.

Thus, if $\beta \leq 1$ or $\beta > 1$ and $\alpha \geq 1$ then for function (II) two-tier hierarchy is optimal (see Theorem 3 and corollary). Fig 8 b illustrates optimal hierarchies for function (II). In the region $\beta > 1$ and $\alpha < 1$ function (II) is neither widening, nor narrowing even on non-overlapping groups (Mishin, 2005). Therefore, for this region Theorems 2 and 3 can not help to obtain optimal hierarchy. However, function (II) is group-monotonic. Thus a tree with minimal cost is optimal (see Theorem 1).

Proposition 4. *Function* (III) *is narrowing for* $\beta \geq 1$.

Proposition 5. *Function* (IV) *is narrowing for* $\beta \geq 1$.

Propositions 4 and 5 allow to obtain optimal hierarchy for functions (III), (IV) and $\beta \geq 1$. In this case 2-hierarchy with minimal cost is optimal (see Theo-

---

[31] Consider a manager m that controls group $s_H(m)$. In process of individual interaction with his or her immediate subordinate m_1 the manager m may inform m_1 about the part of the group $s_H(m)$, which is not controlled by m_1. The volume of this information may depend on difference of complexities $\mu(s_H(m))$ and $\mu(s_H(m_1))$. Manager's cost (IV) depends on the sum of such volumes of information for all immediate subordinates.
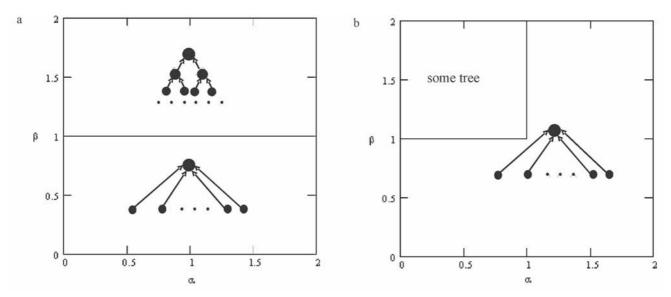
Fig 8. Forms of optimal hierarchy for function (I) (Fig a) and function (II) (Fig b)

rem 2).[32] For $\beta < 1$ we can find the tree with minimal cost using algorithms. But this tree may be non-optimal because functions (III) and (IV) are not group-monotonic. By now methods to solve optimal hierarchy problem for functions (III), (IV) and $\beta < 1$ are unknown.

Widening and narrowing functions imply optimality of two extreme hierarchies: two-tier hierarchy and 2-hierarchy. Usually in real firms there are some "intermediate" hierarchies with span of control $2 < r < +\infty$. Therefore, to model many real firms we have to examine neither widening nor narrowing cost functions. Thus, it is important to solve optimal hierarchy problem for this case. Below we describe a method of searching out the tree with minimal cost. If the cost function is group-monotonic then this tree is optimal (see Theorem 1). For other functions this is the best tree.

Optimal hierarchy problem is discrete optimization problem. Therefore, it is difficult to solve it analytically. One possible way of solution is to consider corresponding continuous problem with continuum set of workers. The exploration of continuous problem of searching out the minimal cost tree for sectional cost functions was pioneered by Goubko (2002). In some cases after the continuous problem is solved we can prove that corresponding tree minimizes cost for the original discrete problem.

Suppose we have to obtain minimal cost tree and cost function $c(s_1, \dots, s_k)$ depends only on complexities

of groups $s_1, \dots, s_k$. Thus, the cost function is given by $c(\mu(s_1), \dots, \mu(s_k))$ (see, for example, functions (I)-(V))[33]. Consider only *homogeneous* cost functions satisfying the following condition. For any $y > 0$ the equality $c(y\mu(s_1), \dots, y\mu(s_k)) = \varphi(y)c(\mu(s_1), \dots, \mu(s_k))$ holds, where $\varphi(\cdot)$ is some continuously increasing function. It can be proven (Goubko, 2002) that $\varphi(y) = y^\gamma$, where γ is *homogeneity coefficient*. If a cost function is homogeneous, then scale of complexity is of no importance. If we multiply all workers' complexities by the same multiplier y, then costs of all hierarchies are multiplied by y^γ. Therefore, scale of complexity does not affect on optimality of hierarchies. Let us define continuous problem corresponding with the discrete problem.

Let $x = \mu(w_1) + \dots + \mu(w_n)$ be total complexity of workers in the discrete problem. Suppose in the continuous problem the set of workers equals to the segment $N = [0;x]$. An individual worker is a point of this segment. The top manager m controls all segment N (all workers). The segment is divided into parts among managers $m_1, \dots, m_k$ immediately subordinated to the top manager. Each of the managers $m_1, \dots, m_k$ controls some part of the segment N. Thus, the segment N is divided into smaller segments with lengths $x_1, \dots, x_k > 0$ controlled by managers $m_1, \dots, m_k$ correspondingly, $x_1 + \dots + x_k = x$. The segment with length x_i controlled by the manager m_i is divided into smaller segments controlled by his or her immediate subordinates, $1 \le i \le k$. These segments are divided again,

---

[32] Mishin and Voronin (2003) obtain 2-hierarchy with minimal cost for (III) and $\beta \ge 1$.

[33] For any tree the groups $s_1, \dots, s_k$ are non-overlapping. So, $\mu(s_1 \cup \dots \cup s_k) = \mu(s_1) + \dots + \mu(s_k)$ and we may suppose that functions (I)-(V) depend only on $\mu(s_1), \dots, \mu(s_k)$.

etc. The tree infinitely "grows". In the tree each manager corresponds with a segment. The length of the segment equals to complexity of the group subordinated to the manager. If manager's immediate subordinates control segments with lengths $x_1, ..., x_k$ then manager's cost equals $c(x_1, ..., x_k)$. Cost of a tree equals to total cost of all managers in the tree. It is necessary to obtain infinite tree with minimal cost.

Goubko (2002) proves that for any homogeneous cost function there exists self-similarly tree H with minimal cost. In H each segment is divided in the same proportion $y_1, ..., y_k > 0$ regardless of hierarchical tier, $y_1 + ... + y_k = 1$. The top piece of self-similarly tree is shown in Fig 9. Controlled segments are shown instead of managers. Immediate subordinates $m_1, ..., m_k$ of the manager m control segments with lengths $y_1 x, ..., y_k x$. Therefore, manager's m cost equals $x^\gamma c(y_1, ..., y_k)$. Total cost of managers $m_1, ..., m_k$ equals $x^\gamma c(y_1, ..., y_k)(y_1^\gamma + ... + y_k^\gamma)$. Expression in the brackets are squares for the managers of the next tier, cubes – for the manager of the next tier, etc. For $\gamma > 1$ such expressions are geometric series with multiplier $y_1^\gamma + ... + y_k^\gamma < 1$ (this inequality follows from inequality (8) because $y_1 + ... + y_k = 1$). Thus, the cost of self-similarly tree H equals to the sum of infinitely decreasing geometric series:

$$c(H) = x^\gamma c(y_1, ..., y_k) / (1 - \sum_{i=\overline{1,k}} y_i^\gamma). \qquad (10)$$

One of such trees minimizes cost. So, it is enough to find $k \geq 2$ and proportion $y_1, ..., y_k$ minimizing expression (10). Corresponding tree is the desired infinite tree with minimal cost.

Let us obtain the tree with minimal cost for function (V). In any tree immediate subordinates of common manager control non-overlapping groups (segments). For any non-overlapping groups $s_1, ..., s_k$ equality $\mu(s_1 \cup ... \cup s_k) = \mu(s_1) + ... + \mu(s_k)$ holds. Therefore, function (V) is given by:

$$c(\mu(s_1), ..., \mu(s_k)) = (\mu(s_1) + ... + \mu(s_k))^\alpha / \min(\mu(s_1)^\beta, ..., \mu(s_k)^\beta) \qquad (11)$$

Expression (11) implies that function (V) is homogeneous. Homogeneity coefficient γ equals $\alpha - \beta$. Thus, we can minimize the cost (10) and obtain infinite tree with minimal cost.

Proposition 6. *Let r_* denote one of two integer numbers closest to the value $r_0 = ((\alpha-1)/\beta)^{1/(\alpha-\beta-1)}$. For continuous problem with cost function (V) and $\alpha - \beta > 1$ symmetric r_*-tree minimizes cost. In this tree any manager has exactly r_* immediate subordinates controlling groups with equal complexity.*

In the proof of Proposition 6 we show that for function (V) values $y_1 = ... = y_k = 1/k$ minimize expression (10). So, symmetric tree minimizes cost. Thus, it is enough to find k minimizing expression (10). The minimum point $r_0 = ((\alpha-1)/\beta)^{1/(\alpha-\beta-1)}$ may be non-integer value. Therefore, r_* is maximal integer less than or equal to r_0, or r_* is minimal integer greater than or equal to r_0 (to define r_* it is enough to substitute these two values in expressions (10) and (11)).

For function (V) with $\alpha - \beta > 1$ Proposition 6 solves the continuous problem. Consider corresponding discrete problem with number of workers $n = r_*^j$ (n is some power of r_*) and the same workers' complexities $\mu(w_1) = ... = \mu(w_n) = 1/n$. In this case top j tiers of the infinite symmetric r_*-tree are just discrete tree controlling workers $w_1, ..., w_n$ (these workers correspond with the tier $j+1$). And cost of this part of the infinite tree equals to cost of discrete tree. Therefore, for $n = r_*^j$ and workers with the same complexity symmetric r_*-tree minimizes cost for the discrete problem[34]. Thus, in this case we solve the discrete problem using continuous approximation method.

In Fig 10 the line $\beta = \alpha - 1$ is shown. The region below this line is divided into regions with the same r_*. In each of these regions optimal span of control does not change. In the top right region symmetric 2-tree minimizes cost. In the region below symmetric 3-tree minimizes cost. In the next region symmetric

---

[34] Otherwise we can reduce the cost of the infinite tree using the discrete tree with less cost to construct top j tiers of the infinite tree.
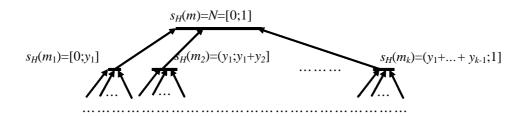

Fig 9. The top piece of self-similarly tree with proportion $y_1, ..., y_k$ and $= 1$

4-tree minimizes cost, etc. If parameters tend to the point (1;0) then r_* grows infinitely (for $r_*<10$ in the figure regions are denoted by numbers). If α increases then in Fig 10 curves exponentially decrease. In Fig 10, thus, 2-tree and 3-tree are shown. In these trees the group controlled by a manager is "divided" into subgroups with the same complexity among manager's subordinates. Trees for more r_* may be shown similarly.

Parameter β may be interpreted as degree of unfavorable influence of little qualification. If β tends to zero then we can subordinate low-qualified employees (controlling groups with low complexity) to the manager with no his or her cost sufficiently increase (see expression (11)). Therefore, if β tends to zero then optimal span of control r_* tends to $+\infty$. Thus, for sufficiently small β two-tier hierarchy with single manager minimizes cost (if $\beta = 0$ then function (V) is widening and two-tier hierarchy is optimal for any number of workers).

There exists the limit of the value r_0 (see Proposition 6) by parameters tending to the critical line $\beta = \alpha-1$. This limit equals $e^{1/\beta}$. So, parameter regions with fixed r_* "reach" the critical line. For special cost function Qian (1994) also considers the problem of searching out the minimal cost tree. If real number of immediate subordinates are possible, then Qian (1994) proves that optimal span of control equals e (each manager has e immediate subordinates). This result coincides with the result for function (V) with $\alpha = 2$ and $\beta = 1$ ($\lim r_0 = e^{1/\beta} = e$).

Fig 10 shows that for any $r \geq 2$ there exists such region of parameters α and β that symmetric r-tree has minimal cost. In many real firms span of control ranges from several immediate subordinates to hundreds immediate subordinates (Mintzberg, 1979). The

values $2 < r < +\infty$ may be interesting to model such firms.

5. Concluding remarks

Further development of the methods of the optimal hierarchy search for sectional cost functions seems perspective, among the following other general directions of future research.

1. Mechanism design. It is important to construct control mechanisms that minimize total wage of employees, which equals to the cost of the optimal hierarchy (this is minimal possible cost). Mishin (2004) constructs such mechanism in a complete information framework. For the case of incomplete information it is necessary to take the "worst case" into consideration. For example, it may be necessary to compensate maximal total cost of all managers, which depends on information available for some metacenter, for instance, the owner of the firm. But in some cases excess incentives provide stability with respect to cost increase (Mishin (2004)). If managers' cost increases, then a manager can restructure the subordinated part of the hierarchy with no assistance (at the expense of manager's own resources). It allows to "adapt" the firm to the cost modifications.

2. Dynamical models of the optimal hierarchy. Parameters of the cost function, the number of workers, certain workers, interaction schemes can change with time. Therefore, the initially optimal hierarchy can later become non-optimal. However, the reconstruction of the hierarchy is associated with large cost. So, in dynamical models one has to compromise the total cost of all managers and the reconstruction cost. Mishin and Voronin (2003) introduce a metric on the set of hierarchies. This metric is one of possible ways
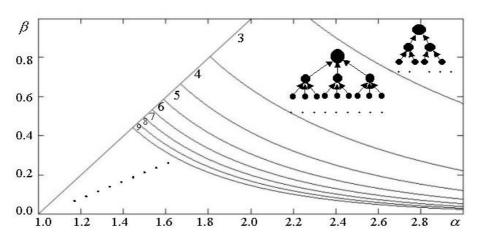

Fig 10. Minimal cost trees for function (V)

to define mathematically the reconstruction cost. The metric allows to model the restructuring effects numerically (Mishin and Voronin, 2002a, 2003). Analytical methods for solving the dynamical problem of the optimal hierarchy are unknown so far.

Acknowledgements

Appendix A. Proofs

Proof of Lemma 1. v is subordinated to m. So, any worker $w \in s_H(v)$ is subordinated to m ($w \in s_H(m)$) because the path from w to v can be extended up to the path from w to m . Therefore, $s_H(v) \subseteq s_H(m)$. Consider $w \in s_H(m)$. The path from w to m contains the node v_j for some $1 \le j \le k$ as $(v_1, m), \ldots, (v_k, m)$ are the only edges incoming to m. So, $w \in s_H(v_j)$ and $s_H(m) \subseteq s_H(v_1) \cup \ldots \cup s_H(v_k)$.

Moreover, $s_H(v_j) \subseteq s_H(m)$ as v_j is subordinated to m for each $1 \le j \le k$. Thus, the equality $s_H(m) = s_H(v_1) \cup \ldots \cup s_H(v_k)$ holds. ∎

Proof of Lemma 2. Consider a tree H. Assume $s_H(v_1) \cap s_H(v_2) \ne \varnothing$ for some manager m and two of his or her immediate subordinates v_1 and v_2. Any worker $w \in s_H(v_1) \cap s_H(v_2)$ is subordinated to the employees v_1 and v_2. So, there are two different paths from w to m (the first path contains the node v_1 and the second path contains the node v_2). These paths diverge at some node $v \in N \cup M$. Thus, the employee v has more than one immediate superior. It contradicts Definition 2. Thus, in the tree H any manager's immediate subordinates control non-overlapping groups of workers. ∎

Proof of Proposition 1: Consider an optimal hierarchy $H \in \Omega(N)$. Let two employees v_1 and v_2 control the same group $s_H(v_1) = s_H(v_2)$. Acyclicity of the hierarchy implies that the employee v_1 does not control the employee v_2 or vice versa. Suppose v_1 does not control v_2. Then consider the immediate superior m_1 of the employee v_2. If v_1 is immediately subordinated to m_1 then the edge (v_2, m_1) can be removed with

no hierarchy cost increase (see inequality (3)). If v_1 is not immediately subordinated to m_1 then the edge (v_2, m_1) can be replaced to the edge (v_1, m_1). The equality $s_H(v_1) = s_H(v_2)$ implies that the cost of the manager m_1 does not change. So, the cost of total hierarchy also does not change. Thus, in both cases the edge (v_2, m_1) can be removed. Similarly, we can remove all edges outcoming from v_2. After that the employee v_2 has no superiors and the employee v_2 can be removed with no hierarchy cost increase[35]. If in the obtained hierarchy some employees control the same group then we can repeat the removal described above. Finally we obtain the optimal hierarchy H' with employees controlling differing groups. Thus, condition (i) holds for H' .

If some manager m_2 in the hierarchy H' has no superiors and controls the group $s_{H'}(m_2) \ne N$ then this manager can be removed with no hierarchy cost increase. We can repeat such removal. As a result, we obtain the optimal hierarchy H'' . In H'' any manager without superiors controls the group N. Definition 1 and condition (i)[36] imply that there is the single such manager m in the hierarchy H''.[37] At least one edge outcomes from any node $v \ne m$ in the hierarchy H'' . Acyclicity implies that we can construct the path from v to m. So, all employees are subordinated to m. Thus, conditions (i) and (ii) hold for the optimal hierarchy H''.

Let the employees v_3 and v_4 be immediately subordinated to the common manager m_3 in the hierarchy H'' and the employee v_3 be subordinated to the employee v_4. Then $s_{H''}(v_3) \subseteq s_{H''}(v_4)$ (see Lemma 1). Inequality (3) implies that the edge (v_3, m_3) can be removed with no hierarchy cost increase. After removal the employee v_3 has at least one immediate superior because v_3 is subordinated to v_4. We can repeat such removal. As a result, we obtain the optimal hierarchy H^* in which condition (iii) holds. The modifications described above do not change groups controlled by the managers. The manager m is the only manager without superiors. Thus, conditions (i), (ii) and (iii) hold for the optimal hierarchy H^*. ∎

Proof of Theorem 1. Proposition 1 implies that there exists an optimal hierarchy $H = (N \cup M, E) \in \Omega(N)$,

---

[35] Definition 1 is fulfilled because the maximal group N is controlled by some manager (if v_2 controls the group N in the hierarchy H then v_1 controls this group too).

[36] We have removed some managers without violation of condition (i).

[37] We cannot remove this manager because in this case Definition 1 is violated and the graph is not a hierarchy controlling the set of workers N.

which satisfies conditions (i)-(iii). If each of the employees except the top manager has exactly one immediate superior then H is an optimal tree (see Definition 2). Otherwise there exists an employee $v \in N \cup M$ with two or more immediate superiors. If there are several such employees then let us consider the employee on the highest tier. So, each of the superiors of the employee v except the top manager has exactly one immediate superior.

Let v_1 and u_1 be some different immediate superiors of the employee v. Condition (ii) of Proposition 1 implies that the employees v_1 and u_1 are subordinated to the top manager m. Thus, there exists the path from v_1 to m and the path from u_1 to m.[38] Therefore, there exist two different paths from v to m. These paths diverge in common node v and converge in other node u (in m or one of subordinates of the manager m). Let $v - v_1 - \ldots - v_{n_1}$ and $v - u_1 - \ldots - u_{n_2}$ be the parts of these paths from v to u. These parts have common first node v, common last node $v_{n_1} = u_{n_2} = u$ and different intermediate nodes. It follows from choice of the node v that each of the managers $v_1, \ldots, v_{n_1-1}$ has exactly one immediate superior – the next node in the path. This is true for the managers $u_1, \ldots, u_{n_2-1}$ too. Corresponding fragment of the hierarchy is shown in Fig 11.

Initial hierarchy H satisfies conditions (i)-(iii) of Proposition 1. Below we describe the reconstruction that does not increase the cost of the hierarchy. After each reconstruction obtained hierarchy will be denoted H just as the initial hierarchy. All reconstructed hierarchies satisfy condition (ii) of Proposition 1. So, all employees are subordinated to the top manager m. Therefore, different paths from v converge and the fragment of the reconstructed hierarchy looks
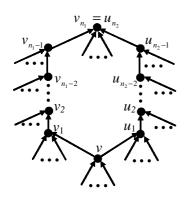

Fig 11. Optimal hierarchy reconstruction with group-monotonic cost function

like the fragment in Fig 11. There are two possible options of hierarchy H reconstruction (Fig 11 explains these options).

a) Suppose $s_H(v) = s_H(v_1)$.[39] So, the employees v and v_1 control the same group of workers. Let us remove the manager v_1. If v is not immediately subordinated to the manager v_2 then let us immediately subordinate the employee v to the manager v_2 instead of the manager v_1. After removal the groups controlled by the managers are not modified. So, only the cost of the manager v_2 can be modified. This cost does not increase because of group-monotony. Thus, the obtained hierarchy is optimal.

After v_1 removal some employees may have no superiors. Such employee is not a worker because all workers are subordinated to the top manager. So, after v_1 removal in addition to the top manager some other managers may have no superiors. Such managers can be removed. The obtained graph is an optimal hierarchy. After removal new managers may have no superiors. These managers can be removed too, etc. Finiteness of M implies that we obtain the optimal hierarchy with only top manager having no superiors. Thus, condition (ii) of Proposition 1 holds.

b) Suppose $s_H(v) \neq s_H(v_1)$. So, the manager v_1 controls a wider group than the employee v: $s_H(v) \subset s_H(v_1)$. Thus, v_1 has at least two immediate subordinates. Let us remove the edge (v, v_1). After removal the manager v_1 still has subordinates. The group $s_1 = s_H(v_1)$ controlled by the manager v_1 can be changed to the new group s_1' if some workers from the group $s_H(v)$ are not controlled by the manager v_1 after removal. However, v_1 controls workers from the group s_1 which are not part of the group $s_H(v)$. Thus, $s_1' \subseteq s_1$, $(s_1 \setminus s_1') \subseteq s_H(v)$. There is exactly one edge outgoing from the node v_1. The modification of the group $s_1 = s_H(v_1)$ can cause the modification of the group $s_2 = s_H(v_2)$ controlled by the manager v_2. Let s_2' be the modified group. As described above only workers from $s_H(v)$ can be removed from the group s_1. So, only such workers can be removed from the group s_2. Thus, $s_2' \subseteq s_2$, $(s_2 \setminus s_2') \subseteq s_H(v)$. Similarly for each $i = \overline{3, n_1 - 1}$ the group $s_i = s_H(v_i)$ controlled by the manager v_i changes to the group s_i', $s_i' \subseteq s_i$, $(s_i \setminus s_i') \subseteq s_H(v)$.

Consider the group $s_H(v_{n_1})$. This group equals to the union of the groups controlled by all the immediate

[38] One of these paths can contain one node if $v_1 = m$ or $u_1 = m$.

[39] In some cases reconstructed hierarchies do not satisfy condition (i) of Proposition 1. So, the equality $s_H(v) = s_H(v_1)$ can hold.

95

subordinates of the manager v_{n_1} (see Lemma 1). Among these groups only the group s_{n_1-1} controlled by the manager v_{n_1-1} can be changed after the edge (v,v_1) removal[40]. It follows from $(s_{n_1-1} \setminus s'_{n_1-1}) \subseteq s_H(v)$ that only workers from the group $s_H(v)$ can be removed from the group s_{n_1-1}. However, these workers are the part of the group $s_H(u_{n_2-1})$. Thus, the group $s_H(v_{n_1})$ is not changed. Therefore, the groups controlled by the superiors of the manager v_{n_1} are not changed too.

So, removal of the edge (v,v_1) can change the groups $s_H(v_1),...,s_H(v_{n_1-1})$ only. Thus, the top manager still controls all the workers, each manager has subordinates and the obtained graph is acyclic (edge removal cannot cause cycles). Therefore, the obtained graph satisfies all conditions of Definition 1. We obtain the hierarchy. Moreover, each employee except the top manager has at least one immediate superior. All employees are subordinated to the top manager because of acyclicity. So, the hierarchy satisfies condition (ii) of Proposition 1.

The number of employees immediately subordinated to the manager v_1 decreases by one. The number of employees subordinated to each of the managers $v_2,...,v_{n_1}$ does not change. However, the group controlled by immediate subordinate of the manager v_i can be reduced, $i = \overline{2,n_1}$. So, costs of managers $v_1,...,v_{n_1}$ do not increase because of group-monotony. Thus, the obtained hierarchy is optimal.

Both in the option a) and in the option b) we obtain the optimal hierarchy satisfying condition (ii) of Proposition 1. Therefore, we can repeat the reconstruction a) or b) while there is an employee with two or more immediate superiors. After each reconstruction the number of edges decreases at least by one. Finiteness of the edge set E implies that the reconstructions come to an end after finite number of steps. In the obtained optimal hierarchy H^* only top manager has no superiors. Each of the other employees in H^* has exactly one immediate superior. So, H^* is an optimal tree. ∎

Proof of Theorem 2. Consider an optimal hierarchy $H \in \Omega(N)$. Let k be the maximal number of employees immediately subordinated to common manager. If $k = 2$ then H is the required optimal 2-hierarchy. If $k>2$ then consider a manager m with k immediately subordinated employees $v_1,...,v_k$.

Let $s_1=s_H(v_1),...,s_k=s_H(v_k)$ be the groups controlled by the employees $v_1,...,v_k$. As the cost function is narrowing there exist a number of employees $1 < j < k$ and permutation $(i_1,...,i_k)$ satisfying inequality (6). Let us reconstruct the hierarchy H: hire new manager m_1 and immediately subordinate the employees $v_{i_1},...,v_{i_j}$ to m_1 instead of m, immediately subordinate m_1 to m (see the example in Fig. 6). Inequality (6) implies that the cost of the hierarchy does not increase. Thus, the obtained hierarchy is optimal. The manager m_1 has $j<k$ immediate subordinates. The manager m has $k-j+1<k$ immediate subordinates. So, in the obtained hierarchy the number of managers with k immediate subordinates decreases by one. We can repeat such reconstruction while there exists the manager with k immediate subordinates. As a result, we obtain the optimal hierarchy with maximal number $k'<k$ of employees immediately subordinated to common manager. If $k'>2$ then we can repeat reconstructions. As a result we obtain the optimal 2-hierarchy. ∎

Proof of corollary (from Theorems 1 and 2). Theorem 1 implies that there exists an optimal tree because the cost function is group-monotonic. In the proof of Theorem 2 we can consider this tree as initial optimal hierarchy H. Lemma 2 implies that immediate subordinates of any manager control non-overlapping groups of workers. Therefore, there are no overlapping groups among the groups $s_1,...,s_k$ in the proof of Theorem 2. So, we can reconstruct the hierarchy because the cost function is narrowing on non-overlapping groups. After the reconstruction we obtain some tree (new manager and each of other employees except the top manager have exactly one immediate superior). After all reconstructions we obtain the optimal 2-tree. ∎

Proof of Theorem 3. Proposition 1 implies that there exists an optimal hierarchy $H \in \Omega(N)$, which satisfies conditions (i)-(iii). According to condition (ii) there exists a manager m controlling all other employees. If m is a single manager then H is the optimal two-tier hierarchy. Otherwise there exists a manager m_1 immediately subordinated to the manager m. Let $v_1,...,v_j$ be all immediate subordinates of the managers m_1. Let $s_1 = s_H(v_1),...,s_j = s_H(v_j)$ be the groups controlled by the employees $v_1,...,v_j$. As the hierarchy H satisfies condition (i) of Proposition 1 each manager has at least two immediate subordinates. So, $j >1$ and the manager m has other immediate subordinates besides m_1. Let $v_{j+1},...,v_k$, $k \geq 3$ be all such immediate subordinates. Let $s_{j+1} = s_H(v_{j+1}),...,s_k = s_H(v_k)$ be the groups controlled by the employees $v_{j+1},...,v_k$.

---

40 Among manager's v_{n_1} immediate subordinates only v_{n_1-1} controls the managers $v_1,...v_{n_1-2}$ because each of them has only one immediate superior.

Suppose the manager m_1 has some immediate superiors m' besides m. So, there exist two different paths from m_1 to m: the first path contains only two nodes m_1 and m, the second path contains the manager m'. Besides m_1 the second path contains one of the employees $v_{j+1},\ldots,v_k$ immediately subordinated to m. So, this employee controls the manager m_1. It contradicts condition (iii) of Proposition 1. Therefore, top manager m is the single immediate superior of the manager m_1.

Condition (iii) of Proposition 1 implies that there are no immediate subordinates of the manager m among the employees $v_1,\ldots,v_j$ (otherwise immediate subordinate m_1 controls other immediate subordinate). So, there are no identical employees among $v_{j+1},\ldots,v_k$ and $v_1,\ldots,v_j$. Thus, the described fragment of the hierarchy looks like the fragment in Fig 6 b.

Inequality (7) holds for any groups $s_1,\ldots,s_k$, $k \geq 3$, any number $1 < j < k$ and any permutation $(i_1,\ldots,i_k)$ because the cost function is widening. If $(i_1,\ldots,i_k)=(1,\ldots,k)$ then inequality (7) is given by:

$$c(s_1,\ldots,s_k) \leq c(s_1,\ldots,s_j) +$$
$$c(s_1 \cup \ldots \cup s_j, s_{j+1},\ldots,s_k) \quad . \qquad (*)$$

Let us reconstruct the hierarchy: immediately subordinate the employees $v_1,\ldots,v_j$ to the manager m instead of the manager m_1 and remove the manager m_1. Obtained fragment of the graph looks like the fragment in Fig. 6 a). The manager m controls all workers as before the reconstruction. So, the obtained graph is a hierarchy. The groups controlled by other managers do not change too. In the obtained hierarchy the cost of the manger m (first member of the inequality (*)) is less than or equal to costs of the managers m and m_1 in the initial hierarchy (right-hand member of inequality (*)). Thus, the obtained hierarchy is optimal and satisfies conditions (i) and (ii) of Proposition 1. But condition (iii) may be violated because some of the employees $v_1,\ldots,v_j$ may be subordinated to some of the employees $v_{j+1},\ldots,v_k$. Suppose the employee v_{j_1} is subordinated to the employee v_{j_2}, $1 \leq j_1 \leq j$, $j+1 \leq j_2 \leq k$. Lemma 1 leads to $s_{j_1} \subseteq s_{j_2}$. Inequality (3) implies that "excess" edge (v_{j_1}, m) can be removed with no hierarchy cost increase. After removal the employee v_{j_1} is subordinated to the top manager but not immediately (through the employee v_{j_2}). We can repeat such removal. As a result, we obtain the optimal hierarchy satisfying conditions (i), (ii) and (iii) of Proposition 1.

The obtained optimal hierarchy contains less managers

than the initial hierarchy because the manager m_1 has been removed. We can repeat similarly reconstructions while there are two or more managers. As a result, we obtain the optimal two-tier hierarchy with the single manager m. ∎

Proof of corollary (from Theorems 1 and 3). Theorem 1 implies that there exists an optimal tree because the cost function is group-monotonic. In the proof of Theorem 3 we can consider this tree as initial optimal hierarchy H. Lemma 2 implies that immediate subordinates of any manager control non-overlapping groups of workers. Therefore, there are no overlapping groups among the groups $s_1,\ldots,s_k$ in the proof of Theorem 3. So, we can reconstruct the hierarchy because the cost function is widening on non-overlapping groups. After the reconstruction we obtain some tree. After all reconstructions we obtain the optimal two-tier hierarchy. ∎

Proof of Proposition 2. Consider the groups $s_1,\ldots,s_k$, $k \geq 3$. Let z_1 and z_2 be the left-hand member and the right-hand member of inequalities (6) and (7) (these inequalities correspond with narrowing and widening cost functions). Suppose $\beta \leq 1$. Let us prove inequality (7) for any $1 < j < k$ and any permutation $(i_1,\ldots,i_k)$. Inequality (7) is given by $c(s_1,\ldots,s_k) \leq c(s_{i_1},\ldots,s_{i_j}) + c(s_{i_1} \cup \ldots \cup s_{i_j}, s_{i_{j+1}},\ldots,s_{i_k})$. Let us define the following values: $x_1 = \mu(s_{i_1})^\alpha$, ..., $x_j = \mu(s_{i_j})^\alpha$, $x' = \max(x_1, \ldots, x_j)$, $x = x_1 + \ldots + x_j$, $y_{j+1} = \mu(s_{i_{j+1}})^\alpha$, $y_{j+2} = \mu(s_{i_{j+2}})^\alpha$, ..., $y_k = \mu(s_{i_k})^\alpha$, $y' = \max(y_{j+1},\ldots,y_k)$, $y = y_{j+1} + \ldots + y_k$, $s = s_{i_1} \cup \ldots \cup s_{i_j}$. Then the left-hand member and the right-hand member of inequality (7) are given by:

$$z_1 = (x + y - \max(x', y'))^\beta,$$

$$z_2 = (x - x')^\beta + (\mu(s)^\alpha + y - \max(y', \mu(s)^\alpha))^\beta.$$

Inequality (9) and $\beta \leq 1$ imply that the inequality $z_2 \geq (x + y + \mu(s)^\alpha - x' - \max(y', \mu(s)))^\beta$ holds. To prove inequality (7) ($z_2 \geq z_1$) it is enough to prove:

$$x + y - \max(x', y') \leq x + y +$$
$$\mu(s)^\alpha - x' - \max(y', \mu(s)^\alpha) \quad .$$

This inequality is given by:

$x' + \max(y', \mu(s)^\alpha) \leq \mu(s)^\alpha + \max(x', y')$. If $y' \leq \mu(s)^\alpha$ then the inequality is simplified: $x' \leq \max(x', y')$. So, the inequality holds. If $y' > \mu(s)^\alpha$ then the inequality is given by $x' + y' \leq \mu(s)^\alpha + \max(x', y')$. The inequalities $y' \leq \max(x', y')$ and $x' \leq \mu(s)^\alpha$ hold because $s = s_{i_1} \cup \ldots \cup s_{i_j}$. Thus, inequality (7) holds. So, if $\beta \leq 1$ then function (I) is widening.

Suppose $\beta \geq 1$. Let s_1 be the group with maximal complexity: $\mu(s_1) = \max(\mu(s_1),\ldots,\mu(s_k))$ (otherwise we can renumber the groups $s_1,\ldots,s_k$). Consider the groups s_1, s_2 ($j=2$) and the permutation $(1,2,\ldots,k)$. Let us prove inequality (6) which is given by: $c(s_1,\ldots,s_k) \geq c(s_1,s_2) + c(s_1 \cup s_2, s_3,\ldots,s_k)$. The left-hand member and the right-hand member of inequality (6) are given by: $z_1 = (\mu(s_2)^\alpha + \ldots + \mu(s_k)^\alpha)^\beta$ and $z_2 = \mu(s_2)^{\alpha\beta} + (\mu(s_3)^\alpha + \ldots + \mu(s_k)^\alpha)^\beta$. Inequality (8) and $\beta \geq 1$ lead to $z_1 \geq z_2$. Thus, inequality (6) holds. So, if $\beta \geq 1$ then function (I) is narrowing. ∎

Proof of Proposition 3. Consider the groups $s_1,\ldots,s_k$, $k \geq 3$. Let z_1 and z_2 be the left-hand member and the right-hand member of inequality (7). The inequality corresponds with widening cost function. Let us prove inequality (7) for any $1 < j < k$ and any permutation $(i_1,\ldots,i_k)$:

$$c(s_1,\ldots,s_k) \leq c(s_{i_1},\ldots,s_{i_j}) +$$
$$c(s_{i_1} \cup \ldots \cup s_{i_j}, s_{i_{j+1}},\ldots,s_{i_k}).$$

Let us define the following values: $s = s_{i_1} \cup \ldots \cup s_{i_j}$, $x = \mu(s_{i_1})^\alpha + \ldots + \mu(s_{i_j})^\alpha$, $y = \mu(s_{i_{j+1}})^\alpha + \ldots + \mu(s_{i_k})^\alpha$. Then the left-hand member and the right-hand member of inequality (7) are given by: $z_1 = (x + y)^\beta$ and $z_2 = x^\beta + (\mu(s)^\alpha + y)^\beta$. If $\beta \leq 1$ then (9) leads to $z_1 \leq x^\beta + y^\beta \leq z_2$. Thus, inequality (7) holds. So, if $\beta \leq 1$ then function (II) is widening.

If the groups $s_1,\ldots,s_k$ are non-overlapping then $\mu(s) = \mu(s_{i_1}) + \ldots + \mu(s_{i_j})$. If $\alpha \geq 1$ then (8) leads to $\mu(s)^\alpha \geq \mu(s_{i_1})^\alpha + \ldots + \mu(s_{i_j})^\alpha = x$. Thus, the inequality $z_2 \geq (x + y)^\beta = z_1$ (inequality (7)) holds too. So, if $\beta > 1$ and $\alpha \geq 1$ then function (II) is widening on non-overlapping groups. ∎

Proof of Proposition 4. Consider the groups $s_1,\ldots,s_k$, $k \geq 3$. Let z_1 and z_2 be the left-hand member and the right-hand member of inequality (6) corresponding with narrowing cost function. Let s_1 be the group with maximal complexity: $\mu(s_1) = \max(\mu(s_1),\ldots,\mu(s_k))$ (otherwise we can renumber the groups $s_1,\ldots,s_k$). Consider the groups s_1, s_2 ($j=2$) and the permutation $(1,2,\ldots,k)$. Let us prove inequality (6) which is given by $c(s_1,\ldots,s_k) \geq c(s_1,s_2) + c(s_1 \cup s_2, s_3,\ldots,s_k)$.

Let us define the values $x = \mu(s_1 \cup \ldots \cup s_k)^\alpha$, $y = \mu(s_1 \cup s_2)^\alpha$, $z = \mu(s_1)^\alpha$. Then $z \leq y \leq x$. The left-hand member and the right-hand member of inequality (6) is given by $z_1 = (x/z - 1)^\beta$, $z_2 = (y/z - 1)^\beta + (x/y - 1)^\beta$. Inequality (8) and $\beta \geq 1$ imply that the inequality $z_2 \leq (y/z - 1 + x/y - 1)^\beta$ holds. Using this estimation we can prove inequality (6) ($z_2 \leq z_1$) with the help of proving the inequality

$x/z - 1 - y/z + 1 - x/y + 1 \geq 0$. This inequality is given by:

$$(xy + yz - y^2 - xz)/yz = (x - y)(y - z)/yz \geq 0.$$

Thus, inequality (6) holds. So, if $\beta \geq 1$ then function (III) is narrowing. ∎

Proof of Proposition 5. Consider the groups $s_1,\ldots,s_k$, $k \geq 3$. Let z_1 and z_2 be the left-hand member and the right-hand member of inequality (6) corresponding with narrowing cost function. Consider the groups s_1, s_2 ($j=2$) and the permutation $(1,2,\ldots,k)$. Let us prove inequality (6) which is given by $c(s_1,\ldots,s_k) \geq c(s_1,s_2) + c(s_1 \cup s_2, s_3,\ldots,s_k)$.

Let us define the values $x = \mu(s_1 \cup \ldots \cup s_k)^\alpha$, $y = \mu(s_1 \cup s_2)^\alpha$, $x_1 = \mu(s_1)^\alpha,\ldots,x_k = \mu(s_k)^\alpha$. The left-hand member and the right-hand member of inequality (6) are given by $z_1 = (kx - x_1 - \ldots - x_k)^\beta$, $z_2 = (2y - x_1 - x_2)^\beta + ((k-1)x - y - x_3 - \ldots - x_k)^\beta$. Inequality (8) and $\beta \geq 1$ imply that the inequality $z_2 \leq ((k-1)x + y - x_1 - \ldots - x_k)^\beta$ holds. The right-hand member is less than or equal to z_1 because $y \leq x$. Thus, inequality (6) ($z_2 \leq z_1$) holds. So, if $\beta \geq 1$ then function (IV) is narrowing. ∎

Proof of Proposition 6. The equality $\gamma=\alpha-\beta$ holds for function (V). Let us substitute expression (11) ((V) for non-overlapping groups) in expression (10). Then the cost of infinite tree is given by:

$$x^\gamma(y_1 + \ldots + y_k)^\alpha / \qquad\qquad (*)$$
$$[\min(y_1^\beta,\ldots,y_k^\beta)(1 - \textstyle\sum_{i=1,k} y_i^\gamma)].$$

The numerator in the brackets equals 1 ($y_1+\ldots+y_k=1$). To minimize the expression it is enough to maximize the denominator. It is obvious that the expression $\min(y_1^\beta,\ldots,y_k^\beta)$ reaches maximum when $y_1=\ldots=y_k=1/k$. With the help of the simplest mathematical analysis methods we can prove that for $\gamma>1$ the expression $(1 - \sum_{i=1,k} y_i^\gamma)$ reaches maximum when $y_1=\ldots=y_k=1/k$.

Thus, the symmetric k-tree minimizes cost function (V). In this tree each manager has exactly k immediate subordinates. These subordinates control the groups with the same complexity. So, we have to find optimal k. Without the constant x^γ expression (*) with $y_1=\ldots=y_k=1/k$ is given by the function $\xi(k)$:

$$\xi(k) = k^\beta / (1 - k/k^\gamma) = k^{\beta+\gamma-1}/(k^{\gamma-1} - 1) =$$
$$k^{\alpha-1}/(k^{\alpha-\beta-1} - 1).$$

Let us differentiate the function by k and ignore the positive multiplier:

$$\xi(k) = (\alpha-1)k^{\alpha-2}\left(k^{\alpha-\beta-1}-1\right) - (\alpha-\beta-1)k^{\alpha-\beta-2}k^{\alpha-1} =$$
$$k^{\alpha-2}\left[(\alpha-1)\left(k^{\alpha-\beta-1}-1\right) - (\alpha-\beta-1)k^{\alpha-\beta-1}\right] =$$
$$k^{\alpha-2}\left[\beta k^{\alpha-\beta-1} - (\alpha-1)\right]$$

The sign of the derivative depends only on the sign of the expression in the brackets. The derivative equals to zero when $k = r_o = ((\alpha-1)/\beta)^{1/(\alpha-\beta-1)}$. If $k < r_0$ then the derivative is negative (the cost decreases) because of $\alpha - \beta - 1 > 0$. If $k > r_0$ then the derivative is positive (the cost increases). Thus, r_0 is minimal point. If r_0 is not an integer then one of the nearest two integers is minimal point. ∎

References

BOLTON, P.; DEWATRIPONT, M. (1994) The firm as a communication network. *Quarterly Journal of Economics*, CIX, p. 809–839.

CALVO, G.; WELLISZ, S. (1978) Supervision, loss of control and the optimal size of the firm. *Journal of Political Economy*, 86, p. 943–952.

CALVO, G.; WELLISZ, S. (1979) Hierarchy, ability and income distribution. *Journal of Political Economy*, 87, p. 991–1010.

DAVIES, G.; SMITH, M.; TWIGGER, W. (1991) Leading people: a model of choice and fate for leadership development. *Leadership & Organization Development*, 12, p. 7–11.

GOUBKO, M. (2002) Structure of the optimal organization of continuum workers. *Automation and Remote Control*, 63, p. 1966–1979.

GROSSMAN, S.; HART, O. (1982) Implicit contracts under asymmetric information. *Quarterly Journal of Economics*, 1, p. 110–124.

GROSSMAN, S.; HART, O. (1983) An analysis of the principal-agent problem. *Econometrica*, 51, p. 7–45.

HARDY, G. H.; LITTLEWOOD, J. E.; POLYA, G. (1934) *Inequalities*. Cambridge University Press, London.

HART, O. D.; HOLMSTROM, B. (1987) Theory of contracts. In *Advances in economic theory*. Proceedings of 5-th World Congress. Cambridge University Press, Cambridge, p. 71–155.

JAGO, A. G.; VROOM, V. H. (1975) Perceptions of leadership style: superior and subordinate descriptions of decision-making behavior. In: Hunt, J. G.; Larson, L. L. (Eds.). *Leadership Frontiers*. Southern Illinois University Press, Carbondale, p. 103–120.

KEREN, M.; LEVHARI, D. (1983) The internal organization of the firm and the shape of average costs. *Bell Journal of Economics*, 14, p. 474–486.

KEREN, M.; LEVHARI, D. (1979) The optimal span of control in a pure hierarchy. *Management Science*, 25, p. 1162–1172.

MARSCHAK, T. A.; RADNER, R. (1972) *Economic theory of teams.* Yale University Press, New Haven, CT.

MASKIN, E.; QIAN, Y.; XU, C. (2000) Incentives, information and organizational form. *Review of Economic Studies*, 67, p. 359–378.

MELUMAD, D. N.; MOOKHERJEE, D.; REICHELSTEIN, S. (1995) Hierarchical decentralization of incentive contracts. *The Rand Journal of Economics*, 26, p. 654–672.

MINTZBERG, H. (1979) *The structuring of organizations*. Prentice-Hall, Englewood Cliffs, NJ.

MISHIN, S.; VORONIN, A. (2002a) A model of the optimal control of structural changes in an organizational system. *Automation and Remote Control*, 63, p. 1329–1342.

MISHIN, S.; VORONIN, A. (2002b) Algorithms of searching for an optimal structure of organizational system. *Automation and Remote Control*, 63, p. 803–814.

MISHIN, S.; VORONIN, A. (2003) *Optimal hierarchical structures*. Preprint, Institute of Control Sciences, Moscow, Russia.

MISHIN, S. (2004) Optimal incentive in multilevel hierarchical structures. *Automation and Remote Control*, 65, p. 768–789.

21. MISHIN, S. (2005) *Optimal organizational hierarchies in firms*. Preprint, Institute of Control Sciences, Moscow, Russia.

22. PETERS, T. (1987) *Thriving on chaos.* Knopf, New York.

23. RADNER, R. (1992) Hierarchy: The economics of managing. *Journal of Economic Literature*, 30, p. 1382–1415.

24. RADNER, R. (1993) The organization of decentralized information processing. *Econometrica*, 61, p. 1109–1146.

25. SIMON, H. A. (1957) The compensation of executives. *Sociometry*, 20, p. 32–35.

26. VAN ZANDT, T. (1995) Continuous approximation in the study of hierarchies. *The Rand Journal of Economics*, 26, p. 575–590.

27. VAN ZANDT, T. (1996) *Organizations with an endogenous number of information processing agents*. Organizations with Incomplete Information, Cambridge University Press, Cambridge.

28. WILLIAMSON, O. (1967) Hierarchical control and optimal firm size. *Journal of Political Economy*, 75, p. 123–138.

29. WILLIAMSON, O. (1975) *Markets and hierarchies*. Free Press, New York.

30. QIAN, Y. (1994) Incentives and loss of control in an optimal hierarchy. *The Review of Economic Studies*, 61, p. 527–544.

31. QIAN, Y.; ROLAND, G.; XU, C. (1997) *Coordinating changes in M-form and U-form organizations*. Mimeo, Stanford University.